Automated leak localization performance without detailed demand distribution data

Moors, Janneke; Scholten, Lisa; van der Hoek, Jan Peter; den Besten, Jurjen

DOI
10.1080/1573062X.2017.1414272

Publication date
2018

Document Version
Final published version

Published in
Urban Water Journal

Citation (APA)
Moors, J., Scholten, L., van der Hoek, J. P., & den Besten, J. (2018). Automated leak localization performance without detailed demand distribution data. *Urban Water Journal*, *15*(2), 116-123. https://doi.org/10.1080/1573062X.2017.1414272

Urban Water Journal

ISSN: 1573-062X (Print) 1744-9006 (Online) Journal homepage: http://www.tandfonline.com/loi/nurw20

Automated leak localization performance without detailed demand distribution data

J. Moors, L. Scholten, J. P. van der Hoek & J. den Besten

To cite this article: J. Moors, L. Scholten, J. P. van der Hoek & J. den Besten (2018) Automated leak localization performance without detailed demand distribution data, Urban Water Journal, 15:2, 116-123, DOI: 10.1080/1573062X.2017.1414272

To link to this article: https://doi.org/10.1080/1573062X.2017.1414272

URBAN WATER JOURNAL, 2018
VOL. 15, NO. 2, 116–123
https://doi.org/10.1080/1573062X.2017.1414272

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Automated leak localization performance without detailed demand distribution data

J. Moors[a,b], L. Scholten[a], J. P. van der Hoek[a,c] and J. den Besten[b]

[a]Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, The Netherlands; [b]Assetmanagement, Oasen N.V., Gouda, The Netherlands; [c]Waternet, Strategic Centre, Amsterdam, The Netherlands

ABSTRACT
Automatic leak localization has been suggested to reduce the time and personnel efforts needed to localize (small) leaks. Yet, the available methods require a detailed demand distribution model for successful calibration and good leak localization performance. The main aim of this work was to analyze whether such a detailed demand distribution is needed. Two demand distributions were used: a factorized distribution that distributes the inflow demand proportionally across the consumption nodes according to individual billing data, and a uniform distribution that equally distributes demand across all consumption nodes. The performance of the automatic leak localization method, using both demand distribution models, was compared. A new measure for leak localization performance that is based on the percentage of false positive nodes is proposed. It was possible to localize the leaks with both demand distribution models, although performance varied depending on the timing and duration of the measurement.

ARTICLE HISTORY
Received 12 October 2016
Accepted 29 November 2017

KEYWORDS
Model-based leak localization; demand distribution; district metered area

Introduction

Consumers normally report bursts that cause visible implications such as surface flooding. Leaks in water supply networks that do not cause water to come up to the surface can continue unreported for a long time, resulting in potentially large volumes of lost water. Additionally, moving ground particles around a small leak tend to enlarge a leak (Puust et al. 2010). Earlier and automatic leak localization in addition to fast leak detection can save water and prevent small leaks turning into bursts. In order to detect leaks automatically, the network is divided into smaller areas whose inflows are individually metered ('District Metered Areas'; DMAs). Leaks in DMAs are detected by making a simple water balance that compares expected demand and actual water use (Bakker 2014; Romano, Kapelan, and Savić 2012). Additionally, an increase of the minimum night flow (MNF) is used for the detection of new leaks (Farley et al. 2008). Reducing the search area from the whole network to a DMA does not solve the problem of finding the exact location of a leak and therefore, apart from leak detection, an additional leak localization technique is needed to localize the leak (Bakker 2014).

Currently, leak localization is frequently done with acoustic equipment such as listening rods, leak correlators, leak noise loggers and non-acoustic methods like gas injection, ground penetrating radar technology and infrared photography (Li et al. 2015). These methods are very accurate, but it takes a long time to find a leak in a large search area (Li et al. 2015; Puust et al. 2010). Additionally, acoustic equipment is less effective for new pipe materials (for example polyvinyl chloride, PVC). Noise is transmitted less far in these (non-metallic) materials, resulting in a longer searching time (Gao et al. 2005; Li et al. 2015). Even with small DMAs, leak localization results in a time-consuming and labor-intensive process.

To reduce the search area and time, software-based methods are used next to the previously described hardware based methods (Li et al. 2015). These can be divided into non-numerical and numerical modeling methods. Non-numerical models use Artificial Intelligence techniques which need historical pressure data for training (Li et al. 2015). The drawback of these techniques is that the required large amounts of training data are not always available (yet) (Li et al. 2015; Tao et al. 2013). The numerical modeling methods use a model and compare simulated results with field data (Li et al. 2015). One of these methods is based on pressure transients in pipes. After the appearance of a leak, the pressure wave can accurately be localized by the use of sensors with high sampling intervals. This method is very sensitive to the exact configuration of the whole network, which is often unknown in real systems. Together with, amongst others, the perceived cost of high-frequency pressure loggers and the unwillingness to generate water hammer in the system, its practical application is scarce or limited to individual large pipelines (Li et al. 2015; Puust et al. 2010; Savic, Kapelan, and Jonkergouw 2008). Model-based approaches that also consider changes and are applicable to a network of pipes is the model-based leak localization method by Pérez et al. (2011a) and Farley, Mounce, and Boxall (2013). While the approach by Farley et al. (2013) aims to determine subareas within which a leak is likely located using a minimum of measurement devices, the method by Pérez et al. (2011a) employs fault isolation to identify the most probable location of the leak within the network. This latter method seeks for pressure

CONTACT J. Moors janneke.moors@oasen.nl
Supplemental material for this paper is available online at https://doi.org/10.1080/1573062X.2017.1414272.

anomalies between field measurements and simulated values from a hydraulic model in a 'leak-free situation' above a certain threshold. Quevedo et al. (2011) improved the method by using the residual between field measurements and simulated values directly without binarising, implying information loss. Hence, it also compares the difference between modeled and observed pressure, but instead of focusing on transients, it considers the spatially distributed changes in pressure. The method of Quevedo et al. (2011) was validated by Mirats-Tur et al. (2014) for two leaks in the DMA Nova Icària in Barcelona. The leaks were successfully localized within a 150 m radius from the real leak (Mirats-Tur et al. 2014). This approach is reported to require accurate estimates of the spatial distribution of customers' water demand, because of its influence on pressure variations and therefore also leak localization performance (Cugueró-Escofet et al. 2015; Meseguer et al. 2014; Mirats-Tur et al. 2014; Pérez et al. 2011b; Sanz and Pérez 2014; Sanz et al. 2015). For the same reasons, Pérez et al. (2009) have recommended to perform leak localization during night hours when there is less consumption and hence less noise in the pressure data. These high data demands are a serious limitation for the practical applicability of this approach, as such detailed demand data are often unavailable. At the same time, it seems somewhat contradictory that very detailed demand data should be obtained only to apply the method at a time where almost no demand occurs. Consequently, this work aims to identify whether a much simpler, uniform demand distribution would be sufficient for leak localization as compared to a factorized demand distribution model (based on individual billing data), using the method of Quevedo et al. (2011). Furthermore, an accurate spatial demand distribution may not be needed when the leakage is the same or even three times the MNF, as typical for smaller networks (or DMAs). Given that the influence of different factors on pressure variation – and hence leak localization performance – are not well-studied, we also investigated the impact of demand fluctuations at different times of the day, different leak sizes, and different locations within the topography of the network. These three research questions were studied in the DMA Leimuiden, the Netherlands, with a MNF of 4.5 m^3/h, this is approximately 7% of the amount of the already studied DMA Nova Icària (Pérez et al. 2014a). In order to compare the leak localization performance across experimental designs (different leak sizes, locations, and times of the day), we propose a novel performance indicator that is based on the percentage of incorrectly detected leaks as opposed to the distance to the real leak.

Methods

Model-based leak localization

The leak localization method of Quevedo et al. (2011) is based on the comparison of the observed pressure changes in the network and the simulated pressure changes of multiple simulations, it was applied sequentially, assuming a leak at every node. The model allows simulating a 'leak-free situation' in order to calculate the fault indicator vector ϕ, obtained by subtracting the measurements, p, from their model-estimates in a leak free scenario, $\hat{p}_0$, with ns being number of sensors (Meseguer et al. 2014):

$$\phi(t) = \begin{bmatrix} p_1(t) - \hat{p}_{10}(t) \\ \vdots \\ p_{ns}(t) - \hat{p}_{ns0}(t) \end{bmatrix} \tag{1}$$

Subsequently the model is used to simulate sequentially a leak in all nodes, resulting in different pressures compared to the modeled pressures without a leak. This is stored in the theoretical fault signature matrix (FSM). The FSM contains the modeled differences between the leak-free pressures ($\hat{p}_0$) and the pressures ($\hat{p}$) after a modeled leak, with ns being number of sensors and nn number of nodes:

$$\mathrm{FSM}(t) = \begin{bmatrix} \hat{p}_{1,1}(t) - \hat{p}_{10}(t) & \cdots & \hat{p}_{1,nn}(t) - \hat{p}_{10}(t) \\ \vdots & \ddots & \vdots \\ \hat{p}_{ns,nn}(t) - \hat{p}_{ns0}(t) & \cdots & \hat{p}_{ns,nn}(t) - \hat{p}_{ns0}(t) \end{bmatrix} \tag{2}$$

All the columns of the FSM are correlated (using Pearson's correlation coefficient) with the fault indicator vector, resulting in a correlation value c for every node. By sorting the correlations in descending order, a ranking of the nodes with the most probable leak locations is created. The maximum values in this vector represents the most probable nodes to have a leak. In Supplemental Material 1 the time dependency is explained.

Hydraulic model set up, calibration and leak simulation

In order to mathematically represent the flows and pressure changes in the real network of DMA Leimuiden, a hydraulic model was built with corresponding links (3243, representing pipes) and nodes (3218, representing consumption points and pipe connections). The hydraulic modeling software EPANET 2.0 was used with the default values for all model parameters (e.g. specific gravity, relative viscosity and maximum number of trials) except for the formula to calculate the head loss (Darcy-Weisbach instead of Hazen-Williams) and are described in the EPANET 2 User's Manual (Rossman 2000). The EPANET Programmer's Toolkit was used in Python to automate model calibration and leak simulation for all 3218 nodes (van Rossum 1995). Leaks are modeled as a constant demand. The geometrical and structural pipe data and customer connections (pipe and node identifier (ID), inner diameter, length, roughness values as used by the water company) were derived from the geographic information system (GIS)-database and the 'basic administration of addresses and buildings'. To study the influence of the demand distribution on the performance of the leak localization method two demand distributions were used. The less data-demanding approach was the uniform model. It distributes the DMA inflow equally across all customers. The second model was the factorized model. This model divides the demand in the same way as proposed in Quevedo et al. (2011): the customers get a factor relative to the overall consumption based on the billing information. Hereby the assumption is that a customer who uses more water in a year, also uses more water in every period of time, proportionally to the yearly consumption. For the factorized model the billing information of the DMA Leimuiden in 2014 was used. The model was calibrated for both, once assuming the uniform and once the factorized demand distribution.

This is important, because the demand model used for calibration directly affects the performance of leak localization. The flow measurements at the inlet are allocated to the individual nodes across the network using the particular demand distribution (uniform or factorized). Pressure head measurements at the inlet are used to model reservoir levels at the inlet of the model (average piezometric head is 35.7 m). These form the boundary conditions under which the model is calibrated and leaks are simulated (Sanz and Pérez 2014). The simulated pressures and the measured pressures in the low flow conditions were used for the best estimate of the measurement node elevations. First, all node elevations in the model are set to zero. During calibration, observed pressures and simulated pressures are compared. The difference between the observed pressures and the simulated pressures during low consumption are used to set the elevation of the nodes in the model that have a pressure sensor in the real network. The elevations were set during minimum consumption when the influence of the roughness of the pipe is small. The root mean squared error (RMSE) is used as the objective function in our calibration for our calibration parameter pipe roughness. The Darcy–Weisbach roughness coefficient was adjusted to minimize the RMSE during extended period simulation calibration (one day, 96 time steps).

Design of the artificial leaking campaign

The DMA Leimuiden spans 1.46 km^2 with 4285 inhabitants (1835 households) (CBS 2015). The network length is 26 km and it consists mainly of PVC (54%). The delivered monthly volume is 20,000 m^3 to 1932 customer connections. During this research, the water of DMA Leimuiden was supplied by one inlet. For being able to compare simulated pressures with observed pressures in the leak free and leak situation, a measurement campaign, in which artificial leaks were created at hydrants in the network, was set up. This required defining the measurement devices and where to place them, the leak locations, leakage sizes (flows) and durations.

Sensor placement and deployment

The performance of the leak localization method and the location of the sensors are interdependent (Bonada and Meseguer 2014a; Bonada, Meseguer, and Mirats-Tur 2014b; Meseguer et al. 2014). Initially it was planned to install all 15 smart meters that Oasen, the water company servicing DMA Leimuiden, had acquired. The smart meters comprised a pressure sensor, temperature sensor and a pulse counter that can be attached to and installed next to an original water meter to derive the inflow. The pressure device included a pressure transmitter (JUMO MIDAS C08, measurement error 0.35% of full scale: 0–4 bar, 20°C), one measurement per five seconds, reported resolution 0.01 bar). The locations were selected based on the highest flow velocities and preferred flow direction as identified by modeling the current hydraulic situation. At these points the change of pressure was expected to be the most sensitive. In our case, the maximum velocity ranged from 0.01 to 0.15 m/s. This amounted to simulated nighttime pressure drops between 0.20 kPa in location I and 10.98 kPa in location III. Furthermore, the sensors were planned at the outer areas and on long pipe segments. They were installed in front of the existing water meters at the customer connections, from which inflow measurements were obtained. When the customer does not use water, the pressure at the house is the same as the pressure in the network. Due to unacceptance by customers, holidays, practical limitations of the location and firmware problems, only six of the 15 planned smart meters were installed and in operation before the start of the artificial leak campaign. Therefore, seven pressure loggers (PrimeLog+, measurement error 0.1% of full scale: 0–10 bar, one measurement per second) were additionally installed on fire hydrants. Their location was defined with the optimal model-based sensor location deployment procedure and used a Genetic Algorithm (Bonada, Meseguer, and Mirats-Tur 2014b) implemented in the Python package DEAP (Fortin et al. 2012). To perform this procedure a uniform demand model is calibrated with pressure measurement at nine locations across the DMA, in the same way as explained in the section 'Hydraulic model set up, calibration and leak simulation'. The optimization objective was to minimize the largest group with the same leak response (leak response group) after truncating the pressure response to one decimal. The parameters were; population size: 4000, crossover fraction: 0.5, individual mutation probability: 0.2, attribute mutation probability: 0.1, number of generations: 20.

Furthermore, flow and pressure measurements were obtained at the inlet. Pressure at the inlet was measured with the same pressure device as used at the fire hydrants (PrimeLog+). The flow meter at the inlet (KROHNE) had a measurement uncertainty of 0.2% of the measured value and a logging interval of one in 15 min. In Supplemental Material 2 the sensor deployment of smart meters and pressure loggers as well as the DMA inlet with flow and pressure measurements is shown.

Leak location, size, timing and duration

Sixteen artificial leaks, with a duration of 15 min, were created by opening fire hydrants at three different locations, for two leak sizes (i.e. two opening degrees of the hydrants), and at three different times of the day. The locations of the artificial leaks were chosen in a different leak response group (according to the uniform model at MNF) and based on their accessibility. Locations are shown in Supplemental Material 2. For being able to assess the influence of demand variations on leak localization performance, artificial leaks were created during the night, morning, and day. No measurements were taken at location II during the morning peak. The measurements were performed on one day and the morning peak was too short to open fire hydrants at all three locations with two leak sizes. Assuming that leak size will also affect leak localization performance, at each location first a larger and then a smaller leakage of 15 and 7.5 m^3/h, respectively, were performed (leak schedule in Supplemental Material 3). This gives an indication since a fair comparison should have the same inflow but these may differ within half an hour, especially in the morning peak. By incident, a leakage of approximately 5.2 m^3/h occurred during the time when the smart sensors were placed in the network (location 'RL' in Supplemental Material 2). For the real leak, measurements of a whole day were available (see inflow hydrograph later). This provided the opportunity to test both leak localization approaches for a real leak. The leak localization of the real leak was performed with the smart meters only since the pressure loggers were already removed when the leak occurred.

Applied time steps

To execute the leak localization method three different time steps are needed (Figure S1). These are the sensor sample time (T_s), scenario analysis time step (T_a) and the diagnosis sliding window (T_w) (Meseguer et al. 2014) (Supplemental Material 1). The T_s for the pressure measurements were sampled every second (pressure loggers) and every 5 s (smart meters). The mean of 15 min for these values is calculated to retrieve the T_a. This is determined by the larger sensor sample time (T_s) of the flow measurements of the inlet of 15 min. Both, T_a and T_w were 15 min for the artificial leaks. One fault indicator vector ϕ and FSM were computed resulting in one leak localization result with correlations for all nodes. For the leak localization of the real leak extended period simulation was carried out, due to variation in the inflow (see later), whereas for the artificial leak a steady state was assumed. Thus, T_a was again 15 min and both a T_w of 15 min and 60 min were compared. For the real leak 96 (when $T_a = T_w$ was used) or 24 (when $T_w = 60$ min) leak localization results were obtained for one day. In addition to this, one leak localization result was created for one day with the accumulation of the correlations above 0.5 as proposed in Pérez et al. (2014a).

Leak localization and localization performance

In this research the performance of the leak localization with the two demand distribution models was not compared by the distance to the leak, but rather the percentage of false positive (FP) nodes. FP nodes are the nodes that have a higher correlation value than the node in the model where the actual leak was. In practice, this group of FP nodes should be as small as possible since these are locations that are searched first without finding the leak. The exact percentages of FP nodes cannot be compared across locations, because the nodes are not evenly distributed over the area. Hence, the number of FP nodes may be higher in areas with many nodes rather than remote areas. The magnitudes remain relatively comparable, where e.g. a FP percentage below 5% can be interpreted as good across locations.

Results

Calibration

Sensor deployment

The calibration result of the model that is used for sensor deployment of the pressure loggers was a RMSE of 6.73 kPa. This was for nine sensors and 96 time steps and the uniform model.

Leak localization

The RMSE during calibration of the factorized model was smaller than the RMSE of the uniform model (RMSE = 9.02 kPa for the factorized model, RMSE = 9.71 kPa for the uniform model).

Artificial leaks

An overview of the leak localization performance executed with the uniform and factorized distributions is shown in Table 1. Both models demonstrated comparable behavior and performance during nighttime measurements as well as for location III during the morning peak and location II and III during daytime. The factorized model was only slightly better than the uniform model. Overall, the performance was better for larger leaks than smaller leaks (15 and 7.5 m^3/h, respectively). The worst performance was observed for a small leak during nighttime measurements (Table 1, for location I the performance was 74.4% and 72.2% for the uniform and factorized models, respectively). During the morning peak and with a small leak size, the factorized demand model performs much better than the uniform demand model for location I. Both models perform relatively poorly for location I and II during daytime, although the uniform model performs better than the factorized model, except for the larger leak at location I. For all three times, the performance of the uniform and factorized model was under 2.5% for location III.

The models show very different behavior during the time when the hydrants were opened to create the artificial leaks in the morning peak for location I, as shown in Figure 1. Figure 1 shows the relative correlation for each network node. A high correlation with reference to the other nodes in the same network (black nodes) indicates a high probability of the existence of a leak,

Table 1. Leak localization performance (%FP nodes) of the artificial leaks during night, morning peak and daytime and the performance when the correlations are accumulated (accumulated C) for the real leak.

Time	Location	Leakage size(m^3/h)	Q_{leak}/Q_{total}	Uniform demand(FP nodes)	Factorized demand(FP nodes)
Night	I	7.5	0.64	74.4%	72.2%
		15	0.72	41.7%	41.1%
	II	7.5	0.78	2.5%	4.0%
		15	0.96	1.0%	1.2%
	III	7.5	0.82	1.0%	1.3%
		15	0.91	0.8%	1.7%
Morning peak	I	7.5	0.13	14.7%	5.0%
		15	0.27	41.0%	14.0%
	III	7.5	0.17	2.2%	0.8%
		15	0.28	2.0%	1.9%
Daytime	I	7.5	0.20	26.1%	40.7%
		15	0.34	22.1%	7.8%
	II	7.5	0.20	20.1%	40.7%
		15	0.39	18.1%	19.8%
	III	7.5	0.28	2.4%	1.9%
		15	0.42	1.2%	0.8%
Accumulated C, T_w = 15 min	Real Leak	5.2	0.14 to 0.96	0.5%	0.2%
Accumulated C, T_w = 16 min	Real Leak	5.2		11.8%	0.3%

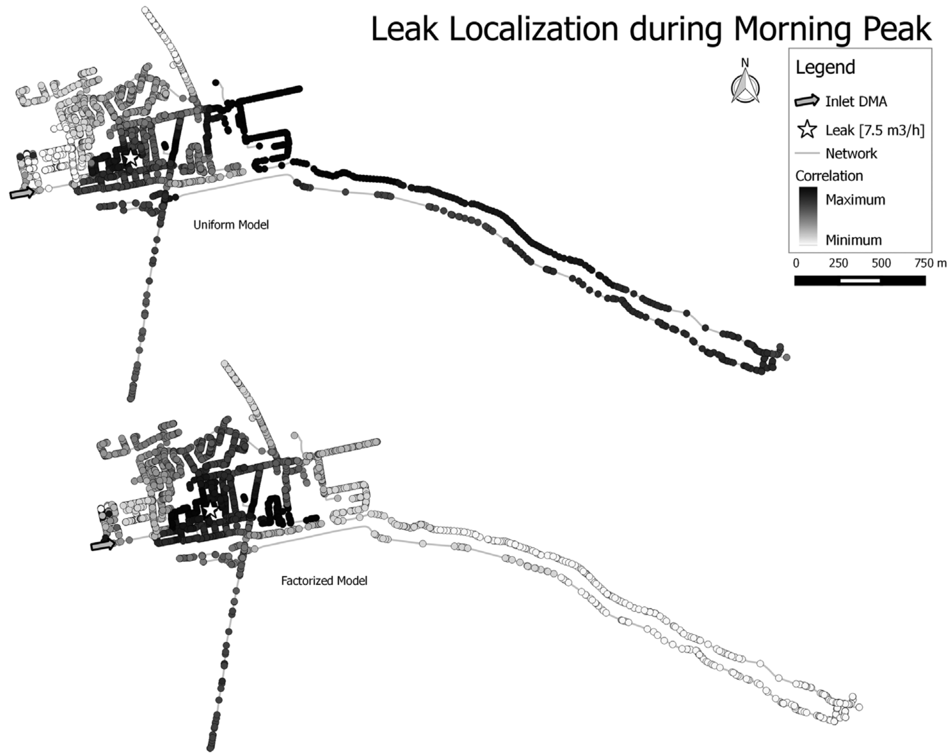

Figure 1. Correlation between observed and simulated leaks with a uniform (top) and factorized demand (bottom) distribution and a leakage size of 7.5 m^3/h. The higher the correlation (darker), the better the result. See section 'Model-based leak localization' in the methods section for details.

while a low correlation with reference to the other nodes in the same network (white nodes) indicates a low probability of a leak at that location. The leak at location I was poorly localized with the uniform model. In contrast, for location I during the morning peak, the factorized model localized the 7.5 m^3/h leakage with 5.0% FP nodes (bottom Figure 1 and Table 1).

Real leak

Figure 2 shows the distance of the accumulated correlations (accumulated over 60 min and correlations above 0.5) from the real leak in Leimuiden. The location of the highest correlation (maximum correlation: C_{max}) is represented by a big star. Smaller stars represent nodes that have correlations larger than 0.99 × C_{max}. Both models had their maximum near the real leak (uniform model: 35 m, factorized model: 11 m).

Because the real leak lasted for more than one day, it was possible to study the leak performance over the day. In Figure 3 the calculated localization performance (black line) is shown. The DMA inflow hydrograph (Figure 3, gray line) is shown as well to clearly visualize the time when the MNF flow and morning peak flow occurs. Notable was that the percentage of FP nodes showed strong fluctuations, ranging from 0% (the real location of the leak) to 41.05% of FP nodes at 16:15 h in the uniform model (Figure 3(A)). The minimum and maximum performance of the leak localizations calculated with the factorized model (Figure 3(B)) are 0% (9:45 h and 12:45 h) and 57.49% (06:45 h) FP nodes respectively. During the time the MNF occurred (between 02:00 h and 04:30 h) the percentage of FP nodes was approximately 10% for both models.

The performance of the leak localization for the same day but with diagnosis sliding window of 60 min (T_w) (combination of four time analysis steps of 15 min) are shown in the two bottom figures of Figure 3. The performance of the uniform model fluctuates between 0.19% and 31.67% FP nodes (minimum 17:00 h and maximum 08:00 h in Figure 3(C)). The performance is rather poor during daytime (> 15% FP nodes), and good during the night (<10% FP nodes). The factorized model fluctuates between 0.03% and 18.77% FP nodes (minimum 09:00 h, maximum 14:00 h in Figure 3(D)), with good performance also during the night and in addition between 09:00–12:00 h in the morning and 15:00–18:00 h in the afternoon.

Discussion

The performance of both demand distribution models was almost the same during the night. At night, the leakage size is large compared to the flow in the network due to low water consumption by customers (Q_{leak}/Q_{total} = 0.64–0.96, Table 1) whereas during the morning and daytime, the leakage size compared to the total flow is smaller (0.13–0.42, Table 1). Because the consumption is small compared to the leakage (large relative leakage volume to demand ratio), the different demand distributions have a small impact on pressure during the night. The leak localization performance was mainly poor for location I. The poor localization of the leaks at location I at night, can probably be explained by the lower ratio of leakage size to total inflow compared to the other two locations (column Q_{leak}/Q_{total} in Table 1). The leaks at location I were created earlier than the leaks at the other two locations (location I start: 00:30 h, location II start

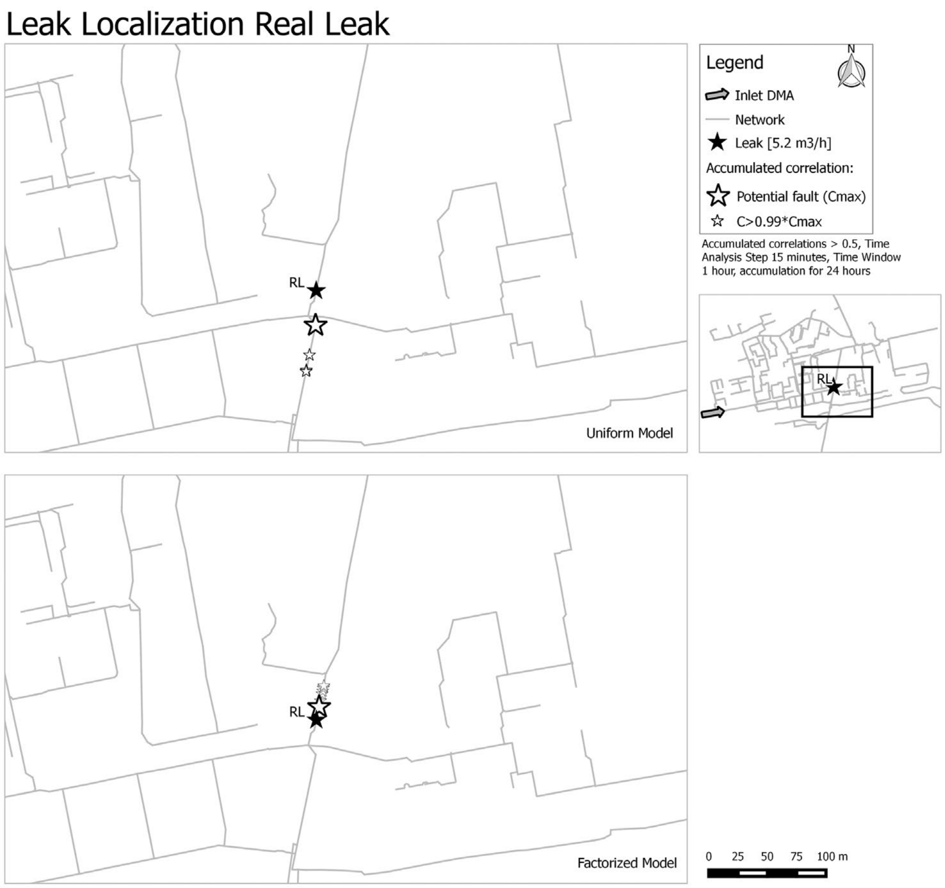

Figure 2. Highest accumulated correlations in the uniform (top) and factorized model (bottom) from the real leak.

02:00 h, location III start 03:00 h, Supplemental Material 3) and the ratio of leakage size to total flow was lower (0.72, 0.96, 0.91 for location I, II, III, respectively, see Table 1). A smaller leakage size to total flow ratio means that there is more inflow due to consumption rather than the leakage, which is distributed over the customers. Possibly, both models distributed the consumption and leakage flow at location I incorrectly during daytime. Since we did not measure the consumption of all the customers in the DMA, other reasons cannot fully be excluded. A wrong demand distribution can result in pressure differences that mask the pressure difference caused by a leak (Pérez et al. 2011a). The head loss caused by a leak is larger at higher flows. This explains why the factorized model was able to localize the leak in the looped part of the network during the morning peak, while it was unable to do so during other times.

Another reason why the leaks at leak location I were poorly localized during the night may be the location of the leak in the network. Leak location I is located in a strongly looped part of the network while location III is localized well and is situated in a branch (Supplemental Material 2). Pérez et al. (2011a) stated that ‘The main handicap of the methodology is that in a highly looped network pressure drops due to a leak are not very significant’ (337). During the night, the maximum simulated pressure response due to a leak at location I with a leakage size of 7.5 m^3/h was low (‑0.29 and ‑0.20 kPa for the uniform and factorized model, respectively). This is in agreement with the literature which states that the precision of the sensor resolution is important for this method (Meseguer et al. 2014). To the authors’ knowledge, the influence of the leak location on localization performance is not widely discussed in the literature, let alone in sufficient detail to derive conclusions. Benchmarks for being able to compare localization performance across networks and methods are missing. The structure of the DMA Leimuiden is a combination of branched and looped parts and hence has a very different structure than for example the (very) regularly looped block-layout as in DMA Nova Icària (Bonada and Meseguer 2014a; Meseguer et al. 2014; Mirats-Tur et al. 2014; Pérez et al. 2014a, 2014b; Quevedo et al. 2011). Although not completely independent of the spatial distribution of the nodes across the network, the leak localization performance measure used in this research allows better cross-network comparison than common distance measures which are often network configuration dependent. When the nodes are evenly distributed (e.g. every 10 m), the leak localization performance measure becomes spatially independent.

The opportunity of the real leak showed a high fluctuation of leak localization performance during the day, both with a diagnosis sliding window (T_w) of 15 min and 60 min. The maximum percentage of FP nodes was higher when the leak localization was performed directly after every time analysis step, thus T_a = T_w = 15 min, than when a diagnosis sliding window (T_w) of 60 min is used. This is consistent with the observation of Meseguer et al. (2014) that the use of multiple time analysis steps adds robustness to the performance of the leak localization. The leak localization result could otherwise be affected by limited sensor resolution or the uncertainties regarding the demand (Meseguer et al. 2014).

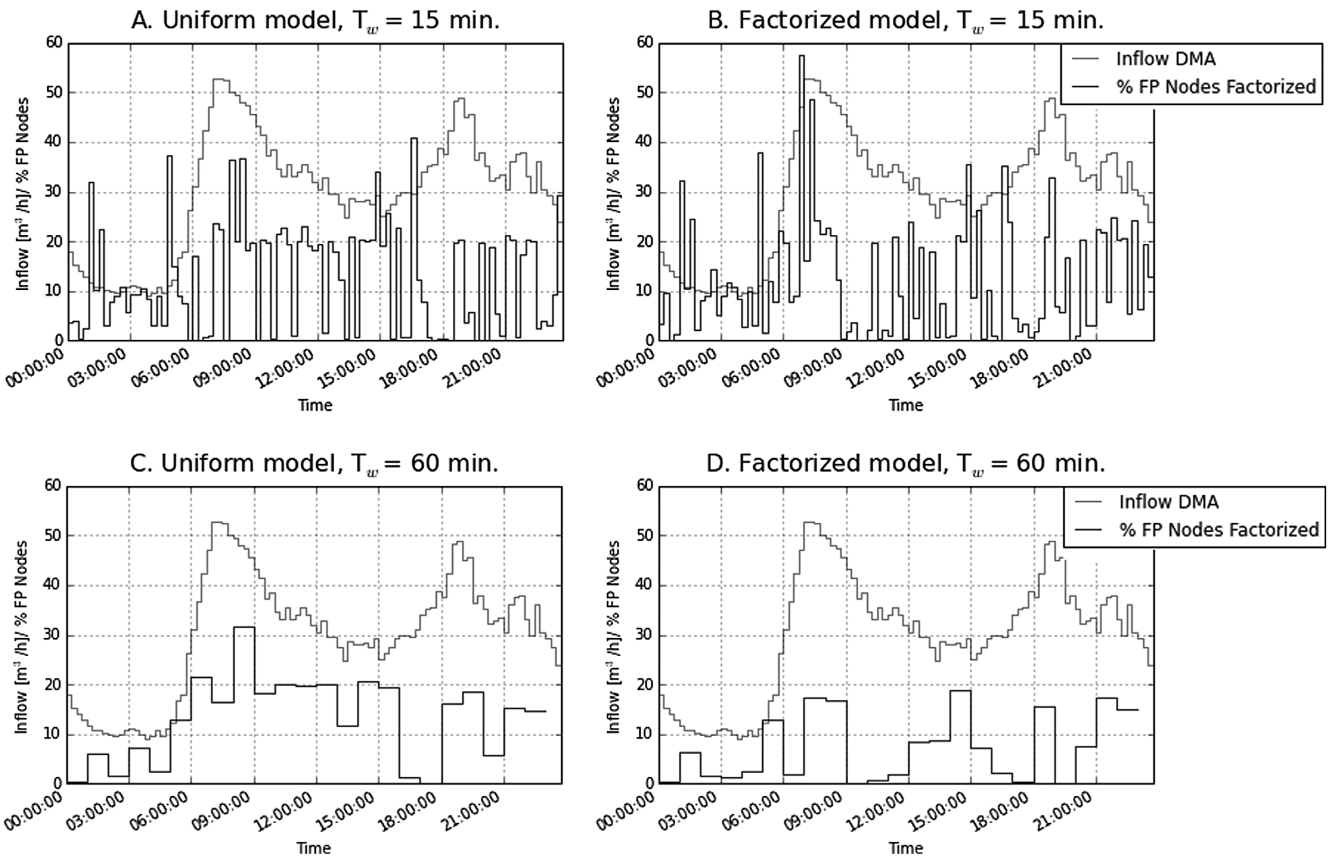

Figure 3. Percentage of false positive (FP) nodes (in black) versus District Metered Area (DMA) inflow (gray) on 8 December 2015 with uniform (graphs A and C) and factorized demand distribution (graphs B and D), time analysis step and time window 15 min (graphs A and B), the time analysis step 15 min and time window 60 min (graphs C and D).

Given the observed dependency of leak localization performance on the relative leakage size (leak-to-demand volume ratio), but also the network topography, sensor placement, and diagnosis sliding window, we believe that the percentage of FP nodes is a more suitable indicator to compare the performance of leak localization approaches in our case than the distance indicators (like distance between the nodes with the highest leak correlation and the leak or the mean distance to the gravity center of those nodes with correlations over 99% of the maximum correlation) as used by e.g. Pérez et al. (2011b, 2011c). Even so, the dependency of the leak localization performance on the conditions of the specific case remains an issue if the ambition were to compare the leak localization results of different methods across cases. In light of the few case studies reported in the literature, it is thus far only possible to derive qualitative insights about the dependencies between different influence factors and leak localization performance. For being able to obtain generalizable results about the influence of other factors and their quantitative implications, systematic studies of the factors that determine leak localization performance in a larger set of cases are needed.

Conclusions

The leak localization method of Quevedo et al. (2011) was validated for 16 artificial leaks of 15 min and one real leak using a factorized and a uniform demand distribution. Although the factorized demand distribution yielded slightly better results than the uniform demand distribution, the results were satisfactory for both distributions and all leaks, the leak being within a search area of less than 15% of the potential leak locations (nodes in the DMA) in most experimental settings. The leak localization method can thus be applied in practice even when the data situation does not allow to use a factorized demand distribution and in DMAs with a small MNF. The number of sensors was relatively high (six smart meters and seven pressure loggers), but the good results may encourage water utilities to invest in such devices. To get more insight in general applicability of the method in practice, more research is needed on the performance of leak localization of different locations in other DMAs with different pipe network structures. Additionally more research is needed on the leak size dependency on the accumulation time step (in this research one day).

The factorized model yields better results than the uniform model, particularly when the leakage sizes are small compared to the flow in the network, for example during the morning peak. With large relative leakage sizes, the difference in localization performance of both demand distributions is negligible, as observed during the night. As a result, using a factorized demand distribution is preferred for artificial leak localization campaigns, particularly when the objective is to localize small leaks or determine leak localization performance. The use of the percentage of FP nodes showed to be a good indicator for the performance of the leak localization method in our study case. Since leak localization performance depends not only on the modeling method, however, but also on the configuration of the network and location of measurement devices, future studies that aim to benchmark approaches across different networks or sensor placement strategies need to explicitly consider such dependencies to avoid potentially biased results.

Acknowledgements

The authors would like to thank drinking water company Oasen and Oasen employees Patrick van Bekkum, Peter van Houwelingen, Jeroen Kegge and Kevin Holthuijsen in particular for their support in placing the sensors, as well as performing the measurement campaigns.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Bakker, M. 2014. "Optimised Control and Pipe Burst Detection by Water Demand Forecasting." Doctoral dissertation, TU Delft, Delft, The Netherlands.

Bonada, E., J. Meseguer, and J. M. Mirats-Tur. 2014a. "On the Structure of the Objective Function for a Pressure Sensor Placement Optimizing Methodology Based on Genetic Algorithms Applied to Model-Based Leakage Localization in Distribution Water Networks." 11th International Conference on Hydroinformatics Proceedings, Session R57, New York.

Bonada, E., J. Meseguer, and J. M. Mirats-Tur. 2014b. "Practical-Oriented Pressure Sensor Placement for Model-Based Leakage Location in Water Distribution Networks." 11th International Conference on Hydroinformatics Proceedings, Session R53, New York.

CBS. 2015. *Kerncijfers wijken en buurten 2014* (Dutch, Translation: Key Figures of Districts and Neighbourhoods). Accessed October 1, 2016. http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=82931NED&D1=0,3,22-26,45-61,99,121&D2=0,802-840,14224-14227&HDR=T&STB=G1&VW=T

Cugueró-Escofet, P., J. Blesa, R. Pérez, M. A. Cuguero-Escofet, and G. Sanz. 2015. "Assessment of a Leak Localization Algorithm in Water Networks under Demand Uncertainty." *IFAC-PapersOnLine* 48 (21): 226–231.

Farley, M., G. Wyeth, Z. M. Ghazali, A. Istandar, S. Singh, N. Dijk, V. Raksakulthai, and E. Kirkwood. 2008. *The Manager's Non-Revenue Water Handbook: A Guide to Understanding Water Losses*. Bangkok: Ranhill Utilities Berhad and the United States Agency for International Developement.

Farley, B., S. R. Mounce, and J. B. Boxall. 2013. "Development and Field Validation of a Burst Localization Methodology." *Journal of Water Resources Planning and Management* 139 (6): 604–613.

Fortin, F.-A., D. Rainville, M.-A. G. Gardner, M. Parizeau, and C. Gagné. 2012. "DEAP: Evolutionary Algorithms Made Easy." *The Journal of Machine Learning Research* 13 (1): 2171–2175.

Gao, Y., M. Brennan, P. Joseph, J. Muggleton, and O. Hunaidi. 2005. "On the Selection of Acoustic/Vibration Sensors for Leak Detection in Plastic Water Pipes." *Journal of Sound and Vibration* 283 (3-5): 927–941.

Li, R., H. Huang, K. Xin, and T. Tao. 2015. "A Review of Methods for Burst/Leakage Detection and Location in Water Distribution Systems." *Water Science and Technology: Water Supply* 15 (3): 429–441.

Meseguer, J., J. M. Mirats-Tur, G. Cembrano, V. Puig, J. Quevedo, R. Pérez, G. Sanz, and D. Ibarra. 2014. "A Decision Support System for on-Line Leakage Localization." *Environmental Modelling & Software* 60: 331–345.

Mirats-Tur, J. M., P.-A. Jarrige, J. Meseguer, and G. Cembrano. 2014. "Leak Detection and Localization Using Models: Field Results." *Procedia Engineering* 70: 1157–1165.

Pérez, R., G. Sanz, V. Puig, J. Quevedo, M. A. Cuguero Escofet, F. Nejjari, J. Meseguer, G. Cembrano, J. M. Mirats Tur, and R. Sarrate. 2014a. "Leak Localization in Water Networks: A Model-Based Methodology Using Pressure Sensors Applied to a Real Network in Barcelona [Applications of Control]." *Control Systems, IEEE* 34 (4): 24–36.

Pérez, R., V. Puig, J. Pascual, A. Peralta, E. Landeros, and L. Jordanas. 2009. "Pressure Sensor Distribution for Leak Detection in Barcelona Water Distribution Network." *Water Science and Technology: Water Supply* 9 (6): 715–721.

Pérez, R., V. Puig, J. Pascual, J. Quevedo, E. Landeros, and A. Peralta. 2011a. "Methodology for Leakage Isolation Using Pressure Sensitivity Analysis in Water Distribution Networks." *Control Engineering Practice* 19 (10): 1157–1167.

Pérez, R., F. Nejjari, V. Puig, J. Quevedo, G. Sanz, M. Cugueró, and A. Peralta. 2011b. "Study of the Isolability of Leaks in a Network Depending on Calibration of Demands." 11th International Conference on Computing and Control for the Water Industry, Exeter. Vol. 455460.

Pérez, R., J. Quevedo, V. Puig, F. Nejjari, M. Cugueró, G. Sanz, and J. Mirats. 2011c. "Leakage Isolation in Water Distribution Networks: A Comparative Study of Two Methodologies on a Real Case Study." Control & Automation (MED), 2011 19th Mediterranean Conference on, IEEE, Corfu, 138–143.

Pérez, R., M.-A. Cugueró, J. Cugueró, and G. Sanz. 2014b. "Accuracy Assessment of Leak Localisation Method Depending on Available Measurements." *Procedia Engineering* 70: 1304–1313.

Puust, R., Z. Kapelan, D. Savić, and T. Koppel. 2010. "A Review of Methods for Leakage Management in Pipe Networks." *Urban Water Journal* 7 (1): 25–45.

Quevedo, J., M. Cugueró, R. Pérez, F. Nejjari, V. Puig, and J. Mirats. 2011. "Leakage Location in Water Distribution Networks Based on Correlation Measurement of Pressure Sensors." IWA Symposium on Systems Analysis and Integrated Assessment, San Sebastián, 290–297.

Romano, M., Z. Kapelan, and D. A. Savić. 2012. "Automated Detection of Pipe Bursts and Other Events in Water Distribution Systems." *Journal of Water Resources Planning and Management* 140 (4): 457–467.

Rossman, L. A. 2000. "EPANET 2 Users Manual, US Environmental Protection Agency." Water Supply and Water Resources Division, National Risk Management Research Laboratory, Cincinnati, OH, 45268.

van Rossum, G. 1995. "Python Tutorial." In *Centrum voor Wiskunde en Informatica* (Dutch Translation: Centre for Mathematics and Informatics). Amsterdam. Technical Report CS-R9526.

Sanz, G., and R. Pérez. 2014. "Comparison of Demand Pattern Calibration in Water Distribution Networks with Geographic and Non-Geographic Parameterization." 11th International Conference on Hydroinformatics Proceedings, Session R50, New York.

Sanz, G., R. Pérez, Z. Kapelan, and D. Savić. 2015. "Leak Detection and Localization through Demand Components Calibration." *Journal of Water Resources Planning and Management* 142 (2): 04015057.

Savic, D. A., Z. Kapelan, and P. Jonkergouw. 2008. "Quo Vadis Water Distribution Model Calibration?" *Urban Water Journal* 6 (1): 3–22.

Tao, T., H. Huang, F. Li, and K. Xin. 2013. "Burst Detection Using an Artificial Immune Network in Water-Distribution Systems." *Journal of Water Resources Planning and Management* 140 (10): 04014027.