Detecting leak regions through model falsification

GAUDENZ MOSER AND IAN F.C. SMITH
Applied Computing and Mechanics Laboratory (IMAC)
School of Architecture, Civil and Environmental Engineering (ENAC)
École Polytechnique Fédérale de Lausanne (EPFL)
Lausanne, Switzerland
gaudenz.moser@epfl.ch

Abstract Pressurized fluid-distribution networks are strategic elements of infrastructure. In the case of fresh-water distribution networks, advanced sensor-based diagnostic methodologies have the potential to provide enhanced management support. Since a significant percentage of fresh water is lost globally due to leaks in these networks, the challenge to improve performance is compatible with goals of sustainable development. The scope of this research includes the diagnosis of water-distribution networks and more generally, pressurized fluid-distribution networks through development of model-based data-interpretation methods to assess performance. The strategy of model falsification is combined with network reduction techniques to obtain reliable and computationally efficient diagnoses. A case study involves the detection of leaks from an initial set of 263 leak scenarios. Preliminary results show that this methodology has the potential to detect leak regions, even with a small number of sensors.

1. Introduction

Pressurized fluid-distribution networks such as water distribution networks are key strategic elements of infrastructure. Drinking water is a precious resource that is necessary to preserve. Preservation involves reducing losses in the water distribution networks. A study carried out by the World Bank (Kingdom et al., 2006) has shown that each year, 32 billion cubic meters of water are lost through leaks around the world and 30% of this has occurred in developed countries. These numbers justify why efficient monitoring systems for the detection and localization of leaks are needed to reduce the volume of water that is lost. Advanced sensor-based diagnostic methodologies have the potential to provide enhanced management support.

Leaks in fresh-water distribution are not a recent challenge. Water-distribution networks are important and vulnerable elements of infrastructure. As early as 1892, Hope studied water losses in public supplies (Hope and Bircumshaw, 1996). Examples of leak-detection methods were described by Babbitt et al. (1920). Descriptions of basic inspection methods such as visual observation and sounding through the soil with a steel rod were discussed as well as more advanced techniques such as water-hammer techniques and acoustic measurements.

There are several categories of leak detection techniques. One is transient-based techniques those use pressure measurement. These techniques use the measured transient signal to detect leaks. Colombo et al. (2009) completed a review of transient-based leak detection methods and sorted them into three types: inverse-transient analysis (Vítkovský et al., 2000, 2007), frequency-domain techniques and direct transient analysis. The uncertainty associated with the system affects the results of transient-based techniques. Therefore they are primarily used on single, underground pipelines (Puust et al., 2010). These techniques need further development before they can be applied to complex pipe networks such as urban water-distribution systems.

Other technics are based on the comparison of measurements with predictions obtained from hydraulic models. This challenge is often formulated as an optimisation task. The goal is to minimize the differences between the measurements taken on the network and predicted values from flow models. Such techniques are based on minimization of least-squares and they were developed by Pudar (1992) and Liggett and Andersen and Powell (2000). Another methodology used to solve this problem is to use Bayesian inference. Poulakis et al. (2003) have proposed a Bayesian system-identification methodology for leakage detection. Other studies that used Bayesian inference for leak detection were presented by Rougier (2005), Puust et al. (2006) and Barandouzi et al. (2012).

The applicability of these methodologies to real networks may be limited under certain circumstances. For example, transient-based techniques may not appropriate for large complex networks. Furthermore, the hypotheses made when using traditional residual minimization and Bayesian inference techniques are usually hard to meet because of the systematic modelling errors and unknown relationships between uncertainties.

This paper presents a model-based methodology that accommodates systematic uncertainties and is robust in the presence of unrecognized correlations. Section 2 describes the diagnosis methodology in more detail. Section 3 explains how this methodology has been adapted to leak detection. Section 4 illustrates the usefulness of the methodology through an application to a part of the Lausanne water distribution network and finally section 5 contains conclusions and a discussion of more general impact of this work.

2. Model Falsification

Goulet and Smith (2012) developed a model falsification methodology for diagnosis called error-domain model falsification. This methodology is most useful in cases where little information is available to describe the relationships between uncertainties at several prediction locations. The methodology uses only the marginal uncertainty distribution at each location where predictions and measurements are compared. Prior knowledge is used to define bounds for the parameters values to identify and to build sets of possible scenarios. A scenario corresponds to a set of parameter values describing the state of the system (e.g. a leak location). The goal is to have enough scenarios to cover all possible behaviours of the system.

Figure 1 shows the principle of model falsification. Measurements ($\boldsymbol{y}$) are compared with predictions ($\boldsymbol{g}(\boldsymbol{s})$) from each scenario ($\boldsymbol{s}$) that have been obtained using the model of the system ($\boldsymbol{g}(\,)$). This comparison involves modelling errors and measurement errors. Measurement errors are mainly due to sensor resolution since noise and sensor bias are usually negligible. Modelling errors are due to the model simplification and to the errors included in the model parameters. The parameters are not exactly known, they are based either on the network plans or on measurements or on estimations.

Modelling errors and measurement errors may be represented by random variables ($\boldsymbol{U_{model}}, \boldsymbol{U_{meas}}$). The random variable $\boldsymbol{U_c}$ corresponds to the combined uncertainty obtained by subtracting $\boldsymbol{U_{meas}}$ from $\boldsymbol{U_{model}}$. The probability density function (pdf) of $\boldsymbol{U_c}$ describes the probability for the possible outcomes of the difference between predictions and measurements. Threshold bounds ($\boldsymbol{T_{low}}, \boldsymbol{T_{high}}$) are defined using this probability density function by taking the shortest interval including a probability of φ. Threshold bounds are used as criterion to falsify or keep a scenario. If the difference between measured and predicted values ($\boldsymbol{g}(\boldsymbol{s}) - \boldsymbol{y}$) is outside the interval defined by the threshold bounds, the scenario is falsified. Otherwise, the scenario is a candidate solution.

In the case when multiple measurements are used, the target probability is computed using the Šidák correction and becomes $\varphi^{1/n}$ where n is the number of measurements used (Abdi, 2007).

Using error-domain model falsification, Goulet et al. (2013) studied the applicability of model falsification for the detection of leaks in water distribution network. The study showed good results for leak of 100 litres per minute. However, for full-scale applications, the diagnostic methodology must be able to locate leaks smaller than 25 litres per minute. This paper proposes to develop this methodology further in order to obtain an improved and more sensitive diagnostic methodology for water distribution systems.

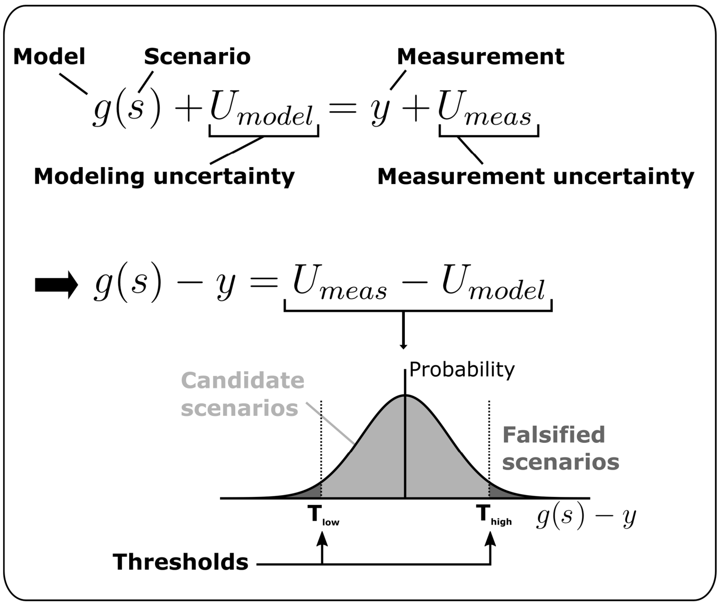

Figure 1 - Scheme of the falsification process

3. Application to leak detection

The objective is to provide a general diagnostic methodology -- for water distribution networks, and more generally, for pressurized fluid distribution networks -- that is able to locate leak regions. The methodology finds an area, or areas, in which the leak must be found. The size of the region as well as the number of regions depend on the number of sensors that are used and on the prior knowledge of the system. A secondary goal is to have a diagnostic methodology that is useful even when the number of sensors is small.

The leak detection methodology that will be developed is based on comparison between flow measurements and flow predicted by numerical simulations. The numerical simulations are done using the water distribution network simulation software EPANET (Rossman, 2000). This methodology includes three steps as shown in Figure 2. The first step is to obtain a simpler equivalent configuration of the network in order to reduce the complexity of the numerical model. Based on the similarities between water distribution networks and electrical networks, Ulanicki et al. (1996) developed an algorithm to simplify water distribution networks. This approach uses the Gaussian elimination process to remove certain nodes and to allocate their demand to the remaining nodes.

The second step is to compare *in situ* flow measurements with predictions obtained from a population of leak scenarios. For this study, a leak scenario is a possible configuration of the system in the presence of water loss at one node. In the case of leak detection each scenario represents the system through a different leak configuration. These configurations are obtained by varying the parameter of the leak such as its position and its intensity.

Finally, the last step is to eliminate the scenarios that are not compatible with the measurements. This operation is done by falsification. Scenarios are falsified using threshold values obtained by combining measurement and modelling uncertainties. Finally, scenarios that are not falsified show the situations that could explain the measurements. Therefore, they are considered to be candidate scenarios.

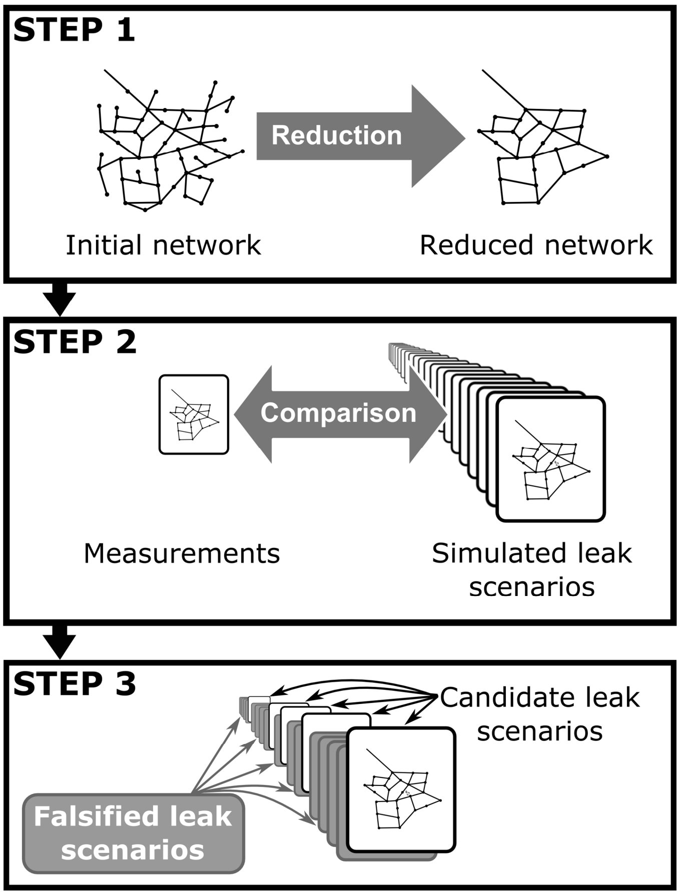

Figure 2 - Steps of the leak detection methodology

4. Case study

This section presents results that have been obtained with a study done on a part of the Lausanne fresh-water distribution-network. This network contains 263 nodes and 295 pipes (Figure 3). The network-reduction step of the diagnostic methodology was not applied; the falsification process was directly applied on the initial network. In this preliminary study, measurements are simulated. Simulations of measurements and leak scenarios are performed based on the minimum water demand. Analysis of water distribution networks is generally conducted during minimum demand hours because uncertainties related to the consumption is

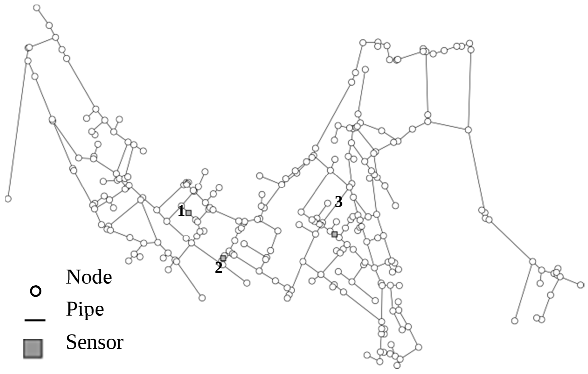

Figure 3 - Optimized sensor location for three sensors

minimal in this time period. In this case, this minimal demand is 25m^3/h. This value was then divided by the number of nodes to obtain the mean consumption for each node. For the simulations, the nodal demand is described by an exponential distribution. The exponential distribution is a standard representation for this water demand; a high probability to have a low consumption and a low probability to have a high consumption (Goulet et al., 2013).

For this application, the number of sensors is chosen to be three and they were placed using a greedy algorithm. The principle of a greedy algorithm is to begin by searching the optimal configuration for one sensor. The second step is to find the best configuration of two sensors including the optimal one found in the previous step without changing its position. This process is repeated to find the optimal configuration for the number of sensors desired. In this case, the algorithm has been stopped once a configuration for three sensors has been found. Figure 3 shows the optimized positions obtained. Figure 4 shows results obtained for four leak scenarios. White circles are the demand nodes. The links between these nodes are the pipes. Squares represent sensor locations on the network. The cross indicates the position of the simulated leak. In each of these four examples, the leak intensity is 100 l/min. The nodes in dark are the candidate leak scenarios i.e. those possibilities that have not been falsified.

The four examples in Figure 4 illustrate two situations. First, in the Cases (1) and (2), the number of candidate scenarios is important. The size of the regions that is defined by the candidate leak scenarios is too large to be able to identify precisely the leak-region. These results show that the methodology may be useful only to falsify one side of the network. In Case (1) all candidate leak scenarios on the right side have been eliminated and in the Case (2) the left side is falsified. These results may be useful in practice if the methodology is combined with local leak detection techniques, such as acoustic methods. Discarding half of the leak locations also divides by two the time necessary to cover the entire network in order to find the leak using a local technique. Nevertheless, more accuracy is desirable. In the Cases (3) and (4) the number of candidate leak scenarios is lower than in the Cases (1) and (2). In such situations, the region defined by the candidate leak scenarios is small enough to obtain information related to the leak location.

Figure 5 gives information about the performance of this sensor configuration. This graph is the cumulative distribution function for the number of candidate leak scenarios that are expected. The horizontal axis gives the number of candidate scenarios as a percentage of the total number of leak scenarios. The vertical axis represents the probability of achieving the percentage on the horizontal axis. This figure shows that with these sensors, for a leak of 100 l/min, the probability to obtain a number of candidate leak scenarios less than 40% of the total number of possible leak scenarios is equal to 95%.

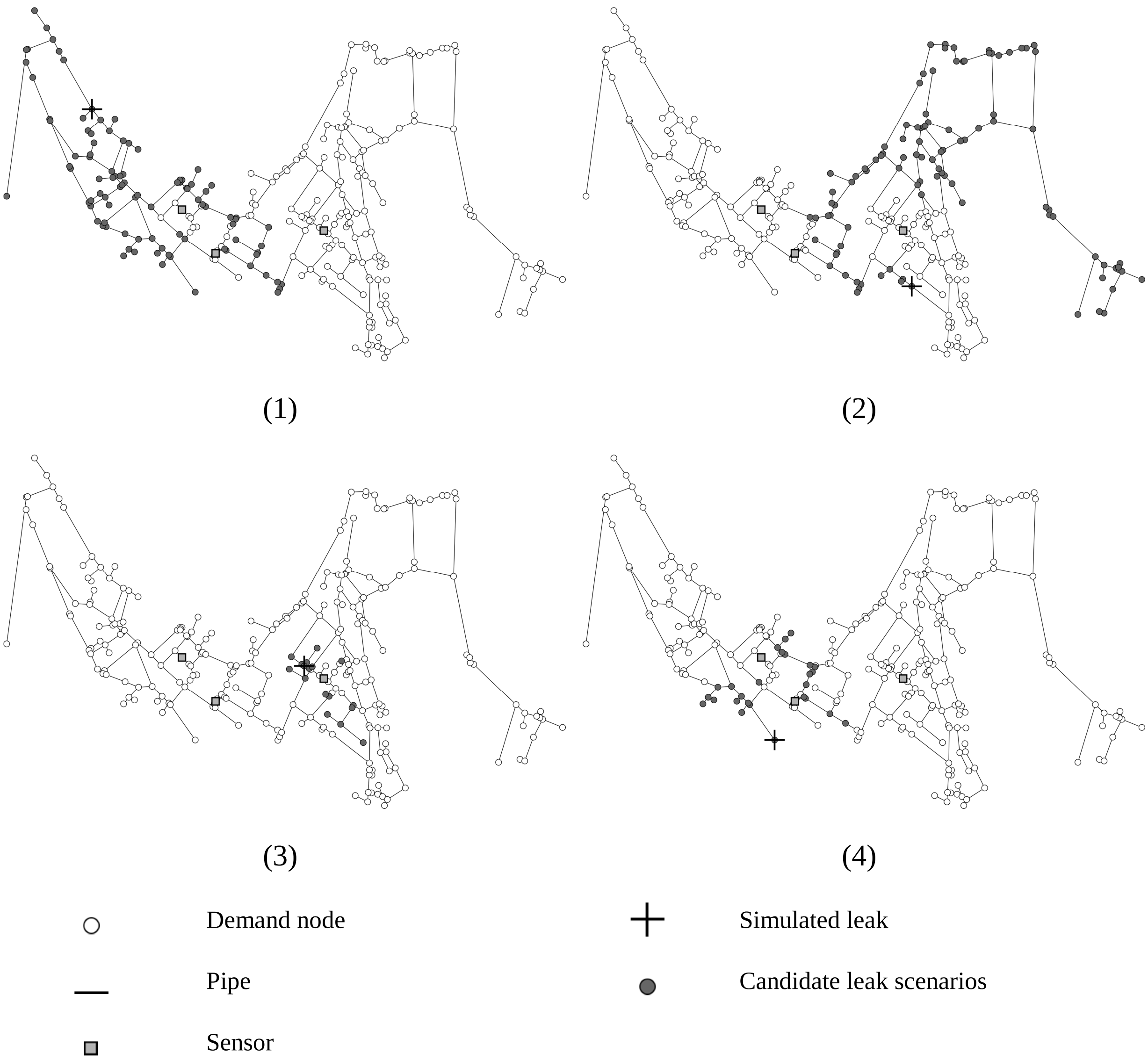

Figure 4 - Examples of results for four cases (1-4)

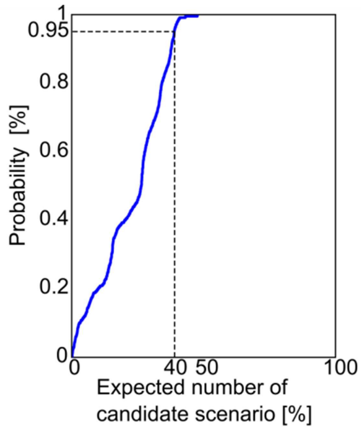

Figure 5 - Expected number of candidate leak scenarios for the three sensors configuration

Table 1 illustrates, for this case study, the importance of water-demand uncertainties in comparison to uncertainties of other parameter of the model. This table gives for the flow predictions at the sensor location, the mean and the standard deviation of the uncertainties due to the model parameters. They are computed using Monte-Carlo approach. In the first line of Table 1 the uncertainty is obtained by considering the uncertainties of all the parameters of the model and in the second line by considering only the uncertainties of the water demand. This results show that the variations of water demand has a huge importance in comparison with the other parameters, because the values with only the water demand are practically equal to the values with all the parameters.

Table 1 - Mean and standard deviation of flow uncertainties at the sensor position

	Sensor 1		**Sensor 2**		**Sensor 3**	
	Mean [l/min]	Standard deviation [l/min]	Mean [l/min]	Standard deviation [l/min]	Mean [l/min]	Standard deviation [l/min]
All parameter	-5.7015	2.7636	42.5402	5.9881	131.6950	12.8587
Demand	-5.6646	2.7118	42.5455	6.1170	131.6982	12.3632

5. Conclusions

Network reduction combined with model falsification has the potential to support leak-region identification. The results show that model falsification has potential to identify a population of possible leak scenarios. However, results show that the methodology needs more development to be able to identify leaks smaller than 100 l/min.

Even a small number of sensors may lead to useful results. Although a small number of sensors may not able to identify precisely the leak region, the results could be used in combination with local methods, such as acoustic techniques. The example of the Lausanne water-distribution network shows that with only three sensors, it is possible to falsify more than 40% of the total number of leak scenarios. More sensors would reduce further the size of the population of candidate scenarios.

The performance of this methodology is related to the degree uncertainty of key parameters. Through reducing these uncertainties by increasing the knowledge of the system, the number of falsified scenarios increases. In this situation, water demand is the most important parameter of the model; variations significantly change the behaviour of the system.

This research is important because advanced tools for measurement-data interpretation facilitate better management of infrastructure, and this leads to reductions in the costs of repair, replacement, network expansion and other interventions.

6. Acknowledgments

This research is funded by EPFL Middle East. The authors acknowledge the support from Eau Service, the water provider of the city of Lausanne.

7. References

Abdi, H. 2007. Encyclopedia of measurement and statistics. chapter The Bonferonni and Sidák corrections for multiple comparisons.

Andersen, J. H. & Powell, R. S. 2000. Implicit state-estimation technique for water network monitoring. *Urban Water,* 2**,** 123-130.

Babbitt, H., Amsbary, F. & Gwinn, D. 1920. THE DETECTION OF LEAKS IN UNDERGROUND PIPES [with DISCUSSION]. *Journal (American Water Works Association),* 7**,** 589-595.

Barandouzi, M. A., Mahinthakumar, G., Brill, E. D. & Ranjithan, R. 2012. Probabilistic Mapping of Water Leakage Characterizations Using a Bayesian Approach. *Proceedings of World Environmental and Water Resources Congress 2012,* Albuquerque, New Mexico, USA. 3248-3256. ASCE (American Society of Civil Engineers).

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.

Goulet, J.-A., Coutu, S. & Smith, I. F. 2013. Model falsification diagnosis and sensor placement for leak detection in pressurized pipe networks. *Advanced Engineering Informatics***,** in press.

Goulet, J. A. & Smith, I. F. C. 2012. Predicting the Usefulness of Monitoring for Identifying the Behavior of Structures. *Journal of Structural Engineering***,** in press.

Hope, W. & Bircumshaw, J. 1996. The waste of water in public supplies, and its prevention. *Proceedings of Institution of Civil Engineers. Municipal Engineer,* London, UK. 68-75.

Kingdom, B., Liemberger, R. & Marin, P. 2006. The challenge of reducing non-revenue water (NRW) in developing countries. How the private sector can help : a look at performance-based service contracting. *Water Supply and Sanitation Sector Board Discussion Paper Series, no 8, World Bank, Washington, DC.*

Poulakis, Z., Valougeorgis, D. & Papadimitriou, C. 2003. Leakage detection in water pipe networks using a Bayesian probabilistic framework. *Probabilistic Engineering Mechanics,* 18**,** 315-327.

Pudar, R. S. & Liggett, J. A. 1992. Leaks in Pipe Networks. *Journal of Hydraulic Engineering,* 118**,** 1031-1046.

Puust, R., Kapelan, Z., Koppel, T. & Savic, D. 2006. Probabilistic Leak Detection in Pipe Networks Using the SCEM-UA Algorithm. *Proceedings of Water Distribution Systems Analysis Symposium 2006*, Cincinnati, Ohio, USA. 1-12. ASCE (American Society of Civil Engineers).

Puust, R., Kapelan, Z., Savic, D. A. & Koppel, T. 2010. A review of methods for leakage management in pipe networks. *Urban Water Journal,* 7**,** 25-45.

Rossman, L. A. 2000. *EPANET 2: users manual,* Cincinati, OH, US Environmental Protection Agency.

Rougier, J. 2005. Probabilistic leak detection in pipelines using the mass imbalance approach. *Journal of Hydraulic Research,* 43**,** 556-566.

Ulanicki, B., Zehnpfund, A. & Martinez, F. 1996. Simplification of water distribution network models. *Proceedings of Second International Conference on Hydroinformatics*, Zürich, Switzerland. 493-500. IAHR (International Association for Hydraulic Research).

Vítkovský, J. P., Lambert, M. F., Simpson, A. R. & Liggett, J. A. 2007. Experimental observation and analysis of inverse transients for pipeline leak detection. *Journal of Water Resources Planning and Management,* 133**,** 519.

Vítkovský, J. P., Simpson, A. R. & Lambert, M. 2000. Leak detection and calibration using transients and genetic algorithms. *Water Resources Planning and Management***,** 258–262.