

A DECISION SUPPORT SYSTEM FOR ON-LINE LEAKAGE LOCALIZATION

**Jordi Meseguer^{a,*}, Josep M. Mirats-Tur^a, Gabriela Cembrano^{b,a}
Vicenç Puig^{b,c}, Joseba Quevedo^c, Ramón Pérez^c, Gerard Sanz^c
David Ibarra^d**

^a*CETaqua, Water Technology Center, Carr. Esplugues 75, Cornellà de Llobregat, 08940, Spain*

^b*Institut de Robòtica i Informàtica Industrial(CSIC-UPC)*

^c*Dpt, Enginyeria de Sistemes, Automàtica i Informàtica Industrial (UPC)*

^d*Aqualogy Aqua Ambiente Servicios Integrales S.A*

phone: +34 93 312 48 29 fax : +34 93 312 48 01

e-mail: jmeseguer@cetaqua.com

Abstract:

This paper describes a model-driven decision-support system (software tool) implementing a model-based methodology for on-line leakage detection and localization which is useful for a large class of water distribution networks. Since these methods present a certain degree of complexity which limits their use to experts, the proposed software tool focuses on the integration of a method emphasizing its use by water network managers as a decision support system. The proposed software tool integrates a model-based leakage localization methodology based on the use of on-line telemetry information, as well as a water network calibrated hydraulic model. The application of the resulting decision support software tool in a district metered area (DMA) of the Barcelona distribution network is provided and discussed. The obtained results show that the leakage detection and localization may be performed efficiently reducing the required time.

Keywords: Software Tools, Fault Diagnosis, Water Networks, Hydraulic Models, Leakage Localization, Pressure Sensitivity.

1. INTRODUCTION

Water utilities provide clean water service to local communities and charge the service by the metered water consumption. However, not every drop of water produced at a water treatment plant reaches customers and generates revenue for water companies. Instead, a significant proportion of drinking water is lost, due to either water leaking from the distribution pipelines or unauthorized water usage. Water loss represents a major fraction of *non-revenue water (NRW)*¹: more than 65% of *NRW* arises from unauthorized water consumption, meter inaccuracies and leaks from the water mains source-to-taps infrastructure (IWA, 2000). On average, more than 15% of water produced is lost in the United Kingdom (Wu et al., 2010). Regarding the overall water loss of Spain, the estimation of the Spanish Statistics

¹ *Non revenue water (NRW)* is water that has been produced and is “lost” before it reaches the customer. Losses can be real losses (through leaks, sometimes also referred to as physical losses) or apparent losses (for example through theft or metering inaccuracies).

National Institute (INE) was around $791 \text{ Hm}^3/\text{year}$ in 2009, while the overall water supply was around $4709 \text{ Hm}^3/\text{year}$. Thus, around 25% of the supplied water was lost. The AWWA² *Water Loss Control Committee* (2003) reports that water loss reduction and the associated revenue loss recovery stand among the most promising areas of water resources improvements in North America.

In general, water leakage³ rate is high which causes financial loss and worsens water utilities' public reputation: it may damage the infrastructure and cause third-party damage, water and financial losses, energy losses and health risks (Wu, 2008). Consequently, a systematic approach is needed to identify likely leakage hotspots so that detection crews can find leaky mains more quickly, leading to quicker repairs. Most of the leakage management related methods developed so far can be broadly classified as follows (Puust et al 2010): (1) leakage assessment methods which focus on quantifying the amount of lost water; (2) leakage detection methods which are primarily concerned with the detection of leakage hotspots and (3) leakage control models which are focused on the effective control of current and future leakage levels.

Continuous improvements on water loss management are being applied, based on the use of new available technologies. Nonetheless, the whole leakage localization process may still require long periods of time (i.e. weeks, months) with an important volume of water wasted before the leak is found (Pérez et al. 2011). To avoid these inconveniences, leakage detection and localization based on mathematical models may be used (Brdys & Ulanicki, 1994) which can “compare” the data gathered by installed sensors in the network with the data obtained by a model of this network⁴. If a meaningful difference is detected between these data sets, a detection of an abnormal event is obtained. This implies that modelling is paramount in order to achieve successful results. It provides the means for detection of possible faults (i.e. leaks) as well as their probable localization in the network. Model-based methodologies are widely used in environmental applications when rational decision-making must be carried out taking into account complex physical processes and existing interlinks among environmental, industrial and social issues. (Wierzbicki et al., 2000): i.e. water resource management and operation (Rani et al., 2010), (Giupponi, 2005), management of complex networks in the urban water cycle (Ocampo et al., 2013), water demand management (Bakker et al., 2013) (Makropoulos et al., 2008), climate change effect on water resources and water supply (Pouget et al., 2012) (Laucelli et al., 2012).

Regarding the specific problem of water loss reduction tackled in this paper, the use of flow and pressure sensors together with hydraulic models of the water network for leak detection and localization is a suitable approach for the on-line monitoring of water balance. A direct approach based on simulation was proposed in (Mandel, 1998) and (Almandoz et al., 2005). An alternative approach based on an inverse approach that formulates the leak detection and localization problem as a parameter estimation approach was presented by Pudar & Ligget (1992) and further inverse approaches were investigated (Ligget & Chen, 1995) (Kapelan et al., 2003) (Wu & Sage, 2006). Additionally, other

² AWWA: American Water Works Association (<http://www.awwa.org/>)

³ Different definitions of leakage in distribution systems exist. The most frequently used one defines the leakage as (amount of) water which escapes from the pipe network by means other than through a controlled action (Ofwat 2008).

⁴ See *Appendix 1.1* for background information about mathematical modelling applied to water distribution networks.

contributions integrating data-driven and model-driven approaches (Farley et al., 2010; Bicik et al., 2013) or based just on a data-driven approach (Romano et al., 2013) were also presented.

Pérez et al. (2011) presents a direct modelling methodology developed to help network operators deal with the detection and localization of leaks in district metered areas (DMAs⁵) of water distribution networks. The leakage detection procedure is performed by comparing pressure data of certain DMA inner nodes⁶ with their estimation using the simulation of the mathematical network model. Taking into account a fault diagnosis context (*Appendix 1.2*), the leakage localization procedure presented in (Pérez et al., 2011) is an application of the binary fault isolation procedure based on the theoretical fault signature matrix (Gertler, 1998), applying a parallel diagnostic inference process between this binary matrix and the binary observed fault signature vector. In (Quevedo et al., 2011), this method was updated to work with non-binary observed fault signatures and theoretical fault signature matrices enhancing the overall performance of the method.

Model-based leakage localization methods are an important tool to detect leaks and reduce water losses. They may also be used as support systems in decision-making processes related to the water network maintenance and investments. In general, as stated by Savić et al. (2011), model-based methodologies for environmental applications (i.e. leakage detection and localization) presents a certain degree of complexity which limits their use to experts in the modelling field involved in each environmental application. Nonetheless, if computer-based models for environmental applications were integrated into a model-driven *Decision Support Systems (DSSs)* (Power, 2004; Power et al., 2007; Savic et al 2011), this type of models should be useful for managers of environmental systems (i.e. water networks) as support systems in decision-making processes without the need for expertise in the mathematical modelling. The field of leakage localization-based software tools (DSSs) is very active; a number of commercial products are in the market, such as the *TaKaDu*⁷ system (Armon et al., 2011) based on a data-driven approach and *WaterGEMS*⁸ (Wu & Sage, 2006) based on an inverse modelling approach using Genetic Algorithms. Focusing on the topic of software tools, the main goal of this paper is to describe a model-driven DSS prototype for leakage detection and localization in water distribution networks. The aim of the presented approach is not to compete with existing commercial tools but to show a straightforward approach which takes benefit from the existing hydraulic models used by the water network operators for operation and planning purposes. A web-based tool has been developed and applied to a DMA in the Barcelona water network in both simulated and real scenarios. Regarding the model-based leakage localization method, this DSS integrates the approach presented by (Quevedo et al., 2011). Thereby, this tool has been developed with the main purpose of widening the use of model-based leakage localization techniques by users with low expertise in this modelling field taking benefit from the existent hydraulic models used by the operators. In general, the tool proposed in this paper is based on the integration of the instrumentation data existing in the DMA network and the hydraulic model of the DMA network with advanced model-

⁵ District Metered Area (DMA) is a defined area of the distribution system that can be isolated by valves and for which the quantities of water entering and leaving can be metered. The subsequent analysis of flow and pressure, especially at night when a high proportion of users are inactive, enables leakage specialists to calculate the level of leaks in the district.

⁶ 'DMA inner nodes' expression refers to all DMA nodes excluding those ones related to the inlets.

⁷ <http://www.takadu.com/>

⁸ <http://www.bentley.com/>

based leakage localization techniques. As a result of this integration, a set of functionalities efficiently supporting the leakage localization process are available for the intended user.

This paper is organized as follows: *Section 2* presents the case study to which the developed web-based leakage localization DSS prototype has been applied. *Section 3* gives an overview of the requirements and architecture of this software tool and its application to the considered case of study. Then, in *Section 4*, the integration of the considered leakage localization methodology into the proposed DSS is tackled: on the one hand, the foundation of the integrated methodology is recalled and, on the other hand, the integration procedure is given. Next, the main features of the proposed software prototype and the results obtained in leakage scenarios considered in the case of study are reported (*Section 5*). Finally, resulting conclusions are given in *Section 6*. Additionally, and for the paper to be self-contained, a short background in water network mathematical modelling and fault diagnosis is included in the *Appendix* given its importance in the considered model-based leakage localization DSS software prototype.

2. CASE STUDY DESCRIPTION

2.1 Description of Nova Icaria District Metered Area (DMA)

The performance of the leakage localization DSS prototype presented in this paper has been tested in one DMA of Barcelona using a past real leakage scenario. The water network which contains this DMA supplies both Barcelona and its metropolitan area with approx. *3 million* inhabitants and is managed by the water company Aigües de Barcelona. The whole water network is composed of *4574 km* of pipes, *65* pumping stations and *72* water tanks with a water storage capacity of *250,542 m³*. This network is segmented into *117* pressure levels, and *214* DMAs. In this pilot implementation, the Barcelona's DMA of *Nova Icària* has been used which is included in pressure level *55* within the city network. This DMA with a network length around *43 km* has two inlets (Alaba and Llull), *1996* nodes and *3442* pipes. Its area is approx. *227 hm²* with a delivered monthly volume of *170,000 m³*, on average. In Figure 1, the water network of Nova Icaria DMA can be seen from the EPANET file which contains the hydraulic model of this network. In this figure, the two DMA inlets have been highlighted using red triangle symbols. Regarding the instrumentation, the Nova Icaria DMA is provided with flow and pressure sensors at both inlets and *six* inner pressure sensors (green star symbols) already deployed in the DMA⁹. The sample time associated with all these sensors is set to *10 minutes*.

⁹ The deployment of these existing sensors was based on a binary approach since the first version of the used leakage localization methodology was also binary (Pérez et al., 2011). Under this framework, the sensors distribution could be understood as optimal but this does not hold for the framework presented in this paper.



Figure 1 Water network of Nova Icaria DMA (EPANET model) highlighting inner pressure sensors (green stars) and DMA inlets (red triangles).

2.2 Nova Icaria DMA Network Modeling Approach

Regarding the water network of Barcelona and its metropolitan area, Aigües de Barcelona uses a *Geographical Information System (GIS)* to register the current state of this network (i.e. volumes in tanks, flows and pressures at pipes, junctions or valves) and all its properties (i.e. roughness, length and diameter of pipes, elevation of junctions) and to manage certain operation and maintenance processes. In order to ensure the accuracy of the *GIS* representation, this system is kept updated whenever there is a change in the network. Additionally, this system is also used to generate the DMA hydraulic model in EPANET format and to keep this model updated. Thereby, in order to carry out isolated simulations of this DMA using this EPANET model, the boundary conditions must be set using the pressure and flow measurements registered in the DMA inlets: the pressure measurement at every time instant k is set at every DMA inlet while the total measured DMA inflow is distributed using a constant coefficient (base demand) in each demand node which has been obtained using the billing information of Aigües de Barcelona. Thereby, the same consumption pattern is assumed for all demand nodes in the DMA.

2.3 Description of Leakage Scenarios in Nova Icaria

In 2011, the Nova Icaria DMA was affected by several leakage scenarios which caused the increase of the DMA minimum night flow¹⁰ in 10 l/s. During the first half of 2011, an increase of 4 l/s was already noticed by DMA SCADA operators. However, since the beginning of September 2011, the DMA minimum night flow started a progressive worsening reaching an overall increase of 10 l/s in October 2011 when a leak was found. Once this leak was repaired, the DMA minimum night flow decreased in 6 l/s. Nonetheless, its value was still abnormal when compared to its values registered in 2010. During the beginning of November 2011, a set of small leaks could be found in an area close to the localization of the first leak. When all these small leaks were repaired, the minimum night flow decreased in 4 l/s reaching again its normal values. The proposed software tool was tested using the historic data of this scenario to assess its efficiency in detecting the leaks. In Figure 2, the localization of the 6 l/s-leakage is signaled using a blue arrow while the set of small leaks (4 l/s) found in the beginning of November 2011 is inside the marked blue square area. In *Section 5.2*, the above described leakage scenarios will be considered to show the performance of the proposed leakage localization software tool prototype.

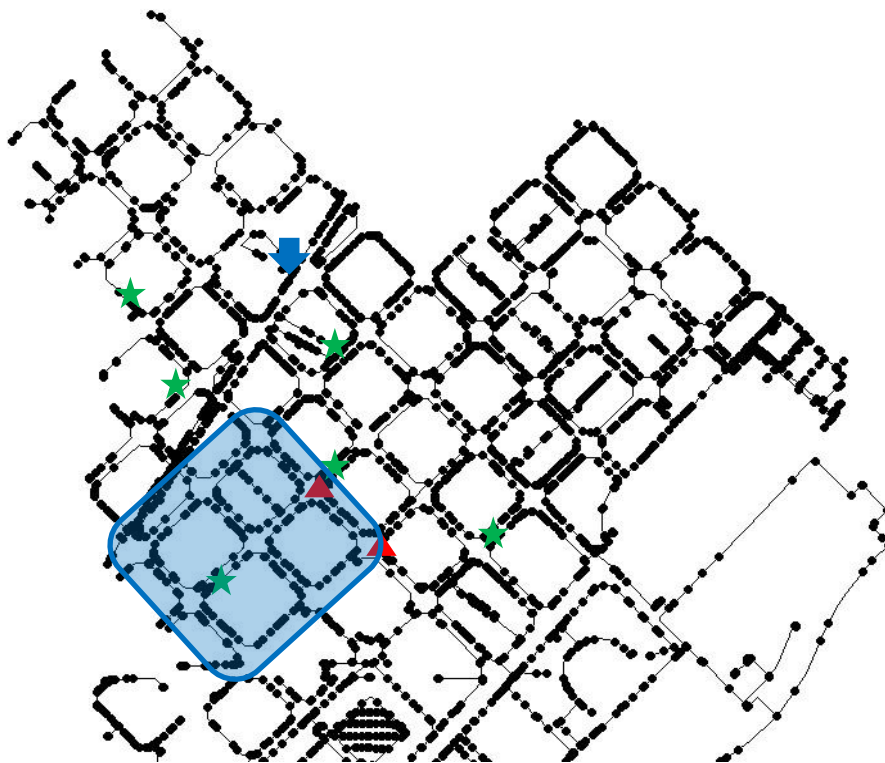


Figure 2 Water network of Nova Icaria DMA highlighting in blue the localization of the two leakage scenarios.

¹⁰ *Minimum night flow* is the amount of water flowing into a discrete zone or district metered area (DMA) during the period of lowest demand, typically during night hours (AWWA, 2009). Given that the consumption into such a zone is at a minimum having a characteristic pattern over the time, unexpected variations of the minimum night flow regarding this pattern may indicate the existence of leaks. As a result, night flow analysis is being used as a basic technique to indicate the existence of leaks in a discrete zone or DMA.

3. OVERVIEW OF THE PROPOSED WEB-BASED SOFTWARE TOOL FOR LEAKAGE LOCALIZATION

3.1 Requirements

In order to bring the power of model-based fault diagnosis techniques (see *Appendix 1.2*) applied to leakage localization in water networks to the intended users, a prototype of a model-driven DSS integrating those techniques has been developed. The general requirements for this new DSS for leakage localization in water networks are (Power et al., 2007):

1. To provide an easy access to model-based leakage localization techniques for users which are non-specialist (i.e. drinking water networks technicians and managers) so that they could locate those leaks affecting the DMA network water losses using their own resources more efficiently and reducing the associated water losses.
2. To provide the capability of an automatic¹¹ or manual leakage localization analysis based on the integration of the measurements coming from the instrumentation deployed in the DMA water network with the leakage localization DMA water network model.
3. To avoid the need of complex coding to set up the automatic or manual leakage localization analysis in a new DMA network or to carry out the updates of an existing leakage localization water network model when the DMA water network or its associated instrumentation suffer changes in their current state.
4. To provide an intuitive Graphical User Interface (GUI) that will allow the easy creation and modification of specific DSS scenarios. A DSS scenario contains the required information to carry out the automatic or manual leakage location analysis in the water network of a specific DMA. Additionally, the DSS GUI also integrates a *Geographical Information System* (GIS) viewer¹² which lets enhance the visualization of both the structure of the DMA water network and its existing instrumentation, and the results obtained in the analyses of every scenario.
5. To provide an easy access to the leakage localization DSS from every device with internet access using a web-based framework.

3.2 Structure

The main unit of work of this DSS for leakage localization in water networks is the scenario. Each scenario is related to a single DMA allowing automatic and manual leakage localization analyses in this DMA. Then, when an analysis of this type has to be carried out in a different DMA, another scenario must be created. Mainly, the foundation of the analysis carried out in a scenario related to a given DMA is

¹¹ Automatic analysis refers to the fact of running automatically a leakage localization analysis when new measurements coming from the instrumentation are available. When these measurements are updated on-line, on-line leakage localization analyses are carried out.

¹² *Microsoft Bing Maps* technology (<http://www.bing.com/maps/>) has been used to build a GIS viewer to display street maps, water network topology, and existing instrumentation and analysis results associated with a given DMA. In this tool prototype, the GIS viewer is compatible with street maps provided by *Microsoft Bing Maps*, EPANET files to display the water network topology and the computed results (i.e. leak locations).

based on the integration of the measurements coming from the existing instrumentation, the DMA hydraulic model and the model-based leakage localization algorithms.

Regarding the DMA instrumentation, pressure and flow sensors at DMA inlets and pressure sensors at certain DMA inner nodes are required which offer on-line measurements. Traditionally, this instrumentation and its corresponding measurements are integrated in the SCADA system used to perform the supervision of the overall network. In general, the sensor measurements are stored in the *raw database* which contains the raw metered data. Nonetheless, this data may be affected by certain errors, such as outliers and missing data which prevent its direct use since it may cause a wrong performance of the SCADA supervision processes. Normally, a validation process is applied to this data to remove those errors and to try to reconstruct the lost information when possible. This validated data is stored in the *validated database*. Regarding the DSS for leakage localization, this tool is linked to this *validated database* in order to obtain the on-line measurements of the DMA instrumentation. By this means, when establishing the settings of a given scenario, database time series which store the validated sensor data of a certain DMA network must be mapped with the hydraulic model of this DMA.

Another key element in the leakage localization analysis of a certain DSS scenario is the DMA hydraulic model. The current version of this DSS is compatible with EPANET (Rossman, 2000) and the DMA hydraulic model is built externally¹³ using this simulation tool and uploaded to the DSS when the settings of a given scenario are being established. In general, the information taken from the EPANET model of the DMA network is not only the network topological structure but also those fixed network parameters: i.e. roughness, length and diameter for pipes; positions and characteristics for valves; level of every junction relative to the sea level; base demand of every junction. Thus, it must be taken into account that the accuracy of this model is an important issue to guarantee a good performance of the overall leakage localization process¹⁴. Apart from the DMA network hydraulic model, other network parameters that must be informed are the location of every sensor and its depth with respect to the ground level. Another important input of the DMA network model are the demand patterns associated with those junctions having a related known demand. Both the DMA hydraulic model and the above mentioned parameters must be informed when the settings of every scenario are established. They are then stored in the *parameter database* of the DSS. Additionally, this DSS offers certain processes to help the user to keep this information updated which is a crucial issue to ensure the goodness of the analysis results.

Finally, both the information coming from the external SCADA *validated database* and the internal DSS *parameter database* are integrated with the model-based leakage localization algorithm in order to perform the automatic or manual analyses in a certain DMA network. The specific features offered by the DSS in every scenario allow the user to set up these analyses. Thereby, a DMA network manual analysis

¹³ The leakage localization software tool prototype presented in this paper is based on integrating the existent hydraulic model used by the operator for water network operation and planning purposes.

¹⁴ The calibration of the hydraulic model involves the calibration of both physical parameters (i.e roughness) and the nodal demands. On one hand, the EPANET model of the DMA obtained from the water company *Operation and Planning Department* should be a suitable representation of the DMA network and on the other hand, nodal demands have been calibrated by means of estimating different demand patterns according to the historical billing information and the type of water usage (i.e. domestic, industrial,...). At the end of this calibration process, the goal is to have an error between pressure model predictions at the monitored nodes and the corresponding pressure measurements of the order of the minimum pressure change that induces the smallest leak to be detected.

can be carried out choosing basically the start and end date of the analysis. Additionally, an automatic analysis can also be set up just selecting this property when establishing the analysis settings. In this case, a DMA network analysis runs every day automatically using the available sensor data of the previous day. In general, different analyses can be carried out at the same time for a given scenario.

In Figure 3, a conceptual scheme of the DSS structure is given, highlighting the above-described integration of the data coming from the DMA instrumentation (*SCADA validated database*), the DMA hydraulic model (*DSS parameter database*) and the model-based leakage localization algorithms (*DSS web server*). This scheme also highlights the web nature of this DSS which renders its availability from a wide range of devices connected to internet or to those intranet networks allowed to access to the DSS web server. Thereby, end-users can access the DSS easily to manage leakage localization scenarios and their corresponding leakage localization analyses.

Regarding the *Information Technology* used to develop this DSS prototype, *.NET Microsoft* technology has been used. This includes the use of *Silverlight 5* and *ASP.NET (C#) Framework 4.0* platforms to develop the DSS user interface, C# programming language to develop the different server services and the *IIS 7.0* platform to build the DSS web server. Moreover, *Microsoft Bing Maps technology* has also been used to build the *GIS viewer* functionality.

3.3 Leakage Localization Web-Based DSS Tool: Nova Icaria DMA Application

In order to apply the leakage localization DSS prototype to Nova Icaria DMA, this tool must be integrated with the SCADA system used to supervise the DMA networks, as described in *Section 3.2*. Using this SCADA system, the pressure and flow at the inlets of every DMA are monitored. This is possible since linked to every inlet there is a multilog data logger which, on the one hand, registers these measurements every *10 minutes* and on the other, is integrated with the DMA SCADA through a GSM network. Thereby, every day at *7 a.m.*, the SCADA system retrieves the inlet measurements of all DMAs from the *00:00 h* to the *23:50 h* of the previous day. Once this raw data is retrieved, a data validation process is carried out to check and correct abnormal values inserting the resulting data in the *validated database*. In the case of Nova Icaria, *six* inner pressure sensors are also available whose measurements have a main importance to carry out the DMA leakage localization process. These sensors are linked to *GSM* dataloggers which register pressure readings every *10 minutes*. Finally, this registered information is also retrieved by the SCADA system following the same procedure as for the DMA inlets.

In order to ease the integration of the DSS prototype with the DMA measurements stored in the *SCADA validated database*, once the previous day's DMA measurements are available, this data is packed in a *XLS* file and sent using an e-mail process to the *internal database* of the DSS web-server. These particularities related to the DSS application to the Nova Icaria DMA can also be seen in Figure 3.

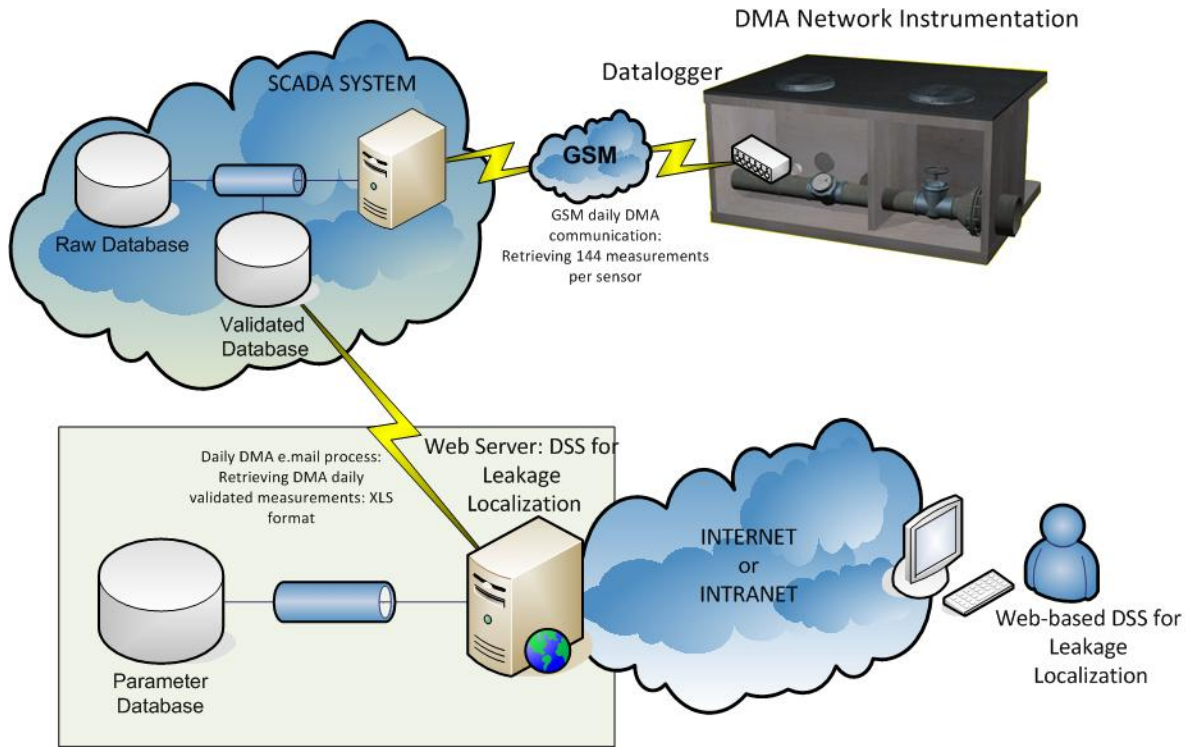


Figure 3 Conceptual scheme of the DSS prototype for leakage localization applied to Nova Icaria DMA.

4. MODEL-BASED LEAKAGE LOCALIZATION METHOD: DSS TOOL INTEGRATION

4.1 Model-based leakage localization methodology

The DSS for leakage localization in water distribution networks presented in this paper integrates the model-based leakage isolation technique proposed by (Quevedo et al., 2011) which is an evolution of the one proposed by (Pérez et al., 2011). These methodologies are based on the standard theory of model-based fault diagnosis described in *Appendix 1.2* which has already been applied to water networks to detect faults in flow meters (Ragot & Maquin, 2006) or in open channels with dynamic models (Bedjaoui & Weyer, 2011).

The analyzed system is a DMA water network, the potential faults are those leaks affecting the network components (i.e. pipes) and the process variables are the pressure at certain nodes of the water network. Hence, the leakage localization method is based on comparing, at every time instant k , the monitored pressure disturbances caused by leaks at certain inner nodes of the DMA network with the theoretical pressure disturbances caused by all potential leaks obtained using the DMA network mathematical model. Thereby, the residual set, $\mathbf{r}(k) \in \mathcal{R}^{ns}$, is determined by the difference between the measured pressure at certain network nodes, $\mathbf{p}(k) \in \mathcal{R}^{ns}$, and the estimated pressure at these nodes obtained using the network model considering a scenario free of leaks, $\hat{\mathbf{p}}_0(k) \in \mathcal{R}^{ns}$:

$$\mathbf{r}(k) = \mathbf{p}(k) - \hat{\mathbf{p}}_0(k) \quad (1)$$

The size of the residual vector \mathbf{r} , ns , depends on the number of inner¹⁵ pressure sensors of the DMA network. In (Pérez et al., 2011), an optimal pressure sensor placement for leakage localization applications was presented to achieve the minimum economical costs (number of sensors) keeping a suitable performance of the leakage localization method.

In the analysis presented in (Quevedo et al., 2011), the number of considered potential leaks, $\mathbf{f} = \{f_1, f_2, \dots, f_{np}\}$, was equal to the number of network nodes, np since from the modelling point of view, leaks were placed in these locations.

On the other hand, the theoretical pressure disturbances caused by all potential leaks are stored in the theoretical fault signature matrix, $\mathbf{FSM}(k) \in \mathfrak{R}^{ns \times np}$ (Appendix 1.2), with as many rows as DMA inner pressure sensors, ns , and as many columns as potential leaks (DMA network nodes), np . This matrix can be obtained from a sensitivity-to-leak analysis which evaluates the theoretical effect of a potential leak f_j in the pressure of all the monitored nodes, p_i , which determine the residual vector \mathbf{r} . If this process is repeated for all potential leaks, the sensitivity matrix $\mathbf{S} \in \mathfrak{R}^{ns \times np}$ (Pudar & Ligget, 1992) is obtained as follows:

$$\mathbf{S} = \begin{pmatrix} \frac{\partial p_1}{\partial f_1} & \dots & \frac{\partial p_1}{\partial f_{np}} \\ \vdots & \ddots & \vdots \\ \frac{\partial p_{ns}}{\partial f_1} & \dots & \frac{\partial p_{ns}}{\partial f_{np}} \end{pmatrix} \quad (2)$$

where each element s_{ij} measures the effect of the leak f_j in the pressure p_i of the node where the inner pressure sensor i is located. It is extremely difficult to calculate \mathbf{S} analytically in a real network because a water network is a large scale problem described by a multivariable non-linear system of equations which may also be non-explicit. Then, (Pérez et al., 2011) proposes to generate the sensitivity matrix by simulation of the network model approximating the sensitivity s_{ij} at time instant k by

$$s_{ij}(k) = \frac{\hat{p}_{if_j}(k) - \hat{p}_{i0}(k)}{f} \quad (3)$$

where $\hat{p}_{if_j}(k)$ is the predicted pressure in the node where the pressure sensor i is placed when a nominal leak of size f is forced in node j and $\hat{p}_{i0}(k)$ is the predicted pressure associated with the sensor i under a scenario free of leaks. Then, repeating this process for all np potential faults the approximation of the sensitivity matrix is obtained. In the initial method, the approximation of the sensitivity matrix \mathbf{S} (Eq. (2)) and the observed residual $\mathbf{r}(k)$ (Eq. (1)) were binarised using a certain threshold carefully selected in order to avoid filtering those pressure disturbances caused by leaks. Then, the binarised sensitivity matrix \mathbf{S} became the theoretical fault signature matrix $\mathbf{FSM}(k)$ while the binarised residual vector $\mathbf{r}(k)$ became the binary fault indicator vector $\phi(k) \in \mathfrak{R}^{ns}$ (Appendix 1.2). Then, using the fault isolation parallel inference approach (Appendix 1.2), the localization of the leak could be determined.

¹⁵ 'DMA inner sensor' expression refers to those sensors deployed in the DMA excluding those ones placed in the DMA inlets.

Nonetheless, (Quevedo et al., 2011) shows that the performance of this method could be enhanced removing those binarization operations since they could cause a loss of meaningful information for the leakage isolation process. In this case, the fault indicator vector $\phi(k)$ is obtained as follows:

$$\phi(k) = \begin{pmatrix} p_1(k) - \hat{p}_{10}(k) \\ \vdots \\ p_{ns}(k) - \hat{p}_{ns0}(k) \end{pmatrix} \quad (4)$$

where p_i is the pressure measured by sensor i and \hat{p}_{i0} is the predicted pressure value associated with pressure sensor i under a scenario free of leaks. Regarding the matrix $FSM(k)$, its expression is computed using the sensitivity matrix definition (Eq. (2)) as follows:

$$FSM(k) = \begin{pmatrix} \frac{\hat{p}_{1f_j}(k) - \hat{p}_{10}(k)}{f} & \dots & \frac{\hat{p}_{1f_{np}}(k) - \hat{p}_{10}(k)}{f} \\ \vdots & \ddots & \vdots \\ \frac{\hat{p}_{nsf_j}(k) - \hat{p}_{ns0}(k)}{f} & \dots & \frac{\hat{p}_{nsf_{np}}(k) - \hat{p}_{ns0}(k)}{f} \end{pmatrix} \quad (5)$$

where \hat{p}_{if_j} is the predicted pressure associated with pressure sensor i when a nominal leak of size f is forced in node j and \hat{p}_{i0} is the predicted pressure associated with pressure sensor i under a scenario free of leaks.

Both the matrix $FSM(k)$ and the vector $\phi(k)$ depend on the working point that is, on the demand and boundary conditions (Vento and Puig, 2009). Therefore, they must be computed at every analysis time step.

Regarding the leakage isolation process, the method proposed by (Quevedo et al., 2011) follows the parallel inference process (Appendix I.2). However, an adaptation was made to handle with the non-binary property of fault indicator vector $\phi(k)$ and matrix $FSM(k)$. In short, the proposed method is based on a correlation process which compares the fault indicator vector $\phi(k)$ (Eq. (4)) with the theoretical signatures of all potential leaks which are stored in the matrix $FSM(k)$ (Eq. (5)). This comparison is obtained applying the correlation function¹⁶. Those potential leaks whose theoretical signatures (columns of $FSM(k)$) have the biggest correlation values with the fault indicator vector $\phi(k)$ point out the most probable nodes to have the leak (Eq. (6)).

$$\max_j \left(\rho_{\phi, FSM_j}(k) \right) \quad , \quad j = 1, \dots, np \quad (6)$$

¹⁶ Pearson's correlation coefficient ρ_{x_i, x_j} between two variables x_i , x_j may be defined as

$\rho_{x_i, x_j} = \text{cov}(x_i, x_j) / \sqrt{\text{cov}(x_i, x_i)\text{cov}(x_j, x_j)}$, where $\text{cov}(\mathbf{a}, \mathbf{b}) = E[(\mathbf{a} - \bar{\mathbf{a}})(\mathbf{b} - \bar{\mathbf{b}})]$ is the covariance function between two variables \mathbf{a} and \mathbf{b} , being $\bar{\mathbf{a}} = E(\mathbf{a})$ and $\bar{\mathbf{b}} = E(\mathbf{b})$ respectively.

where $\rho_{\phi, FSM_j}(k)$ is the obtained correlation between the fault indicator vector, $\phi(k,)$ and the j^{th} -column of the theoretical fault signature matrix, FSM_j , associated with a potential leak in node j .

4.2 Integration of the leakage localization method into the proposed DSS tool

As mentioned in *Section 3.2*, the main unit of work of the proposed model-driven DSS for leakage localization is the scenario which is related to a single DMA allowing the user to run a set of analyses in this DMA. Thus, the structure of a certain analysis depends basically on leakage localization methodology presented in *Section 4.1*. Nonetheless, there are also certain scenario analysis parameters that must be informed when establishing the general DSS settings: *scenario analysis time step* (T_a), *sensor sample time* (T_s) and *diagnosis sliding window* (T_w). Thereby, T_a sets the time step used to compute a leakage localization result, T_s is the sample time of the DMA instrumentation and T_w determines the past time window considered to compute the leakage localization result. Nonetheless, it must be taken into account that there are certain constraints on the scenario analysis parameters that must be fulfilled:

$$T_a \geq T_s \quad (7)$$

when $T_a > T_s$, the sensor measurements are smoothed out during the time window determined by T_a computing their average value in order to reduce the effect of the nuisance noise affecting the measurements. Thereby, k is the discrete time considered in the leakage localization analysis being derived from the time step T_a . Another constraint is:

$$T_w = \tau T_a \quad (8)$$

where $\tau \in \mathbb{N}$ and $\tau > 1$. Here, τ can be seen as the discrete time value of the past time window T_w regarding the scenario analysis time step T_a . In this sense, at every time instant the fault indicator vector $\phi(k)$ (Eq. (4)) and the theoretical fault signature matrix $FSM(k)$ (Eq. (5)) are computed. However, when applying the procedure given by Eq. (6) to determine those leakage locations with the highest correlation values, instead of using directly these computed values at time instant k , $\phi^*(k)$ and $FSM^*(k)$ are computed using the τ -past values of $\phi(k)$ and $FSM(k)$:

$$\phi^*(k) = \begin{pmatrix} \phi(k) \\ \vdots \\ \phi(k - \tau) \end{pmatrix} \quad (9)$$

$$FSM^*(k) = \begin{pmatrix} FSM(k) \\ \vdots \\ FSM(k - \tau) \end{pmatrix} \quad (10)$$

The fact of using the τ -past values of $\phi(k)$ and $FSM(k)$, instead of their values at time instant k , adds a certain degree of robustness to the leakage localization method against certain anomalies affecting the

sensor measurements (i.e. limited sensor resolution¹⁷) or the existing uncertainty regarding the DMA water demand and hydraulic model¹⁸.

In the current version of this leakage localization DSS prototype, every planned leakage localization analysis is related to a certain date analysing the whole day. In this sense, a common value for the *scenario analysis time step* (T_a) is 1 hour. This means that a leakage localization result is given hourly for $k \geq \tau$ based on the sensor measurements and the DMA hydraulic model predictions corresponding to a τ hour - past time window. Regarding the value of the *diagnosis sliding window* parameter (T_w) and the derived τ parameter, its value depends mainly on the sensor resolution and the accuracy of the known information about the DMA hydraulic model and the DMA water demand distribution.

Regarding the structure of a given leakage localization analysis, the general sequence of steps performed at every time instant k to obtain the diagnosis result at this time instant is:

Step 1 - Calibration of the DMA hydraulic model. The main goal of this step is to obtain a calibrated DMA hydraulic model at time instant k . The main inputs used by this process are: the DMA EPANET hydraulic model, the pressure and flow measurements in the DMA inlets between time instant k and $k-1$, the base demand values of the DMA inner nodes and the value of the demand patterns at time instant k . As an output, a calibrated model of the DMA hydraulic network at time instant k is obtained following the procedure described in *Section 4.1*. Thereby, it must be taken into account that the resulting DMA hydraulic model estimates the DMA water network in a leakage-free scenario since all the flows measured in the DMA inlets are assumed to be caused by all the known DMA consumers.

Step 2 - DMA hydraulic model simulation considering potential DMA leaks. The main goal of this step is to simulate the effect of all potential leaks at time instant k in the pressure values of those monitored DMA inner nodes. The main inputs used by this step are: the calibrated DMA hydraulic model at time instant k (*Step 1*) and the locations of all the potential leaks that may exist in the DMA. *Step 2* follows an iterative process with as many iterations as considered potential leaks. In each iteration, the DMA hydraulic model is calibrated assuming the existence of a leak in a certain inner node modelled using a leak reference magnitude, specified by an emitter coefficient value in EPANET, for this node. A previous knowledge about the size of the leakage affecting the DMA (i.e. using a *minimum night flow* analysis) can be used to set the size of the reference leak through the DSS tool option related to the scenario analysis settings.

Step 3 - Computing the fault indicator vector ϕ . The main goal of this step is to compute the value of the fault indicator vector $\phi(k)$ at time instant k . The main inputs used in this step are: the pressure sensor measurements, \mathbf{p} , between time instant k and $k-1$ of those monitored DMA inner nodes and the pressure

¹⁷ Pressure sensors may have a limited resolution (i.e. 10 cm for Nova Icaria DMA inner sensors). As a result pressure disturbance caused by leaks in DMA inner nodes may not be measured accurately. The fact of using $\phi^*(k)$ and $FSM^*(k)$ may also help minimizing this negative effect

¹⁸ The use of a model-based leakage localization method based on the analysis of pressure variations produced by leakage in the water distribution network helps to add certain degree of robustness regarding existing uncertainty about DMA water demand and hydraulic model. Pressure at water distribution network nodes is less sensitive to water demand or hydraulic model errors than flows through the pipes. (Pudar & Liggett, 1992). In general, demand and modelling uncertainty can be captured using historical data free of leaks and by trying to adjust uncertainty bounds in demand and in model using a bounding approach as described, for example, in (Blesa et al., 2011)

estimation at these nodes using the calibrated DMA hydraulic model assuming a leakage-free scenario, $\hat{p}_0(k)$ (Step 1).

Step 4 – Computing the theoretical fault signature matrix FSM. The main goal of this step is to compute the value of the theoretical fault signature matrix $FSM(k)$ at time instant k . The main inputs used in this step are: the pressure estimation at time instant k of all those monitored nodes considering all DMA potential leaks, $\hat{p}_{if_j}(k)$ (Step 2) and the estimation at time instant k at these nodes using the calibrated DMA leakage-free hydraulic model, $\hat{p}_{i_0}(k)$ (Step 1). In order to obtain the value of $FSM(k)$, the procedure given by Eq. (5) is applied.

Step 5 – Leakage localization result. The main goal of this step is to give the most probable localizations of the existing leaks. This result is given when time instant $k \geq \tau$ and using τ -past values of ϕ and FSM computed in Step 3 and Step 4, respectively. Thereby, initially, $\phi^*(k)$ (Eq.(9)) and $FSM^*(k)$ (Eq. (10)) are computed and then the procedure given by Eq. (6) is applied to obtain the most probable localizations of the leaks affecting to the DMA water network.

In Figure 4, a conceptual scheme showing the above described five-step procedure can be seen highlighting the inputs and outputs of every step and the existing links among them.

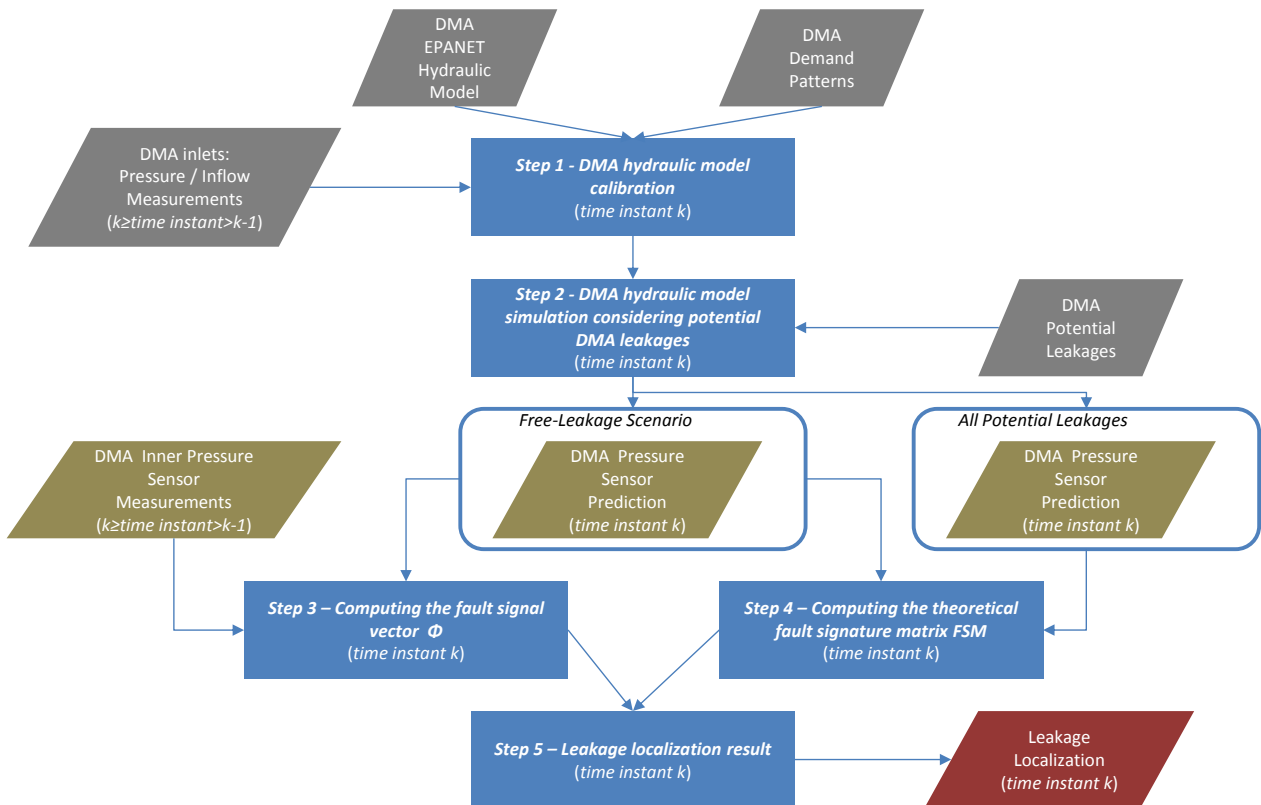


Figure 4 Conceptual scheme of the five-step procedure followed by the leakage localization analysis at time instant k

5. LEAKAGE LOCALIZATION: WEB-BASED SOFTWARE TOOL

5.1 Main Features

Considering the architecture of the web-based leakage localization DSS prototype reviewed in *Section 3*, this section describes the main features of this tool prototype offered to the end-user. Once the user has been logged in the DSS through the web browser, the first menu options that can be used are related to the *creation, edition and management* of scenarios.

The scenario creation process starts setting the scenario name and importing the EPANET file related to the hydraulic model of the DMA water network which will be analyzed in this scenario. Once the EPANET file is imported, the DMA hydraulic network is automatically displayed with the GIS viewer as a GIS layer on top of a geographic layer which shows the street map of the DMA. . A general screenshot of the *Graphical User Interface (GUI)* available when the scenario is created can be seen in Figure 5 highlighting the toolbar (label '1'), GIS options (label '2'), GIS viewer (label '3'), area displaying the last available scenario analyses (label '4') and area displaying the last diagnosed leaks (label '5').

Regarding the new scenario toolbar (Figure 5, label '1'), the following options are offered:

- *Settings*: this menu option offers different sub-options which mainly let the user set the specific scenario settings. These settings let the user fill the number of sensors placed in the DMA water network (inlets and inner nodes) mapping the nodes where they are with the variables of the SCADA *validated database* which stores the measurements metered by these sensors. Additionally, the visualization properties of the network pipes can also be edited by changing their color according to their diameter. Once these specific scenario settings have been entered, the network GIS layer is updated to show the parameterized pipe colors and a new GIS layer is added on top showing the sensor placements and highlighting the type of sensors. Finally, the scenario analysis parameters (T_a , T_s , T_w) (*Section 4.2*) can also be filled in. In Figure 6, an example of the GIS viewer associated with a given scenario can be seen where the following layers are shown: street layer, DMA network wrapping area (label '1'), DMA network pipes colored according to their diameter, DMA network sensors distinguishing between the inner pressure sensors (label '2') and the pressure / flow sensors of the DMA inlets (label '3').
- *See*: this is the toolbar option which is active by default. This menu option enables a set of GIS functions which can be applied through the different GIS layers (Figure 5, label '2'). Among these options, the following can be highlighted: '*Hide/show DMA network layer*', '*Hide/show DMA sensor layer*', '*Hide/show DMA network wrapping area*' and '*Hide/show DMA leaking pipes*'. Additionally, when clicking on top of a certain sensor, a new window is launched enabling the visualization of its measurements.
- *Projects*: this toolbar option enables the creation and management of analyses for a given scenario. Mainly, the creation process let the user set the main analysis properties: analysis name, selection of the DMA network inner pressure sensors used in this analysis, emitter coefficient related to the *reference leak size*, '*automatic*' property. Regarding the analysis management

feature, this option enables a set of possibilities to edit the analysis properties, to program daily analysis runs and to visualize the results obtained in every run. When the 'automatic' property of the analysis is marked, a daily automatic run is performed using the sensor measurements of the previous day. Otherwise, when this property is not set, manual runs must be programmed using the 'Program' option of the analysis management feature and specifying the start and end date. Finally, the 'Result' option let access to the results of all the analysis runs associated with a given scenario. This option leads to a window like the one shown in Figure 7. On the *project area* (Figure 7, label '1'), the list of available runs can be seen. Then, when a certain run is chosen, a new GIS layer is added on the top highlighting the most likely leak locations (Figure 7, label '2'). Thereby, this option also enables a shift bar in order to visualize the results obtained in every time instant of the selected run (Figure 7, label '3').

Regarding the *scenario management* and the *scenario edition* options available after the login, the former allows the user to choose one of the available scenarios in the DSS. Once a scenario has been chosen, the scenario toolbar (Figure 5, label '1') and all their features are available. Finally, regarding the *scenario edition* menu option, this feature lets the user update the DMA hydraulic model associated with a given scenario enabling an import process of the DMA water network EPANET file. Once this import process has ended, the scenario and its associated analysis settings can be updated using the *scenario management* option (Figure 5, label '1').

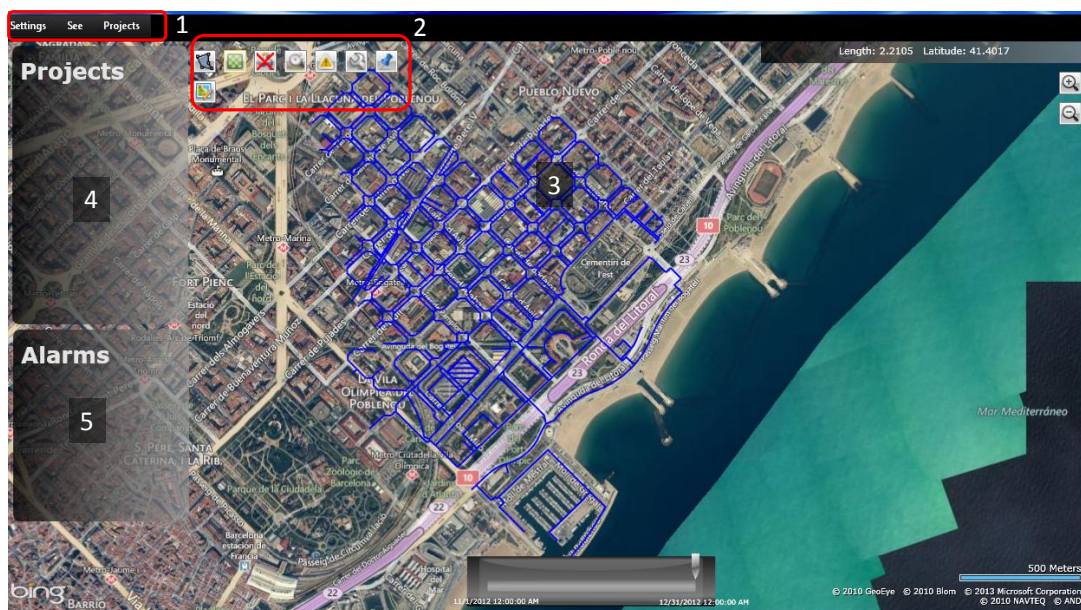


Figure 5 Graphical User Interface available when the scenario is created highlighting the toolbar (label '1'), GIS options (label '2'), GIS viewer (label '3'), area displaying last available scenario analyses (label '4') and area displaying last diagnosed leaks (label '5').

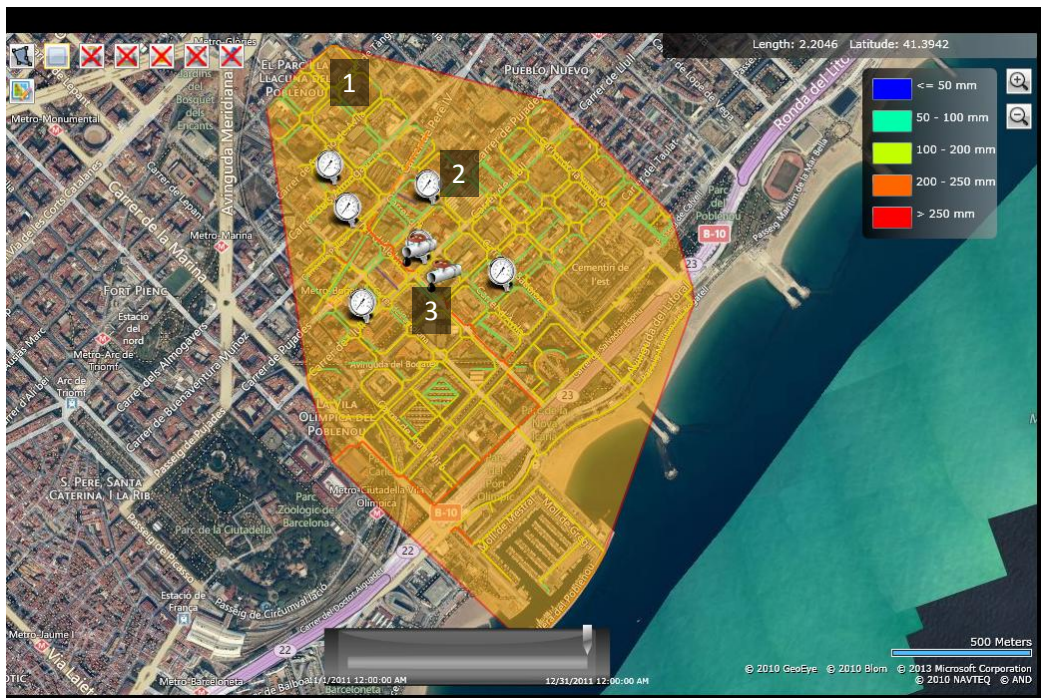


Figure 6 Example of the GIS viewer displaying the following layers : street layer, DMA network wrapping area (label ‘1’), DMA network pipes coloured according to their diameter, DMA network sensors distinguishing between the inner pressure sensors (label ‘2’) and the pressure / flow sensors of the DMA inlets (label ‘3’).



Figure 7 Visualization of the analysis results obtained for a given run: list of available runs (label ‘1’), resulting most likely leak locations for a given time instant (label ‘2’), shift bar to visualize the results obtained in every time instant of the selected run (label ‘3’).

5.2 Leakage Localization DSS Prototype: Leakage Scenario Analysis applied to Nova Icaria DMA

This section shows the results obtained by the leakage localization DSS prototype when applied to the study of the historical leakage scenarios described in *Section 2.3*. Mainly, two set of analyses were carried

out: one using the instrumentation data corresponding to a few days before the 6 l/s-leak was found and another one, using the data corresponding to a few days after this first leak was found. In these analyses, the *scenario analysis time step* (T_a) parameter was set to 1 hour while the *sensor sampling time* (T_s) parameter was set to 10 minutes . This means that a whole daily analysis (*Section 4.2*) was structured in 24 iterations (one for each hour) where in every iteration the hydraulic model was calibrated using the average value of the sensor measurements corresponding to the last hour. Additionally, the *diagnosis sliding window* (T_w) was set to 5 hours (i.e. $\tau=5$ in (Eq. (8))) and consequently, the scenario analyses were set up to provide an hourly leakage localization result as of the 5^{th} iteration since the constraint $k \geq \tau$ must be satisfied in order to apply the procedure given by Eq. (6) . Regarding the *reference leak size* used in the analyses, this parameter was set to 6 l/s given that this value could be easily derived from the minimum night flow analysis conducted by the operator.

Figure 8 summarizes the results obtained in the first set of analyses carried out with instrumentation data corresponding to a few days before the 6 l/s-leak was found. This figure corresponds to the DSS GIS viewer where the DMA water network is colored in blue, the inner pressure sensors are plotted using meter icons (Figure 6, label ‘2’) while the DMA inlets are signaled using valve icons (Figure 6, label ‘3’). Additionally, round marker icons (Figure 6, label ‘2’) can also be seen which determine the top rated possible leakage localizations proposed by the DSS once an hourly result has been computed. On the other hand, on the top of the GIS layers, the known localization of 6 l/s-leakage and the set of remaining leaks (4 l/s) have been highlighted using a red arrow icon and a red square area, respectively. In between these two leakage spots, a green square area has also been plotted highlighting the top rated localizations proposed by the DSS prototype in this first group of analyses. The leakage localization area proposed by the DSS may be seen as quite acceptable since this prototype is assuming the existence of just a single leak and in this case, the DMA hydraulic network was affected by two (or more) simultaneous leaks and consequently, the DMA sensors were registering the conjunctive effect of these two abnormalities. Additionally, it must be taken into account that the fact of just using few DMA inner sensors does not let the user distinguish the effect of all the potential leaks and as a result, an area which may contain the leakage localization must be always given as a result.

Similarly, Figure 9 summarizes the results obtained in the second set of analyses carried out with instrumentation data corresponding to a few days after the 6 l/s-leak was found. In this case, this leak was already repaired but a set of small leaks whose overall value was around 4 l/s remained whose known localizations were inside the highlighted red square area (Figure 9). As can be seen in Figure 9, once the 6 l/s-leakage was repaired the green square area indicating the DSS top rated leakage localizations move into the red square area producing a satisfactory result. Unlike the case presented above in Figure 8, in this case just a single but distributed leak was affecting the DMA network satisfying the DSS prototype assumption of the existence of just a single leak.

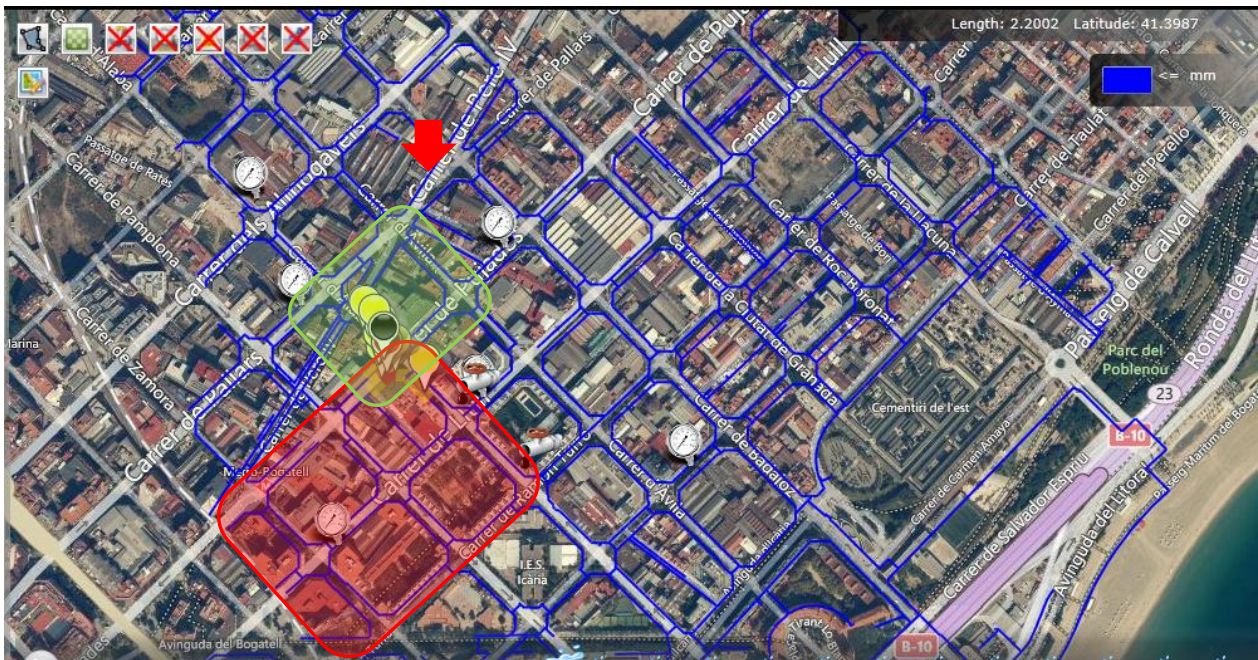


Figure 8 DSS GIS viewer: leakage localization analysis before the 6 l/s leakage was found (red arrow) highlighting in a green square area the most rated localizations proposed by the DSS and in red square area the known localizations of the remaining leaks (4 l/s).

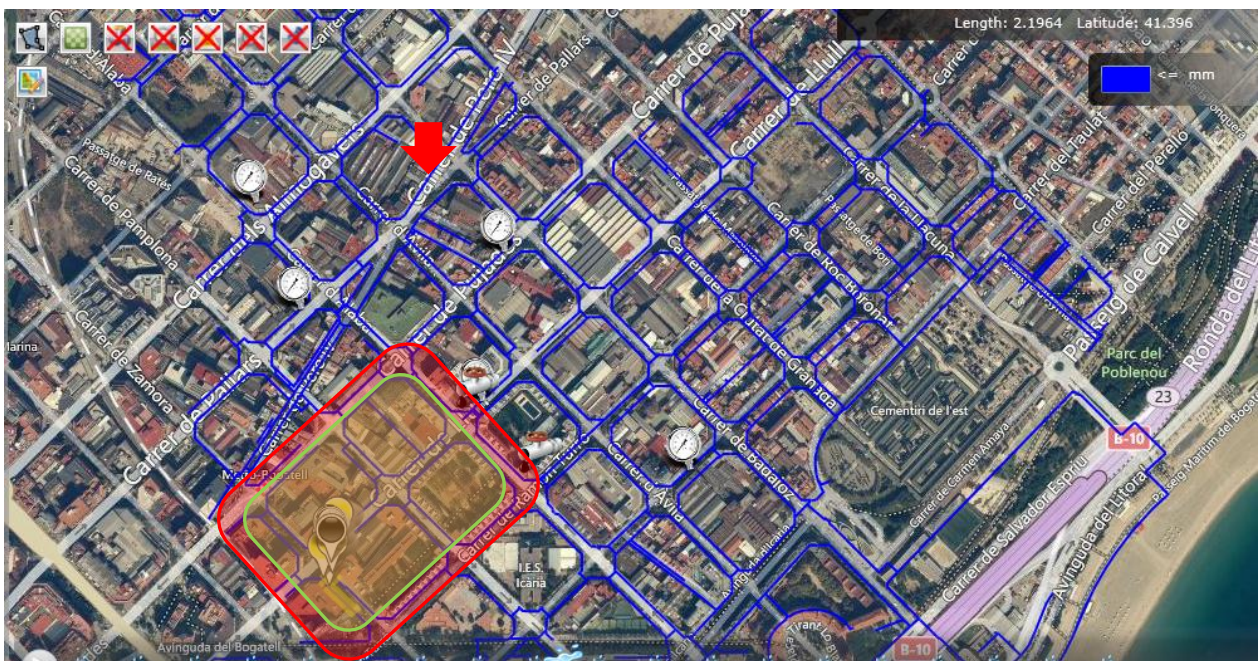


Figure 9 DSS GIS viewer: leakage localization analysis after the 6 l/s leakage was found (red arrow) highlighting in a green square area the most rated localizations proposed by the DSS and in red square area the known localizations of the remaining leaks (4 l/s).

Regarding the goodness of the leakage localization results, the experience gained in the Nova Icaria shows certain important aspects that must be taken into account and which may have a deep influence on the goodness of the obtained results. The most important one is the need of using an accurate hydraulic model not just in its topological structure but also in its parameters (i.e. pipe roughness, length, diameter;

junction and sensor levels). Additionally, having an accurate estimation of the DMA demand distribution (i.e. junction demand base and demand pattern) is also relevant, because certain errors affecting the hydraulic model or the demand distribution estimation may result in inaccurate leakage localization results. In the Nova Icaria pilot, this was obtained through the Aigües de Barcelona billing information. Furthermore, there may exist certain sensor distributions that may add some degree of robustness to minimize those above mentioned negative effects. It must also be taken into account that sensors have a constrained resolution which may cause failure to register the disturbances caused by leaks. This fact may also have a negative effect in the given leakage localization result. However, a careful selection of the *diagnosis sliding window* (T_w) may help overcome this problem.

6. CONCLUSIONS

A web-based model-driven DSS tool integrating an efficient model-based leakage localization method has been described in this paper. This general-purpose software tool makes it easier to implement real-time monitoring in large class of water utilities and may be used by water network managers as a support system without the need for any special expertise in leakage localization models. A pilot was implemented in a DMA of the Barcelona distribution network with 1996 nodes, 3442 pipes and two measured inputs: its network length is approx. 43 km distributed in an area around 227 hm². Six pressure sensors were installed inside the DMA. Data were sampled at 10-minutes intervals and sent on a daily basis to a central station where the mathematical models run. This was achieved through the integration of the web-based model-driven DSS tool implementing the leakage localization method with the SCADA system. The paper provides the results of a validation period using this tool in off-line simulation with real data of the pilot DMA, from September through October 2011. The results show that, using this tool, the leakage detection and localization may be performed efficiently and the leak detection/localization time is significantly reduced. In general, this result is a consequence of the structure of the proposed model-driven DSS tool which allows the integration of the DMA instrumentation (SCADA system), the hydraulic model of the DMA network and the model-based leakage localization algorithms with an intuitive Graphical User Interface (GUI) that will allow the easy creation and modification of on-line automated or manual leakage location analyses and the visualization of both the structure of the DMA water network and the obtained results using a GIS viewer. On the other hand, the web nature of the proposed DSS tool eases the access of the users from a wide range of devices and its deployment process which may speed the decision-making process related to every unexpected event affecting the water network.

Additionally, certain limiting factors on the ability of the methodology to detect and locate leaks are reported, namely: the need for a well-calibrated hydraulic model (topological structure and parameters), the accuracy in estimating the spatial demand distribution within the DMA and the precision of the sensors installed in the network.

The acquired knowledge, the mathematical methods and the software tool developed in this work will allow water operators to detect and locate leaks events earlier and with better accuracy than using today's systems. The software tool has been designed to work automatically based on real sensor data providing information to network managers which will help their decision making. Finally, regarding the sustainability, the use of this system is expected to produce important savings in water.

Furthermore, the diagnosis methodology and the software platform are amenable for use in quality monitoring as well. In an ongoing research project, the authors will address the problem of quality-breach detection in a DMA, as well as maintaining the required quality inside the margins stipulated by the law in each location where the method it is applied.

ACKNOWLEDGMENTS

The authors wish to thank the support received by the AM0901 project funded by R+i Alliance (Suez Environnement) and by the EFFINET grant FP7-ICT-2012-318556 of the European Commission.

APPENDIX: BACKGROUND

1.1 Water Distribution system: Mathematical Modeling

In general, a water distribution system consists of four major types of components: pumps, valves, distribution storage tanks, and pipes. Most systems require booster pumps to overcome differences in elevation, and energy losses caused by friction. Pipes may contain flow or pressure-control devices, such as regulating or pressure-reducing valves (Brdys & Ulanicki, 1994). The different elements of the water distribution system are interconnected to produce a network of nodes and branches whose main purpose is to supply final 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. 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. Therefore, in a pipe, or a junction node, the inflow and the outflow must balance. For a junction node

$$\sum q_{in} - \sum q_{out} = q_{ext} \quad (A.1)$$

where q_{in} and q_{out} are the pipe flow rates into and out of the node, respectively, and q_{ext} is the external demand or supply. Conservation of energy states that the difference in energy (usually named head loss) between two points is equal to any 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

$$\sum_{u \in U_p} h_{p,u} - \sum_{v \in V_p} h_{L,v} = \Delta E \quad (\text{A.2})$$

where $h_{L,v}$ is the head loss across component v along the path, $h_{p,u}$ is the head added by pump u , 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:

$$h_L = Kq^\gamma \quad (\text{A.3})$$

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

In general, hydraulic models for water networks are widely known and used (Brdys & Ulanicki, 1994) and in several approaches, hydraulic modelling has also been used for leakage localization techniques: (Ligget & Chen, 1995; Mandel, 1998; Kapelan et al., 2003; Almandoz et al., 2005; Brown, 2007; Wu et al. 2010; Goulet et al. 2013). In these cases, the dependence of leaks on pressure (Greyvenstein, 2007; Lambert, 1997) is the keystone in the detection, localisation and control.

Model-based leakage localization methods are based on assuming that hydraulic models describe correctly the real hydraulic response of the network. Model representativeness relies strongly on the correct identification of the constants or parameters characterizing the equations of the hydraulic water network model. Calibration of models is itself a main inverse problem previous to any application of the model (Carpentier et al., 1991) (Savić et al., 2009). The total demand in a DMA can usually be approximated in real time by adding flow-meter readings into the area.

In model-based leakage localization techniques, a common procedure (Pérez et al. 2011; Wu, 2008) for analysing the effect of a certain leak is to assign part of the total metered inflow to the leak, and account for this as a pressure-dependent demand. In water network analysis, a leak is often based on the equation of flow through an orifice (Rossman, 2000):

$$q_l(k) = C(p_i(k) - p_o(k))^\beta \quad (\text{A.4})$$

where p_i and p_o are pressures in and out of the pipe which are in general unknown or partially known; q_l the flow of the leakage; C and β are constants.

Different approaches exist to calibrate such hydraulic models depending on whether deterministic or heuristic methods are used. Deterministic methods seek the minimum of an objective function. Their efficiency depends on different factors such as the starting point, the accuracy of the descent direction evaluation given by the gradient and the method used. The obtained solution is generally a local minimum, this being of its main drawbacks. Another disadvantage is the need for gradient evaluations while the calibration process is engaged. Quite a number of works can be found in the literature about these methods, see for instance (Lansley et al., 2001) describing a General Reduced Gradient method to solve the steady-state calibration hydraulic problem; (Datta & Sridharan, 1994) describing the use of a *Sensitivity Analysis Technique* belonging to the Newton methods family; (Pudar & Ligget, 1992) using the

Levenberg-Marquardt approach; (Di Cristo et al., 2003) presented a steady-state hydraulic parameter calibration procedure using a gradient method; (Greco et al., 1999) formulated the minimizing problem as a quadratic programming problem or (Todini, 1999) who used a Kalman filtering approach. On the other hand, one of the main characteristics of adaptive heuristic methods is their ability to move out of local optima and globally sample the search space. Among them, Genetic Algorithms (*GAs*) and Bayesian methods have been widely used in recent years for a variety of water distribution system problems. (Savic & Walters, 1995) were the first to apply *GA* to solve a hydraulic model calibration problem formulated as an inverse problem. This method might be applied to large network with different parameters (demands, pipe diameters...) other than the roughness factors. (Lingireddy & Ormsbee, 1998) presented a *GA* to calibrate water network hydraulic models for demand adjustments. The presented approach was successfully applied to a case study network. (Kapelan et al., 2007) used a Bayesian recursive method to calibrate the hydraulic model and estimate uncertainties.

1.2 Model-Based Fault Detection and Isolation: Leakage Localization in Water Networks

Model-based leakage localization techniques like the ones presented by Pérez et al. (2011) and Quevedo et al. (2011) can be seen as a specific application of fault detection and isolation theory based on analytical models to complex industrial systems. In this case, the analysed system is the DMA water network and, the potential faults are those leaks affecting the network components (i.e. pipes). The fault detection process determines the existence of leaks affecting the water network and then, the fault isolation process locates the place where those leaks are.

In a general framework, fault detection is the process of generating fault indicators ($\phi(k) \in \mathfrak{R}^{n_\phi}$) on the grounds of the process variables ($x(k) \in \mathfrak{R}^{n_x}$) in order to detect faults. Thereby, detection algorithms should generate fault indicators which ought to contain information about faults. The mapping of the space of process variables into the space of fault indicators as well as the evaluation of these signals in order to detect and indicate the fault takes place during the detection stage.

On the other hand, fault isolation is carried out on the basis of fault indicators generated by the detection module. The result of isolation is a diagnosis showing the faults or states of the system. The knowledge of the relationship between the fault indicators and the faults or the technical states of the system is necessary in order to perform the fault isolation. However, a completely reliable and unequivocal presentation of the existing faults or the definition of the diagnosed system state is not always possible due to incomplete and uncertain knowledge of the system, limited isolation among faults or states, uncertainty of fault indicators, etc.

Fault detection based on analytical model is based on generating a residual comparing the measurements of physical variables $y(k) \in \mathfrak{R}^{n_y}$ of the process with their estimation $\hat{y}(k) \in \mathfrak{R}^{n_y}$ provided by the associated system model:

$$r(k) = y(k) - \hat{y}(k) \quad (\text{A.5})$$

where $\mathbf{r}(k) \in \mathcal{R}^{ny}$ is the residual set and is computed at every time instant using the measurements of the system inputs ($\mathbf{u}(k)$) and outputs ($\mathbf{y}(k)$). Then, the generation of residuals needs to be followed by residual evaluation, in order to arrive at detection and isolation decisions. Because of the presence of noise and model errors, the residuals are never zero, even if there is no fault. Therefore, the residual evaluation requires testing the residuals against thresholds, obtained empirically or by theoretical considerations. The fault detection task consists in deciding if there is a fault affecting the monitored system by checking each residual $r_i(k)$ of the residual set against a threshold that takes into account model uncertainty, noise and the unknown disturbances. The result of this test applied to every residual $r_i(k)$ produces an *observed fault signature* $\phi(k)$. A basic way of obtaining these observed fault indicators could be through a binary evaluation of every residual $r_i(k)$ against a threshold τ_i (Gertler, 1998):

$$\phi_i(k) = \begin{cases} 0 & \text{if } |r_i(k)| \leq \tau_i \\ 1 & \text{if } |r_i(k)| > \tau_i \end{cases} \quad (\text{A.6})$$

According to (Kościelny et al., 2004), fault isolation is carried out on the basis of fault indicators, ϕ , generated by the detection module. Thus, the result of fault isolation consists in showing the faults affecting to the system, which requires knowing the relationship between the fault indicators, ϕ , and the faults, $\mathbf{f}(k) \in \mathcal{R}^{nf}$. Models applied to fault isolation or system state recognition should therefore map the space of fault indicators ϕ into the discrete space of faults \mathbf{f} . In general, in fault isolation based on analytical model, the binary *theoretical fault signature matrix* (\mathbf{FSM}) (Gertler, 1998) is used to store the relationship between the fault indicators, ϕ , and the faults, \mathbf{f} . This matrix is defined on the Cartesian product of the set of faults \mathbf{f} and fault indicators ϕ :

$$\mathbf{FSM} \subset \phi \times \mathbf{f} \quad (\text{A.7})$$

This matrix stores the binary influence of a given fault f_j (column of \mathbf{FSM}) on a given fault indicator ϕ_i (row of \mathbf{FSM}). Thereby, if the element \mathbf{FSM}_{ij} of this matrix is equal to '1', it means the fault f_j causes the occurrence of the fault indicator ϕ_i . Otherwise, when the fault f_j has no effect on the fault indicator ϕ_i , the element \mathbf{FSM}_{ij} is equal to '0'. Such as the \mathbf{FSM} matrix is defined, its j^{th} -column contains the binary effect of fault f_j on the fault indicator set ϕ . In consequence, this column is known as the theoretical fault signature of f_j . Fault isolation methods using the binary fault isolation matrix can isolate faults comparing the value of the observed fault indicator with the information stored in that matrix \mathbf{FSM} . In general, a parallel inference approach (Gertler, 1998) assuming a single fault scenario is used to isolate the fault. Parallel diagnostic inference based on the binary diagnostic matrix consists in formulating a fault isolation result comparing the observed binary fault indicators ϕ ($\phi_i \in \{0,1\}$) with the theoretical fault signature associated with all considered fault hypotheses determined by the set \mathbf{f} and stored in the \mathbf{FSM} matrix.

Thereby, the inference procedure ($\phi \rightarrow f$) consists in comparing the binary fault indicators $\phi_i(k)$ (Eq. (A.6)) computed at every time instant by the fault detection module with the theoretical binary fault signature related to every fault hypothesis f_j of the set f which is stored in the j^{th} -column of the binary matrix FSM . If all fault indicators are zero-valued, the fault isolation module shows a lack of faults:

$$\forall \phi_i \in \phi : \phi_i = 0 \Rightarrow F = \mathbf{0} \quad (\text{A.8})$$

where F is the set of fault hypothesis f_j which are consistent with the observed fault indicators.

When some fault indicator values equal one according to the residual evaluation carried out by the fault detection module, the fault isolation algorithm gives as diagnosis result a subset of the fault hypothesis set f whose signatures are consistent with the observations:

$$F = \{f_j \in f : \phi_i = FSM_{ij}, \forall \phi_i \in \phi\} \quad (\text{A.9})$$

Then, a general approach of carrying out the comparison between the observed fault signature $\phi(k)$ and the theoretical one related to every fault hypotheses is calculating the distance between both vectors: $\phi(k)$ and the j^{th} -column of matrix FSM for the hypothesis f_j , *e.g.* using the Hamming distance measurement. As a result of this comparison, a distance measurement $d_j(k)$ is obtained for every fault hypothesis f_j , being $d(k)$ the vector of all the computed distances at time instant k : $d(k) = [d_1(k), \dots, d_j(k), \dots, d_{n_f}(k)]$. If the Hamming distance approach is applied, then

$$d_j(k) = \sum_{i=1}^{n_\phi} (FSM_{ij} \oplus \phi_i(k)) \quad (\text{A.10})$$

where \oplus is the XOR logic operator. Then, the fault hypotheses with the shortest distance regarding the current observed fault signature ϕ are considered as the fault isolation result:

$$F = \left\{ f_j \in f : \forall j \text{ where } d_j(k) = \min_{v \in \{1, \dots, n_f\}} [d_v(k)] \right\} \quad (\text{A.11})$$

This approach gives a simple idea of the fault isolation problem but it may be affected by certain drawbacks which can lead towards a wrong diagnosis result (Puig et al., 2005).

REFERENCES

- Almandoz J., Arregui F., Cabrera E., Cobacho R. (2005). Leakage Assessment through Water Distribution Network Simulation. J. Water Resour. Plng. and Mgmt., Volume 131, Issue 6, pp. 458-466, November/December.
- Armon, A, Gutner, S, Rosenberg, A, & Scolnicov, H. (2011). Algorithmic network monitoring for a modern water utility: a case study in Jerusalem. Water science and technology 63(2) 233-239.
- AWWA (2009) Manual of water supply practices – M36, Third Edition; Water Audits and Loss Control Programs.
- Bakker, M., Vreeburg, J.H.G., van Schagen, J.H.G., Rietveld, L.C. (2013) A fully adaptive forecasting model for short-term drinking water demand, Environmental Modelling & Software, Volume 48, October 2013, Pages 141-151, ISSN 1364-8152.
- Bedjaoui, N., & Weyer, E. (2011). Algorithms for leak detection, estimation, isolation and localization in open water channels. Control Engineering Practice, 19(6), 564-573.

- Bicik, J., Kapelan, Z., Makropoulos, C. and Savic, D. (2013) A Pipe Burst Diagnostics Using Evidence Theory, *Journal of Hydroinformatics*, pp 596-608, vol. 13(4)
- Blesa, J., Puig, V., Saludes, J. (2011) "Identification for passive robust fault detection using zonotope-based set-membership approaches". *International Journal of Adaptive Control and Signal Processing*, Volume 25, Issue 9, pages 788–812, September 2011.
- Brdys, M. A., Ulanicki, B., (1994) State and parameter estimation, *Operational Control of Water Systems: Structures, Algorithms and Applications*. Prentice Hall International (UK) Ltd., pp. 96-131.
- Brown K. (2007) Modeling Leakage in Water Distribution Systems Florida State University Thesis.
- Carpentier P., Cohen G. (1991) State Estimation and Leak Detection in Water Distribution Networks
- Datta, R. S. N., Sridharan, K., (1994). Parameter estimation in water distribution systems by least squares. *Journal of Water Resources Planning and Management* 120 (4), 405-422.
- Di Cristo C., Leopardi A. (2003) Inverse problems in hydraulic network modelling Proceedings of PEDS2003 Conference, Valencia, Balkema, pp.163-171.
- Farley, B., Mounce, S. R., Boxall, J. B., (2010) Field Testing of an Optimal Sensor Placement Methodology in an Urban Water Distribution Network for Event Detection *Urban Water Journal* Vol 7 (6), pp 345-356
- Gertler J.J. (1998) "Fault Detection and Diagnosis in Engineering Systems." Marcel Dekker. 1998
- Carlo Giupponi, Decision Support Systems for implementing the European Water Framework Directive: The MULINO approach, *Environmental Modelling & Software*, Volume 22, Issue 2, February 2007, Pages 248-258, ISSN 1364-8152.
- Greco, M., Del Giudice, G., (1999) New approach to water distribution network calibration *Journal of Hydraulic Engineering* 125 (8), 849-854.
- Goulet J.-A., Coutu S., Smith I.F.C., Model falsification diagnosis and sensor placement for leak detection in pressurized pipe networks, *Journal of Advanced Engineering Informatics*, Volume 27, Issue 2, April 2013, Pages 261-269.
- Greyvenstein B. (2007) An experimental Investigation into the Pressure Leakage Relationship of Some Failed Water Pipes *Journal of Water Supply, Research and Technology - AQUA, IWA*.
- Kapelan, Z., Savic, D. A., Walters, G. A., (2003). A hybrid inverse transient model for leakage detection and roughness calibration in pipe networks. *Journal of Hydraulic Research* 41 (5), 481–492.
- Kapelan, Z., Savic, D.A. and Walters, G.A., (2007). Calibration of WDS Hydraulic Models using the Bayesian Recursive Procedure, *Journal of Hydraulic Engineering*, 133(8), 927-936.
- Kościelny, J.M., Korbicz, J., Kowalczyk, Z., Cholewa, W. (2004). "Fault Diagnosis". Springer-Verlag,
- International Water Association (IWA). (2000). Losses from Water Supply System: Standard Terminology and Recommended Performance Measure. IWA Task Force on Water Loss, London.
- Lansley, K. E., El-Shorbagy, W., Ahmed, I., Araujo, J., Haan, C. T., (2001). Calibration assessment and data collection for water distribution networks *Journal of Hydraulic Engineering* 127 (4), 270-279.
- Laucelli, D., Berardi, L., Giustolisi, O. (2012) Assessing climate change and asset deterioration impacts on water distribution networks: Demand-driven or pressure-driven network modeling?, *Environmental Modelling & Software*, Volume 37, November 2012, Pages 206-216, ISSN 1364-8152,
- Liggett, J. A., Chen, L. C., (1995). Monitoring water distribution systems. the inverse method as a tool for calibration and leak detection. In: Cabrera, E., Vela, A. F. (Eds.), *Improving Efficiency and Reliability in Water Distribution Systems*. Vol. 14. Kluwer Academic, Boston, USA, pp. 107–134.
- Lingireddy, S., Ormsbee, L. E., (1998). Optimal network calibration model based on genetic algorithms. Tech. rep., University of Kentucky, USA.
- Makropoulos, C.K., Natsis, K., Liu, S., Mittas, K., Butler, D. (2008) Decision support for sustainable option selection in integrated urban water management, *Environmental Modelling & Software*, Volume 23, Issue 12, December 2008, Pages 1448-1460, ISSN 1364-8152
- Mandel, D. (1998). Diagnostic a base de redondance analytique: Application à un réseau urbain de distribution d'eau potable. Thèse de Doctorat. CRAN. Nancy. France.

- Ocampo-Martínez, C., Puig, V., Cembrano, G., and Quevedo, J. (2013) Application of predictive control strategies to the management of complex networks in the urban water cycle, *IEEE Control Systems Magazine*, 33:15-45.
- Ofwat, 2008. International comparison – leakage. Available from: http://www.ofwat.gov.uk/regulating/reporting/rpt_int_08leakageintro.
- Pérez, R., Puig, V., Pascual, J., Quevedo, J., Landeros, E., Peralta, A., (2011). Methodology for leakage isolation using pressure sensitivity analysis in water distribution networks. *Control Engineering Practice* 19, 1157–1167.
- Pudar, R.S., & Liggett, J.A. (1992). Leaks in pipe networks. *Journal of Hydraulic Engineering*, 118(7), 1031–1046.
- Puig, V. J. Quevedo, T. Escobet, and B. Pulido. “A New Fault Diagnosis Algorithm that Improves the Integration of Fault Detection and Isolation” in *Proceedings of ECC-CDC’05*, Sevilla, Spain. 2005.
- Puust, R., Kapelan, Z., Savic, D. and Koppel, T. (2010): A review of methods for leakage management in pipe networks, *Urban Water Journal*, 7:1, 25-45
- Quevedo, J., Cugueró, M., Pérez, R., Nejjari, F., Puig, V., Mirats, J. (2011) Leakage location in water distribution networks based on correlation measurement of pressure sensors. *IWA Symposium on Systems Analysis and Integrated Assessment*, San Sebastián.
- Pouget, L., Escaler, I., Guiu, R., Mc Ennis, S., Versini, P. (2012) Global Change adaptation in water resources management: The Water Change project, *Science of The Total Environment*, Volume 440, 1 December 2012, Pages 186-193, ISSN 0048-9697
- Power, D.J. (2004). *Free Decision Support Systems Glossary*, <http://dssresources.com/glossary/>
- Power, D.J., Sharda, R., (2007). Model-driven decision support systems: Concepts and research directions. *Decision Support Systems* 43, 1044–1061.
- Ragot, J., Maquin, D., 2006. Fault measurement detection in an urban water supply network. *Journal of Process Control* 16, 887–902.
- Rani, D. and Moreira, M. M., (2010). ‘Simulation–Optimization Modeling: A Survey and Potential Application in Reservoir Systems Operation’, *Water Resources Management*, vol. 24, no. 6, pp. 1107–1138, Jul. 2009.
- Romano, M., Kapelan, Z. and Savic, D.A., (2013), "Geostatistical Techniques for Approximate Location Of Pipe Failure Events in Water Distribution Systems", *Journal of Hydroinformatics* Vol 15 No 3 pp 634-651 IWA Publishing 2013
- Rossman, L. (2000). *EPANET 2 - User's Manual*. Water Supply and Water Resources Division, National Risk Management Research Laboratory.
- Savić, D. A., Walters, G. A., (1995). Genetic algorithm techniques for calibrating network models. *Tech. Rep.95/12*, Centre for Systems and Control Engineering, University of Exeter, UK.
- Savić, D.A., Bicik, J., Morley, M.S., (2011). A DSS generator for multiobjective optimisation of spreadsheet-based models. *Environmental Modelling & Software* 26, 551–561.
- Savić, D. A., Kapelan, Z., and Jonkergouw, P. (2009). "Quo vadis water distribution model calibration?" *Urban Water Journal*, 6(1), 3-22.
- Todini, E., (1999) Using a kalman filter approach for looped water distribution network calibration In: Savić, D. A., Walters, G. A. (Eds.), *Proceedings Water Industry Systems: Modelling and Optimization Applications*. Vol. 1. University of Exeter, UK, pp. 327-336.
- 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.
- Wierzbicki, A. P., Makowski, M., Wessels, J. (2000) *Model-Based Decision Support Methodology with Environmental Applications*. *Mathematical Modelling: Theory and Applications*, Vol. 9, (Eds.) 2000, XVII, 475 p (Springer).
- Wu, Z. Y., Sage, P. (2006) Water Loss detection via genetic algorithm optimization-based model calibration. *8th Annual International Symposium on Water Distribution System Analysis*, Cincinnati, Ohio, August 27-30.
- Wu, Z. Y. (2008) *Innovative Optimization Model for Water Distribution Leakage Detection*. Watertown, USA: Bentley Systems Inc.
- Wu, Z. Y., Sage, P., Turtle, D. (2010) Pressure-Dependent Leak Detection Model and Its Application to a District Water System. *Journal of Water Resources Planning and Management*; ASCE / January – February 2010.