

Available online at www.sciencedirect.com



Control Engineering Practice 00 (2016) 1-14



Optimal Pressure Sensor Placement and Assessment for Leak Location¹ Using a Relaxed Isolation Index: Application to the Barcelona Water Network[☆]

Miquel À. Cugueró-Escofet, Vicenç Puig, Joseba Quevedo^a

^aSupervision, Safety and Automatic Control Research Center (CS2AC), Polytechnic University of Catalonia (UPC), Terrassa Campus, Gaia Research Bldg., Rambla Sant Nebridi, 22. 08222 Terrassa, Barcelona, Spain (e-mail: {miquel.angel.cuguero,vicenc.puig,joseba.quevedo}@ upc.edu).

Abstract

Water distribution networks are large complex systems that are affected by leaks, which often entail high costs and may severely jeopardise the overall water distribution performance. Successful leak location¹ is paramount in order to minimize the impact of these leaks when occurring. Sensor placement is a key issue in the leak location² process, since the overall performance and success of this process highly depends on the choice of the sensors gathering data from the network. Common problems when isolating leaks in large scale highly-gridded real water distribution networks include leak mislabelling and large location areas obtention due to similarity of leak effect in the measurements, which may be caused by topological issues and led to incomplete coverage of the whole network. The sensor placement strategy may minimize these undesired effects by setting the sensor placement optimisation problem with the appropriate assumptions (e.g. geographically cluster alike leak behaviors) and taking into account real aspects of the practical application such as the acceptable leak location distance. In this paper, a sensor placement methodology considering these aspects and a general sensor distribution assessment method for leak diagnosis in water distribution systems is presented and exemplified with a small illustrative case study. Finally, the proposed method is applied to two real District Metered Areas (DMAs) located within the Barcelona water distribution network.

© 2011 Published by Elsevier Ltd.

Keywords: Sensor placement, fault detection and isolation, leak location³, correlation coefficient, water distribution networks.

1 1. Introduction

An issue of great concern in water drinking networks is the existence of leaks at the distribution stage, highly related with water resource savings and management costs. The traditional approach to leak control is a passive one, whereby the leak is repaired when it becomes visible. Recently developed acoustic instruments also allow non-visible

⁵ leak location [1], but their application in large-scale water networks is very expensive and time-consuming. A viable

^AThis work has been partially funded by the Spanish Ministry of Science and Technology through the Project ECOCIS (Ref. DPI2013-48243-C2-1-R) and Project HARCRICS (Ref. DPI2014-58104-R), and by EFFINET grant FP7-ICT-2012-318556 of the European Commission.

¹R4-C4

¹R4-C4

²R4-C4 ³R4-C4

^{*} K4-C4

2

solution to this problem is to divide the network into District Metered Areas (DMAs), where the *flow* and the *pressure* 6 are measured [2, 3], and to maintain a permanent leak control-system. Leaks in fact increase the flow and decrease the pressure measurements at the DMA inputs. Several empirical studies propose mathematical models to describe 8 the leak flow with respect to the pressure at the leak location [4, 5]. Best practice in the analysis of DMA flows 9 consists in the estimation of the leak when the flow is minimum. This typically occurs at night time, when customers' 10 demand is low and the leak component is at its highest percentage over the flow [3]. Therefore, an accepted approach 11 by the practitioners² is to monitor the DMA or groups of DMAs in order to detect and repair the leaks occurring by 12 analyzing the minimum night flow, and also to employ techniques to estimate the corresponding leak magnitude [3]. 13 However, leak detection may not be easy to perform, since unpredictable variations in the customers' demand and 14 measurements noise may occur, as well as long-term trends and seasonal effects.³ 15

Several works in the literature have addressed the leak location problem in DMAs. In [6], a review of transient-16 based leak detection methods is summarized. In the seminal work [7], a model-based leak detection and location⁴ is 17 solved by means of a least-squares estimation problem. The latter problem is, however, not easy to solve when con-18 sidering the non-linear models involved. Alternatively, a method based on pressure measurements and leak sensitivity 19 analysis is proposed in [8], where a set of residuals (generated as the difference between pressure measurements pro-20 vided by several sensors installed within the DMA and their estimations by the network hydraulic model) is analysed 21 considering a certain threshold which takes into account practical factors e.g. the model uncertainty and the mea-22 surement noise. This approach shows satisfactory results under ideal conditions, but its performance degrades when 23 considering nodal demand uncertainty and measurement noise. This technique is improved in [9], where an extended 24 time horizon analysis is considered and a comparison of the performance using different metrics is performed. 25

The performance of the leak location⁵ approach is highly dependent on the sensor number and placement within 26 the DMA. Hence, the sensor placement strategy is a key issue to consider in the overall process. There is an important 27 trade-off between the number of sensors and the subsequent cost which prevents the use of a high number of sensors 28 for leak location purposes. Consequently, this number should be optimised at the sensor placement stage in order to 29 produce the highest possible benefit, that is, maximize the leak location performance at the minimum cost. According 30 to these constraints, the sensors considered here are pressure sensors since they are a cheaper alternative to flow 31 meters for the company managing the network, but the methodology presented might also be applied using different 32 sensor setups if required e.g. combining pressure and flow meters as in [10] or chlorine meters for water quality fault 33 diagnosis. Hence, the methodology may be arranged with minor modifications to different fault diagnosis purposes 34 and schemes.⁶ 35

Regarding sensor placement for fault detection and isolation (FDI) purposes, several works may be found in the 36 literature concerning this subject. Some approaches consider the study of structural matrices in order to locate sensors 37 based on isolability criteria [11]. In [12], an optimal set of sensors for model-based FDI is sought by means of an 38 optimisation method based on binary linear programming. These works are embraced in the general framework of FDI 39 of dynamic systems. However, they are not specially suited to consider the non-explicit non-linear set of equations 40 describing a water distribution network. Alternatively, several works treated the sensor placement problem when 41 applied to water distribution networks, most of them addressing the water contamination monitoring (e.g. [13, 14]), 42 43 where sensor placement is considered in a large water distribution network in order to detect malicious introduction of contaminants. Regarding leak location⁷, less contributions addressed the problem of sensor placement. This problem 44 is studied in [15], where an strategy based on the leak isolability maximization is considered to optimally place the 45 sensors based on the water network structural model, and in [8], where an optimal sensor placement is formulated 46 as an integer programming problem, similarly as presented here. Also, an entropy-based approach for efficient water 47 loss incident detection is introduced in [16]. 48

Furthermore, leak location⁸ in real water networks involves discrimination among a high number of possible leak locations (typically, the DMA nodes) leading to mislabel the right one due to the limited number of sensors available.

⁴R4-C4

- ⁵R4-C4
- ⁶R4-C7 ⁷R4-C4
- ⁸R4-C4

2

[/] Control Engineering Practice 00 (2016) 1-14

²R4-C6 ³R2-C3

⁵¹ However, in practice it is not needed to locate the leak at the exact place since final on-the-ground leak location⁹ ⁵² techniques (e.g. ground-penetrating radar, acoustic listening devices [17]) may locate leaks in a precise way starting ⁵³ from an area close to where the actual leak is occurring. Hence, this calls for a methodology of sensor placement ⁵⁴ trying to cluster similar leak behaviors geographically in order to minimize the number of installed sensors and locate

⁵⁵ the leak within a certain cluster distance precision.

Having all this into account, here a new approach for sensor placement focused on leak location¹⁰ in DMAs is 56 proposed, based on the method introduced in [18]. Alternatively to [8], the approach presented here does not binarize 57 the sensitivity matrix, hence the complete numerical precision of this matrix is used, leading to better leak location¹¹ 58 performance as pointed out in [18, 9]. This approach requires the reformulation of the optimisation problem introduced 59 in [8], since even both approaches are formulated as an integer optimization problem, isolability conditions considered 60 in the former do not apply here. The novel aspects of the methodology are, first, the use of the nodal distances together 61 with the sensitivity matrix at the sensor placement stage, in order to face the problem of mislabelling between leak 62 signatures, which occurs in case of DMAs with a high number of nodes and a low number of sensors. Also, the 63 sensitivities used to obtain the correlation between leak signatures are non-binary in order to avoid loss of information. The main aim is to reduce the effect of the leak mislabelling at the sensor placement stage, trying to geographically 65 cluster nodes with similar leak signature. Hence, the sensor distribution promoting this behaviour is selected, and 66 the rest are discarded. Work in the same direction has been done in the evolution between [19] and [18] at the leak 67 location stage. In contrast with [19], in [18] the binarisation of the leak signature has been avoided in order to prevent 68 the aforementioned loss of information. However, for a reduced number of sensors, the problem of mislabelling in 69 large DMAs is still present. This is the reason why in this work this problem is targeted in a previous stage, i.e. sensor 70 placement stage.¹² The second novel aspect presented here is the proposal of an assessment methodology using new 71 figures of merit in order to provide the goodness of a certain sensor set from the leak location point of view, which is 72 the next step after the sensors are placed. The assessment indices proposed assume that the leak location algorithm 73 will be based on the correlation between leak signatures, but are independent of the methodology used in order to 74 place the sensors, taking into account the intrinsic leak mislabelling that may occur in case of real DMAs with a low 75 ratio between available sensors and network nodes, which may jeopardise the leak location. To the knowledge of the 76 authors, the use of a general assessment in terms of potential number of isolated leaks is not present in the literature. 77 In [10], an assessment based on the isolation distance is presented in a real DMA, but this do not include the goodness 78 of the sensor distribution regarding the number of isolable leaks for the whole network.¹³¹⁴ Furthermore, the non-79 linear integer nature and the large dimension of the resulting optimisation problem calls for the use of an optimisation 80 tool able to handle a problem with such features. A well-suited approach to handle this problem is the one based 81 on Genetic Algorithms (GAs) [20, 21]. GAs are widely-used optimisation methods based on heuristics which mimic 82 the natural evolution, such as crossover, mutation or inheritance. This is performed by means of a fitness function 83 which selects the best individuals among different generations in order to provide an optimal solution to an specific 84 problem.¹⁵ The methodology presented is first illustrated in a small example and then evaluated in several DMAs, 85 located in the Barcelona network. 86

The paper is organized as follows: the leak location¹⁶ methodology used as the basis for this work is introduced in Section 2. The sensor placement methodology is presented in Section 3, and the isolability assessment used to evaluate the goodness of the sensor set proposed is introduced in Section 4. The application case studies, based on several DMAs, and the results obtained applying the methodology proposed are shown in Section 5. Finally, in Section 6, some concluding remarks and future work are given.

⁹R4-C4 ¹⁰R4-C4 ¹¹R4-C4 ¹²R2-C5 ¹³R2-C8 ¹⁴R2-C2, R2-C14, R5-C2 ¹⁵R4-C8 ¹⁶R4-C4 3

3

92 **2. Leak Location**¹⁷ **Problem**

The leak location¹⁸ problem may be separated in two different stages, which correspond to the sensor placement and the leak location¹⁹ itself, given a set of sensors. The leak location approach is summarised in this section, since it is the basis of the sensor placement algorithm formulation proposed in this work.

The leak location methodology considered here aims to locate leaks within a DMA by means of some pressure measurements gathered from the network and their estimations obtained by a network hydraulic model. For a given DMA with *N* demand nodes and *M* pressure sensors, the leak detection methodology relies on the computation of the residuals $\mathbf{r} = [r_1 \dots r_M]^T$, where $r_i \in \mathbf{r}$ is the difference between the pressure measurement p_i and its corresponding estimation \hat{p}_i obtained from a leakless simulation using the corresponding network hydraulic model as follows

$$r_i = p_i - \hat{p}_i, \quad i = 1, \dots, M$$
 (1)

¹⁰¹ having one residual per each available pressure measurement within the DMA.

On the other hand, the leak location²⁰ method relies on the study of the residual vector in (1) by means of sensitivity analysis, aiming to determine the effect of each particular leak on every available pressure sensor measurement at a certain time [7]

$$\mathbf{S} = \begin{pmatrix} s_{11} & \cdots & s_{1N} \\ \vdots & \ddots & \vdots \\ s_{M1} & \cdots & s_{MN} \end{pmatrix}$$
(2)

given $M \le N$ sensors within the network and N possible faults (assuming leaks only in nodes) with

$$s_{ij} = \frac{\hat{p}_{ij} - \hat{p}_i}{f_j}, \quad i = 1...N, j = 1...N$$
 (3)

where \hat{p}_i is the leakless scenario pressure estimation in node *i* and \hat{p}_{ij} is the pressure estimation in node *i* due to leak f_j scenario occurring in node *j*.

To obtain the sensitivity matrix \mathbf{S} , a leak scenario per each node is generated by numerical simulation using EPANET hydraulic solver [22], obtaining the sensitivity vector corresponding to one column of the sensitivity matrix \mathbf{S} as follows

$$\mathbf{s}_{j} = \begin{bmatrix} s_{1j} \\ \vdots \\ s_{Mj} \end{bmatrix}, \quad j = 1, \cdots, N$$
(4)

which is also known as leak signature²¹. Each simulated fault scenario is performed by setting a leak of magnitude f_i in 111 the j^{th} DMA network node. This procedure is repeated for all the N existing network nodes. Then, matching both the 112 residual vector in (1) and the sensitivity vectors in (4), leak location²² may be performed by checking which node has 113 the highest potential to present a leak. This analysis may be performed by using different metrics [23]. Here, a method 114 presented in [18, 10], based on the correlation between residual and sensitivity vectors, is considered. According to 115 the study in [9], this²³ method presents the best performance for leak location, even it should be remarked that the 116 sensor placement method presented in this paper could be applied with alternative leak location methods exploiting 117 sensitivity analysis. 118

¹⁷ R4-C4
¹⁸ R4-C4
¹⁹ R4-C4
²⁰ R4-C4
²¹ R4-C10
²² R4-C4
²³ R4-C9

The current metric considered here for leak location²⁴ is based on the correlation function given by the inner product of the regressor vector in (1) and the sensitivity vector in (4), for each particular fault in node j

$$\gamma_j = \frac{\mathbf{s}_j^T \mathbf{r}}{|\mathbf{s}_j||\mathbf{r}|}.$$
(5)

Then, the highest correlation determines the candidate leaky node k

$$\gamma_k = \max(\gamma_1, \cdots, \gamma_N). \tag{6}$$

The objective here is to develop a methodology to place a given number of sensors, M, within a DMA in order 122 to obtain a sensor set maximizing leak isolability under realistic conditions. In DMAs with a large number of nodes, 123 the sensitivity to different leaks occurring in different nodes may be very similar. This sensitivity similarity may lead 124 to confusion between different leaks when a low number of sensors is available and uncertainty in the measurements 125 and in the model is present, which is generally the actual situation. This situation may be solved e.g. by increasing 126 the number of sensors, in order to increase the dimension of the leak signature, or by selecting these measured nodes 127 with a methodology preventing sensitivity similarity for the different leak scenarios considered, as suggested by the 128 methodology presented here. This methodology relies on the leak location scheme presented in this section.²⁵ This 129 is the first stage of the twofold leak location²⁶ problem, where leaks are located given a set of sensors at the second 130 stage. The methodology to obtain this sensor set, based on the correlation-based method presented here, is introduced 131 in the next section. 132

3. Sensor Placement Methodology

¹³⁴ 3.1. Sensor Placement as an Optimisation Problem

The goal here is to place the best sensor set in order to locate the leak as precisely as possible within the consid-135 ered water network. The sensor distribution method is based on the system sensitivity matrix (2). As discussed in 136 the introduction, a former methodology is presented in [8], where the residuals are binarized by a certain threshold 137 value. In the approach presented here, the complete information of the residual is used in order to avoid data loss and 138 hence to increase leak discriminability [18]. Also, the sensor placement method uses a relaxed isolation index to better 139 handle some real-world effects affecting water network systems, such as system non-linearity, sensor measurements 140 resolution and model uncertainty (e.g. in the demands or network element parameters). These real-world effects cause 141 deviation between the modelled and the actual system behavior, which may lead to mislabel the latter, and the con-142 fusion between different leak scenarios (sensitivity vectors in (4)). However, if the confusion involves geographically 143 close behaviors, these undesired effects do not severely impact the final leak location²⁷. Hence, the optimal sensor 144 distribution takes into account that the leak location²⁸ distance may be relaxed and places the sensors accordingly 145 in order to geographically cluster leaks with similar signature (4). In order to perform the sensor placement of M146 sensors, let us define the binary decision vector that represents the selected sensors 147

$$\mathbf{x} = \begin{pmatrix} x_1 & \cdots & x_N \end{pmatrix}^T \tag{7}$$

where $x_i = 1$ if the pressure sensor in node *i* is installed and 0 otherwise. Defining

$$\mathbf{X}(\mathbf{x}) = diag(x_1, \cdots, x_N) \tag{8}$$

the corresponding sensitivity vectors can be represented as follows

$$\bar{\mathbf{s}}_{j}(\mathbf{x}) = \mathbf{X}(\mathbf{x})\mathbf{s}_{j}, \quad j = 1, \cdots, N \tag{9}$$

- ²⁴R4-C4 ²⁵R2-C15 ²⁶R4-C4 ²⁷R4-C4
- ²⁸R4-C4

where \mathbf{s}_j is the sensitivity matrix obtained when all the *N* sensors are available, i.e. M = N. Hence, the projection between two different leak signatures²⁹ *i* and *j* for a given subset of sensors **x** is introduced by their inner product as follows

$$\gamma_{ij}(\mathbf{x}) = \frac{\bar{\mathbf{s}}_i^T(\mathbf{x})\bar{\mathbf{s}}_j(\mathbf{x})}{|\bar{\mathbf{s}}_i(\mathbf{x})||\bar{\mathbf{s}}_j(\mathbf{x})|} = \frac{\mathbf{s}_i^T\mathbf{X}(\mathbf{x})\mathbf{s}_j}{|\mathbf{X}(\mathbf{x})\mathbf{s}_i||\mathbf{X}(\mathbf{x})\mathbf{s}_j|}, \quad i, j = 1, \cdots, N$$
(10)

where $\bar{\mathbf{s}}_i, \bar{\mathbf{s}}_j$ are vectors corresponding to two different fault signatures (columns) for each class (leak) in the sensitivity matrix (2) and γ_{ij} is a measure of similarity between these two classes. From (10), the projection matrix may be stated

155 as follows

$$\Gamma(\mathbf{x}) = \begin{pmatrix} \gamma_{11}(\mathbf{x}) & \cdots & \gamma_{1N}(\mathbf{x}) \\ \vdots & \ddots & \vdots \\ \gamma_{N1}(\mathbf{x}) & \cdots & \gamma_{NN}(\mathbf{x}) \end{pmatrix}.$$
 (11)

Regarding the nature of its elements, the matrix derived in (11) is called cross-correlation matrix. It may be noted that the latter is symmetric, so $\Gamma = \Gamma^{\top}$.

In order to evaluate the quality of a sensor allocation setup, $\rho_{ii}(\mathbf{x})$ is defined

$$\rho_{ij}(\mathbf{x}) = \left(\gamma_{ij}(\mathbf{x}) \left(1 - \frac{d_{ij}}{d_{\max}}\right)\right)^{d_c} + \left(\left(1 - \gamma_{ij}(\mathbf{x})\right) \frac{d_{ij}}{d_{\max}}\right)^{d_f}, \quad i, j = 1 \dots N$$
(12)

where γ_{ij} is the cross-correlation between leak i and leak j signature vectors, d_{ij} is the topological (pipe) distance 159 between leaky nodes i and j, d_{max} is the maximum pipe distance for the whole network and d_c and d_f are tuning 160 parameters related with the included high-correlated close leaks cluster and the excluded high-correlated distant leaks 161 for a given i-j leak pair, respectively.³⁰ This particular cost function aims to obtain the best sensor set in order to locate 162 the leaky node as precisely as possible, grouping leaks with similar correlation as geographically close as possible, 163 whilst discarding sensor sets promoting leaks with similar signature in distant locations. On the one hand, parameter 164 d_c is related with the isolation zone, i.e. the zone in which we allow leaks to be mislabelled. An acceptable value 165 for the perimeter is about 200 m in a real DMA, and is provided by the company managing the network. Inside this 166 perimeter, other on-the-ground techniques (e.g. ground penetrating radar) are used for finer isolation. Hence, d_c is 167 selected accordingly, i.e. making the cost function decrease its value when distance to most correlated leak is below 168 the selected perimeter, also taking into account the corresponding DMA d_{max} . On the other hand, parameter d_f is 169 related with the exclusion zone, i.e. the zone where leaks should not be mislabelled with the leaks in the isolation 170 zone. Hence, d_f is selected such that outside the inclusion zone leak mislabelling is penalized, i.e. the cost function 171 decreases its value when distance is above the selected perimeter and correlation with the potential leak occurring 172 increases. The values of d_c and d_f parameters should be selected such that (12) range from zero to one for each 173 *i-j* leak pair. These parameters may also be used to adjust the target of the optimisation. The bigger d_c , d_f , the 174 narrower the related zone. Generally d_c is chosen bigger than d_f so the slope of the exclusion term is lower, since the 175 exclusion zone embraces all the nodes in the network outside the inclusion perimeter. However, if one wants to get 176 focused on the leak isolation zone, d_f may be chosen arbitrarily high in order to penalise arbitrarily distant nodes.³¹ 177

¹⁷⁸ Considering (12), the sensor placement may be stated as an optimisation problem, with the following cost function

²⁹R4-C10

³⁰R2-C16, R4-C11

³¹R2-C12, R2-C13, R4-C11

/ Control Engineering Practice 00 (2016) 1-14

$$\rho(\mathbf{x}) = 1 - \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \rho_{ij}(\mathbf{x}).$$
(13)

7

As shown in (13), (12) is obtained for all the N^2 *i-j* node pairs and normalized, so (13) range from zero to one.³² Then, the optimisation problem may be formulated as follows

minimize
$$\rho(\mathbf{x})$$

subject to $\sum_{i=1}^{N} \mathbf{x}_{i} = M$ (14)

where $\rho(\mathbf{x})$ is to be optimised over the full N sensors set available, and M is a predefined restriction on the number of

¹⁸² sensors to install. The cost function in (13) for a single i-j leak pair is depicted in Figure 1 for illustrative purposes.³³

¹⁸³ The criterion to select the parameters d_c and d_f may be illustrated with Figure 1b, for a DMA with $d_{\text{max}} = 1000$ m.

Parameter d_c is selected so γ_{ij} starts decreasing at a normalised distance $d_{ij}/d_{max} = 0.2$, corresponding to a distance

of 200 m. Similar criterion is applied for the selection of d_f , related with the leak exclusion area.³⁴ Hence, the use of

this cost function aims to achieve a sensor distribution obtaining high-correlation/low-distance (first term in (12)) and

¹⁸⁷ low-correlation/high-distance (second term in (12)) leak scenario combinations.

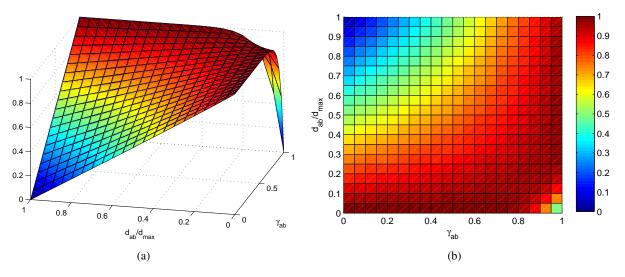


Figure 1: Cost function for a single i - j pair

The sensor placement optimisation problem (14) is solved using GA, which is a suitable approach for large-scale binary non-linear problems as the one considered here [24]. Further details on the GA parameters utilised to solve this particular problem are given in Section 5.2^{35} .

4. Isolability Assessment³⁶

¹⁹² In order to assess the fault isolability capabilities of a fault isolation method considering a particular set of mea-¹⁹³ surement points and a given topology, a metric based on the confusion matrix is used [25]. The confusion matrix is

³²R4-C13 ³³R2-C6, R4-C12 ³⁴R2-C12 ³⁵R2-C17 ³⁶R2-C18

a specific table layout which allows the visualisation of the performance achieved by a certain fault diagnosis layout, 194 i.e. a certain sensor set and its corresponding sensitivity model. Each column of this matrix represents instances in a 195 predicted class/fault, whilst each row stands for instances in an actual class/fault. The name stems from the fact that 196 this representation allows to check when the fault diagnosis method is confusing two different classes, commonly by 197 mislabelling one as another. A confusion matrix displays the number of correct and incorrect predictions made by the 198 fault isolation model compared with the actual class occurring in the test data. Here, a variation of the confusion ma-199 trix is presented in (15) in order to show the mislabelling between different leaks by comparing the predicted classes 200 against themselves37 201

$$C = \begin{pmatrix} \kappa_{11} & \cdots & \kappa_{1N} \\ \vdots & \ddots & \vdots \\ \kappa_{N1} & \cdots & \kappa_{NN} \end{pmatrix}$$
(15)

where $\kappa_{ij} \in \{0, 1\}$ for i, j = 1...N. Matrix in (15) shows how the fault isolation model obtained by a certain sensor set is mislabelling different faults between two different nodes *i* and *j*, which could be confused according to the considered metric. The values of κ_{ij} depend on each particular isolation criterion used. Here, a criterion based on the cross-correlation (11) is used to obtain the maximum correlation for each actual fault

$$\gamma_{ij_{\max}} = \max_{j \in 1...N} \gamma_{ij}, \quad i = 1...N$$
(16)

²⁰⁶ being κ_{ij} as follows,

$$\kappa_{ij} = \begin{cases} 1 & \gamma_{ij} = \gamma_{ij_{\text{max}}} \\ 0 & \text{otherwise} \end{cases}, \quad i, j = 1 \dots N.$$
(17)

Hence, the matrix (15) is called confusion cross-correlation matrix here. In order to provide less conservative isolation results while still realistic and well suited to the optimisation criterion stated in $(14)^{38}$, the matrix of pipe distances among nodes of the network may be presented

$$D = \begin{pmatrix} d_{11} & \cdots & d_{1N} \\ \vdots & \ddots & \vdots \\ d_{N1} & \cdots & d_{NN} \end{pmatrix}$$
(18)

and the isolation condition in (17) may be relaxed by a certain fault isolation cluster distance d_{cluster} as follows

$$\kappa_{ij} = \begin{cases} 1 & \max \mathbf{d}_{ij_{\max}} < d_{\text{cluster}} & \text{and} & d_{ij} < \max \mathbf{d}_{ij_{\max}} \\ 0 & \text{otherwise} \end{cases}, \quad i, j = 1 \dots N$$
(19)

where $\mathbf{d}_{ij_{\text{max}}}$ is the distance between the actual faulty node *i* and the node (or nodes) with highest correlation $\gamma_{ij_{\text{max}}}$

(i.e. predicted faulty nodes), and d_{cluster} is the maximum allowed distance between the actual faulty node *i* and the predicted faulty nodes, in order to consider the leak in *i* is well isolated. When several predicted faulty nodes are obtained, the worst case (i.e. max $\mathbf{d}_{ij_{\text{max}}}$) is considered.³⁹

The number of correctly isolated faults is given by the isolation index⁴⁰ 40

$$\zeta = \operatorname{tr}(C) \tag{20}$$

³⁷R2-C19 ³⁸R4-C14 ³⁹R4-C15, R2-C20 ⁴⁰R4-C17

so the correct isolated faults are those which are assigned to its own class and not to any other possible fault occurring 216 in the system. The best isolation index⁴¹ (ζ_{best}) for a given $d_{cluster}$ is obtained when sensors in all nodes are available 217 i.e. when M = N, which states a topological limit 218

$$0 \le \zeta_{opt} \le \zeta_{best} \le N \tag{21}$$

where ζ_{opt} is the isolation index obtained with the corresponding optimal sensor placement, for a given $d_{cluster}$. Let 219 us also define a particular ζ_{opt} and ζ_{best} considering (17), i.e. ζ_{opt_0} and ζ_{best_0} , respectively. Then, a more general 220 topological limit which does not depend on the distance between nodes may be given by 221

$$0 \le \operatorname{rank} \mathbf{S} \le \zeta_{opt_0} \le \zeta_{best_0} \le N \tag{22}$$

where S is the sensitivity matrix obtained when all the N sensors are available, i.e. M = N. Relation in (22) is 222 meaningful since ζ_{best_0} computation may be infeasible for DMAs with a high number of nodes N. Then, rank S may 223 provide a useful computationally efficient approximation, specially when this magnitude is close to the DMA number 224 of nodes N. It must be noted that, since matrix S is affected by the pressure sensor resolution, confusion between 225 leaks may be induced (e.g. linear dependency between columns of \mathbf{S}) as the DMA size increases.⁴² 226

It may also be noted that the ratio $\phi_{best} = \frac{\zeta_{best}}{\zeta_{best_0}} \ge 1$ suggests the benefit obtained by the geographic relaxation when 227 all the sensors are available (the bigger the better), whilst the ratio $\phi_{opt} = \frac{\zeta_{opt}}{\zeta_{opt_0}} \ge 1$ suggests the geographical relaxation benefit for the sensor subset considered. This benefit may be also obtained from an extra coverage percentage over 228

229

 ζ_{best} as follows 230

$$\delta = \frac{\zeta_{opt} - \zeta_{opt_0}}{\zeta_{best}} 100 \tag{23}$$

where δ is the percentage of extra coverage over ζ_{best} obtained when geographically relaxing the assessment. 231

5. Application Examples: Hanoi and Barcelona Drinking Water Networks 232

5.1. Description 233

Several DMAs of different level of complexity are used here in order to show the performance of the method 234 presented. First, a reduced DMA is considered to illustrate the method. The Hanoi DMA, an existing benchmark 235 network widely used in the literature (see, e.g. [26]), is considered for this purpose (Figure 2). This DMA has 31 236 nodes and 34 links, and delivers water to the end consumers by means of a single input point. Also, two different 237 DMAs located in the Barcelona area, with higher nodal density, are used as case studies (Figure 3).⁴³ On the one 238 hand, the Canyars DMA (Figure 4) is located at the pressure level 80 within the Barcelona water supply network. 239 This DMA has 694 nodes and 719 links, and delivers water to the end consumers by means of a single input point. On 240 the other hand, the Castelldefels Platja DMA (Figure 5) is located at the pressure level 50 within the Barcelona water 241 supply network. This DMA has 4952 nodes and 5116 links, and covers an area of 606 ha. The DMA has two inputs 242 (Ferrocarril and Pi Tort) delivering water to the end consumers. The current DMA size motivated its reduction to an 243 equivalent hydraulic model of 2828 nodes using skeletonization techniques [27], more suitable for high demanding 244 computation algorithms involved here. 245

In order to simulate these DMAs isolated from the water supply network, the boundary conditions (i.e. pressure 246 and flow measurements from the network) are fixed. Generally, pressure is fixed using a reservoir and the overall 247 demand is obtained as the sum of the inflow distributed through the DMA using a demand pattern model. The total 248 inflow is distributed using a constant coefficient (base demand) in each consumption node. Hence, all the consump-249 tions are assumed to share the same profile, whilst the billing information is used to determine the base demand of 250 each particular consumption. A good estimation of the demand model is paramount for the real case application. 251

⁴¹R4-C17 ⁴²R4-C18 43R2-C10

252 5.2. Results

In this section, the results achieved applying the sensor placement methodology described in Section 3 are presented. The sensors considered here are pressure sensors which may be installed in any node of the network. The maximum isolation distance $d_{cluster}$, which is a parameter given by the company managing the network, is assumed of 200 m for the DMAs in Section 5.2.2 and Section 5.2.3, whilst is assumed of 2000 m for the illustrative DMA in Section 5.2.1 due to its particularly low nodal density. For distances below $d_{cluster}$, there exist alternative more precise methods to isolate the leak e.g. ground penetrating radar. Scenarios have been generated using EPANET hydraulic simulation software, as introduced in Section 2.

Regarding GA parameters, an initial population of 100 random sensor sets, including the potential sensors to be 260 used, is employed to seed the GA algorithm. At this stage, already installed DMA sensors may be included to seed 26 the GA. The number of individuals in each generation is set to 100, the maximum number of generations allowed 262 is set to 30, the termination tolerance on the fitness function value is set to 1×10^{-6} and the number of generations 263 over which cumulative change in fitness function value is less than the termination tolerance (stall generations limit) 264 is set to eight. Since the optimum obtained by the GA is not global, consecutive GA optimisations are conducted until 265 fitness function value do not improve between two overall optimisations, aiming to achieve the best possible solution. 266 The selection of these parameters takes into account that the optimisation is dealing with real high dimension DMAs 267 and the problem may be computationally intensive. In order to face such computational issues, the use of local parallel 268 computing is used when multiple labs are available in the host PC, in order to increase computation power. The host 269 PC implemented Intel[®] CoreTM i7 Quad-Core processors and 8 GB of 1600 MHz Dual Channel DDR3 memory, 270 which allowed the use of such technique.⁴⁴ 271

272 5.2.1. Hanoi DMA

The sensor placement results obtained considering Hanoi DMA (Figure 2) are depicted in Figures 7a to 7d. The 273 sensitivity matrix \mathbf{S} is obtained for a 24 h scenario using an emitter coefficient (i.e. discharge coefficient for emitter 27 placed at junction, representing the flow in liters per second (LPS)⁴⁵ occurring at a pressure drop of 1 psi [22]) of 275 5 LPS/psi^{0.5}. The sensitivity S is concatenated for the 24 hours available, leading to a dimension of 744×31 . The 276 distance used here is the topological distance among nodes, i.e. minimum pipe distance between these elements. This 277 network has a low density of nodes per squared meter, being the minimum distance among the closest nodes of 484 m. 27 Hence, d_{cluster} should be increased in comparison to a regular network, in order to provide realistic results. For this 279 particular network, the maximum number of isolable faults considering all the sensors available (ζ_{best_0}) is 31, and the 280 maximum number of isolable faults considering all the sensors available and $d_{\text{cluster}} = 2000 \text{ m} (\zeta_{\text{best}})$ is 28 (90 % of 281 N = 31 nodes forming the network), respectively. According to the specified d_{cluster} and $d_{\text{max}} = 16426$ m, the cost 282 function parameters have been chosen of $d_c = 5.36$ and $d_f = 0.57$.⁴⁶ The evolution of the GA optimisation for each 283 sensor distribution is depicted in Figures 6a to 6d. In the lower row, the latter figures show the evolution of the average 284 distance between individuals among generations in the bottom-left subplot, and the fitness of each individual in the 285 last generation in the bottom-right subplot. In the upper row, the evolution of the best and mean fitness value per 286 generation is depicted in the upper-left subplot, and the GA stopping criteria is depicted in the upper-right subplot. 287 These include the generations limit (30) i.e. the maximum number of generations per optimisation, the time limit 288 (unspecified) i.e. the maximum time in seconds whilst the GA runs before stopping, the stall generations limit (eight) 289 i.e. the number of maximum consecutive generations without improving the average relative change in the best fitness 290 function over a given function tolerance (1×10^{-6}) and the stall time limit (unspecified), i.e. the time interval in 291 seconds after the GA stops if no improvement is obtained in the best fitness value. Time constraints have not been 292 specified since they were not critical parameters in the optimisations. As it is shown in the Figures 6a to 6d, the 293 stopping criteria met in all the optimisations for this particular case is the stall generation limit. Isolation assessment 294 results concerning sensor distribution for different number of sensors (from two to five) are detailed in Table 1. It 295 may be observed how the results obtained between four and five sensors do not improve in terms of ζ_{opt} , even a 296 better ρ is achieved for five sensors at the optimisation stage. In this case, the benefit of installing extra sensors may 297

⁴⁴R1-C1, R2-C12, R2-C17 ⁴⁵R4-C19

⁴⁶R2-C12

Number of sensors	2	3	4	5
ζ _{opt}	14	20	22	22
% of <i>N</i>	45	64.5	70.97	70.97
% of ζ_{best}	50	71.43	78.57	78.57
% of ζ_{best_0}	45	64.52	70.96	70.96
ζ_{opt_0}	30	31	31	31
ρ	0.6804	0.6586	0.6445	0.6426

Table 1: Isolation assessment results, Hanoi DMA

²⁹⁸ obtain reduced isolation clusters, but still bigger than d_{cluster} . Hence, the optimal sensor distribution is obtained for ²⁹⁹ four sensors (Figure 7c) since is the one achieving best ζ_{opt} with the minimum number of sensors. For the particular ³⁰⁰ layout of this DMA, which is geographically large ($d_{\text{max}} = 16426$ m) but with low nodal density (N = 31 nodes, with

minimum distance among closest nodes of 484 m), the geographical relaxation is not providing any particular benefit

 $_{302}$ (δ is negative for all the distributions considered). However, the methodology presented here is useful when there

exists leak signature confusion, which is not the case in this network ($\zeta_{opt} = N$ for almost all the sensor distributions

considered). Hence, a network of this characteristics is useful for illustrative purposes, but it is not a target network

³⁰⁵ for the proposed methodology, more intended to be used in larger DMAs found in real water networks as the ones

³⁰⁶ presented in the following sections.⁴⁷

307 5.2.2. Canyars DMA

The sensor placement results obtained when considering Canyars DMA (Figure 4) are depicted in Figures 8a to 30 8c. The sensitivity matrix **S** is obtained for a fixed leak of 6 LPS, in an hourly sampled scenario comprised between 309 24/02/2014 9h and 25/02/2014 9h. Thus, the sensitivity S is concatenated for the 24 hours available leading to a 310 dimension of 16656×694 . Also, the information in this matrix considers sensor resolution of 0.1 m in order to 311 take into account current technological constraints in the simulated scenario. The distance used here is the topological 312 distance among nodes, i.e. minimum pipe distance between these elements. For this particular network, the maximum 313 number of isolable faults considering all the sensors available (ζ_{best_0}) is 399, and the maximum number of isolable 314 faults considering all the sensors available and d_{cluster} (ζ_{best}) is 398 (57 % of N = 694 nodes forming the network). 315 According to the specified $d_{\text{cluster}} = 200 \text{ m}$ and $d_{\text{max}} = 888 \text{ m}$, the cost function parameters have been chosen of 316 $d_c = 8.31$ and $d_f = 1.04$, respectively.⁴⁸ 317

Isolation assessment results concerning sensor distribution for different number of sensors (from two to four) are 318 detailed in Table 2. It may be observed how the relaxation by d_{cluster} does not have much effect when having all the 319 sensors available (i.e. $\phi_{best} \approx 1$), but it does for a limited sensor set (see Table 2) e.g. for two sensors, with $\zeta_{opt} = 267$ 320 and $\zeta_{opt_0} = 116$, $\delta = 38$ % extra coverage over ζ_{best} is achieved when geographically relaxing the assessment. It may 321 also be observed how the results obtained between three and four sensors do not improve in terms of ζ_{opt} , even a better 322 ρ is achieved for four sensors at the optimisation stage. In this case, the benefit of installing extra sensors may obtain 323 reduced isolation clusters, but still bigger than d_{cluster} . Hence, since the coverage of the network is high (97 % of ζ_{best}), 324 the optimal sensor distribution is obtained for three sensors (Figure 8b) since is the one achieving best ζ_{opt} with the 325 minimum number of sensors. 326

The impact of sensors resolution is also worth to be noted. Although it does not have impact on the maximum number of isolable faults $\zeta_{\text{best}} = 398$ (hence, the maximum achievable coverage is not limited by the sensors resolution but by the topological network setup, when sufficient number of sensors are available), it does have impact on ζ_{opt} for different sensor setups (hence, for limited information gathered from the network, sensor resolution effect is noticeable). For example, considering five full-resolution sensors setup, almost complete coverage of the network is achieved ($\zeta_{\text{opt}} = 395$), against the 388 isolable faults achieved by the five limited-resolution (0.1 m) sensors setup counterpart.

⁴⁷AE-C1, R2-C10 ⁴⁸R2-C12

Table 2: Isolati	ion assessment	results, Ca	anyars DMA
------------------	----------------	-------------	------------

Number of sensors	2	3	4
ζopt	267	388	388
% of N	38	56	56
% of ζ_{best}	67	97	97
% of ζ_{best_0}	67	97	97
ζ_{opt_0}	116	242	245
ρ	0.7375	0.7342	0.7321

Table 3: Isolation results, Castelldefels Platja DMA

Number of sensors	4	5	6
ζopt	2649	2665	2665
% of <i>N</i>	93.67	94.24	94.24
% of ζ_{best}	93.8	94.37	94.37
% of rank S	93.67	94.24	94.24
ζ_{opt_0}	902	982	1067
ρ	0.5116	0.5086	0.5071

334 5.2.3. Castelldefels Platja DMA

The sensor placement results obtained considering Castelldefels Platja network (Figure 4) are depicted in Fig-335 ures 9a to 9c. The sensitivity matrix S is obtained for an emitter coefficient of 0.92 LPS/psi^{0.5}, in an hourly sampled 336 scenario comprised between 24/02/2014 9h and 25/02/2014 9h, so S is concatenated for the 24 hours available and is 337 of dimension 67872 x 2828. Also, the information in this matrix is truncated by sensor resolution (i.e. 0.1 m). The 338 distance used here is the topological distance among nodes, i.e. minimum pipe distance between these elements. For 339 this particular network, the computation of ζ_{best_0} is not possible due to computational issues related with network size 340 N, as introduced in Section 4. Alternatively, the rank of S is used, providing a maximum number of isolable faults 341 approximation considering all the sensors available, that is 2828. Since this value is close to N, it may be considered 342 a feasible approximation of the maximum number of isolable faults. Also, the maximum number of isolable faults 343 considering all the sensors available and d_{cluster} (ζ_{best}) is 2824 (the 99.9 % of N = 2828 nodes forming the network). 344 It may be observed how the relaxation by d_{cluster} does not either have much impact in this DMA when having all the 345 sensors available (i.e. $\phi_{best} \cong$ 1), but it does as in Canyars DMA when limited number of sensors are available (see 346 Table 3) e.g. for four sensors, with $\zeta_{opt} = 2649$ and $\zeta_{opt_0} = 902$, $\delta = 62\%$ extra coverage over ζ_{best} is achieved when 347 geographically relaxing the assessment. According to the specified $d_{\text{cluster}} = 200 \text{ m}$ and $d_{\text{max}} = 7222 \text{ m}$, the cost 348 function parameters have been chosen of $d_c = 6.41$ and $d_f = 0.46$, respectively.⁴⁹ 349

Isolation assessment results concerning sensor distribution for different number of sensors considered (from four to six) are detailed in Table 3. It may be seen how for five and six sensors, the number of isolable faults for the optimal sensor set (ζ_{opt}) equals 2665, so according to the criterion no advantage is obtained from the usage of this extra sensor. Hence, the number of suggested sensors for this network is five (Figure 9b), achieving a theoretical coverage of the 94.24 % of the total possible faults.

355 6. Conclusions

In this paper, a successful sensor placement and leak location⁵⁰ assessment methodologies are proposed in order to improve the performance of leak location⁵¹ in water distribution networks, which may have severe impact on maintenance costs and performance of the water distribution along DMAs. Common problems arising on the leak

⁴⁹R2-C12 ⁵⁰R4-C4

⁵¹R4-C4

diagnosis in large real water networks can be addressed at the sensor placement stage, e.g. leak discriminability and 359 large location⁵² areas, when taking into account real world leak diagnosis trade-offs related with geographic location⁵³ 360 precision. Hence, a general method of sensor placement is proposed, taking into account these trade-offs by clustering 361 similar leaks geographically within an acceptable location⁵⁴ area from the application point of view. The proposed 362 method achieved promising leak location⁵⁵ results, evaluated by an also proposed general assessment method for leak 363 diagnosis in water distribution systems, in a small illustrative DMA in Hanoi and two DMAs situated in the Barcelona 364 365 urban area. These results motivate the use of the proposed methodology in the actual and similar water networks. Further work involves the inclusion of the number of sensors to install as part of the optimisation problem, as well as 366 the consideration of uncertainty (e.g. in sensor measurements and demand model) in the sensor placement algorithm 367 to cope with more realistic assumptions. Also, the extension to multiple leak scenarios may be considered in future 368 steps of this work, by e.g. developing further methods in order to expand the sensitivity matrix accordingly, taking 369 into account that the selection of these new scenarios should be performed carefully in order to avoid computational 370 issues derived from handling high dimension matrices.⁵⁶ 371

372 Acknowledgement

This work has been partially funded by the Spanish Ministry of Science and Technology through the Project ECOCIS (Ref. DPI2013-48243-C2-1-R) and Project HARCRICS (Ref. DPI2014-58104-R), and by EFFINET grant FP7-ICT-2012-318556 of the European Commission. The authors also wish to thank the support received by the

³⁷⁶ Water Technological Center (CETAQUA) of the company managing the Barcelona water network (AGBAR).

377 **References**

- Y. Khulief, A. Khalifa, R. Mansour, M. Habib, Acoustic detection of leaks in water pipelines using measurements inside pipe, Journal of
 Pipeline Systems Engineering and Practice 3 (2) (2012) 47–54.
- M. Lambert, A. Simpson, J. Vitkovsky, X.-J. Wang, P. Lee, A review of leading-edge leak detection techniques for water distribution systems,
 in: 20th AWA Convention, Perth, Australia, 2003.
- [3] R. Puust, Z. Kapelan, D. A. Savic, T. Koppel, A review of methods for leakage management in pipe networks, Urban Water Journal 7 (1)
 (2010) 25–45.
- [4] A. Lambert, What do we know about pressure leakage relationships in distribution systems?, in: IWA Conference System Approach to
 leakage control and water distribution system management, no. May, Brno, Czech Republic, 2000, pp. 1–9.
- J. Thornton, A. Lambert, Progress in practical prediction of pressure: leakage, pressure: burst frequency and pressure: consumption relation ships, in: Leakage Conference Proceedings, Halifax, Canada, 2005, pp. 1–10.
- [6] A. F. Colombo, P. Lee, B. W. Karney, A selective literature review of transient-based leak detection methods, Journal of Hydro-environment Research (2009) 212–227.
- [7] R. S. Pudar, J. A. Liggett, Leaks in pipe networks, Journal of Hydraulic Engineering 118 (7) (1992) 1031–1046.
- [8] R. Pérez, V. Puig, J. Pascual, J. Quevedo, E. Landeros, A. Peralta, Methodology for leakage isolation using pressure sensitivity analysis in
 water distribution networks, Control Engineering Practice 19 (10) (2011) 1157 1167. doi:10.1016/j.conengprac.2011.06.004.
- [9] M. V. Casillas, L. Garza-Castañón, V. Puig, Model-based leak detection and location in water distribution networks considering an extended horizon analysis of pressure sensitivities, Journal of Hydroinformatics 16 (2014) 649–670.
- [10] R. Pérez, M. A. Cugueró, J. Cugueró, G. Sanz, Accuracy assessment of leak localisation method depending on available measurements,
 Procedia Engineering 70 (2014) 1304–1313. doi:10.1016/j.proeng.2014.02.144.
- [11] A. Yassine, S. Ploix, J. M. Flaus, A method for sensor placement taking into account diagnosability criteria, International Journal of Applied Mathematics and Computer Science 18 (4) (2008) 497–512. doi:10.2478/v10006-008-0044-5.
- [12] A. Rosich, R. Sarrate, F. Nejjari, Optimal sensor placement for FDI using binary integer linear programming, in: 20th International Workshop
 on Principles of Diagnosis, DX09, 2009, pp. 235–242.
- [13] A. Krause, J. Leskovec, C. Guestrin, J. Vanbriesen, C. Faloutsos, Efficient sensor placement optimization for securing large water distribution
 networks, Journal of Water Resources Planning and Management 134 (6) (2008) 516–526.
- [14] M. Aral, J. Guan, M. Maslia, Optimal design of sensor placement in water distribution networks, Journal of Water Resources Planning and
 Management 136 (1) (2010) 5–18.
- [15] R. Sarrate, F. Nejjari, A. Rosich, Sensor placement for fault diagnosis performance maximization in distribution networks, in: Control &
 Automation (MED), 2012 20th Mediterranean Conference on, 2012, pp. 110–115.

⁵²R4-C4

⁵³R4-C4

⁵⁴R4-C4

⁵⁵R4-C4

⁵⁶R2-C4, R4-C5

- [16] S. Christodoulou, A. Gagatsis, A. Xanthos, S. Kranioti, A. Agathokleous, M. Fragiadakis, Entropy-based sensor placement optimization for
 waterloss detection in water distribution networks, Water Resources Management 27 (13) (2013) 4443–4468.
- ⁴⁰⁹ [17] M. Farley, S. Trow, Losses in Water Distribution Networks, IWA publishing, London, 2003.
- [18] J. Quevedo, M. A. Cugueró, R. Pérez, F. Nejjari, V. Puig, J. M. Mirats, Leakage location in water distribution networks based on correlation
 measurement of pressure sensors, in: 8th IWA Symposium on System Analysis and Integrated Assessment (Watermatex), International Water
 Association (IWA), San Sebastián, 2011, pp. 290–297.
- [19] R. Pérez, V. Puig, J. Pascual, A. Peralta, E. Landeros, L. Jordanas, Pressure sensor distribution for leak detection in Barcelona water distribu tion network, Water Science & Technology: Water Supply 9 (6) (2009) 715 721. doi:10.2166/ws.2009.372.
- 415 [20] C. R. Reeves (Ed.), Modern Heuristic Techniques for Combinatorial Problems, Mcgraw Hill, UK, 1995.
- 416 [21] J. R. Koza, Survey of genetic algorithms and genetic programming, Proceedings of 1995 WESCON Conference (1995) 589–594.
- 417 [22] L. A. Rossman, EPANET 2 Users Manual, Environmental Protection Agency (EPA), U.S. (September 2000).
- 418 [23] L. Rokach, O. Maimon, Clustering methods, in: Data Mining and Knowledge Discovery Handbook, 2005, pp. 321–352.
- [24] K. Gallagher, M. Sambridge, Genetic algorithms: A powerful tool for large-scale nonlinear optimization problems, Computers & Geosciences
 20 (7-8) (1994) 1229–1236.
- 421 [25] T. Fawcett, An introduction to ROC analysis, Pattern Recognition Letters 27 (8) (2006) 861–874.
- 422 [26] M. V. Casillas-Ponce, L. E. Garza-Castañón, V. Puig-Cayuela, Model-based leak detection and location in water distribution networks considering an extended-horizon analysis of pressure sensitivities, Journal of Hydroinformatics 16 (3) (2014) 649–670. doi:10.2166/hydro.2013.019.
- 424 [27] T. Walski, D. Chase, D. Savic, W. Grayman, S. Beckwith, E. Koelle, Advanced Water Distribution Modeling and Management, Haestad 425 Press, 2003.

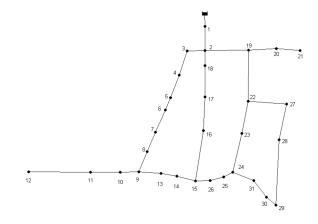


Figure 2: Hanoi DMA

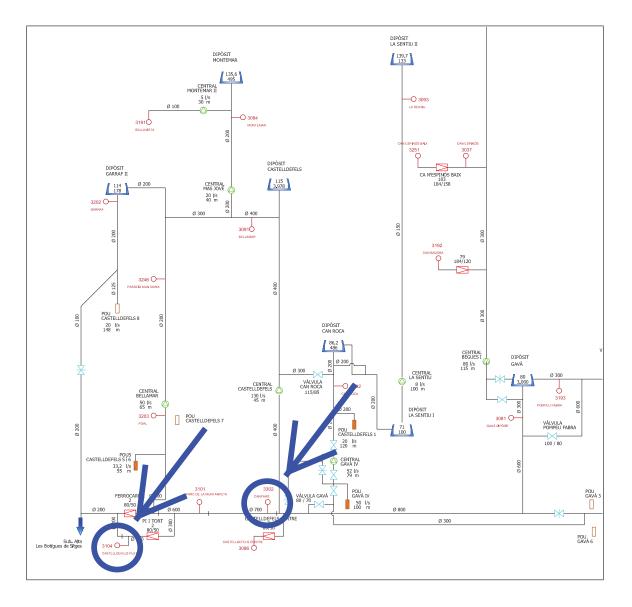


Figure 3: Barcelona Drinking Water Supply Network detail (arrows: Castelldefels Platja and Canyars DMAs, respectively)

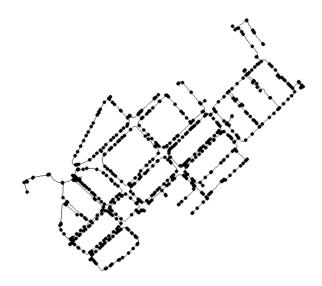


Figure 4: Canyars DMA



Figure 5: Castelldefels Platja DMA

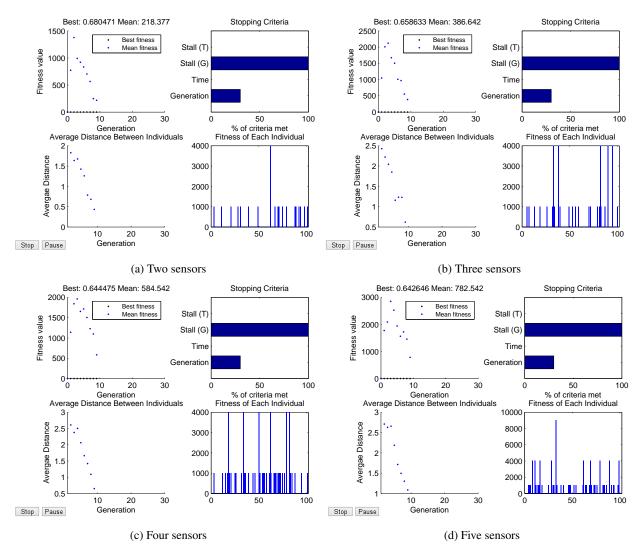


Figure 6: Genetic Algorithms optimisation evolution in Hanoi DMA

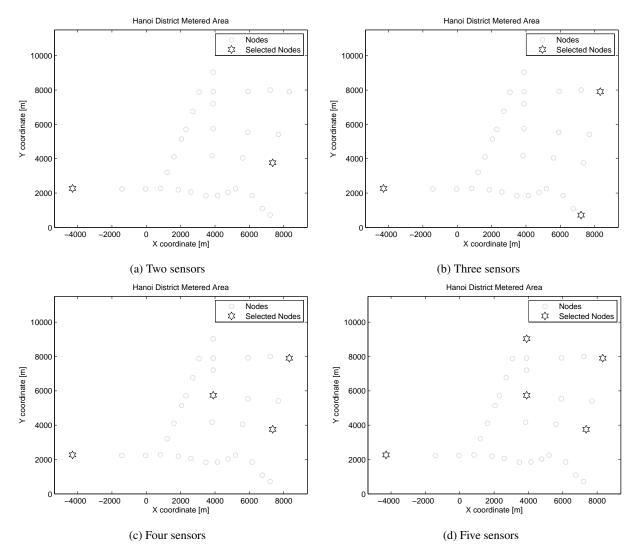


Figure 7: Sensor placement in Hanoi DMA

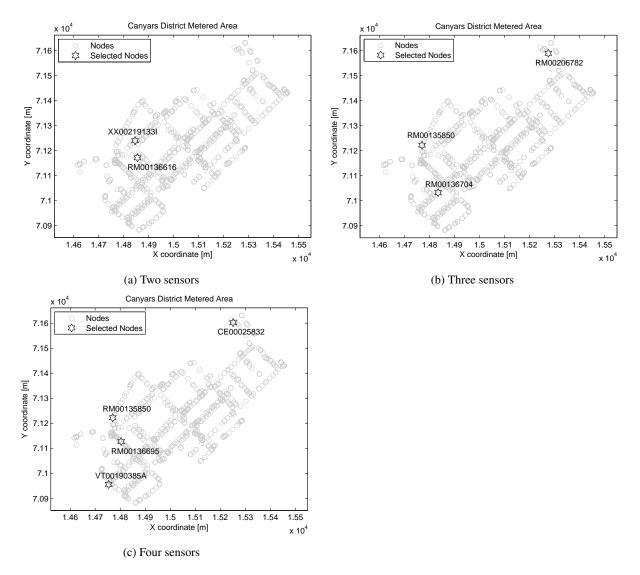


Figure 8: Sensor placement in Canyars DMA

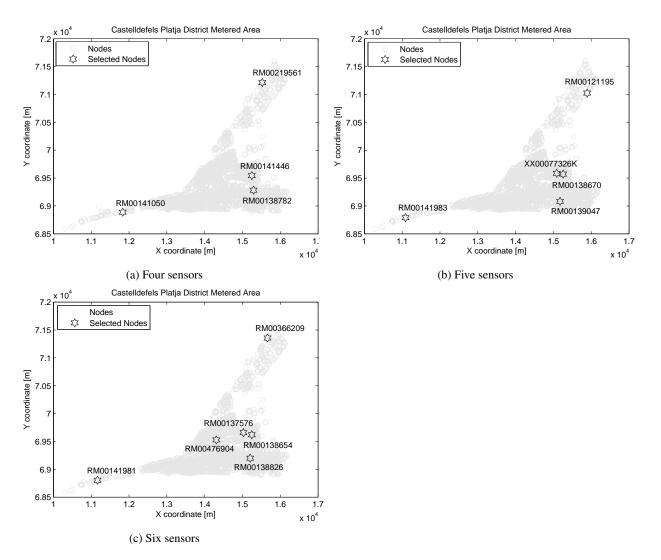


Figure 9: Sensor placement in Castelldefels Platja DMA

21