Advanced Engineering Informatics xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Advanced Engineering Informatics

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

Performance comparison of reduced models for leak detection in water distribution networks

Gaudenz Moser *, Stephanie German Paal, Ian F.C. Smith

Applied Computing and Mechanics Laboratory, School of Architecture, Civil and Environmental Engineering (ENAC), Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland

ARTICLE INFO

Article history:
Received 16 April 2014
Received in revised form 7 July 2015
Accepted 16 July 2015
Available online xxxx

Keywords:
Water distribution network
Network reduction
Leak detection
Model falsification
Leak intensity
Computational time

ABSTRACT

This paper presents a methodology for comparing the performance of model-reduction strategies to be used with a diagnostic methodology for leak detection in water distribution networks. The goal is to find reduction strategies that are suitable for error-domain model falsification, a model based data interpretation methodology. Twelve reduction strategies are derived from five strategy categories. Categories differ according to the manner in which nodes are selected for deletion. A node is selected for deletion according to: (1) the diameter of the pipes; (2) the number of pipes linked to a node; (3) the angle of the pipes in the case of two-pipe nodes; (4) the distribution of the water demand; and, (5) a pair-wise combination of some categories.

The methodology is illustrated using part of a real network. Performance is evaluated first by judging the equivalency of the reduced network with the initial network (before the application of any reduction procedure) and secondly, by assessing the compatibility with the diagnostic methodology. The results show that for each reduction strategy the equivalency of networks is verified. Computational time can be reduced to less than 20% of the non-reduced network in the best case. Results of diagnostic performance show that the performance decreases when using reduced networks. The reduction strategy with the best diagnostic performance is that based on the angle of two-pipe nodes, with an angle threshold of 165°. In addition, the sensitivity of the performance of the reduced networks to variation in leak intensity is evaluated. Results show that the reduction strategies where the number of nodes is significantly reduced are the most sensitive.

Finally this paper describes a Pareto analysis that is used to select the reduction strategy that is a good compromise between reduction of computational time and performance of the diagnosis. In this context, the extension strategy is the most attractive.

1. Introduction

Drinking water is one of the most precious resources for humanity. Annually, 184 billion USD are spent on clean water supply worldwide: however, collectively, water utilities lose an estimated 9.6 billion USD each year due to water leakage [42]. In addition, one third of reporting countries lose more than 40% of clean water pumped through distribution systems due to leaks, and worldwide, countries lose 20% of their clean water on average. Through reducing these leaks by just 5% and pipe bursts by 10%, utilities could save up to 4.6 billion USD.

The Sensus survey also includes a prediction that leak reductions can also lead to economies related to producing and purchasing water as well as reduced energy consumption required to pump and treat water for distribution. According to this survey, the need for leak detection services has been recognized by most global water utilities. However, only 40% of utilities reported having leak detection services. At this time, most utilities react to leakage on an ad-hoc basis, responding to obvious leaks and bursts and repairing infrastructure as required. Therefore, there is a need for more rational and systematic strategies for managing this infrastructure. This leads to requirements for efficient monitoring of water-supply networks. Advanced sensor-based diagnostic methodologies have the potential to provide enhanced management support.

Several studies have involved leak detection in fresh-water supply networks. Hope [15] studied water losses in public supplies. Babbitt et al. [3] described examples of leak-detection methods such as visual observation and sounding through the soil with a steel rod. Other more advanced techniques, including

* Corresponding author.
E-mail address: gaudenz.moser@epfl.ch (G. Moser).

http://dx.doi.org/10.1016/j.aei.2015.07.003
1474-0346/

water-hammer techniques and acoustic measurements were also examined nearly one hundred years ago.

There are both direct and indirect leak detection techniques. Various direct techniques were developed such as leak-noise correlation [13,8,9], pig-mounted acoustic sensing [19], and ground penetrating radar [7]. Although these techniques are considered the most accurate for leak detection, they are not appropriate for monitoring large networks due to their high cost. These methods complement other methods by precisely locating leaks in network segments that have already been identified.

There are several categories of indirect leak detection techniques. Two common methods are water balance [16] and night flow measurement at district metered areas (DMA) [20]. The principle of water balance is to audit the network in order to force equality between water placed into the distribution system and water taken out. In the night flow DMA method, the network is separated into areas and the water that comes in and out is metered. Water loss is estimated by taking these measurements when the demand is minimal, at night.

Another category is the transient-based techniques which use pressure measurement. These techniques use measured transient signals to detect leaks. Colombo et al. [6] completed a review of transient-based leak detection methods and sorted them into three types: inverse-transient analysis [46,45], frequency-domain techniques and direct transient analysis [48,47,43]. Uncertainties associated with these systems affect the accuracy of results. Many techniques within this category are primarily used on single, underground pipelines [34]. Most are currently not available to be used on complex water distribution networks. An exception is the study presented by Whittle et al. [47]. However, in this case, slow leak development requires other detection methods.

Other techniques are based on comparisons of measurements with predictions obtained from hydraulic models. This challenge is often formulated as an optimization task. The goal is to minimize the differences between the measurements taken on the network and predicted values from flow models. Such techniques are often based on minimization of least-squares [32,2]. Mounce et al. [22,23] developed a methodology using machine learning and fuzzy inference. Another methodology is Bayesian inference. Poulakis et al. [29] have proposed a Bayesian system-identification methodology for leakage detection. Other studies were presented by Rougier [41], Puust et al. [33] and Barandouzi et al. [5]. Romano et al. [37,38,36] used Bayesian inference in a pipe burst detection framework. The applicability of these methodologies to real networks may be limited under certain circumstances. Hypotheses made when using either traditional residual minimization or Bayesian inference techniques are usually impossible to meet due to systematic modelling errors and the unknown values of correlations that are induced [12].

Due to the size and the complexity of water distribution networks in cities, it is advantageous to include network reduction techniques in diagnostic methodologies. The principle of reducing a network to a simpler equivalent network is common in electrical engineering. Analogies between electrical and hydraulic networks were used to develop an algorithm to simplify water-distribution networks [44,18].

As has been done for electrical networks by Balabanian and Bickart [4], the theory of linear graphs has been used to build mathematical models of water networks. The principle of the methodology is to linearize the non-linear system and then apply Gaussian elimination to perform the reduction [14]. The last step involves transforming the linear reduced system to retrieve non-linearity. In this way, the reduced equivalent system preserves the hydraulic behavior and the non-linearity of the initial system.

The reduction algorithm developed by Ulanicki has been used by several researchers, each varying according to the strategy that was used to choose which nodes and pipes to eliminate. Preis et al. [30,31] used the algorithm to estimate hydraulic state in urban water networks by deleting pipes under a given diameter. The reduction algorithm has also been used for water quality analysis [26–28]. A graph-search algorithm reduced networks by eliminating the nodes in such a way that the reduced network maintains water quality properties.

Currently, studies have described only one reduction strategy at a time. Comparisons among reduction strategies have yet to be completed. In addition, the gain in computational time when using a reduced model has not been quantified except in the paper from Preis et al. [31] (again, for one strategy). Moreover, Ulanicki's reduction techniques have not been combined with a data-interpretation technique to develop a leak detection methodology.

The task of finding a good compromise between two or more goals involves multi-criteria decision making. A simple way of solving this challenge is to first find a set of Pareto-optimal solutions [25] and then perform further analysis on a smaller set of solutions. This type of multi-criteria decision making is used in many researches. Nouiri [24] developed a tool to optimize water resource management using the Pareto optimality concept. Mala-Jetmarova et al. [17] studied the trade-offs between water quality and pumping cost objectives. No study was found that used Pareto optimization for selecting network-reduction strategies.

Model falsification for leak detection was developed initially by Robert-Nicoud et al. [35]. A model-based system-identification method originally proposed for structures was applied to leak detection in hydraulic networks. In a subsequent study, Goulet and Smith [11] developed a model falsification method for infrastructure diagnosis. The methodology, called error-domain model falsification, was developed principally for bridge diagnosis. Using this methodology, a preliminary study has been carried out on leak detection [10]. A follow-on study using error-domain falsification has been performed by Moser and Smith [21]. None of these studies involve network reduction.

This paper describes a methodology for evaluating network reduction strategies. The goal is to choose the strategy which is most compatible with the error-domain model falsification framework. Twelve network reduction strategies for water-network management are compared using part of the water supply network in Lausanne, Switzerland for illustration. The reduced network is then used with a model falsification methodology for detecting leaks. Gains in computation time are quantified and compared. In addition, the effect of the leak severity on the effectiveness of reduction is evaluated. Finally this paper identifies, using Pareto analysis, strategies that provide good compromises between performance and computational time.

Section 2 describes the error-domain model falsification methodology and the principle of the network reduction. Section 3 presents the reduction strategies studied in this paper. Finally Section 4 includes an analysis of the results obtained using the reduction strategies.

2. Methodology

In this section the strategies used for network reduction are explained. The principle of error-domain model falsification is also described. Finally, a description of the leak-detection methodology obtained by combining these two principles is provided.

2.1. Network reduction

The network reduction technique used for this study was developed by Ulanicki et al. [44]. This section explains the principle of this technique. More precise explanations, such as the complete mathematical formulation of the methodology can be found in Ulanicki's paper. This reduction technique is based on similarities between electrical networks and hydraulic networks. In the same way that the Ohm's law gives a potential difference in the function of current and resistance, the Hazen-Williams hydraulic model predicts the head-loss in a pipe as a function of the flow and the "resistance" of the pipe. The Hazen-Williams relation may also be given in the inverse form (1), the flow (q) as a function of conductance (g) and headloss (Δh). Conductance is function of pipe length, pipe diameter and the Hazen-Williams pipe-friction coefficient.

$$q = g|\Delta h|^{0.54} sign(\Delta h) \tag{1}$$

With the Hazen-Williams relation and node-branch incidence matrix (Λ), a mathematical model of the entire network can be built (2). The incidence matrix, a concept taken from linear graph theory, represents the topology of the network. For a network of m nodes and n pipes, the incidence matrix size is $m \times n$. This matrix (Λ) provides the link between the pipe flow vector ($\boldsymbol{Q}(\boldsymbol{\Delta h}) = (q_1(\Delta h_1), \ldots, q_n(\Delta h_n))^T$) and the nodal demand vector ($\boldsymbol{q}^{nod} = (q_1^{nod}, \ldots, q_m^{nod})^T$). Each element of the pipe flow vector can be written as a function of head loss using the Hazen-Williams relation (1). The resulting mathematical model represents a relation between the head loss and nodal demand.

$$\Lambda \boldsymbol{Q}(\boldsymbol{\Delta h}) = \boldsymbol{q}^{nod} \tag{2}$$

The second step of the reduction technique is to perform a linearization of the model. In order to linearize the model, the assumption of small variations under a given operation point, defined by nodal head ($\boldsymbol{h^0}$) and nodal demand ($\boldsymbol{q}^{nod\,0}$), is made. This leads to a linear model (3) of the system represented by the linearized branch conductance matrix ($\boldsymbol{A}$) that multiplies the vector of the nodal head variations ($\delta \boldsymbol{h} = \boldsymbol{h} - \boldsymbol{h^0}$) to obtain the vector of nodal demand variations ($\delta \boldsymbol{q}^{nod} = \boldsymbol{q}^{nod} - \boldsymbol{q}^{nod\,0}$).

$$\boldsymbol{A}\delta \boldsymbol{h} = \delta \boldsymbol{q}^{nod} \tag{3}$$

The linearized branch conductance matrix is a symmetric matrix with as many rows and columns as nodes. The elements [k, l] of this matrix ($k \neq l$) represent the linearized conductance of each pipe between Node k and Node l; if it is null then there is no connection between these two nodes. The elements (k,k) of the diagonal represent the node conductance which is equal to the sum of the conductance of all pipes connected to Node k.

The third step is to remove the desired nodes by eliminating the corresponding rows and columns from the linearized model by using the Gauss elimination algorithm. For example, when eliminating Node k, each row of the matrix corresponding to a neighboring node (node connected to Node k) is subtracted from a multiple of the row k. The constant used in this multiplication is chosen for each row such that the k-th element of the row becomes zero. Due to this, the multiple is equal to the linearized conductance of the pipe between the two nodes divided by the conductance of Node k. This gives the following relation for each element a_{ij} of the matrix $\boldsymbol{A}$ (4).

$$a_{ij} = a_{ij} - a_{kj}\, a_{ik}/a_{kk} \tag{4}$$

In the same way, the demand of Node k is redistributed to its neighboring nodes. For each neighboring node the demand is subtracted from a multiple of the nodal demand of k. Likewise, the linearized conductance of the pipes connected to Node k are either assigned to the remaining pipes or to a new one. Therefore, for each demand δq_i the following relation is used (5).

$$\delta q_i = \delta q_i - \delta q_k\, a_{ik}/a_{kk} \tag{5}$$

The last step is to return to a non-linear model by transforming the linearized conductance for each pipe into a non-linear conductance. The length of each pipe is determined by the distance between the nodes which are connected to each other. For the diameter and friction coefficients, one of these parameters is fixed and the other is computed from the definition of the conductance. For this study, the friction coefficient has been fixed because all the pipes are considered the same material.

The model of the network used is only constituted of the main pipes of the network. For this network, the majority of the main pipes are composed of cast iron. For this reason, it is admissible to model the network with the same material for all pipes. Since the non-reduced network is modelled using the same material, pipes of reduced networks were modelled in the same way by maintaining the Hazen-Williams friction factor constant.

2.2. Error-domain model falsification

Fig. 1 shows the principle of model falsification. Measurements of system quantities ($\boldsymbol{y}$) are compared with predictions of the same quantities ($\boldsymbol{g}(\boldsymbol{s})$). Predictions are obtained by simulating scenarios ($\boldsymbol{s}$) using the model of the system ($\boldsymbol{g}(\,)$). Each scenario is a representation of a possible state of the system. Scenarios chosen have to cover the entire range of behavior that the system-identification method should be able to recognize. To compare measurements with prediction involves modelling errors and measurement errors. Measurement errors are mainly due to sensor resolution (precision of the measure) since noise and sensor bias are usually negligible. In practice, noise may be reduced by filtering and sensor bias by sensor calibration. Modelling errors are due to the model simplification and to the errors included in the model parameters. Values of these parameters that are usually not known precisely are based either on the network plans, measurements or 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. This pdf is calculated by using a Monte-Carlo approach. In this way, the combined uncertainty is obtained by computing a high number of samples with varied random variables ($\boldsymbol{U_{model}}, \boldsymbol{U_{meas}}$).

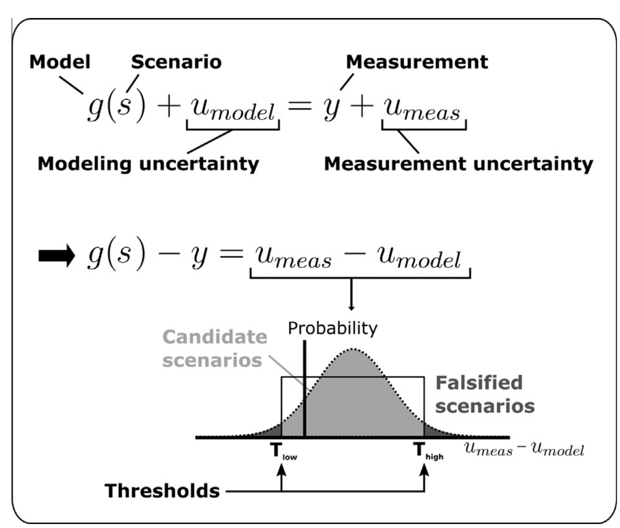


Fig. 1. Scheme of the falsification process.

Please cite this article in press as: G. Moser et al., Performance comparison of reduced models for leak detection in water distribution networks, Adv. Eng. Informat. (2015), http://dx.doi.org/10.1016/j.aei.2015.07.003

Threshold bounds ($\boldsymbol{T_{low}}, \boldsymbol{T_{high}}$) are defined using this pdf by taking the shortest interval including a probability of φ (for example, 95%). In Fig. 1 a simplified case is illustrated for one measurement; however, multiple measurements are generally used. For these cases, error-domain model falsification involves multidimensional pdfs. To ensure a probability of φ on the multidimensional pdf, the target probability, used for computing threshold bounds for each measurement, is obtained using the Šidák correction and becomes $\varphi^{1/n}$ where n is the number of measurements obtained [1].

Threshold bounds are used as criterion to falsify or keep a scenario. The difference between measured and predicted values ($\boldsymbol{g}(\boldsymbol{s}) - \boldsymbol{y}$) is computed for each scenario. If this number (vector, if multiple measurements) is outside the interval defined by the threshold bounds, the scenario is falsified. Otherwise, if this difference is within the bounds for each measurement, then the scenario is deemed a candidate solution. Since likelihood distributions are not well known, no candidate solution is considered to be more likely than another. This means that each candidate scenario is considered to have the same probability to be the solution of the diagnosis. The methodology does not lead to the most probable solution.

2.3. Application to leak detection

The objective of this research is to combine error-domain model falsification with a network-reduction strategy in order to develop an efficient leak-detection methodology for complex water supply networks. This methodology is capable of considering biased uncertainties which are typically present in modelling challenges. In addition by using the reduction process, the leak detection methodology is applicable to complex water distribution networks.

This methodology includes three steps (Fig. 2). The first step is to obtain a simpler equivalent configuration of the network in order to reduce the complexity of the numerical model. The second step is to compare *in situ* flow measurements with flow predictions obtained from a population of leak scenarios. This is done by observing the difference obtained by subtracting measured values from predicted values.

Each leak scenario represents a different leak configuration of the system. For this study, scenarios are constructed following two hypotheses: (1) there is only one leak; and, (2) it occurs at the nodes. The configurations are obtained by varying leak position (the node where the leak occurs) and the leak intensity (the flow going out through the leak). This means that the number of scenarios is, for this case, equal to the number of nodes multiplied by the number of intensities considered. It is not necessary to consider leaks that occur at intermediate points of pipes because due to uncertainties only leak regions will be identified. In order to compare results from reduced networks with those from the initial network, the leaks are modelled for all the networks at the nodes of the non-reduced network. This provides a consistent number of scenarios for each network. When a leak occurs on an eliminated node, the leakage is distributed to the remaining nodes, following the reduction technique.

Since leaks occur at the nodes, they are modelled, in the simulation software (EPANET), by varying nodal demands. Due to this, as the uncertainty of the demand increases, it becomes increasingly difficult to differentiate a leak from a change in nodal demand. To reduce this error, the measurements are taken when consumption is the smallest, during the night. Other parameters may be considered such as tank level, water demand and income flow at the pumps. Consideration of all parameters is necessary in a practical case. However, to keep from unnecessarily increasing the number of scenarios, only leak position and intensity are considered in this study.

The last step is to eliminate scenarios which are incompatible with the measurements. Scenarios are falsified using threshold

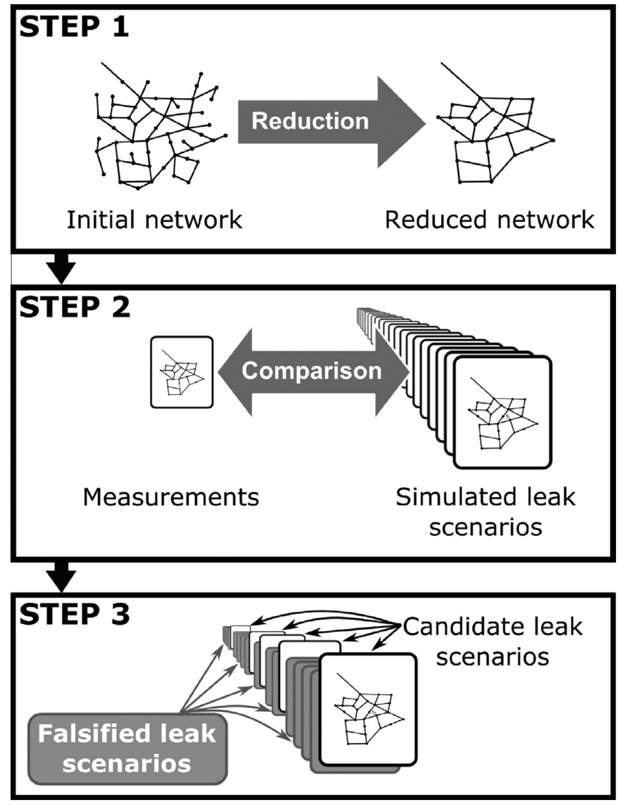


Fig. 2. Step of the leak detection strategy.

values obtained by combining measurement and modelling uncertainties. If the difference between flow measurements and flow predictions of leak scenarios is outside the threshold then the scenario is falsified. Finally, scenarios that are not falsified are leak configurations that are capable of explaining the measurements. Therefore, they are considered to be candidate scenarios.

3. Reduction strategies

In this paper, five categories of water-supply-network-reduction strategies are presented. Several processes can be designed to reduce a network according to certain criteria. These criteria are used to determine the nodes in the network which are eliminated. This section focuses on the reduction strategies that are used for this study.

3.1. Study case

In this paper, the reduction processes are tested on one of the water supply networks of the city of Lausanne. This network is not connected to the other networks of the city of Lausanne; it is totally independent. All the networks in Lausanne are isolated from one another.

This network (Fig. 3) contains 295 pipes and 265 nodes and is equipped with three flow-meters. In the figure, the demand nodes are represented by the white circles, the pipes by the black lines and each of three sensor locations by an ‘x’. A pipe with a sensor cannot be removed in the reduction process. For this reason, nodes attached to these pipes are labeled ‘irremovable’. All the studies in this paper are performed using this network and this sensor configuration. It assumed that this network contains sufficient complexity to be able to provide a meaningful test of reduction-strategy performance.

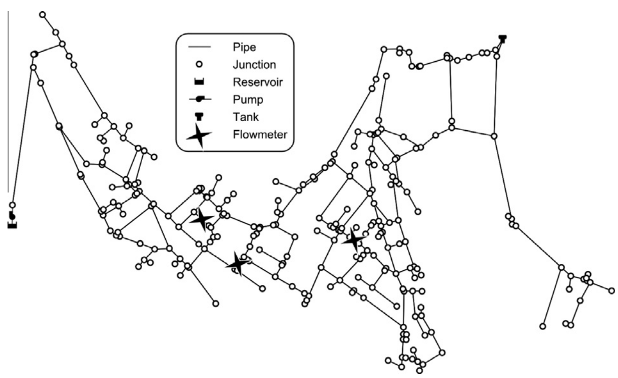


Fig. 3. Initial diagram of the city of Lausanne water supply network.

For this study case, the distribution of the demand on each node (nodal demand) is not known; only the demand of the entire network (global demand) is known. Therefore, the nodal demand is modelled, for each node, using an exponential distribution with the mean equal to the average nodal consumption. The average nodal consumption is the global consumption divided by the number of nodes. The exponential distribution is a good representation for water demand since there is a high probability to have low consumption and low probability to have high consumption. The predictions are computed by performing steady state simulations.

The global demand used in this study is the minimum demand obtained from the hourly averaged demand of the network; this represents a global demand of 416 l/min.

3.2. Pipe diameter

The first reduction process is based on the diameter of the pipes. Initially, a pipe diameter limit is defined, and then all the pipes having a diameter higher than this specified limit are selected. All nodes which are not linked to the selected pipes are then eliminated. The use of the Gaussian process to eliminate the nodes ensures the connectivity of the reduced network by creating fictitious pipes if necessary. Fig. 4 shows the result of this reduction procedure, using a pipe diameter limit of 150 mm. For a better comparison, pipes of the non-reduced network are represented in gray on the same figure. The reduced network has 224 pipes and 196 nodes.

3.3. Extension

The second reduction process is based on the elimination of the extremity pipes. The decision criterion for this process is the number of pipes to which each node is connected. Each node that is linked with only one pipe is eliminated. After the application of this principle, some other nodes will be linked with only one pipe. For this reason, this process has to be applied iteratively, as long as nodes that fill the criterion are found. The principle of this reduction strategy is described in Fig. 5. This example shows that in the first step Node 1 is eliminated because it is connected to only on pipe. After that, Node 2 becomes a node connected to only one pipe; for this reason, Node 2 is eliminated in the second step. The result is the elimination of the entire extension. Fig. 6 shows the result for this reduction procedure, with the non-reduced network represented in gray. The reduced network has 175 nodes and 204 pipes.

3.4. Angle

The third reduction process is based on the angle between two pipes connected to the same node. The goal of this procedure is to eliminate the nodes that are in series while maintaining the general topology of the network. If the reduced network diverges too much from the initial network (regarding the topology), some regions of the network may be neglected in the leak detection process. The topology is maintained to ensure that the identification process covers the main regions of the network. First the nodes connected with only two pipes are selected. Following this, each node specified by an angle between its pipes larger than a pre-determined limit is eliminated. The principle of this reduction strategy is described in Fig. 7. In this scheme, the pipes attached to the central node form an angle (α) larger than the angle limit. For this reason, this node is then eliminated. Fig. 8 shows the result of this reduction process with an angle limit of 150°, with the non-reduced network represented in gray. The reduced network has 230 pipes and 201 nodes.

3.5. Consumption

The fourth reduction process presented in this paper considers the yearly consumption values throughout the water supply network. The goal is to design a procedure which eliminates the nodes associated with low consumption. Each of the consumption values are assigned to the nearest node. The highest consumption value is deemed 100%, and the remainder of the consumption values is adjusted *pro-rata*. For this reduction process the criterion for selecting the nodes to eliminate is the consumption percentage. All the nodes under a specified limit are deleted. Fig. 10 shows the result for this reduction process with a limit of 50%, with the non-reduced network represented in gray. The reduced network is constituted of 295 pipes and 180 nodes.

This example shows that the reduction strategy does not lead, in each case, to a reduction of the number of pipes. In this case

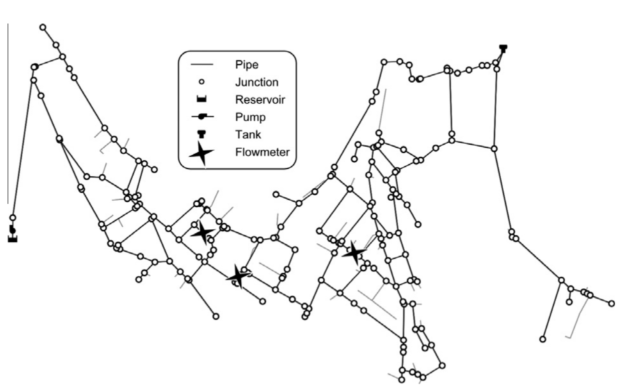


Fig. 4. Network reduced through simplifying the pipes with a diameter smaller than 150 mm.

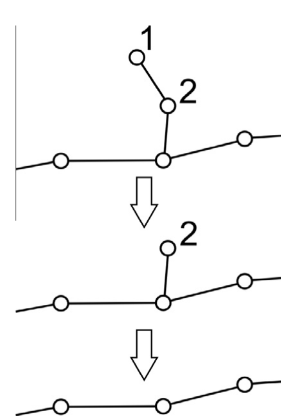


Fig. 5. Example of extension elimination.

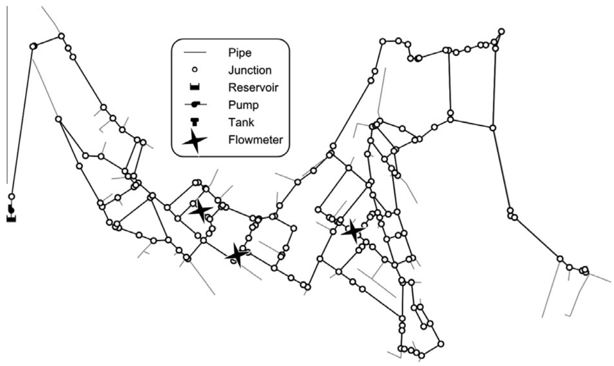


Fig. 6. Network reduced by eliminating all the extension nodes.

the number of pipes is equal to the non-reduced network. In addition the network is chaotic with some pipes crossing over each other. The reason for this behavior is that if a node that is connected with more than three nodes is eliminated, then the number of pipes increases.

Fig. 9 shows how a node is eliminated in each of the following four cases: (a) two-pipe nodes, (b) three-pipe nodes, (c) four-pipe nodes, and (d) five-pipe nodes. In each case, the central node is deleted. For the two-pipe nodes, the node elimination reduces the number of pipes by one. In the case of the three-pipe nodes, the number of pipes remains the same. For the case of four-pipe nodes, the number of pipes increases from four to six. Finally for the five-pipe nodes, the number of pipes increases from five to ten.

This explains how it is possible to increase the number of pipes in instances when the number of nodes is reduced. The same behavior is observed for reduction using the pipe diameter when the specified diameter limit is substantially high. Physically, when a node is eliminated, all the nodes connected to that node have to be connected to one another in order to maintain the equivalency of the system. It is similar to the process in electrical engineering known as the star-mesh transformation [39].

3.6. Combination extension and angle

In the fifth reduction process, the second (extension) and third (angle) processes are combined. Fig. 11 shows the result of this combination using the angle limit of 150° on the network of the city of Lausanne, with the non-reduced network represented in gray. The reduced network is constituted of 123 pipes and 94 nodes. In this case the extension strategy is applied before the angle strategy. For the results, both cases are studied, the case when extension is applied first (Extension & Angle 150°) and the case when angle is applied first (Angle 150° & Extension).

4. Results

4.1. Reduction

For each of the five reduction strategies described above, the magnitude of the reduction in size of the network is quantified

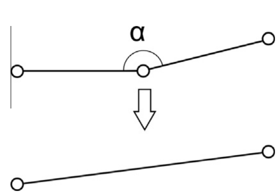


Fig. 7. Example of node elimination by angle limit.

based on the resultant number of nodes and pipes. These values are displayed in Table 1. These results show that the reduction strategy that is the most efficient considering only the number of pipes and nodes eliminated is the *Extension & Angle 150°*. It suppresses 64.5% of the nodes and decreases the number of pipes by 58.3%. In comparison, the second best strategy in terms of node reduction, with 60.8%, is not as strong for the number pipes – only a 37.3% reduction. The reason is the same as for reduction strategies based on consumption. These two categories of reduction strategies lead to elimination of nodes with more than three connections and this increases the number of pipes (Fig. 9).

4.2. Hydraulic equivalency

Table 2 shows simulation results for flows at sensor positions in terms of the difference between the initial network and reduced networks. The numerical simulations have been carried out using the water distribution network simulation software EPANET [40]. The flow calculated at each of the three sensor locations is extracted, and those for the reduced networks are compared with those for the initial network. These results show that, for most of the strategies presented in this paper, the relative error is less than 1%. Only the strategies *Consumption 50%* and *Diameter 200* create an error that is greater than 1%. These two strategies are cases where many new pipes are added due to elimination of nodes with more than three connections. The errors present from computing the conductance of these fictitious pipes influences the pipe measurement predictions due to the way in which the flow is distributed in the network.

4.3. Computational time

The principal motivation for the use of a reduced network is to decrease computation time. Table 3 gives the relative computation time, in comparison with the non-reduced network, for the twelve reduction strategies. These represent the time necessary to compute the threshold for each pipe of the system. The thresholds are computed using 10^5 Monte-Carlo simulations to combine modelling and measurement uncertainties. The modelling uncertainty is a combination of errors due to the model simplification and the model parameters. These parameter uncertainties (i.e., pipe diameter, pipe roughness and node elevation) are computed for each reduced network using Monte-Carlo simulation. This means that the errors introduced by using the reduced network will further influence the threshold values. Then, the pdf of the combined uncertainty for each pipe is obtained by simulating a total of 10^5 samples with varied random variables. Thresholds are determined by taking 95% of this pdf.

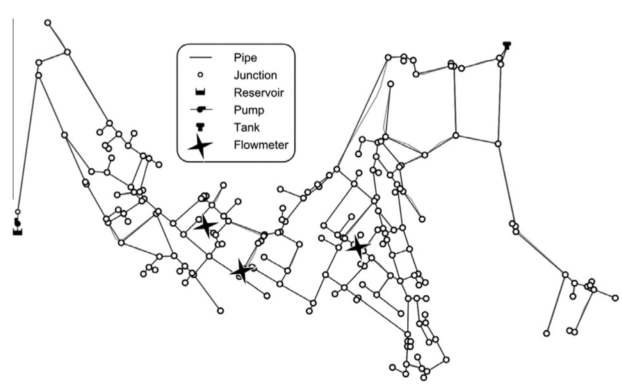


Fig. 8. Network reduced through simplifying two-pipe nodes when the angle between them is greater than 150°.

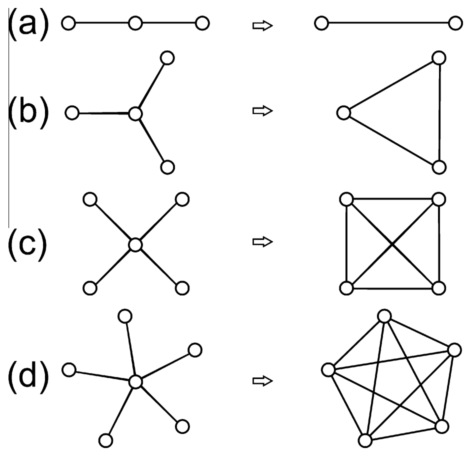


Fig. 9. Example of node elimination for: (a) two pipe node, (b) three pipe node, (c) four pipe node and (d) five pipe node.

This is done for each possible sensor location. Results show that simplifying the network can lead to a computational time as low as 18.2% of that of the initial network.

For reduction strategies that are based on consumption, computation time may increase. This is due to an increase in the number of pipes for reduction strategies that lead to elimination of nodes that are connected to more than three pipes, see Fig. 9.

4.4. Expected identifiability

The performance of the reduced networks is compared using a cumulative distribution (CDF) function for the expected number of candidate scenarios. This CDF is built by testing a large number of simulated leaks on the water supply network. For each leak, the number of candidate scenarios is computed using the error-domain model falsification procedure presented above. To have the same leak scenarios for all the networks, the leaks were simulated on the same number of nodes as the non-reduced network. When the leak occurs on an eliminated node, it is redistributed to remaining nodes in the same way as the demand. The leak scenarios are simulated using the same reduced model for each network strategy. The hypothesis is made that when adding a leak to the network, the model parameters remain within the range of validity for the reduced model.

The CDF for the non-reduced network (Fig. 12) provides a reference for comparison. This graph shows that there is a 95% probability to identify less than 127 candidate leak scenarios (or to falsify more than 138 leak scenarios). This means that, for this three sensor configuration, in 95% of cases it is possible to reduce the population of candidate leak scenarios to half, for a leak intensity of 100 l/min. With a 75% probability, it is possible to falsify less than 93 candidate leak scenarios while with a 50% probability, less than 72 candidate leak scenarios are falsified.

In practice, this means that the utility manager only needs to search for the leak location on half of the network. Even if the network is equipped with only three sensors, these results show that it should be possible when combined with a pinpointing method (acoustic correlation) that utility managers already use, to reduce the search time by half on this network.

The performance of the reduced network is lower than the reference case if its CDF is positioned to the right of the reference network CDF. More specifically, when considering the same probability, the number of candidate leak scenarios becomes larger.

In Figs. 13 and 14 the CDFs are given for each reduction strategy studied in this paper for a leak intensity of 100 l/min. As before, the horizontal axis represents the number of candidate leak scenarios and the vertical axis is the probability. The values obtained for the probabilities of 0.95, 0.75 and 0.5 are also given on the horizontal axis. Furthermore, the cumulative distribution function for the non-reduced network is displayed on each graph for comparison.

The CDFs for the reduced networks show that the performance decreases in every case in comparison with the non-reduced network. For each reduced network the expected number of candidate leak scenarios increases, and thus, the number of scenarios that can be falsified decreases. Overall, this indicates a decrease in the identifiability of leaks in the reduced networks in comparison with the non-reduced network. Such behavior is understandable since a reduced network leads to some loss of information and consequently, the uncertainty increases. Higher uncertainties imply that the threshold interval is larger than in the non-reduced case. This results in fewer falsified scenarios.

The results show that information loss is not detrimental to the overall performance of the method. When looking at the 95% probability, the expected number of candidate scenarios is 141 for the worst case (*Diameter 200*). In comparison with 127 candidate scenarios for the non-reduced network, such performance is acceptable in all cases. Since this study is concerned directly with safety aspects, it may be unnecessary to consider a 95% probability; a 75% probability can be considered as a good indicator of performance.

For the reduction strategies based on consumption, the number of expected candidate models at 75% probability increases as the demand limit increases. For the reduction strategy, *Consumption 5%* the expected number of candidate scenarios is 101, and for *Consumption 50%*, this value rises to 122. This indicates a decrease

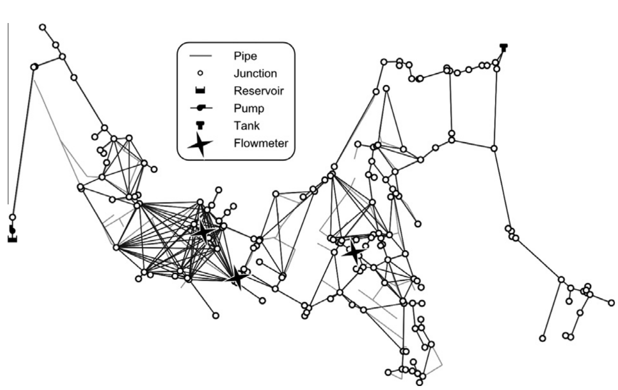


Fig. 10. Network reduced by eliminating nodes with a yearly consumption smaller that 50% in comparison with the highest one.

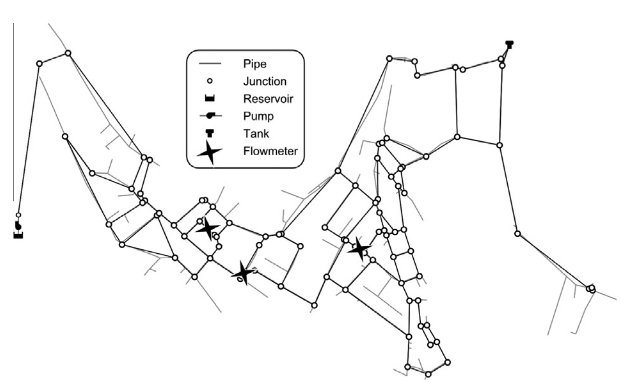


Fig. 11. Network reduced using a combination of angle and extension.

Table 1
Comparison of the number of nodes and pipes obtained after the reduction following 5 reductions strategies.

Reduction procedure	Number of nodes	Node reduction (%)	Number of pipes	Pipe reduction (%)
Initial network	265	–	295	–
Consumption 5%	222	16.2	278	5.76
Consumption 10%	207	21.9	281	4.75
Consumption 25%	189	28.7	274	7.12
Consumption 50%	180	32.1	295	0
Diameter 150	196	26.0	224	24.1
Diameter 200	104	60.8	185	37.3
Extension	175	34.0	204	30.9
Angle 135°	196	26.0	225	23.7
Angle 150°	201	24.2	230	22.0
Angle 165°	216	18.5	245	17.0
Extension & Angle 150°	94	64.5	123	58.3
Angle 150° & Extension	129	51.3	158	46.4

in the performance that is inversely related to the increase in the number of eliminated nodes. The same behavior is observed for all reduction strategies.

When considering the performance alone, by comparison of the results of the reduced networks with that of the non-reduced network, the reduction strategy, *Angle 165°*, appears the most effective. For 75% probability, the number of candidate scenarios is 93 for the non-reduced network and 97 for this reduction strategy.

In order to choose a reduction strategy that performs well when using error-domain model falsification for leak detection, the sensitivity of the performance must be analyzed for different leak intensities, especially for smaller leaks. To reach this goal, the CDFs described previously have been computed for leak intensities: 25 l/min, 50 l/min, 75 l/min, 100 l/min, 150 l/min and 200 l/min. Fig. 15 provides the evolution of the expected number of candidate leak scenarios for 0.5 (graph (a)), 0.75 (graph (b)), and 0.95 (graph (c)), probability, respectively, when varying the leak intensity according to this range. In these three graphs the horizontal axes are the leak intensity in l/min, and the vertical axis provides the number of expected candidate leak scenarios.

These graphs show that for large leak intensity, the curves are parallel. Decreasing the leak intensity from 200 l/min to 100 l/min has the same influence on the performance of each network. However, at 75 l/min a difference is visible. For lower leak intensities, the curves of three reduced networks (*Diameter 200*, *Angle 150 & Extension* and *Extension & Angle 150*) increase at a greater rate than those of the other reduction networks, indicating that the decrease in performance is faster for these three reduction strategies than for the others when the leak severity decreases.

Such decrease in performance with these three networks is due to high reductions in node numbers. Also, the demand is modelled

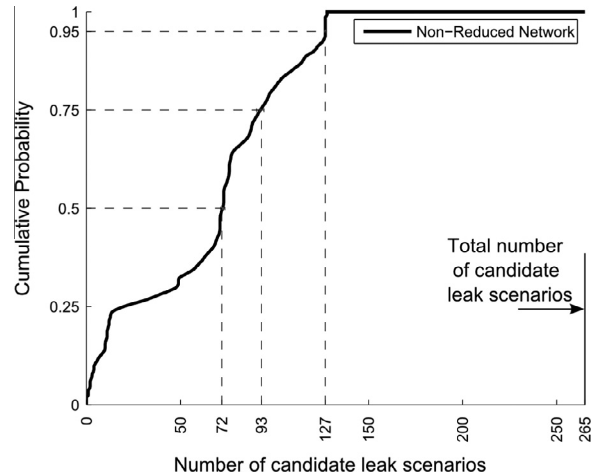


Fig. 12. CDF for the expected number of candidate leak scenario in the case of a leak severity of 100 l/min for the non-reduced network.

at each node using the assumptions described below. For this study case, the nodal demand is modelled, for each node, using an exponential distribution with the mean equal to the average nodal consumption.

The non-reduced network and all reduced networks have the same global consumption. For one network, the mean of nodal demand is equal to the average nodal demand (global demand divided by the number of nodes). Due to this, the more nodes eliminated, the larger the value of the mean nodal demand. This implies that the critical point – when the leak intensity is along the same order of magnitude as the mean nodal demand – is reached faster for the reduced networks with fewer nodes. When this occurs, the diagnostic process is unable to differentiate a leak from a variation of the demand.

The graphs show that the reduction networks that are least sensitive to the leak intensity variation are the following: *Extension*, *Angle* (*135–165*), *Consumption 5%* and *10%*.

4.5. Pareto analysis

To select the reduction strategies that are most suited for leak detection using model falsification, a compromise must be found between the two criteria: (1) reduction of computational time; and, (2) diagnostic performance. Pareto analysis is used to focus the compromise on non-dominated cases. The first step is to find reduction strategies on the Pareto front. Each compromise on this front is dominated by no other compromise. The second step is to select the cases that are the most interesting for this application.

Table 2
Comparison of flows at sensor positions in terms of the difference between the initial network and reduced networks.

Reduction strategy	Flowmeter 1 Difference (%)	Flowmeter 2 Difference (%)	Flowmeter 3 Difference (%)
Consumption 5%	0.18	0.00	0.04
Consumption 10%	0.08	0.01	0.03
Consumption 25%	0.01	0.01	0.04
Consumption 50%	5.54	0.04	0.04
Diameter 150	0.17	0.00	0.00
Diameter 200	4.76	0.03	1.90
Extension	0.18	0.00	0.00
Angle 135°	0.19	0.01	0.02
Angle 150°	0.19	0.01	0.04
Angle 165°	0.18	0.00	0.02
Extension & Angle 150°	0.19	0.01	0.14
Angle 150° & Extension	0.19	0.01	0.04

Table 3
Relative computation time for each network reduction strategies.

	Relative computational time (%)
Initial network	100
Consumption 5%	74.9
Consumption 10%	89.2
Consumption 25%	89.3
Consumption 50%	108.6
Diameter 150	47.9
Diameter 200	32.4
Extension	40.3
Angle 135°	48.7
Angle 150°	67.4
Angle 165°	76.2
Extension & Angle 150°	18.2
Angle 150° & Extension	30.2

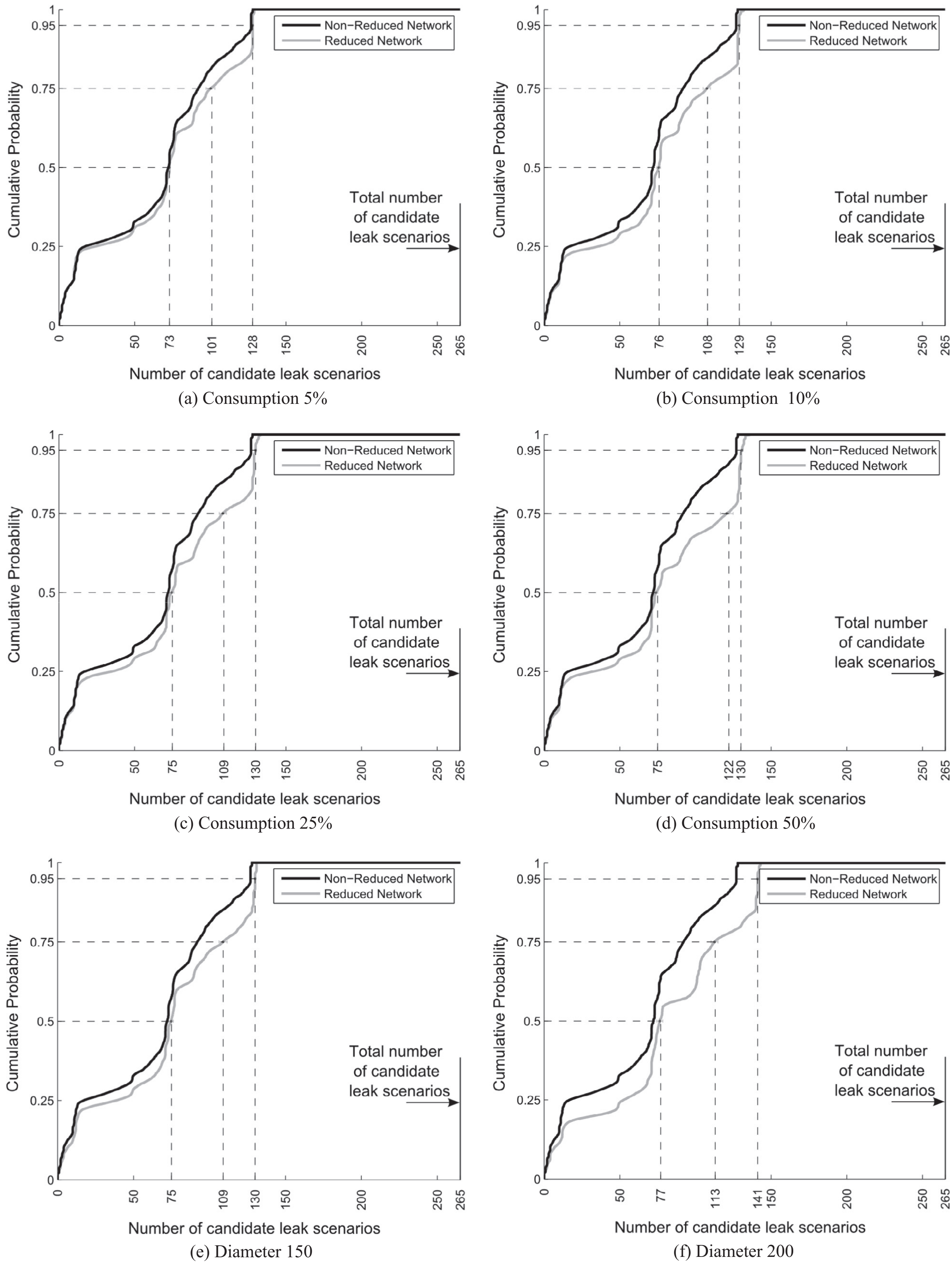
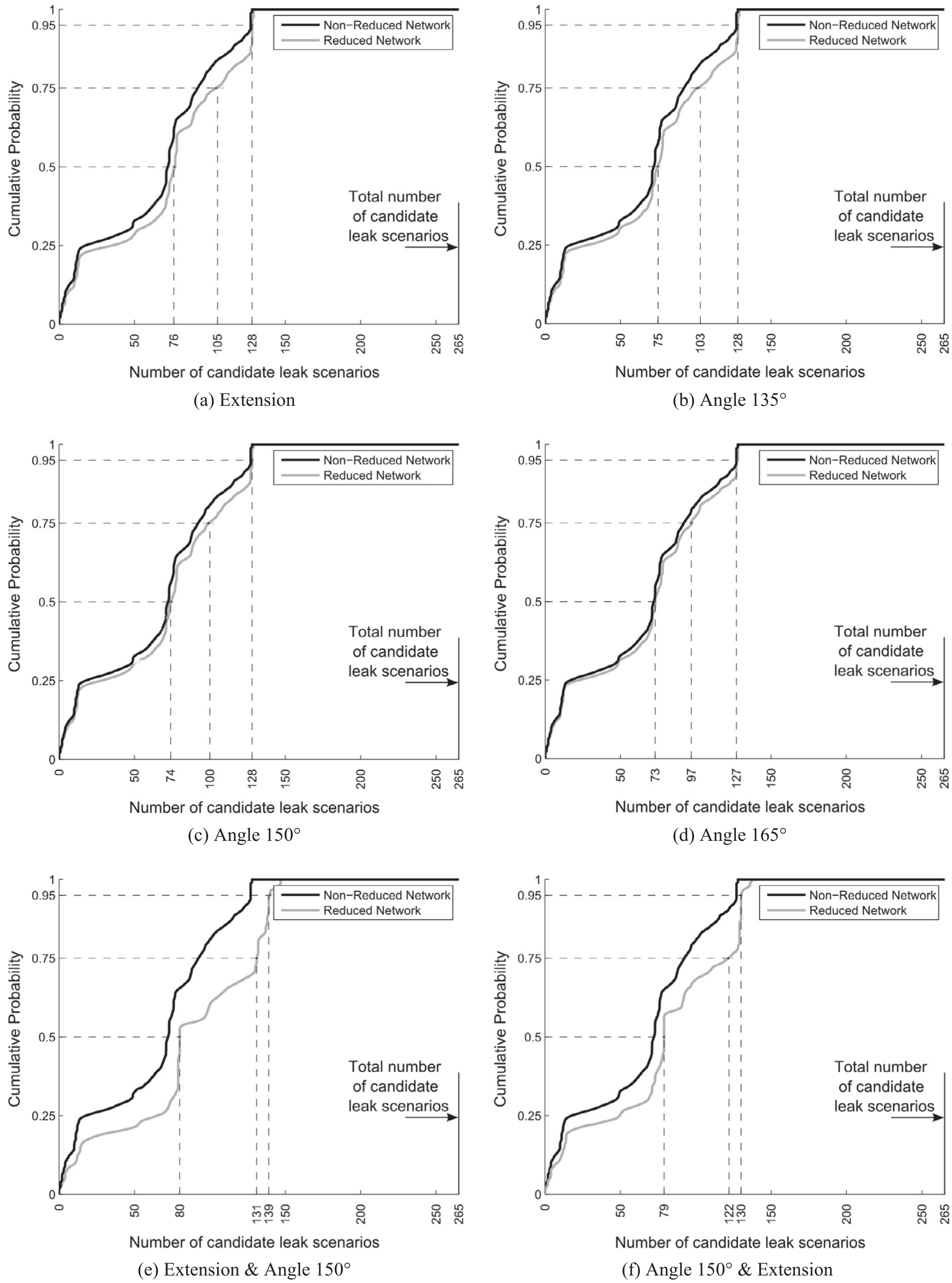


Fig. 13. CDFs for the expected number of candidate leak scenarios in the case of a leak severity of 100 l/min for: Consumption 5% (a), Consumption 10% (b), Consumption 25% (c), Consumption 50% (d), Diameter 150 (e) and Diameter 200 (f).

Please cite this article in press as: G. Moser et al., Performance comparison of reduced models for leak detection in water distribution networks, Adv. Eng. Informat. (2015), http://dx.doi.org/10.1016/j.aei.2015.07.003



(a) Extension
(b) Angle 135°
(c) Angle 150°
(d) Angle 165°
(e) Extension & Angle 150°
(f) Angle 150° & Extension

Fig. 14. CDFs for the expected number of candidate leak scenarios in the case of a leak severity of 100 l/min for: Extension, (a), Angle 135 (b), Angle 155 (c), Angle 165 (d), Extension & Angle 150 (e) and Angle 150° & Extension (f).

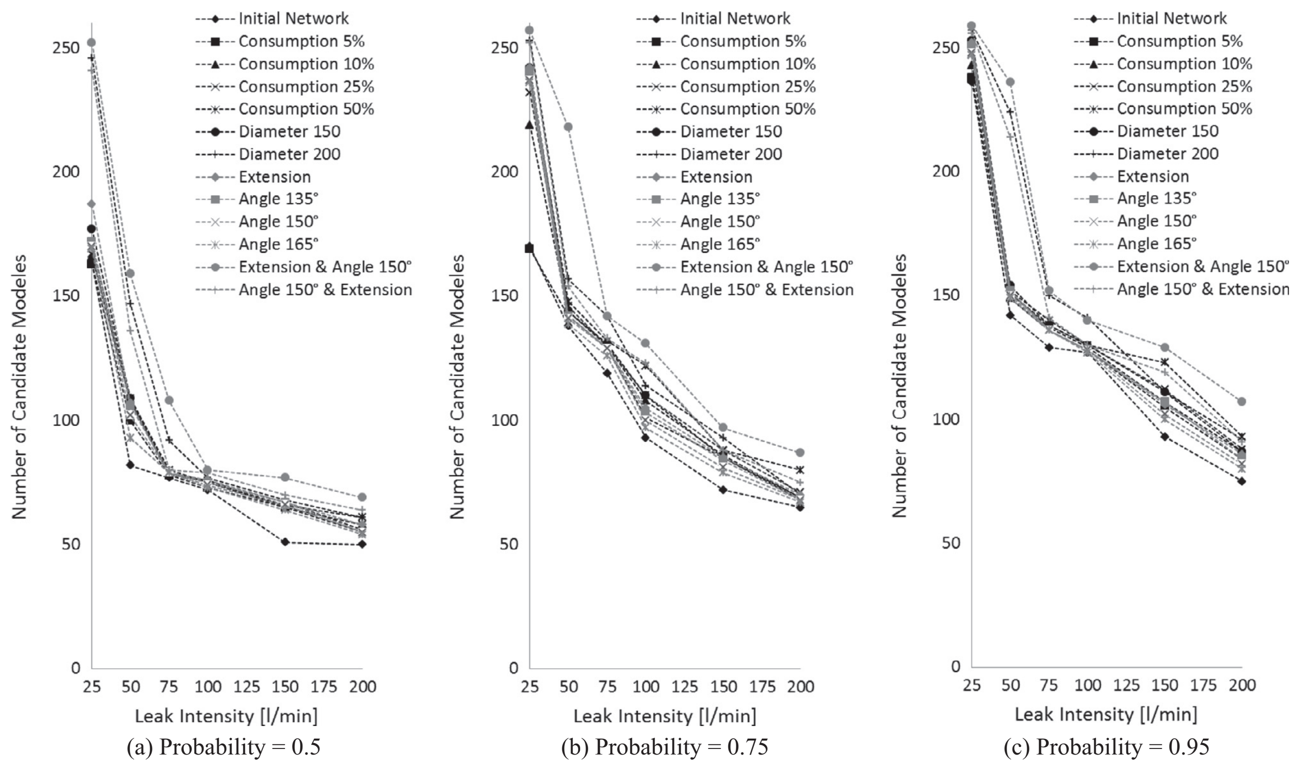


Fig. 15. Evolution of the expected number of candidate leak scenarios for: (a) 0.5 probability, (b) 0.75 probability and (c) 0.95 probability.

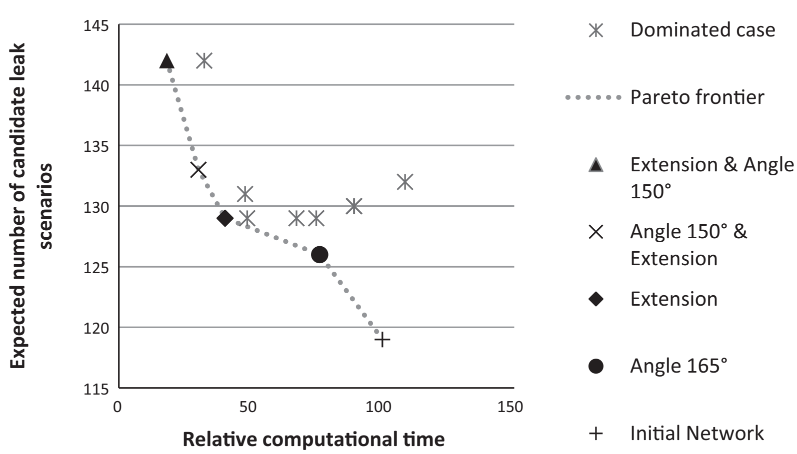


Fig. 16. Pareto front for comparison of computational time with expected identifiability (75% probability and 75 l/min leak intensity).

Considering the performance, the criteria chosen are the following: (1) the expected number of candidate leak scenarios with a probability of 75%; (2) the leak intensity. Instead of computing the Pareto front for the performance at all leak intensities at once, the performances at each intensity are compared separately. Then, all Pareto analyses are considered, and the front which yields the highest number of dominated strategies (Fig. 16) is employed to determine the best reduction strategies.

The case with the lowest number of elements on the Pareto front is the leak intensity of 75 l/min (Fig. 16). The horizontal axis gives the relative computational time and the vertical axis the expected number of candidate leak scenarios. The dashed line is the Pareto front. The strategies located on the front are: *Initial Network*, *Angle 165°*, *Extension*, *Angle 150° & Extension* and *Extension & Angle 150°*. All other reduction strategies are dominated in this case. Pareto analyses thus reduce the choice from twelve strategies to four (excluding the *Initial Network*).

In addition, the *Angle 150° & Extension* and *Extension & Angle 150°* reduction strategies can be eliminated due to the sensitivity to leak severity as displayed in Fig. 15. Although *Angle 165* has good performance, the computational time is too high (76.2% in comparison to the initial network), and thus, this strategy can also be eliminated. Consequently, the result of this study reveals that the *Extension* reduction strategy as well suited for leak detection using model falsification.

5. Discussion

The methodology described in this paper was illustrated using one network with one sensor configuration. Increasing the number of sensors will increase the performance of the diagnosis. This will result in cumulative distribution functions that are situated more to the left of the curves presented in this paper. The conclusions of this paper are therefore not expected to be influenced by the use of more sensors. Such a generalization probably cannot be made for a significantly different network. This paper provides a methodology for determining the best reduction strategy for other networks. It is a tool to help a manager of a water supply network to adapt this leak methodology to his network.

Using other diagnostic methodologies may not lead to the same conclusions. This study was carried out assuming the use of model falsification for structural identification. Further work could involve studies of reduction strategies in combination with other diagnostic methodologies.

Future work will consist of testing the leak detection methodology that combines network reduction and error-domain model falsification with measurements in order to illustrate the strengths and weakness of the methodology through a detailed study of full-scale application. The performance of the error-domain model falsification framework can be improved by increasing the quantity of information that is available. Increasing information can be in the form of additional sensors or a decrease in the unknowns of the system.

6. Conclusions

The analysis of the results leads to the following conclusions.

Reduction strategies used in this paper are useful for reproducing flows with simplified networks.

Since it is possible to reduce computational time to up to 20% of the time for the non-reduced network, gains can be significant.

The strategies that reduce computational time the most are also those which are most sensitive to leak severity. Strong network reduction may lead to decreased performance for small leaks faster than networks with lighter reduction.

The reduction procedures that are most suited for leak detection using model falsification are *Consumption 10%, Extension* and *Angle 135*°. A Pareto analysis shows that a good compromise between the reduction of computational time and diagnostic performance is given by the *Extension* reduction strategy.

Acknowledgments

This research is part of the Water resources Innovation Program of EPFL Middle East and is funded by EPFL Middle East. The authors acknowledge the support from Eau Service, the water provider of the city of Lausanne.

References

[1] H. Abdi, Encyclopedia of Measurement and Statistics, The Bonferonni and Sidák Corrections for Multiple Comparisons, 2007 (Chapter).
[2] J.H. Andersen, R.S. Powell, Implicit state-estimation technique for water network monitoring, Urban Water 2 (2000) 123–130.
[3] H. Babbitt, F. Amsbary, D. Gwinn, The detection of leaks in underground pipes [with discussion], J. Am. Water Works Assoc. 7 (1920) 589–595.
[4] N. Balabanian, T.A. Bickart, Electrical Network Theory, John Wiley & Sons Inc., 1969.
[5] M.A. Barandouzi, G. Mahinthakumar, E.D. Brill, R. Ranjithan, Probabilistic mapping of water leakage characterizations using a Bayesian approach, in: Proceedings of World Environmental and Water Resources Congress 2012, ASCE (American Society of Civil Engineers), Albuquerque, New Mexico, USA, 2012, pp. 3248–3256.
[6] A.F. Colombo, P. Lee, B.W. Karney, A selective literature review of transient-based leak detection methods, J. Hydro-environ. Res. 2 (2009) 212–227.
[7] S. Demirci, E. Yigit, I.H. Eskidemir, C. Ozdemir, Ground penetrating radar imaging of water leaks from buried pipes based on back-projection method, NDT & E Int. 47 (2012) 35–42.
[8] Y. Gao, M.J. Brennan, P.F. Joseph, A comparison of time delay estimators for the detection of leak noise signals in plastic water distribution pipes, J. Sound Vib. 292 (2006) 552–570.
[9] Y. Gao, M.J. Brennan, P.F. Joseph, On the effects of reflections on time delay estimation for leak detection in buried plastic water pipes, J. Sound Vib. 325 (2009) 649–663.
[10] J.-A. Goulet, S. Coutu, I.F.C. Smith, Model falsification diagnosis and sensor placement for leak detection in pressurized pipe networks, Adv. Eng. Inform. 27 (2013) 261–269.
[11] J.-A. Goulet, I.F.C. Smith, Predicting the usefulness of monitoring for identifying the behavior of structures, J. Struct. Eng. 139 (2013) 1716–1727.
[12] J.-A. Goulet, I.F.C. Smith, Structural identification with systematic errors and unknown uncertainty dependencies, Comput. Struct. 128 (2013) 251–258.
[13] D. Grunwell, B. Ratcliffe, Location of Underground Leaks using the Leak Noise Correlator, WRC Technical Report, 1981 (Chapter).
[14] G. Hämmerlin, K.H. Hoffmann, Numerical Mathematics, Springer-Verlag, 1991.
[15] W. Hope, The waste of water in public supplies, and its prevention (including appendix), in: Minutes of the Proceedings, vol. 110, 1892. <http://www.icevirtuallibrary.com/content/article/10.1680/imotp.1892.20225>.
[16] A. Lambert, W. Hirner, Losses from water supply systems: standard terminology and recommended performance measures, in: The Blue Pages the IWA Information Source on Drinking Water Issues, 2000.
[17] H. Mala-Jetmarova, A. Barton, A. Bagirov, Exploration of the trade-offs between water quality and pumping costs in optimal operation of regional multiquality water distribution systems, J. Water Resour. Plann. Manage. (2014).
[18] F. Martinez Alzamora, B. Ulanicki, E. Salomons, Fast and practical method for model reduction of large-scale water-distribution networks, J. Water Resour. Plann. Manage. 140 (2014) 444–456.
[19] B. Mergelas, G. Henrich, Leak locating method for precommissioned transmission pipelines: North American case studies, Leakage 2005 (2005) 12–14.
[20] J. Morrison, Managing leakage by district metered areas: a practical approach, Water 21 (2004) 44–46.
[21] G. Moser, I.F.C. Smith, Detecting leak regions through model falsification, in: Proceedings of 20th International Workshop: Intelligent Computing in Engineering 2013, European Group for Intelligent Computing in Engineering (EG-ICE), Vienna, Austria, 2013.
[22] S. Mounce, J. Boxall, J. Machell, Development and verification of an online artificial intelligence system for detection of bursts and other abnormal flows, J. Water Resour. Plann. Manage. 136 (2009) 309–318.
[23] S.R. Mounce, R.B. Mounce, J.B. Boxall, Novelty detection for time series data analysis in water distribution systems using support vector machines, J. Hydroinform. 13 (2011) 672–686.
[24] I. Nouiri, Multi-objective tool to optimize the water resources management using genetic algorithm and the Pareto optimality concept, Water Resour. Manage. (2014) 1–17.
[25] V. Pareto, Cours d'économie politique, vols I & II, Rouge, Lausanne, Switzerland, 1896.
[26] L. Perelman, M.L. Maslia, A. Ostfeld, J.B. Sautner, Using aggregation/skeletonization network models for water quality simulations in epidemiologic studies, J. Am. Water Works Assoc. 100 (2008) 122–133.
[27] L. Perelman, A. Ostfeld, Water distribution system aggregation for water quality analysis, J. Water Resour. Plann. Manage. 134 (2008) 303–309.
[28] L. Perelman, A. Ostfeld, Water-distribution systems simplifications through clustering, J. Water Resour. Plann. Manage. 138 (2011) 218–229.
[29] Z. Poulakis, D. Valougeorgis, C. Papadimitriou, Leakage detection in water pipe networks using a Bayesian probabilistic framework, Probab. Eng. Mech. 18 (2003) 315–327.
[30] A. Preis, A. Whittle, A. Ostfeld, L. Perelman, On-line hydraulic state estimation in urban water networks using reduced models, in: Proceedings of Tenth International Conference on Computing and Control in the Water Industry, Sheffield, UK, 2009.
[31] A. Preis, A.J. Whittle, A. Ostfeld, L. Perelman, Efficient hydraulic state estimation technique using reduced models of urban water networks, J. Water Resour. Plann. Manage. 137 (2011) 343–351.
[32] R.S. Pudar, J.A. Liggett, Leaks in pipe networks, J. Hydraul. Eng. 118 (1992) 1031–1046.
[33] R. Puust, Z. Kapelan, T. Koppel, D. Savic, Probabilistic leak detection in pipe networks using the SCEM-UA algorithm, in: Proceedings of Water Distribution Systems Analysis Symposium 2006, ASCE (American Society of Civil Engineers), Cincinnati, Ohio, USA, 2006, pp. 1–12.
[34] R. Puust, Z. Kapelan, D.A. Savic, T. Koppel, A review of methods for leakage management in pipe networks, Urban Water J. 7 (2010) 25–45.
[35] Y. Robert-Nicoud, B. Raphael, I. Smith, Configuration of measurement systems using Shannon's entropy function, Comput. Struct. 83 (2005) 599–612.
[36] M. Romano, Z. Kapelan, D. Savić, Evolutionary algorithm and expectation maximization strategies for improved detection of pipe bursts and other events in water distribution systems, J. Water Resour. Plann. Manage. 140 (2014) 572–584.
[37] M. Romano, Z. Kapelan, D.A. Savić, Automated detection of pipe bursts and other events in water distribution systems, J. Water Resour. Plann. Manage. 140 (2012) 457–467.
[38] M. Romano, Z. Kapelan, D.A. Savić, Geostatistical techniques for approximate location of pipe burst events in water distribution systems, J. Hydroinform. 15 (2013) 634–651.
[39] A. Rosen, A new network theorem, J. Inst. Electr. Eng. 62 (1924) 916–918.
[40] L.A. Rossman, EPANET 2: Users Manual, Cincinati, OH, US Environmental Protection Agency, 2000.
[41] J. Rougier, Probabilistic leak detection in pipelines using the mass imbalance approach, J. Hydraul. Res. 43 (2005) 556–566.
[42] Sensus, Water 20/20: Bringing Smart Water Networks into Focus, Technical Report, 2012.

[43] S. Srirangarajan, M. Allen, A. Preis, M. Iqbal, H.B. Lim, A.J. Whittle, Water main burst event detection and localization, in: Proceedings of 12th Water Distribution Systems Analysis Conference (WDSA'10), 2010.

[44] B. Ulanicki, A. Zehnpfund, F. Martinez, Simplification of water distribution network models, in: Proceedings of Second International Conference on Hydroinformatics, IAHR (International Association for Hydraulic Research), Zürich, Switzerland, 1996, pp. 493–500.

[45] J.P. Vítkovský, M.F. Lambert, A.R. Simpson, J.A. Liggett, Experimental observation and analysis of inverse transients for pipeline leak detection, J. Water Resour. Plann. Manage. 133 (2007) 519.

[46] J.P. Vítkovský, A.R. Simpson, M. Lambert, Leak detection and calibration using transients and genetic algorithms, Water Resour. Plann. Manage. (2000) 258–262.

[47] A. Whittle, M. Allen, A. Preis, M. Iqbal, Sensor networks for monitoring and control of water distribution systems, in: Proceedings of the 6th International Conference on Structural Health Monitorinf of Intelligent Infrastructure, Hong Kong, 2013.

[48] A.J. Whittle, L. Girod, A. Preis, M. Allen, H.B. Lim, M. Iqbal, S. Srirangarajan, C. Fu, K.J. Wong, D. Goldsmith, WATERWISE@ SG: a testbed for continuous monitoring of the water distribution system in Singapore, Water Distrib. Syst. Anal. WSDA (2010).