Control Engineering Practice 19 (2011) 1157–1167

Contents lists available at ScienceDirect

Control Engineering Practice

journal homepage: www.elsevier.com/locate/conengprac

Methodology for leakage isolation using pressure sensitivity analysis in water distribution networks

Ramon Pérez [a,*], Vicenç Puig [a,b], Josep Pascual [a], Joseba Quevedo [a], Edson Landeros [c], Antonio Peralta [d]

[a] Advanced Control Systems Group (SAC), Universitat Politècnica de Catalunya (UPC), Rambla Sant Nebridi, 10, 08222 Terrassa, Spain
[b] IRI Institut de Robòtica i Informàtica Industrial (CSIC-UPC), Spain
[c] CETAQUA Water Technological Center, Spain
[d] AGBAR Barcelona Water Company, Spain

ARTICLE INFO

Article history:
Received 6 September 2010
Accepted 2 June 2011
Available online 12 July 2011

Keywords:
Pressure sensitivity
Leakage localisation
Sensor placement

ABSTRACT

Leaks are present to some extent in all water-distribution systems. This paper proposes a leakage localisation method based on the pressure measurements and pressure sensitivity analysis of nodes in a network. The sensitivity analysis using analytical tools is not a trivial job in a real network because of the huge non-explicit non-linear systems of equations that describe its dynamics. Simulations of the network in the presence and the absence of leakage may provide an approximation of this sensitivity. This matrix is binarised using a threshold independent of the node. The binary matrix is assumed as a signature matrix for leakages. However, there is a trade-off between the resolution of the leakage isolation procedure and the number of available pressure sensors. In order to maximise the isolability with a reasonable number of sensors, an optimal sensor placement methodology, based on genetic algorithms, is also proposed. These methodologies have been applied to the Barcelona Network using PICCOLO simulator. The sensor placement and the leakage detection and localisation methodologies are applied to several district management areas (DMA) in simulation and in reality.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Water loss in distribution system networks is an issue of great concern for water utilities, strongly linked with operational costs and water resource savings. Continuous improvements in water loss management are applied and new technologies are developed to achieve higher levels of efficiency. Usually a leakage detection method in a District Metered Area (DMA) starts by analysing input flow data, such as minimum night flows and consumer metering data (Lambert, 1994; MacDonald, 2005). Once the water distribution district is identified to have a leakage, various techniques are used to locate the leakage for pipe replacement or repair. Methods for locating leaks range from ground-penetrating radar to acoustic listening devices or physical inspection (Colombo, Lee, & Karney, 2009; Farley & Trow, 2003). Some of these techniques require isolating and shutting down part of the system. The whole process could take weeks or months with a significant volume of water wasted. Techniques based on locating leaks from pressure monitoring devices allow a more effective and less costly search in situ.

This paper presents a model-based methodology to detect and localise leaks. It has been developed within a project carried out by Aguas Barcelona, Water Technological Centre CETaqua, and the Technical University of Catalonia (UPC). The objective of this project is to develop and apply an efficient system to detect and locate leaks in a water distribution network. It integrates methods and technologies available and in use by water companies, including DMA and flow/pressure sensor data, in conjunction with mathematical hydraulic models. The method is based on the analysis of pressure variations produced by leakage in the water distribution network (Pudar & Ligget, 1992). This technique differs from others in the literature, such as the reflection method (LRM) or the inverse transient analysis (ITA), since it is not based on the transient analysis of pressure waves (Ferrante & Brunone, 2003a, 2003b; Misiunas, Lambert, Simpson, & Olsson, 2005; Verde, Visairo, & Gentil, 2007). Alternatively, the leakage detection procedure is performed by comparing real pressure and flow data with their estimation using the simulation of the mathematical network model. Simulation of the network in presence and absence of leakage provides an approximation of pressure sensitivity of nodes in a network when a leak is present in a node. The approximation is used to generate a sensitivity matrix that is binarised using a threshold independent of the node. In order to successfully apply this methodology, the characterisation of district metered areas and consumers, considered a critical issue for a correct model calibration, should be also addressed but is not described in this paper (see, e.g. Perez, de las Heras, Aguilar,

* Corresponding author. Tel.: +34937398620; fax: +349373928.
E-mail address: ramon.perez@upc.edu (R. Pérez).

0967-0661/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.conengprac.2011.06.004

Pascual, & Peralta, 2009a, for further details). Another critical point is the data validation of DMA sensors that can be addressed as it is described for flowmeters in Quevedo et al. (2010). The paper also proposes a methodology for placing pressure sensors within a DMA that optimises leakage detection using a minimum number of sensors based on the approach proposed in Pérez et al. (2009b). Finally, the leakage detection methodology proposed will be tested with sensors installed in a DMA used as case study.

Section 2 reviews water distribution network modelling and presents the case study used to illustrate the proposed methodologies. Model-based fault detection and isolation techniques described in Section 3 are used for the leakage detection and location. Section 4 presents how the leak signature matrix is obtained from the pressure sensitivity matrix. Since the sensor placement is a critical issue for maximising discriminability, an algorithm is presented in Section 5. The signature matrix is generated for the set of sensors selected. This matrix has to be compared with the signature obtained comparing the model and the real measurements. From this comparison, the leakage is located in a set of possible nodes. This methodology is presented in Section 6 and is illustrated by simulation and real results. Finally, Section 7 summarises the conclusions.

2. Water distribution systems: plaça del diamant case study

A water distribution system consists of three major components: pumps, distribution storage, and distribution piping network. Most systems require pumps to supply lift to overcome differences in elevation, and energy losses caused by friction. Pipes may contain flow-control devices, such as regulating or pressure-reducing valves (Brdys & Ulanicki, 1994). The purpose of a distribution system is to supply the system's users with the amount of water demanded, under adequate pressure for various loading conditions. A loading condition is a spatial pattern of demands that defines the users' flow requirements.

2.1. Mathematical modelling

The governing laws for flow in pipe systems under steady conditions are conservation of mass and energy. The law of conservation of mass states that the rate of storage in a system is equal to the difference between the inflow to and outflow from the system. In pressurised water distribution networks, no storage can occur within the pipe network, although tank storage may change over time. Therefore, in a pipe, or a junction node, the inflow and the outflow must balance. For a junction node

$$\sum q_{in} - \sum q_{out} = q_{ext} \tag{1}$$

where q_{in} and q_{out} are the pipe flow rates into and out of the node and q_{ext} is the external demand or supply. Conservation of energy states that the difference in energy between two points is equal to the energy added to the flow in components between these points minus the frictional losses. An energy balance can be written for paths between the two end points of a single pipe, between two fixed graded nodes (a node for which the total energy is known, such as a tank) through a series of pipes, valves, and pumps, or around a loop that begins and ends at the same point. In a general form for any path

$$\sum_{i \in J_p} h_{P,j} - \sum_{i \in I_p} h_{L,i} = \Delta E \tag{2}$$

where $h_{L,i}$ is the headloss across component i along the path, $h_{P,j}$ is the head added by pump j, and ΔE is the difference in energy between the end points of the path. The primary network component is a pipe. The relationship between pipe flow (q) and energy loss caused by friction (h_L) in individual pipes can be represented by a number of equations, including the Darcy–Weisbach and Hazen–Williams equations. The general relationship is of the following form:

$$h_L = Kq^r \tag{3}$$

where K is a pipe coefficient that depends on the pipe's diameter, length, and material and r is an exponent in the range of 2.

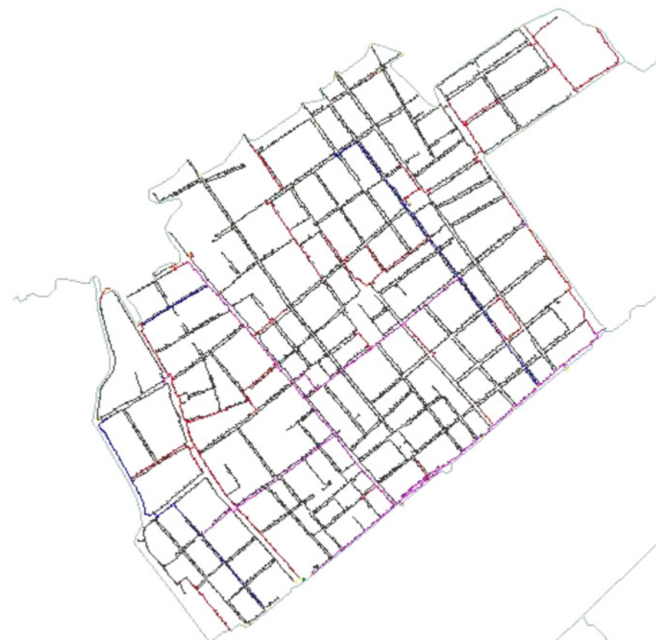


Fig. 1. Case study network: *Plaça del Diamant*.

2.2. Plaça del Diamant DMA case study

The case study used to illustrate the leak localisation methodology presented in this paper is based on Plaça del Diamant DMA at the Barcelona Water Network (see Fig. 1). This DMA is used for illustrating the methodology. Its model contains 1600 nodes and 41.153 m of pipes. This DMA is simulated using PICCOLO software. Demands are assumed to occur in the nodes. In this paper, it will also be assumed that leaks occur at the nodes. Such assumption introduces a minor imprecision compared with those due to the methodology and the uncertainty of the model itself. Distance from the real leakage to the closest junction is much shorter than the diameter of the search zone obtained in the best case. It will be clear with results because the areas obtained include some pipes and nodes. Under such assumption, leaks can be seen as additional demands but with unknown location and quantity.

Simulated leaks introduced in the network are of 1 l/s, more or less 3% of the total demand of the sector (in the nighttime). The demand distribution all over the network is the most variable parameter of the model. Some uncertainty in the demand has also been included in order to test the robustness of the method.

3. Leakage detection and isolation methodology foundations

The methodology of leakage localisation proposed in this paper is mainly based on standard theory of model-based diagnosis described for example in (Gertler, 1998) that has already been applied to water networks to detect faults in flow metres (Ragot & Maquin, 2006) or in open channel with dynamic models (Bedjaoui & Weyer, 2011; Nejjari, Pérez, Escobet, & Traves, 2006).

Model-based diagnosis can be divided in two subtasks: fault detection and fault isolation. The principle of model-based fault detection is to check the consistency of observed behaviour while fault isolation tries to isolate the component that is in fault. The consistency check is based on computing residuals, $\boldsymbol{r}(k)$, obtained from measured input signals $\boldsymbol{u}(k)$ and outputs $\boldsymbol{y}(k)$ using the sensors installed in the monitored system and the analytical relationship which are obtained by system modelling:

$$\mathbf{r}(k) = \boldsymbol{\Psi}(\boldsymbol{y}(k), \boldsymbol{u}(k)) \tag{4}$$

where $\boldsymbol{\Psi}$ is the residual generator function that depends on the type of detection strategy used (parity equation (Gertler, 1998) or observer (Chen & Patton, 1999)). At each time instance, k, the residual is compared with a threshold value (zero in ideal case or almost zero in real case). The threshold value is typically determined using statistical or set-based methods that take into account the effect of noise and model uncertainty (Blanke, Kinnaert, Lunze, & Staroswiecki, 2006). When a residual is bigger than the threshold, it is determined that there is a fault in the system; otherwise, it is considered that the system is working properly. In practice, because of input and output noise, nuisance inputs and modelling errors affecting the considered model, robust residual generators must be used. The robustness of a fault detection system means that it must be only sensitive to faults, even in the presence of model-reality differences (Chen & Patton, 1999).

Robustness can be achieved at residual generation (active) or evaluation phase (passive). Most of the passive robust residual evaluation methods are based on an adaptive threshold changing in time according to the plant input signal and taking into account model uncertainty either in the time or frequency domain (Puig, Quevedo, Escobet, Nejjari, & de las Heras, 2008). In this paper, a passive method in time domain has been proposed for robust fault detection, where the detection threshold has been obtained using the method described in Section 4. Robust residual evaluation allows obtaining a set of observed fault signatures $\boldsymbol{\phi}(k) = [\phi_1(k), \phi_2(k), \cdots, \phi_{n_\phi}(k)]$, where each indicator of fault is obtained as follows:

$$\phi_i(k) = \begin{cases} 0 & \text{if } |r_i(k)| \le \tau_i(k \\ 1 & \text{if } |r_i(k)| > \tau_i(k) \end{cases} \tag{5}$$

where τ_i is the threshold associated to the residual $r_i(k)$ generated from sensor i.

Fault isolation involves identifying the faults affecting the system. It is carried out on the basis of observed fault signatures, $\boldsymbol{\phi}$, generated by the detection module and its relation with all the considered faults, $\boldsymbol{f}(k) = \{f(k)_1, f_2(k), \ldots, f_{n_f}(k)\}$ that are compared with theoretical signature matrix FSM (Gertler, 1998). One element of this matrix FSM_{ij} will be equal to one, if a fault $f_j(k)$ is affected by the residual $r_i(k)$. In this case, the value of the fault indicator $\phi_i(k)$ must be equal to one when the fault appears in the monitored system. Otherwise, the element FSM_{ij} will be zero. A given fault $f_i(k)$ is proposed as a fault candidate when the observed fault signature matches with its theoretical fault signature.

4. Leakage sensitivity analysis

The theoretical signature matrix needed to apply the isolation method presented in previous section can be obtained from a leakage sensitivity analysis. This analysis evaluates the effect of a leakage on the pressure in a node. If this process is repeated for each node and possible leak, the sensitivity matrix (Pudar & Ligget, 1992) is obtained as follows:

$$S = \begin{pmatrix} \frac{\partial p_1}{\partial f_1} & \cdots & \frac{\partial p_1}{\partial f_n} \\ \cdot & \cdots & \cdot \\ \frac{\partial p_n}{\partial f_1} & \cdots & \frac{\partial p_n}{\partial f_n} \end{pmatrix} \tag{6}$$

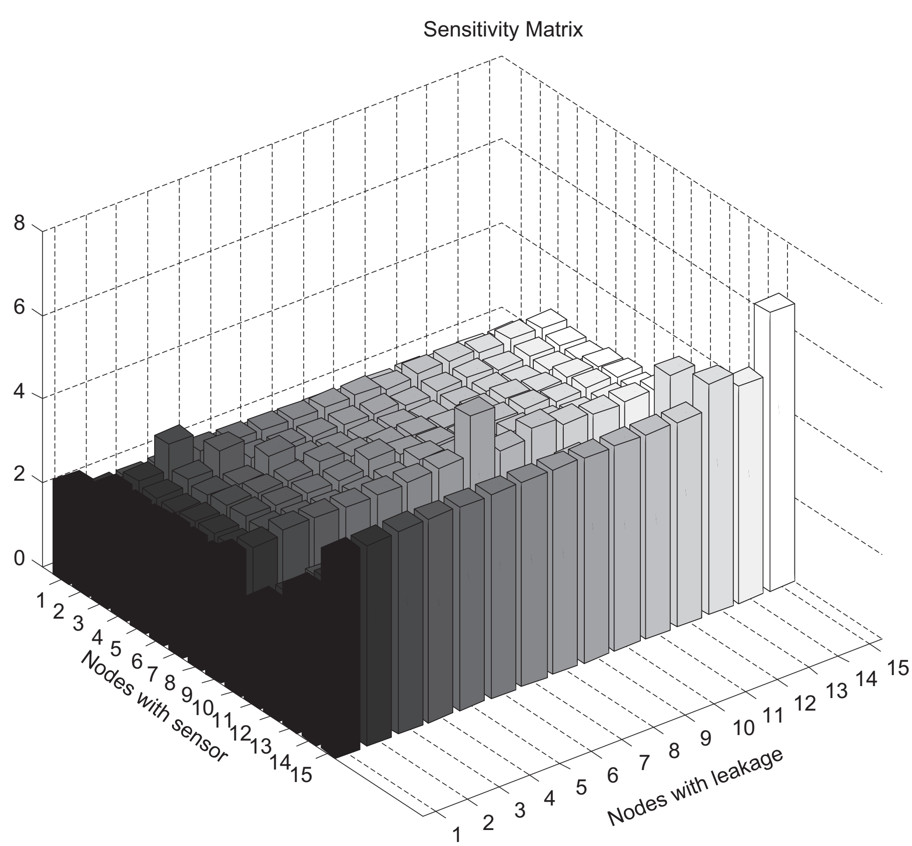


Fig. 2. Sensitivity matrix.

where each element s_{ij} measures the effect of leak f_j in the pressure of node p_i. It is extremely difficult to calculate S analytically in a real network because a water network is a large scale problem described by a multivariable non-linear and non-explicit system of equations as described in Section 2. This work proposes instead generating the sensitivity matrix by simulation as follows: The same leakage is introduced in each node and the increment of pressure is measured in each node. It implies 1600 simulations where 1600 pressures are measured. It has been verified that the analytical and the simulated sensitivity converge for small leakages. The sensitivity matrix depends on the working point that is, on the demand and boundary conditions (Vento & Puig, 2009).

In Fig. 2, the sensitivity matrix for the case study network of Fig. 1 is shown graphically. It has been plotted for 15 nodes distributed homogenously in the DMA as illustration.

Some sensors are much more sensitive to all leakages than others. Thus, a normalisation of sensitivity is needed so that the information provided by any node is comparable. Each row corresponding to a node with a sensor is divided by the maximum value of this row that corresponds to the leakage most important for that node. This procedure leads to the normalised sensitivity matrix:

$$\overline{S} = \begin{pmatrix} \frac{s_{11}}{\sigma_1} & \cdots & \frac{s_{1n}}{\sigma_1} \\ \cdot & \cdots & \cdot \\ \frac{s_{n1}}{\sigma_n} & \cdots & \frac{s_{nn}}{\sigma_n} \end{pmatrix} \tag{7}$$

where $\sigma_i = \max\{s_{i1},\ldots,s_{in}\}$, $i = 1,\ldots,n$. This matrix is shown in Fig. 3 for the considered example. It shows how the most relevant leak is the one on the node itself, the maximum normalised sensitivity is on the diagonal. Columns correspond to nodes with leak and rows correspond to nodes with sensors.

Finally, from the normalised sensitivity matrix (7), the FSM matrix introduced in Section 3 can be derived. Each element FSM_{ij} is equal to zero when leakage j does not affect pressure in node i and it is equal to 1 when leakage j affects node i. The aim is to generate the signature matrix from the normalised sensitivity matrix. In Fig. 3, it can be seen that all leakages affect all pressures. Algorithm 1 presents how the Binarised Sensitivity Matrix ($\overline{S}^b$) is generated.

A process inspired in the ε-method proposed by Sezer and Siljak (1986) is proposed with the aim of identifying the strongest relations between leaks and pressure measurements. In this process, it is absolutely essential to choose conveniently the threshold that controls if a leak has or not an effect on a given pressure. The process proceeds as follows: those leaks that have an effect less than the given threshold are considered as a '0' in the leak signature matrix (5). Otherwise, their effect is considered as a '1'. In this way, the sensitivity matrix is binarised based on the selected threshold. Normalisation allows using a unique threshold for all sensors but the choice of the threshold is most relevant in the process. For small thresholds, all binarised matrix elements are 1 and only detection is possible. As the threshold increases more 0s appear. When threshold approaches 1, then only the diagonal of the signature matrix is 1 and localisation is perfect (or almost perfect, simulation precision makes some nodes equally sensitive to some leakages) but all sensors are needed. Fig. 4 shows how the number of 1s decreases as threshold approaches 1. Number of signatures increases but the significance of each sensor decays.

Algorithm 1. Binarised Sensitivity Matrix Generation **is**

input: n_y is the number of sensors, n_f is the number of leaks
$d(k)$ are the DMA demands
$p(k)$ are the boundary pressures
s_{th} is the binarisation threshold
k time instant when sensitivity matrix is calculated

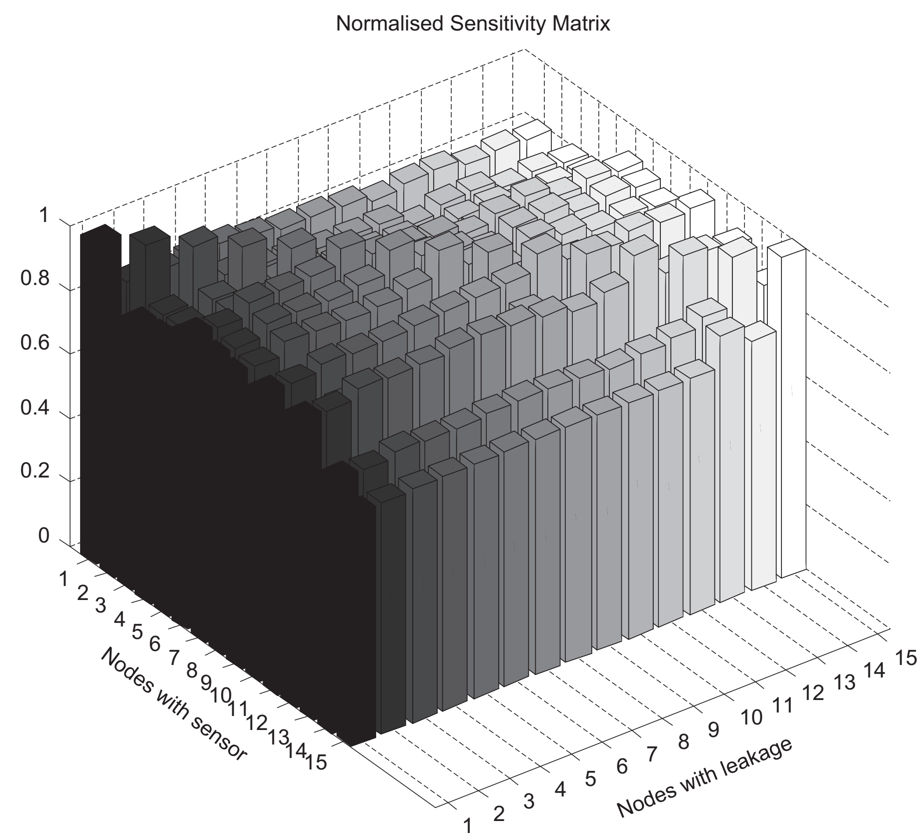


Fig. 3. Normalised sensitivity matrix.

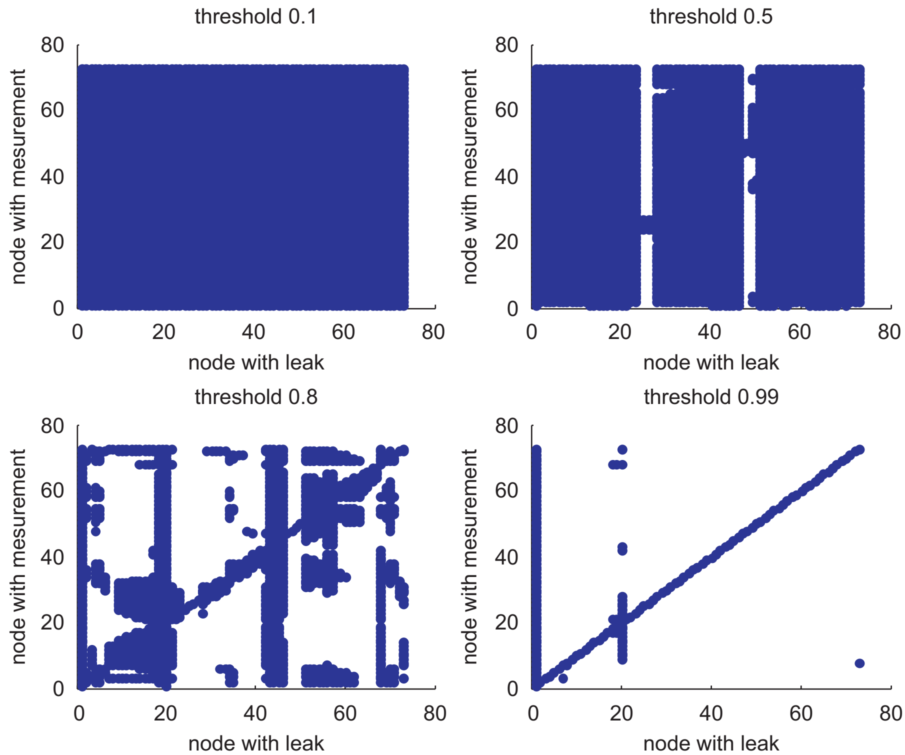


Fig. 4. Number of 1s and 0s depending on threshold.

output: $\overline{S}^b$ and $\sigma_{i\ldots ny}$
for each sensor $i=1,\ldots,n_y$
compute the simulated pressures without leak $\widehat{y}_{i,0}(k)$
for each leak $j=1,\ldots,n_f$
compute the simulated pressures with leak at node j $\widehat{y}_{i,j}(k)$
compute $s_{ij}(k)=\widehat{y}_{i,j}(k)-\widehat{y}_{i,0}(k)$
end
end
for each sensor $i=1,\ldots,n_y$
$\sigma_i(k)=\max\limits_{j=1\ldots nf}(s_{i,j}(k))$
for each leak $j=1,\ldots,n_f$
$\overline{s}_{i,j}(k)=\frac{s_{i,j}(k)}{\sigma_i(k)}$
if $\overline{s}_{i,j}(k)<s_{th}$
$\overline{s}_{i,j}{}^b(k)=0$
else
$\overline{s}_{i,j}{}^b(k)=1$
end
end
end
return $\overline{S}^b(k)$ and $\sigma_{i\ldots ny}(k)$

Fig. 5 shows the evolution of the number of signatures present in the matrix and the maximum number of leakages with the same signature. It corresponds to the 1613 nodes of the network in Fig. 1. Theoretically with 11 sensors (rows) there may be 2047 (that corresponds to $2^{11}-1$ since signature with all 0 is discarded as detection is imposed) different signatures for leakages (columns). In order to get maximum number of signatures, a necessary condition is to have in each column 2^{n-1} 1s, where n is the number of sensors (rows). This necessary condition is fulfilled for the threshold where both lines in Fig. 5 cross ($\sim$0.1). This is the threshold used.

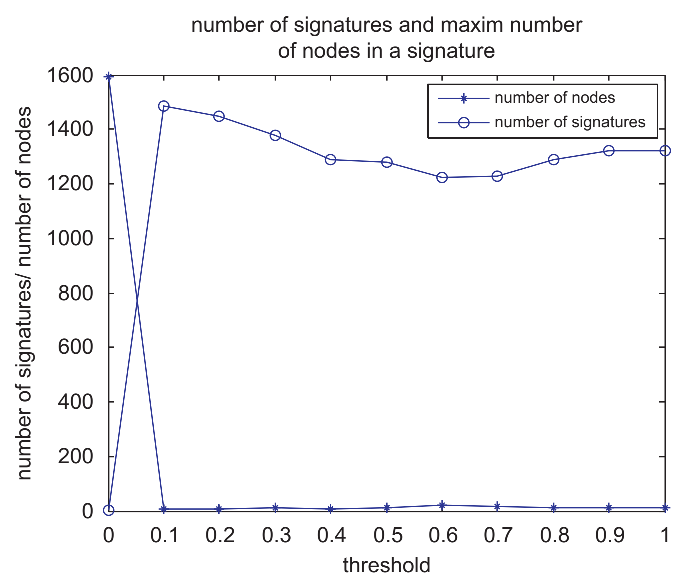


Fig. 5. Evolution of the signature matrix depending on threshold.

Algorithm 2 summarises the leakage detection and isolation procedure using the binarised sensitivity matrix.

Algorithm 2. Leakage Detection and Isolation

input: n_y is the number of sensors, n_f is the number of leaks
N is the time horizon
$d(0)\ldots d(N)$ are the demands
$p(0)\ldots p(N)$ are the boundary pressures
$y_{i\ldots ny}(0)\ldots y_{i\ldots ny}(\mathrm{N})$ are the measured pressures

s_{th} is the binarisation threshold determined using Algorithm 1
output: $f_{1\ldots nf}$ contains the number of incidences of each leak detected in the time horizon N.
initialise $f_{1\ldots nf}=0$
for each instant $k=0,\ldots,N$
  for each sensor $i=1,\ldots,n_y$
   compute simulated pressures without leak $\hat{y}_{i,0}(k)$
   evaluate the residual $r_i(k)=y_i(k)-\hat{y}_{i,0}(k)$
   compute $\overline{S}^b(k)$ and $\sigma_{i\ldots ny}(k)$ from Algorithm 1
   the normalised residual $\overline{r}_i(k)=\frac{r_i(k)}{\sigma_i(k)}$
   if $\overline{r}_i(k)<s_{i,th}$
    $\phi_i(k)=0$
   else
    $\phi_i(k)=1$
   end
  end
  for each leak $j=1,\ldots,n_f$
   if HammingDistance$(\phi(k),\overline{S}^b(:,j)(k))\neq 0$
    $f_j=f_j$
   else
    $f_j=f_j+1$
   end
  end
end
return $f_{1\ldots nf}$

5. Sensor placement algorithm

An optimal sensor placement is defined as a sensor configuration that achieves the minimum economical cost (number of sensors) while observing pre-specified performance criteria (groups of nodes that are not isolable with a minimum number of elements). Since this issue has been addressed in many applications in particular in Song, Chen, Sastry, and Tas (2009) a good literature review is provided.

A model of water network can be represented as a graph $G=(V,E)$, where E is the set of edges that represent the pipes and V is the set of vertices (nodes) where pipes meet. Vertices can represent sources, such as reservoirs or tanks, where water is introduced or sinks (demand points) where water is consumed. Each pipe connects two vertices v_i and v_j and usually is denoted as (v_i,v_j).

Using the graph representation, the problem of optimal sensor placement can be formulated as an integer programming problem, where each decision variable x_i associated to a node v_i of the network can be 1 or 0, meaning that the sensor will be or will not be installed in this node (Bagajewicz, 2000). The starting point of the algorithm is the leakage sensitivity matrix obtained by simulation binarised using the process described in Section 4. Every row corresponds to a hypothetical position of a sensor in a node while every column corresponds to a possible leak in a node. Thus, if a given element of this binary matrix contains a "1", it means that installing a sensor in the node corresponding to this row it would be able to detect the fault associated to the column of this element assuming a single leakage. A particular distribution of sensors (solution) is achieved by instantiating the value of decision variables x_i to "1" (meaning installing the sensor) or to "0" (meaning non installing the sensor). For any particular distribution, a set of groups of indiscernible leaks appear, each group with n_i leaks. The objective of the sensor placement algorithm is to find the sensor distribution that minimises the number of elements for the largest set of leaks with the same signature. The objective (cost) function is therefore

$$J=\min_{x_1,\cdots,x_n}\max\{n_1,\cdots,n_{n_f}\} \tag{8}$$

where $x_1,\ldots,x_n$ are the decision variables that determines a particular sensor distribution and n_i is the number of nodes in group i of indiscernible nodes for a given leakage f_i. In order to increase isolability, this cost should be minimised but at the same time keeping the economical cost reasonable, that is installing the less number of sensor that is possible. The problem is solved for a number of sensors; this number is increased till the cost does not decrease subtantialy. A constraint is included such that all leaks should be detected. It is introduced by forcing that signature with all 0s is not accepted.

This optimisation problem can be solved using either deterministic method based for example in Branch and Bound or heuristic methods based for example in Genetic Algorithms. The first type of methods guarantee the optimal solution but the computation time tends to be exponential with the number of nodes/faults (Sarrate, Puig, Escobet, & Rosich, 2007). On the other hand, the second type of methods just guarantees a suboptimal solution that tends to the optimal one when the size of considered population tends to infinity. Besides the formulation of solutions in series of 1s and 0s are most convenient for a GA. Algorithm 3 describes in detail how the optimal sensors distribution is done.

Algorithm 3. Optimal Sensor Distribution

input: n_f are the number of leaks (nodes), n_y are the number of sensors
 d are the DMA demands
 p are the boundary pressures
output: sensors x, and cost $J(n_y)$
 Solve $\min_x(J)$
 subject:
$\sum_{i=1}^{n_f} x_i=n_y$
 where the cost function J is computed using Algorithm 4
return x

Algorithm 4. Cost function J

input: $\overline{S}^b$ is the binarised sensitivity matrix
  x are the optimisation variables
output: J is the cost of the solution $j=0$
for each node $i=1\ldots n_f$
 if $x(i)==1$
  $\overline{S}^b{}_x(:,j)=\overline{S}^b(:,i)$
  j=j+1
 end
end
for each leak $j=1,\ldots,n_f$
  $m(j)=\mathrm{dec}(\overline{S}^b(:,j))$
 where dec is the conversion of binary to decimal.
end
for $i=1\ldots 2^{ny}$
  n_i=number of i in m
 end
 $J=\max(n_i)$
return J

In Fig. 6, the evolution of cost function is presented. The cost has been taken as the number of nodes in the biggest group of possible leakage isolated with a number of sensors and a threshold between 0.1 and 0.4. The set of sensors should, in the leakage localisation, signal a group of nodes that may include a leak. Optimisation tries that the size of this group is as small as possible. A sharp improvement appears with the first sensors but adding more than 7 or 8 sensors introduce little improvement for any threshold. Therefore only 8 sensors are used.

In Fig. 7, the different groups of nodes with the same leakage signature are shown. There are 39 groups and the biggest contains 190 nodes. The localisation of the sensors after the optimisation process is presented in the last figure.

In an ideal situation with a well calibrated network model, a leakage should be searched in one of these regions instead of the whole sector. It is important to note that regions are connected and geographically coherent. Such coherence is a major issue for further search in situ. For further details see Pérez et al. (2009b).

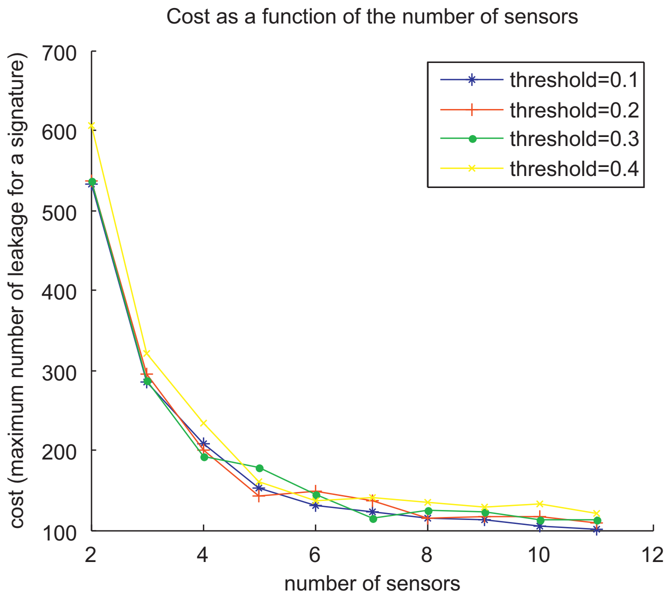


Fig. 6. Evolution of the cost function depending on number of sensors and threshold.

6. Leak isolation results

6.1. Simulation results

The proposed approach of localisation of leakages is first applied in simulation to the *Plaça del Diamant* using the optimal distribution of the sensors obtained in Section 5 consisting in 8 sensors. The process of leak localisation is based on Algorithm 2. If the model were perfect (no uncertainty in demands) and no noise, the leak should be localised with one measurement. However, because of modelling uncertainty and noise, the test has been done during 15 days of simulation (only the lowest consume hour is used each day that corresponds when uncertainty in demands is minimal) and then three options are used to assign the observed leakage signature to a group:

- mean of the sensitivities;
- mean of binarised sensitivities; and
- voting scheme (all days the leak is assigned to a group). The group with more assignations (votes) is the elected.

Results, even without uncertainty/noise, were not good using any of the three decision criteria. It was due to the changing boundary conditions (pressures and flows) that affected very

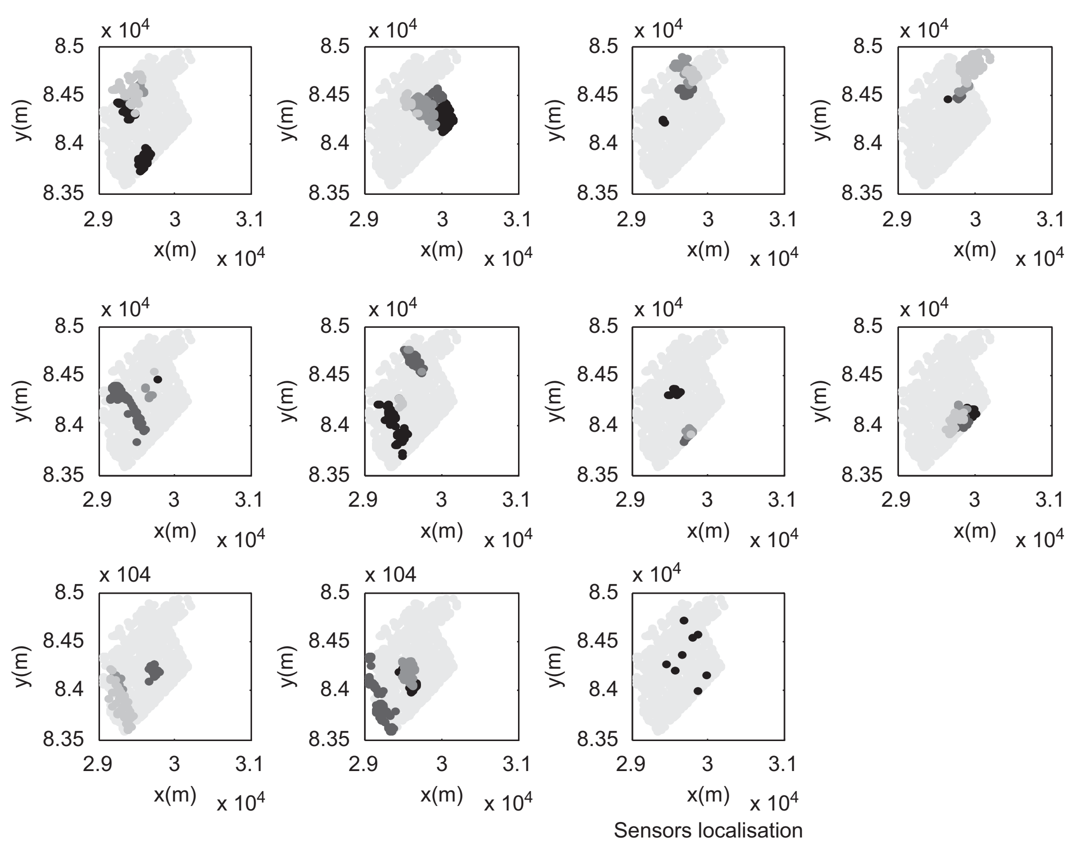


Fig. 7. Groups of nodes with the same leakage signature with 8 sensors and placement of sensors.

much the sensitivity matrix. It is necessary to generate the sensitivity matrix ad hoc for each day with proper boundary conditions that are known. Thus, for each new conditions all the simulations, normalisation and binarisation described in Section 4 are carried out. When a new signature matrix for each day is generated the two first approaches are useless because signature change for each iteration and mean values are meaningless. Thus the third one is tested. It provided perfect results without uncertainty, 100% localisation. It means that each day the group that was signalled suitable to have a leakage contained the node with leakage. These groups were all different each instant and signature matrix is adapted to boundary conditions thus only the voting method had sense. Thus, there are different probabilities of having a leak in a node. This appears in Table 1. It shows the number of nodes that have been signalled 0–15 times (each one for each day). The shadowed line cell corresponds to the one that contains the node that has real leakage. In this case, it corresponds always to the node number 15. It has been done for the 39 groups (one leakage for each) that appeared in sensor distribution (Fig. 7). In Fig. 8, the nodes are presented in grey scale representing the times that have been signalled to be suitable of containing a leakage. The one that contained it appears in the black area.

In order to test the methodology under uncertain parameters in the model, uncertainty in demands was introduced. Uncertainty was estimated using the monthly variation for a demand. It was of 18% of the total demand. Uncertainty was introduced as a coefficient multiplied to the demand of each node generated as a random number between 0.8 and 1.2. The global demand has been kept equal because it is a measured variable and affects greatly the sensitivity.

Table 1
Results using voting criteria adapting signature matrix.

Number of detections	Leak 1	2	3	4	5	6	7	8	9	10	11	12	13
0	529	316	503	316	986	798	782	884	1245	489	1253	1343	363
1	629	505	639	519	438	615	491	518	325	761	131	64	778
2	175	559	311	545	48	147	95	52	64	279	15	9	345
3	126	88	12	88	18	37	76	5	3	9	29	12	59
4	32	57	11	51	17	15	32	22	0	17	1	6	22
5	19	39	21	37	30	0	63	12	0	15	0	7	33
6	54	35	9	11	11	15	30	11	0	9	16	2	13
7	31	8	5	39	7	0	14	5	0	4	117	6	10
8	9	2	15	3	3	0	13	9	0	5	2	9	2
9	1	1	31	0	12	0	2	6	1	5	11	94	4
10	10	2	56	1	18	0	6	17	0	1	5	46	9
11	3	1	15	2	10	2	7	55	0	26	11	7	0
12	1	6	8	1	6	0	6	33	1	0	2	10	0
13	9	18	2	9	0	1	15	10	0	5	4	2	0
14	5	0	1	0	29	6	6	0	0	6	8	20	0
15	7	3	1	18	7	4	2	1	1	9	35	3	2

Number of detections	Leak 14	15	16	17	18	19	20	21	22	23	24	25	26
0	357	1311	1363	354	489	539	477	490	1396	1430	996	959	807
1	869	66	44	853	552	575	778	512	122	92	372	347	533
2	304	21	13	317	156	179	255	182	12	11	82	132	134
3	50	83	5	57	170	59	65	47	55	52	100	90	53
4	14	18	6	14	10	70	4	141	3	5	39	15	50
5	12	119	9	11	7	53	5	4	7	8	2	23	1
6	9	5	5	11	5	95	12	25	3	7	4	8	7
7	9	2	35	7	25	18	33	31	4	1	2	19	1
8	1	0	11	1	29	12	0	17	3	11	5	18	23
9	0	0	23	0	53	9	2	25	11	3	1	11	9
10	2	4	0	3	69	4	1	21	2	0	6	14	14
11	4	8	9	4	11	8	4	46	2	5	15	1	0
12	6	1	0	2	36	3	1	35	0	0	2	1	0
13	2	1	17	4	9	8	0	23	9	11	8	1	6
14	0	0	44	1	15	1	1	24	0	0	2	0	0
15	1	1	56	1	4	7	2	17	11	4	4	1	2

Number of detections	Leak 27	28	29	30	31	32	33	34	35	36	37	38	39
0	596	769	993	594	597	593	544	593	490	490	593	877	568
1	477	585	293	471	597	593	429	593	512	518	595	357	610
2	362	208	90	209	79	109	246	109	182	163	116	43	158
3	109	45	97	55	115	76	127	76	47	5	81	105	48
4	56	2	77	120	121	42	176	42	140	59	40	34	50
5	9	2	34	19	43	16	28	18	5	43	34	11	30
6	0	7	21	32	12	44	24	40	25	76	32	4	52
7	1	5	3	35	33	19	16	26	35	74	13	7	35
8	2	2	8	33	11	51	2	43	18	16	27	0	18
9	6	1	10	14	9	9	2	16	22	101	23	2	8
10	1	1	4	23	13	6	5	6	38	9	9	40	14
11	6	5	2	5	2	4	23	3	25	79	27	139	8
12	1	7	2	16	6	11	6	23	30	3	7	13	30
13	1	0	5	8	0	40	4	22	38	2	27	1	10
14	8	0	0	1	1	12	6	21	7	1	5	6	0
15	5	1	1	5	1	15	2	9	26	1	11	1	1

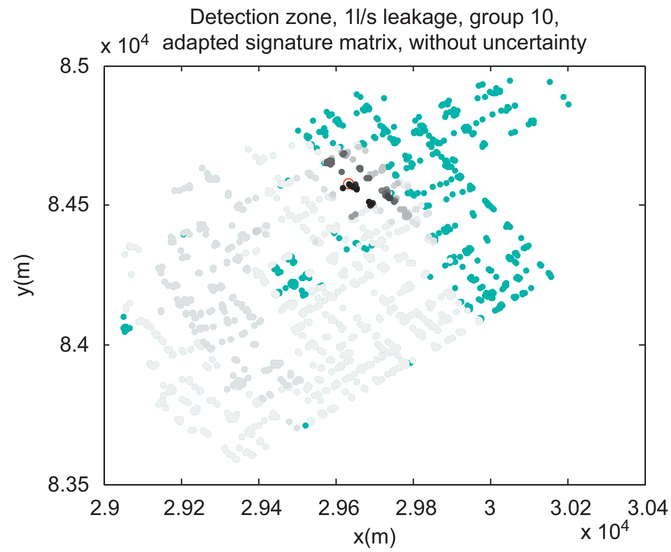


Fig. 8. Localisation of a leak in the correct zone with adapting signature matrix.

Results are presented in Table 2 and Fig. 9. In this case, the leaky node is not always exactly in the most signalled group and the dark grey in the figure does not correspond to 15 but to 9 days. It means that the nodes that more times have been signalled have been signalled thirteen times out of the fifteen. In Fig. 9, the grey scale is lighter than in Fig. 8 because there are less correct detections due to the uncertainty.

Increasing uncertainty interval, the proposed localisation methodology produces poorer results. For a 50% uncertainty, leaks were not well localised but they were localised in a neighbour zone.

The main handicap of the methodology is that in a highly looped network pressure drops due to a leak are not very significant. Therefore it demands high accuracy in transducers. Table 3 show the maximum and minimum pressure drop for leaks 0.5–10 l/s. In high demand hour, the difference is higher but the uncertainty in demand is higher too. Thus, the high cost of sensors may not guarantee good results because of uncertainties in demands.

6.2. Real results

Results from simulation test showed that high accuracy sensors are required. Such sensors exist but represent a major investment. Before such investment is authorised, real test with existing sensors were carried on. Few sensors with non-optimal distribution are available. Measurements have not been taken in best conditions (lower demand time). Nevertheless these results were interesting for the company in order to take further decisions and are presented in this section. A scenario based on a leakage forced in *Enamorats* DMA, in Barcelona network too, is used. This DMA have no qualitative difference with *Plaça del Diamant*. All the steps of methodology exposed so far are applied identically. Only the sensor distribution is not applied because the existing ones are used.

Enamorats DMA model contains 260 nodes and two water input points, where a flow metre and a pressure metre are installed. Input flows in the network and pressures at these points are fixed in the simulation model as boundary conditions. In addition to this information, this DMA has 3 installed pressure sensors, which have been used to apply leakage localisation

Table 2
Results using voting criteria with uncertainty 18%.

Number of detections	Leak 1	2	3	4	5	6	7	8	9	10	11	12	13
0	810	504	688	504	1027	1185	539	688	794	754	1223	1343	765
1	412	646	754	646	207	370	418	708	517	762	164	67	778
2	219	270	15	270	233	28	414	63	326	27	11	6	19
3	34	61	5	61	38	20	73	9	0	23	27	12	37
4	39	83	26	83	17	9	123	29	0	10	17	3	14
5	17	35	25	35	18	0	15	10	0	5	93	9	10
6	71	8	19	8	9	0	18	20	0	11	22	9	9
7	12	2	20	2	4	1	12	26	0	38	6	9	6
8	4	1	42	1	10	17	4	42	1	8	10	94	2
9	17	2	20	2	16	3	6	37	0	2	5	46	0
10	5	22	26	4	25	4	8	8	1	0	23	7	0
11	0	6	0	24	31	3	10	0	1	0	39	10	0
12	0	0	0	0	5	0	0	0	0	0	0	2	0
13	0	0	0	0	0	0	0	0	0	0	0	20	0
14	0	0	0	0	0	0	0	0	0	0	0	3	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0

Number of detections	Leak 14	15	16	17	18	19	20	21	22	23	24	25	26
0	645	1322	1363	623	637	528	1121	638	1272	1272	833	953	811
1	895	66	44	911	554	479	274	561	247	235	535	325	551
2	55	10	5	61	161	187	193	152	14	15	82	119	144
3	11	83	10	11	22	251	42	23	3	12	100	100	58
4	12	18	7	8	3	106	2	3	54	9	39	24	14
5	6	13	9	6	1	34	5	1	4	51	2	44	1
6	1	111	6	5	6	11	3	30	17	13	4	11	8
7	4	1	1	6	28	2	0	40	13	17	2	23	34
8	9	1	36	3	29	1	0	32	1	5	5	20	14
9	1	0	7	0	82	3	0	70	15	11	1	18	2
10	1	4	85	6	94	25	0	64	0	0	6	2	3
11	0	4	58	0	12	10	0	26	0	0	18	1	0
12	0	7	9	0	11	3	0	0	0	0	9	0	0
13	0	0	0	0	0	0	0	0	0	0	4	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0

Number of detections	Leak 27	28	29	30	31	32	33	34	35	36	37	38	39
0	596	596	597	594	597	595	526	593	638	638	595	1165	570
1	477	477	601	478	597	619	464	613	501	511	620	10	636
2	362	367	100	270	85	94	197	99	16	21	106	101	116
3	103	153	118	68	145	85	120	87	80	6	98	106	88
4	29	15	118	50	80	52	192	29	70	0	28	34	54
5	37	3	42	34	42	17	68	32	67	101	19	17	50
6	5	7	23	45	39	18	13	21	40	151	65	4	47
7	3	6	9	30	19	64	8	59	54	23	24	2	12
8	7	1	14	24	15	15	8	21	37	100	9	3	41
9	5	1	8	10	7	5	3	7	44	9	19	38	10
10	8	8	4	10	6	20	2	3	93	73	35	144	16
11	7	6	4	10	4	34	16	39	0	4	20	9	0
12	1	0	2	12	3	20	14	15	0	3	2	7	0
13	0	0	0	5	1	2	9	20	0	0	0	0	0
14	0	0	0	0	0	0	0	2	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0

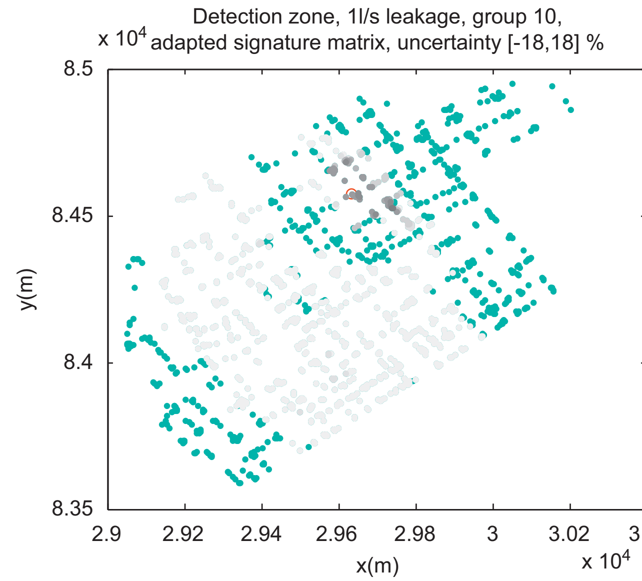


Fig. 9. Localisation of a leak in the correct zone with 18% uncertainty in the demand.

Table 3
Maximum and minimum pressure drop.

Leakage flow (l/s)	Minimum demand hour		Maximum demand hour	
	Minimal ΔP (m)	Maximal ΔP (m)	Minimal ΔP (m)	Maximal ΔP (m)
0.5	0.01	0.02	0.01	0.03
1	0.01	0.04	0.01	0.06
2	0.01	0.09	0.01	0.12
3	0.01	0.14	0.01	0.18
4	0.01	0.19	0.01	0.24
5	0.01	0.24	0.01	0.31
6	0.01	0.29	0.01	0.38
8	0.01	0.37	0.01	0.52
10	0.01	0.44	0.01	0.67

Table 4
Leakage information in *Enamorats* DMA.

Flow (m^3/h)	Flow (l/s)	Leak location	Start time	End time
18	5	Lepant/Aragó	10:20	10:35
14	3.9	Lepant/Aragó	10:37	10:52
9	2.5	Lepant/Aragó	10:53	11:08
6	1.7	Lepant/Aragó	11:10	11:25
16	4.4	Aragó 79	11:53	12:08

methodology. The water company provided boundary conditions (pressure and flow) and pressure inside the DMA (three sensors) data with 10 minute time step. This information was for 5 days in the last day a leakage was forced. Table 4 shows information about this leakage.

The first step is to verify that the hydraulic model provided is correctly calibrated. A four days simulation without any leakage has been done considering pressure values in three internal pressure sensors. The result is the pressure evolution during each day in internal pressure sensors. Differences between the model and reality are important because of demand uncertainty. The worst consequence of these results is that pressure difference caused by a leak can be hidden by the differences due to misfitting of demand model and real demand. To solve this problem, model is corrected with the mean error during no leakage days. Real corrected pressure using these mean errors in each sensor is shown in Fig. 10, compared with the simulation ones.

Although a correction to real pressure has been applied, no difference in the period of leakage can be observed. Thus localisation methodology is applied to see if it is possible to show more information not seen in previous figures. Leakage period duration is about one hour. For leakage period, five second step time data is given by the water company, but only pressures, not flows. If a ten minute step time data is used, in an hour period only 6 samples can be taken. To increase the number of samples a minute time step is proposed. To calibrate the model pressures at the input points are calculated by the mean of the last 30 s data (6 samples) and input flow is taken as a constant during 10 min.

To find discriminable zones obtained with installed sensors, a leak is moved for all 260 possible nodes using the model. For the leakage period two simulations are done: the first one without any leakage and the second one with a leakage moved for 260 nodes. Forced leakage flow is not constant, as it can be seen in Table 3, but only ten minutes data is given for each case. For this reason it is assumed that the leakage flow (5 l/s) is one of them for the whole

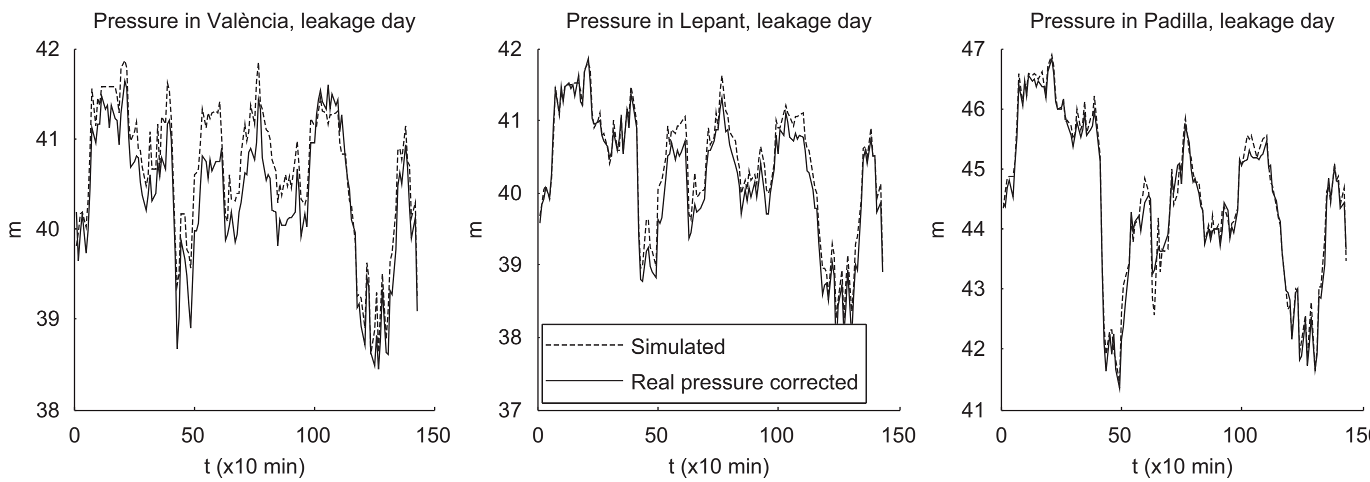
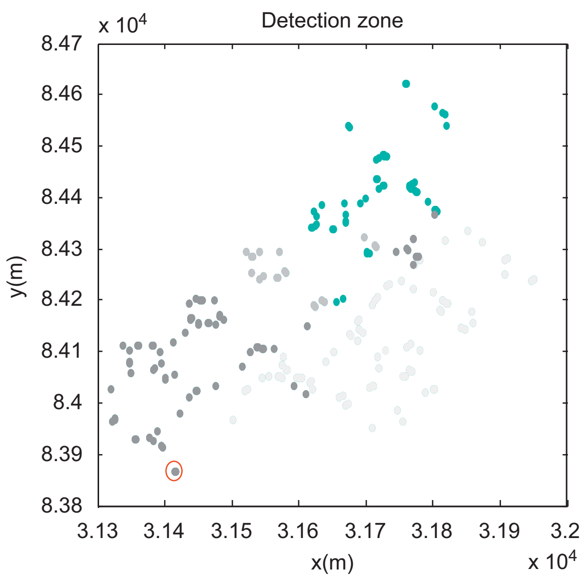


Fig. 10. Corrected real pressures compared with simulated ones.



Threshold = 0.4				
Signature València	Lepant	Padilla	Number of nodes	Number of detections
0	0	0	55	0
1	0	0	23	17
1	1	0	88	31
1	1	1	94	4
Total (max. = 64)				52

Fig. 11. Leakage localisation results with a threshold of 0.4 in a leakage period.

period. This assumption can be justified due the fact that in a real situation the leakage flow may be variable and unknown.

In the same way as in simulation tests, leakage methodology has been applied during more than one time step. In next figures some results are shown. The first case corresponds to the leakage period. As well as in simulation, signature matrix has been calculated depending on boundary conditions. Due to the little quantity of data, the test has been done during the whole leakage period, taking a pressure measurement every minute. Results for a 0.4 threshold are shown in Fig. 11. Although some leakages are not detectable (55 nodes zone), the real leakage is outside this zone. Sensors are not located optimally, so these undetectable leakages were expected. The number of discriminable zones is four, including the non-detectable one. Leakage zone corresponds to the third group, which contains 88 nodes.

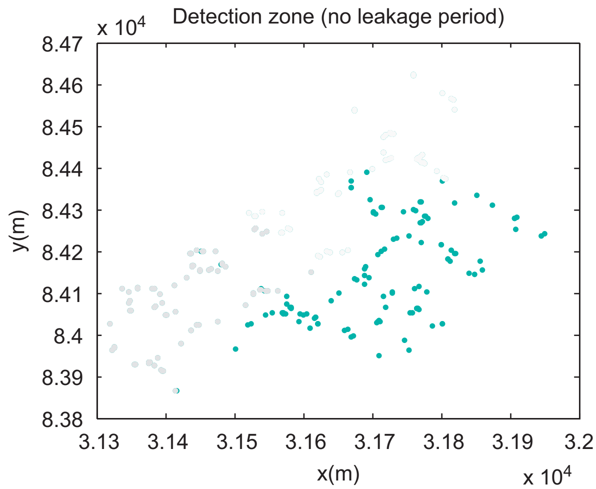
The leakage is given in the circled node: 31 of 64 detections signalled the correct leakage zone. After this test a non leakage period is chosen to apply the methodology. At night discrepancies between reality and the model are smaller than during the day; so it is the best time to do the test. Although an important zone is signalled as a possible leakage zone, the number of detections is only 9 on 42. These results are shown in Fig. 12.

7. Conclusions

A leakage localisation method based on the pressure measurements and sensitivity analysis of nodes in a network has been proposed. The leakage localisation methodology is founded in standard model-based fault diagnosis well established theory.

In order to maximise the isolability with a reasonable number of sensors, an optimal sensor placement methodology based on genetic algorithms is also proposed. The objective function in the minimisation process was the size of the maximum group discriminated. The confidence of the information provided by pressure sensors about leakage could be studied using the Fisher Information Matrix generated using the sensitivity matrix. This new approach is studied as a possible way to define the sensor placement avoiding the optimisation process.

To assess the validity of the proposed approach, it has been applied to a DMA of Barcelona network in real and simulated leak scenarios. Models and information were provided by the water company. For these sectors (DMA), the sensor placement and the leakage detection and localisation methodologies have been applied with successful results even in presence of demand uncertainty in simulation.



Threshold = 0.4				
València	Signature Lepant	Padilla	Number of nodes	Number of detections
0	**0**	**0**	29	0
0	**0**	**1**	94	0
1	**0**	**0**	62	4
1	**1**	**0**	75	5
Total (max. = 42)				9

Fig. 12. Leakage localisation results with a threshold of 0.4 in a non leakage period.

In real test where sensors used where already installed results were poorer. Two main causes are suggested. First the non-optimal distribution of the sensors thus the methodology proposed in Section 4 is currently being applied in an on-going project in order to improve such results. On the other hand, the estimation of demands should be improved and an evaluation of the influence of the misfit of demand model on the methodology has been studied. First results have been published (Pérez et al., 2011).

An issue in the process is to recalculate the sensitivity matrix for each boundary condition using the simulation model because of the high dependence of it to global consumption. This approach is being currently developed using linear parameter varying (LPV) models that consider the consumption as a scheduling variable (Vento & Puig, 2009). Finally, a new approach is being studied that avoids the binarisation of the sensitivity matrix and it is based on correlation of model pressures with leakage and the measurements (Quevedo et al., 2011)

Acknowledgement

The authors wish to thank the support received by WATMAN ref. DPI2009-13744 of the Spanish Ministry of Education and PROFURED project funded by AGBAR-CETAQUA.

References

Bagajewicz, M. (2000). *Design and upgrade of process plant instrumentation*. Lancaster, PA: Technomic Publishers.
Bedjaoui, N., & Weyer, E. (2011). Algorithms for leak detection, estimation, isolation and localization in open water channels. *Control Engineering Practice*, *19*(6), 564–573.
Blanke, M., Kinnaert, M., Lunze, J., & Staroswiecki, M. (2006). *Diagnosis and fault-tolerant control. 2nd ed.* Springer.
Brdys, M. A., & Ulanicki, B. (1994). *Operational control of water systems: Structures, algorithms and applications*. UK: Prentice Hall International.
Chen, J., & Patton, R. J. (1999). *Robust model-based fault diagnosis for dynamic systems*. Kluwer Academic Publishers.
Colombo, A. F., Lee, P., & Karney, B. W. (2009). A selective literature review of transient-based leak detection methods. *Journal of Hydro-environment Research*, *2*, 212–227.
Farley, M., & Trow, S. (2003). *Losses in water distribution networks*. UK: IWA Publishing.
Ferrante, M., & Brunone, B. (2003a). Pipe system diagnosis and leak detection by unsteady-state tests. *1 Harmonic Analysis, Advances in Water Resources*, *26*(1), 95–105.
Ferrante, M., & Brunone, B. (2003b). Pipe system diagnosis and leak detection by unsteady-state tests. *2 Wavelet Analysis, Advances in Water Resources*, *26*(1), 107–116.
Gertler, J. J. (1998). *Fault detection and diagnosis in engineering systems*. Marcel Dekker.
Lambert, A. (1994). Accounting for losses: the Bursa and background concept. *(BABE)IWEM Journal*, *8*(2), 205–214.
MacDonald G. (2005). *DMA design and implementation, a North American Context*. Leakage conference, IWA.
Misiunas, D., Lambert, M. F., Simpson, A. R., & Olsson, G. (2005). Burst detection and location in water distribution systems. *Water Science and Technology: Water Supply*, *Vol. 5*(3–4), 71–80.
Nejjari, F., Pérez, R., Escobet, T., & Traves, L. (2006). Fault diagnosability utilizing quasi-static and structural modelling. *Mathematical and computer modelling*, *Vol. 45*, 606–616.
Perez, R., de las Heras, S., Aguilar, J., Pascual, J., & Peralta, A. (2009a). District management areas characterisation in water network based on clustering. *Water Science Technology*, *9*(5), 591–600.
Pérez, R., Puig, V., Pascual, J., Peralta, A., Landeros, E., & Jordanas, Ll. (2009b). Pressure sensor distribution for leak detection in Barcelona water distribution network. *Water Science & Technology*, *9*(6), 715–721.
Pérez, R., Nejjari, F., Puig, V., Quevedo, J., Sanz, G., Cugueró, M., & Peralta, A. (2011). *Study of the isolability of leaks in a network depending on calibration of demands. Computing and control for the water industry 2011 "Urban water management—Challenges and opportunities"*. University of Exeter.
Puig, V., Quevedo, J., Escobet, T., Nejjari, F., & de las Heras, S. (2008). Passive robust fault detection of dynamic processes using interval models. *IEEE Transactions on Control Systems Technology*, *16*(5), 1083–1089.
Pudar, R. S., & Ligget, J. A. (1992). Leaks in pipe networks. *Journal of Hydraulic Engineering*, *118*(7), 1031–1046.
Quevedo, J., Puig, V., Cembrano, G., Blanch, J., Aguilar, J., Saporta, D., Benito, G., Hedo, M., & Molina, A. (2010). Validation and reconstruction of flow meter data in the Barcelona water distribution network. *Control Engineering Practice*, *18*(6), 640–651. doi:10.1016/j.conengprac.2010.03.003.
Quevedo, J., Cuguero, M. À., Perez, R., Nejjari, F., Puig, V., & Mirats, J. M. Leakage location in water distribution networks based on correlation measurement of pressure sensors. *WATERMATEX*, submitted for publication, San Sebastián.
Ragot, J., & Maquin, D. (2006). Fault measurement detection in an urban water supply network. *Journal of Process Control*, *16*(9), 887–902.
Sarrate, R., Puig, V., Escobet, T., Rosich, A. (2007). Optimal sensor placement for model-based fault detection and isolation. In *Proceedings of the 46th IEEE conference on decision and control*. New Orleans, USA.
Sezer, M. E., & Siljak, V. (1986). Nested epsilon-decomposition and clustering of complex systems. *Automatica*, *22*(3), 321–331.
Song, Z., Chen, Y., Sastry, C. R., & Tas, N. C. (2009). *Optimal observation for cyber-physical systems*. Hanover: Springer.
Vento, J., Puig, V. (2009). Leak detection and isolation in pressurized water networks using interval LPV models. In *Proceedings of the 7th workshop on advanced control and diagnosis (ACD'09)*. Zielona-Gora, Poland.
Verde, C., Visairo, N., & Gentil, S. (2007). *Two leaks isolation in a pipeline by transient response. Advances in water resources*, Vol. 30. Elsevier.