Risk-based sensor placement methods for burst/leak detection in water distribution systems

E. Forconi, Z. Kapelan, M. Ferrante, H. Mahmoud and C. Capponi

ABSTRACT

The optimal placement of sensors for burst/leak detection in water distribution systems is usually formulated as an optimisation problem. In this study three different risk-based functions are used to drive optimal location of a given number of sensors in a water distribution network. A simple function based on likelihood of leak non-detection is compared with two other risk-based functions, where impact and exposure are combined with the leak detection likelihood. The impact is considered proportional to the demand water volume while the exposure is related to the importance of the connections and it is evaluated in social, economic or safety terms. The methods are applied to a district metered area of the Harrogate network by means of a modified EPANET model, to take into account the pressure-driven functioning conditions of the system. The results show that the exposure can lead to a different sensor location ranking with respect to other criteria used and hence the proposed methodology can represent a useful tool for water system managers to distribute the sensors in the network, complying with hydraulic, social and economical requirements.

Key words | leak detection, risk management, sampling, water distribution systems

E. Forconi
Z. Kapelan
H. Mahmoud
Centre for Water Systems,
University of Exeter,
Harrison Building, North Park Road,
Exeter EX4 4QF,
UK

E. Forconi
M. Ferrante (corresponding author)
C. Capponi
Dipartimento di Ingegneria Civile ed Ambientale,
University of Perugia,
Via G. Duranti 93,
06125 Perugia,
Italy
E-mail: *marco.ferrante@unipg.it*

INTRODUCTION

A number of publications in recent years have dealt with the problem of optimal sensor location with respect to various objectives. Many of them addressed the location of sensors in a district metered area (DMA) for the detection of contaminants (Kessler *et al.* 1998; Kumar *et al.* 1999; Ostfeld & Salomons 2004; Shastri & Diwekar 2006; Ostfeld *et al.* 2008; Berry *et al.* 2009). Other authors have investigated the optimal sensor location for solving the inverse, i.e. calibration problem (Kapelan *et al.* 2005).

Farley *et al.* (2010) developed the sensor network design methodology specifically for leak and burst detection in a pipe network. Their methodology is based on the sensitivity of the measured pressures to simulated burst/leak events. More specifically, the sensitivity of a certain location (node) in the pipe network to a leak/burst event is quantified by evaluating the change in pressure at that location from the baseline profile (when there are no bursts/leaks in the network). The likelihood of a burst/leak detection is then estimated by using the instantaneous chi-squared function which maps the aforementioned change in pressure into a 0/1 detection outcome (0 = no detection and 1 = detection). The potential drawback of the Farley *et al.* (2010) approach is that it treats all bursts/leaks in the network equally, i.e. without considering the potential impact they may have on customers. In real-life conditions, a water company may decide to investigate a potential pipe burst/leak event even if the estimated likelihood of this event is not that high, but if that pipe may have a major impact on nearby customers (e.g. cause local road or property damage) and especially if the customers in question are sensitive/critical (e.g. a hospital).

The objective of this paper is to overcome the above deficiency by developing and presenting a new methodology

for sensor location in a water distribution network that is based on the risk (i.e. both likelihood and potential impact) of leak non-detection. In the following, the problem of optimal sensor placement is presented and the sensitivity matrices are introduced and related to the risk of non-detection of a leak. The procedure of the optimal sensor location based on risk is then verified by a case study. Using the risk and relating it to the sensitivity of measured variables (e.g. nodal pressures) to potential burst/leak locations is the key novelty of this approach with respect to previous works (Kapelan *et al.* 2005; Farley *et al.* 2010).

OPTIMAL SENSOR PLACEMENT

The problem of the optimal placement of sensors for burst detection is formulated and solved here as a ranking problem, where potential sensor locations are ranked by minimising the risk of non-detection of bursts/leaks. The optimisation of the sensor network configuration should ideally be done using some optimisation method. However, past work in the field has shown that differences in optimal sensor locations obtained by using optimisation and ranking methods is minimal (see e.g. Kapelan *et al.* 2003). Given this, the relative computational inefficiency of the optimisation method (when compared with the ranking one) and the whole range of uncertainties involved in the selection of optimal sensor locations for burst/leak detection (e.g. uncertain location, timing, size and nature/type of the burst/leak, imperfectly calibrated hydraulic model, not ideally known demands in the network, etc.), we decided, as in many other existing sampling design approaches, to use a ranking type methodology here.

Following the conventional risk definition as a function of hazard, impact on elements at risk and vulnerability, in the outlined procedure these quantities are defined for a water distribution system and related to more conventional sensitivity matrices. To evaluate the effects of these components, three different risk of non-detection functions are used in this paper, considering only the hazard, then hazard and impact, and lastly the combination of hazard, impact and vulnerability. All these risk components were estimated using a calibrated hydraulic model of the analysed water distribution

system, which was assumed to be available. The calibrated model is assumed to contain up to date estimates of pipe friction factors, nodal demands, background (i.e. not burst type) leaks, statuses and characteristics of valves, pumps and other devices and any other model parameter/input values that may affect its predictions of network pressures and flows. This is the starting point in the sensor placement methodology shown below.

The sensitivity matrices

Assuming the use of pressure sensors only (with straightforward extension to consider flow sensors as well), the sensitivity to a burst/leak is here calculated as a difference in pressures between two different states of the system with and without the burst/leak, i.e. as a pressure drop at a given network node due to a burst/leak simulated at some pipe. Hence, as a first step, given a network of N nodes and L links, the pressure heads P_N, and the demands, Q_N, are evaluated at nodes assuming no bursts/leaks in the network. In a demand-driven model the demands do not depend on the pressure distribution in the system while in the used pressure-driven model the demands at nodes vary with pressure. Because extended period simulations are considered, data are recorded at T time steps. The results of the simulations are stored in two matrices, P_{NT} and Q_{NT}, with N rows and T columns. The corresponding vector of the L maximum flows in time at each link, Q_0, is also evaluated.

The second step is to consider a single leak in the system and to evaluate the same quantities as in the previous step. All bursts/leaks are assumed to be located in the middle of the respective pipes. A new set of leak nodes is added to the hydraulic model to achieve this, effectively dividing all existing pipes into two equal parts. The leakage, Q_L, is then modelled with discharge proportional to the pressure at leak location P_L

$$Q_L = C P_L^n \tag{1}$$

Because Equation (1) defines the dependence of leakage flow on pressure, it impacts the evaluation of the sensitivity matrices. Although different relationships can be chosen to evaluate the leakage for a single leak (Greyvenstein &

van Zyl 2007; van Zyl & Clayton 2007; Ferrante *et al.* 2011; Ferrante 2012; van Zyl & Cassa 2013) or for the leaks in a DMA (Ferrante *et al.* 2014b; Schwaller & van Zyl 2015), for the sake of simplicity the orifice equation is used here, with a pressure exponent $n = 0.5$, corresponding to a constant leak area. The value of leak coefficient C is varied from location to location, so that the relative leak size based on the steady state criterion (Ferrante *et al.* 2014a) is constant and equal to 20%, that is for each introduced leak $Q_L = 0.20\ Q_0$. Care has been taken to avoid incoming flow due to a negative pressure at the node.

All pipes in the network are considered as potential burst/leak locations. Bursts/leaks are simulated on all pipes in turn, resulting in a total of L scenarios generated. As a result, for each scenario two matrices, PL_{NTL} and QL_{NTL}, are obtained, containing pressure heads and demands at nodes, respectively, at T time steps.

The final step for the numerical estimate of the sensitivity is the evaluation of the L matrices $SP_{NTL} = PL_{NT} - P_{NTL}$ and $SQ_{NTL} = QL_{NT} - Q_{NTL}$, which express the sensitivity of nodal pressure heads and flows to each one of L simulated leak locations.

To reduce the problem dimension, respective sensitivities are grouped by summing up corresponding values at each network node across time. This corresponds to summing up the matrices SP_{NTL} and SQ_{NTL} along the columns, obtaining L vectors of N rows, SP_{NL} and SQ_{NL}. A further formal simplification was introduced by normalising the terms of the vectors with respect to the maximum and minimum values, resulting in values varying between 0 and 1. Defining SP_{max} and SP_{min} as the maximum and minimum of all the LxN values contained in the L SP_{NL} vectors, the normalised vectors are defined as follows:

$$SP'_{NL} = \frac{(SP_{NL} - SP_{min})}{(SP_{max} - SP_{min})} \tag{2}$$

The SQ'_{NL} vectors are similarly evaluated.

By definition, each one of the N elements of the L vectors SP'_{NL} can be associated with the likelihood of the detection of the leak at position L if a pressure transducer is located at one of the N nodes. The likelihood is maximum where the pressure difference is maximum while it is zero if its variation is lower than the accuracy of the transducer. As a consequence, the likelihood associated with the non-detection of the leak is given by the vectors defined as $L_{NL} = 1 - SP'_{NL}$.

The vectors SQ_{NL} have a different physical meaning. They represent the differences over time of the discharges, and hence differences of volumes, delivered at the nodes due to the leak associated with the L scenarios. Hence, the normalised vectors $I_{NL} = SQ'_{NL}$ measure the impact of the undetected leak on the delivered volumes.

The risk of burst/leak non-detection

In this paper three different functions are used, corresponding to three different criteria for the sensor location to detect leaks. All these functions are based on the definition of risk and require the use of the previously introduced sensitivity matrices.

In the most general sense, the risk, i.e. the possible future effects of a dangerous event, can be evaluated as the function of hazard, impact on the element at risk and vulnerability. The hazard is measured by the likelihood of occurrence of the dangerous event, such as the non-detection of a burst/leak, the impact on the element at risk is measured by the effects of a leak occurrence on the demands, while the vulnerability is measured by the intrinsic importance of the elements that can be damaged, i.e. the customers. In this paper the sensor locations are based on the risk of non-detection minimisation.

The first risk function, $R_1 = L_N$, is defined using the likelihood of burst/leak non-detection only, i.e. assuming that the risk is only associated with the hazard. Note that minimising this risk (i.e. likelihood of non-detection) is equivalent to maximising the likelihood of detection. It is worth noting that past approaches have not really considered the likelihood (i.e. using any value between 0 and 1); they have only considered cases where leaks are either detected or not (i.e. the 0/1 case) such as in Farley *et al.* (2010).

The second risk function, $R_2 = L_N\ I_N$, defines risk of non-detection as the product of the likelihood of non-detection (hazard related) and impact of non-detection. The former is estimated as in the case of the first risk function whilst the latter is estimated here as the normalised volume of water undelivered due to the burst/leak, i.e. using $I_{NL} = SQ'_{NL}$ values (see previous section).

The third risk function, $R_3 = L_N I_N V_N$, is similar to the second one, the main difference being that potentially more vulnerable water users (such as hospitals and schools) are given additional, higher weight (e.g. factor of 3) when estimating the impact relative to other water users (e.g. residential with a factor of 1).

Once the risk function is fixed, potential sensor locations (e.g. network nodes where pressure is monitored) are ranked using the Max-Sum method (Bush & Uber 1998). In this method, the obtained risk of non-detection values are added up across L simulated burst/leak events, resulting in a single risk value associated with each potential sensor location. These values are then ranked from smallest to highest risk of non-detection. Using the three risk functions mentioned above results in three ranked lists of sensor locations, RL_1, RL_2 and RL_3.

A summary of the sensor network design procedure is given in Figure 1.

CASE STUDY

The sensor network design methodology shown in the previous section is applied to the E023 DMA in the Harrogate area of the Yorkshire Water company (Savic *et al.* 2009). The considered DMA (Figure 2) has 16.25 km of pipes made mostly of cast (14.96 km) and ductile iron (0.56 km) and less than 0.75 km of polymeric pipes (MDPE). Diameters

range from 20 to 200 mm. The network implemented in EPANET has $N = 448$ nodes and $L = 468$ links. The nodes with an associated demand are $N_S = 291$ and they are all considered as candidate nodes for the sensor location. The network inflow is at one reservoir node, located in the western part ('Feeding node' in Figure 2). The major metered demand on the south-eastern part corresponds to the connection to another DMA ('Other DMA' in Figure 2).

The node elevations range from 100 m to 124 m and present an irregular distribution, with several hills and valleys. The part of the demands dependent on pressures is modelled using the EPANET hydraulic solver, modified for these purposes (Mahmoud *et al.* 2016). The part of the demands not dependent on pressure follows given water use patterns, such as the typical one shown in Figure 3. The mean demands are shown in Figure 4.

RESULTS AND DISCUSSION

The likelihoods of non-detection, i.e. the values of the first risk of non-detection function, R_1, are shown in Figure 5. The obtained values at nodes are interpolated in the horizontal plane to have a continuous representation of the obtained function.

The second risk function, R_2, values, obtained by the combination of non-detection likelihood and impact, are shown in Figure 6. The comparison with the first risk
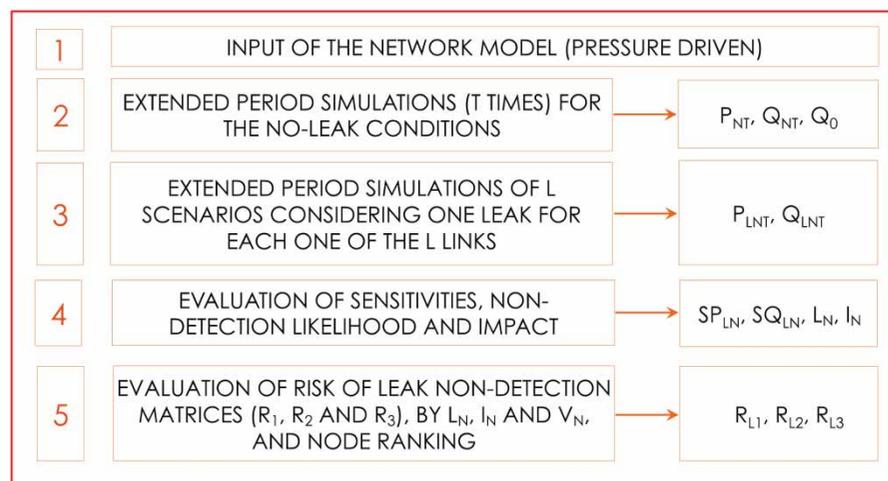

Figure 1 │ Summary of the procedure for the ranking of the nodes based on the non-detection risk.
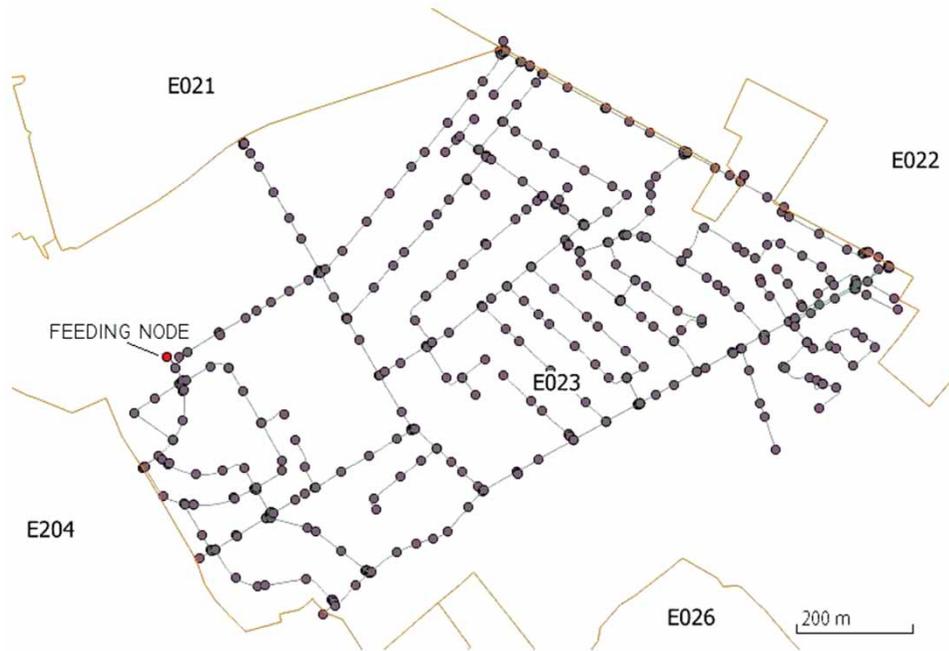
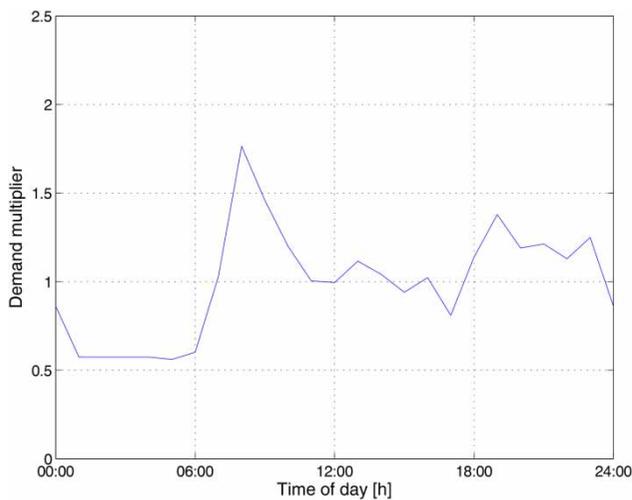**Figure 2** │ Layout of the Harrogate DMA network used for the case study.

Figure 3 │ The typical time pattern associated with the demand not depending on pressure.
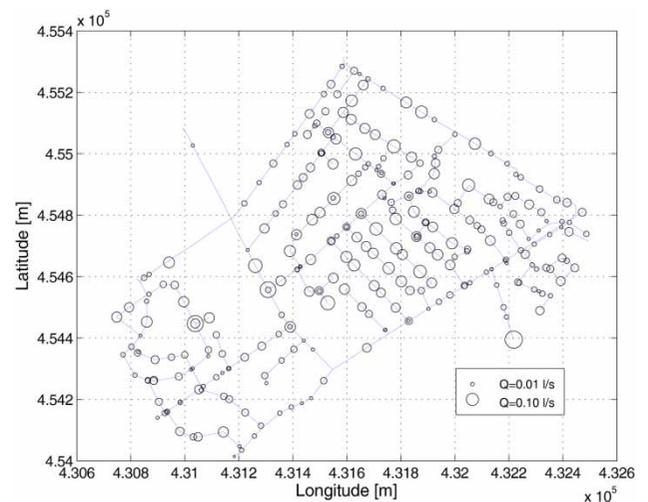

Figure 4 │ Spatial variation of the demand at nodes, if no leaks are occurring in the network. The marker dimensions are proportional to the demands.

function values (based on likelihood only) depicted in Figure 5 clearly shows that the two risk values are substantially different and that consideration of impact of non-detection completely changes the shape of the risk landscape, introducing new local minima and maxima.

The introduction of different weights for sensitive customers results in the third risk function. In the case study here

only two classes of demands are considered, with weights 1 and 2, as shown in Figure 7. The nodes with an exposure 1 correspond to residential customers for which a reduction of the demand is not critical, while nodes with exposure of 2 correspond to special customers where the risk of a demand reduction is considered critical. Other choices of the weight values can be used by decision-makers depending
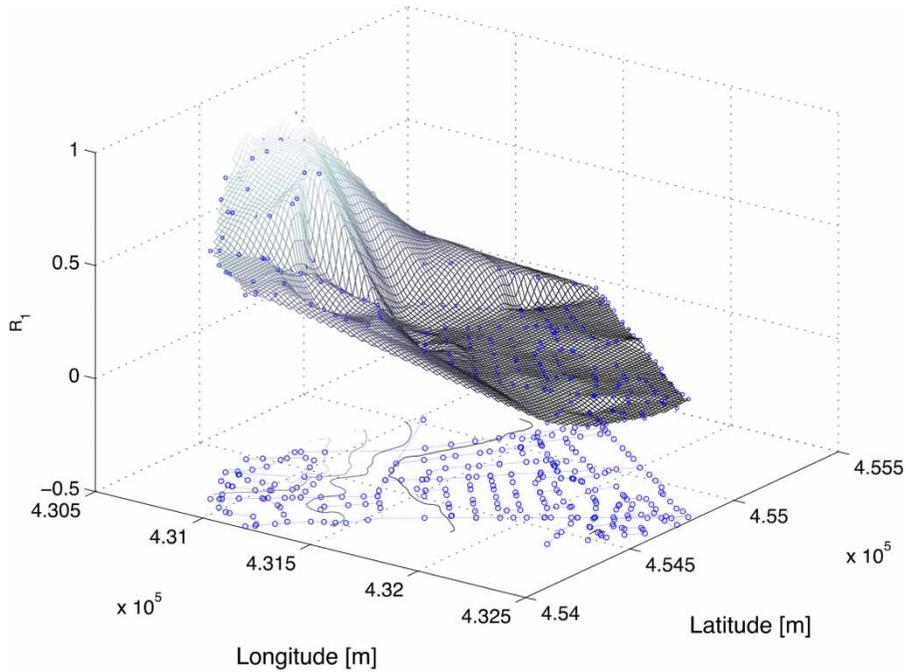
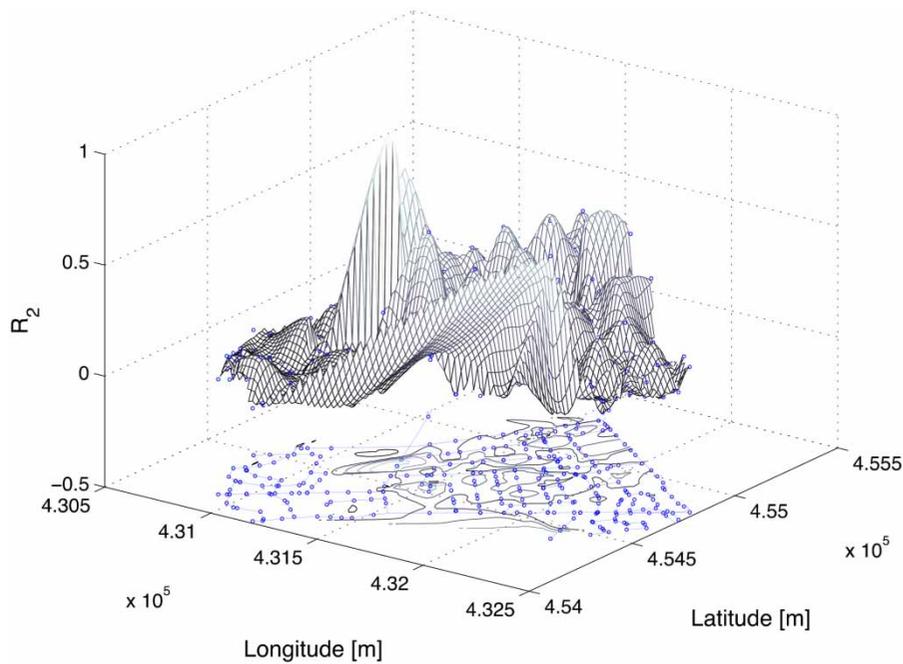**Figure 5** │ Spatial variation of R_1 in the network.

Figure 6 │ Spatial variation of R_2 in the network.

on the available information about the demands, with the highest values corresponding to key customers. The variation in space of the third risk function, R_3, is shown in Figure 8.

The ranking of the nodes by R_1, R_2 and R_3 is shown in Figure 9, where the best five nodes sensing locations are presented (note: the higher the ranking, the larger the marker size). While the rankings based on R_2 and R_3 are similar
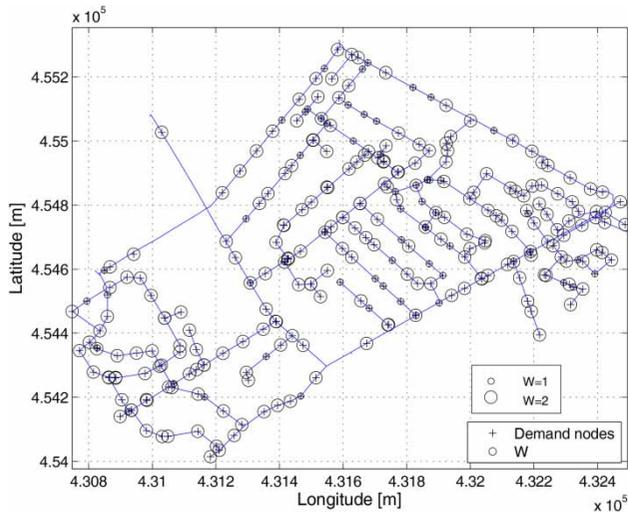
Figure 7 │ Spatial variation of two types of customers used in the third risk function evaluation.

in the sense that four out of the first five ranked nodes coincide, the best five ranked nodes by R_1 are different and are grouped in a completely different area of the network. It is worth noting that the ranking by R_1 is closely related to the usual pressure sensitivity analysis while the other two functions are related to the non-detection risk.
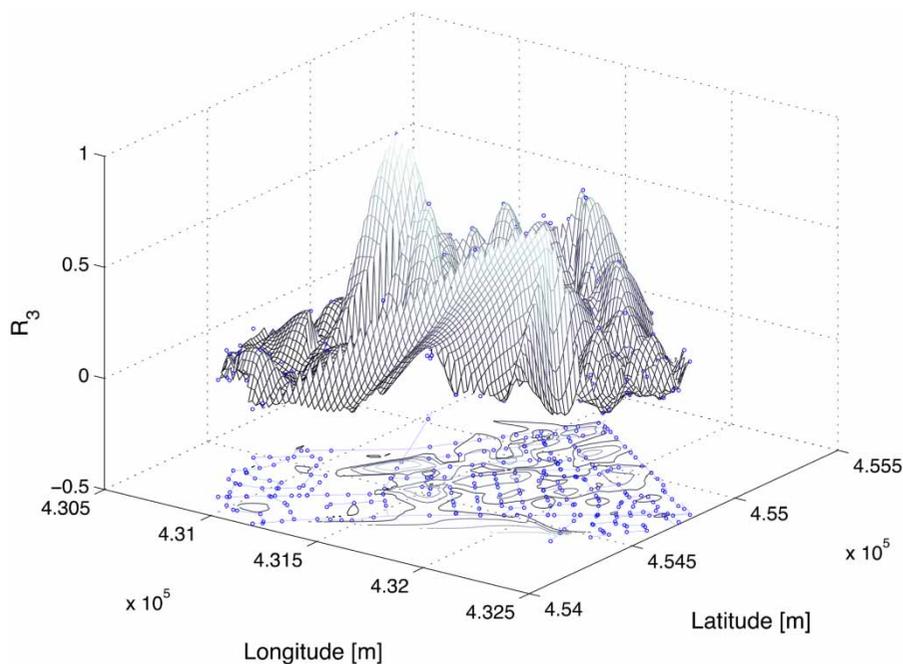
Because the ranking procedure does not depend on the number of sensors used and the deployment of a sensor does not modify the ranking, the N_{AS} available sensors can be placed by selecting the top N_{AS} ranked locations.

The differences in the ranking of the sensor locations are also shown in Figure 10, where the variations of the ranking functions with the node ranking are given. While the values of R_2 and R_3 reduce to 0.5 in less than 20 possible location nodes, R_1 remains above 0.5 for more than 50 nodes.

To investigate the reasons behind the ranking differences, in Figure 11 the values of the risk functions at the first 100 ranked nodes are plotted against the values of a normalised pressure, Pn, obtained as a mean in time of the pressure at nodes in the no-leak conditions and normalised so that 1 corresponds to the maximum value and 0 to the minimum. The correlation between R_1 and pressure values is clearly pointed out, because the highest values of R_1 correspond to the highest values of Pn. This result basically confirms that the sensitivity to pressures is correlated with the pressure values and hence that the best sensor location based on pressure sensitivity tends to prefer nodes where the value of the pressure is high. On the other
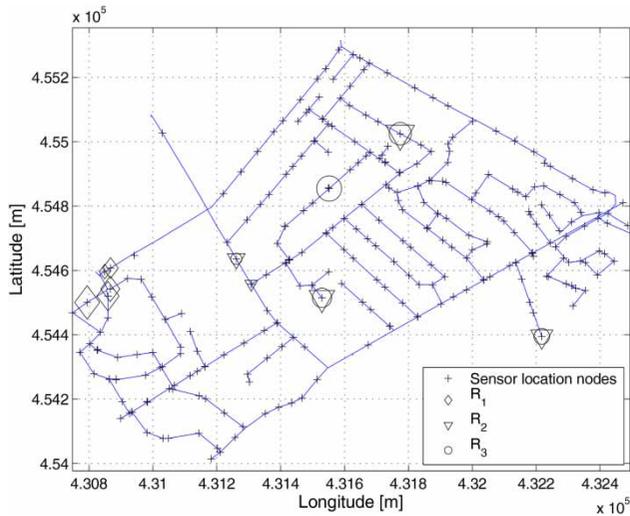

Figure 8 │ Spatial variation of R_3 in the network.

Figure 9 | Top five sensing locations for three proposed risk functions (marker size is proportional to the rank – the higher the rank the larger the size).
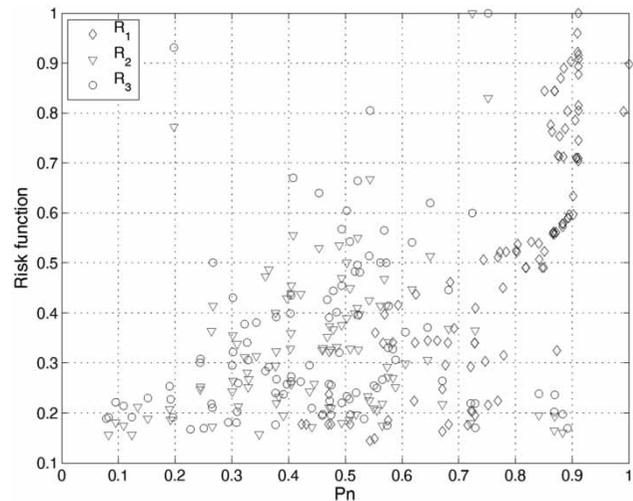

Figure 11 | Correlation of the three risk functions, R_1, R_2 and R_3, with the normalised pressure head.
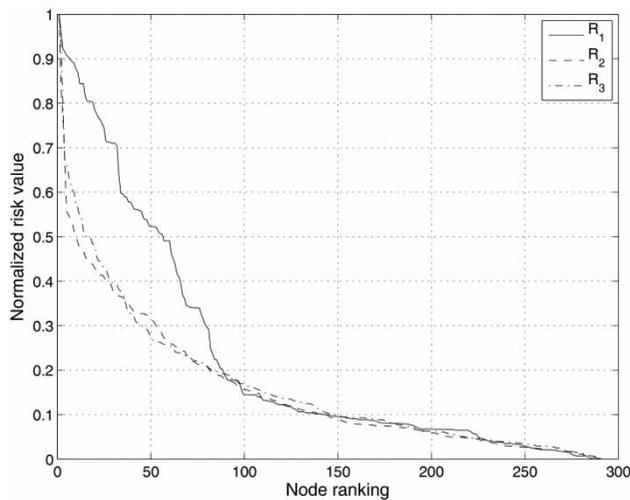

Figure 10 | Sorted values of the three risk functions, R_1, R_2 and R_3, at the network nodes.

hand, the correlation with the pressure values of the true risk functions, i.e. considering impact and vulnerability, is low and the clouds of the R_2 and R_3 data do not show a clear pattern.

CONCLUSIONS

This work developed new risk-based methods for optimal location of pressure sensors in a water distribution system with the aim of detecting pipe bursts/leaks. The sensor

network design problem is formulated as a ranking problem where the sensor locations are ranked by minimising the risk of non-detection of pipe bursts/leaks. The likelihood part of the risk is estimated via normalised sensitivity of nodal pressure to different pipe burst/leak locations whilst the impact part is estimated as normalised undelivered water demand due to a burst/leak. Two different versions of impacts were considered: (a) all water users treated equally and (b) vulnerable customers (such as schools and hospitals) given additional weight proportional to their vulnerability. The potential sensor locations are ranked by minimising the risk of non-detection. Three risk-based methods are proposed and compared with a more conventional approach where the ranking is driven by raw pressure sensitivities. The above sensor network design methods were applied to a real-life DMA in the UK.

The results obtained demonstrate the following:

1. The importance of performing formal sensor network design for the purposes of burst/leak detection, as some network locations are clearly more relevant than others and hence these need to be chosen carefully.
2. A risk-based approach that considers both likelihood and impact of non-detection has the advantage over the likelihood only based (i.e. more conventional) approach as it gives more importance to observing network locations where more water could be potentially lost (i.e. where

more people could be potentially affected) when a burst/leak occurs.

3. The risk method which gives additional weight to potentially sensitive customers (such as hospitals and schools) is preferred to the other risk method, which treats all customers equally.

Future work will focus on further refining the means of estimating the likelihood and impact components of risk of non-detection of pipe bursts/leaks and possibly expanding this approach to address water quality and other issues in distribution systems. Furthermore, the dependence of the sensitivities on the relationship used to link leakage and pressures requires further investigation. The influence of uncertain pipe friction factors, nodal demands and other not as well known hydraulic model parameters and inputs on optimal sensor locations is something that is also worthy of future exploration.

ACKNOWLEDGEMENTS

REFERENCES

Berry, J., Carr, R. D., Hart, W. E., Leung, V. J., Phillips, C. A. & Watson, J.-P. 2009 Designing contamination warning systems for municipal water networks using imperfect sensors. *J. Water Resour. Plann. Manage.* **135**, 253–263. doi:10.1061/(ASCE)0733-9496(2009)135:4(253).

Bush, C. A. & Uber, J. G. 1998 Sampling design methods for water distribution model calibration. *J. Water Resour. Plann. Manage.* **124**, 334–344. doi:10.1061/(ASCE)0733-9496(1998)124:6(334).

Farley, B., Mounce, S. R. & Boxall, J. B. 2010 Field testing of an optimal sensor placement methodology for event detection in an urban water distribution network. *Urban Water J.* **7**, 345–356. doi:10.1080/1573062X.2010.526230.

Ferrante, M. 2012 Experimental investigation of the effects of pipe material on the leak head-discharge relationship. *J. Hydraul. Eng.-ASCE* **138**, 736–743. doi:10.1061/(ASCE)HY.1943-7900.0000578.

Ferrante, M., Massari, C., Brunone, B. & Meniconi, S. 2011 Experimental evidence of hysteresis in the head-discharge relationship for a leak in a polyethylene pipe. *J. Hydraul.*

Eng.-ASCE **137**, 775–780. doi:10.1061/(ASCE)HY.1943-7900.0000360.

Ferrante, M., Brunone, B., Meniconi, S., Karney, B. W. & Massari, C. 2014a Leak size, detectability and test conditions in pressurized pipe systems. *Water Resour. Manage.* **28**, 4583–4598. doi:10.1007/s11269-014-0752-6.

Ferrante, M., Meniconi, S. & Brunone, B. 2014b Local and global leak laws. *Water Resour. Manage.* **28**, 3761–3782. doi:10.1007/s11269-014-0708-x.

Greyvenstein, B. & van Zyl, J. E. 2007 An experimental investigation into the pressure – leakage relationship of some failed water pipes. *J. Water Supply Res. T.* **56**, 117–124. doi:10.2166/aqua.2007.065.

Kapelan, Z., Savic, D. A. & Walters, G. A. 2003 Multi-objective sampling design for water distribution model calibration. *J. Water Resour. Plann. Manage.* **129** (6), 466–479.

Kapelan, Z. S., Savic, D. A. & Walters, G. A. 2005 Optimal sampling design methodologies for water distribution model calibration. *J. Hydraul. Eng.-ASCE* **131**, 190–200. doi:10.1061/(ASCE)0733-9429(2005)131:3(190).

Kessler, A., Ostfeld, A. & Sinai, G. 1998 Detecting accidental contaminations in municipal water networks. *J. Water Resour. Plann. Manage.* **124**, 192–198. doi:10.1061/(ASCE)0733-9496(1998)124:4(192).

Kumar, A., Kansal, M. L., Arora, G., Ostfeld, A. & Kessler, A. 1999 Detecting accidental contaminations in municipal water networks. *J. Water Resour. Plann. Manage.* **125**, 308–310. doi:10.1061/(ASCE)0733-9496(1999)125:5(308).

Mahmoud, H., Kapelan, Z. S. & Savic, D. A. 2016 A response methodology for reducing impacts of failure events in water distribution. *Presented at the 18th International Conference on Water Distribution System Analysis*, Cartagena, Columbia.

Ostfeld, A. & Salomons, E. 2004 Optimal layout of early warning detection stations for water distribution systems security. *J. Water Resour. Plann. Manage.* **130**, 377–385. doi:10.1061/(ASCE)0733-9496(2004)130:5(377).

Ostfeld, A., Uber, J. G., Salomons, E., Berry, J. W., Hart, W. E., Phillips, C. A., Watson, J.-P., Dorini, G., Jonkergouw, P., Kapelan, Z. S., di Pierro, F., Khu, S.-T., Savic, D. A., Eliades, D., Polycarpou, M., Ghimire, S. R., Barkdoll, B. D., Gueli, R., Huang, J. J., McBean, E. A., James, W., Krause, A., Leskovec, J., Isovitsch, S., Xu, J., Guestrin, C., VanBriesen, J., Small, M., Fischbeck, P., Preis, A., Propato, M., Piller, O., Trachtman, G. B., Wu, Z. Y. & Walski, T. 2008 The battle of the water sensor networks (BWSN): a design challenge for engineers and algorithms. *J. Water Resour. Plann. Manage.* **134**, 556–568. doi:10.1061/(ASCE)0733-9496(2008)134:6(556).

Savic, D. A., Boxall, J. B., Ulanicki, B., Kapelan, Z. S., Makropoulos, C., Fenner, R., Soga, K., Marshall, I. W., Maksimovic, C., Postlethwaite, I., Ashley, R. & Graham, N. 2009 Project Neptune: improved operation of water distribution networks. Presented at the Water Distribution Systems Analysis 2008. American Society of Civil Engineers, Reston, VA, pp. 1–16.

Schwaller, J. & van Zyl, J. E. 2015 Modeling the pressure-leakage response of water distribution systems based

on individual leak behavior. *J. Hydraul. Eng.-ASCE* **141**, 04014089–8. doi:10.1061/(ASCE)HY.1943-7900. 0000984.

Shastri, Y. & Diwekar, U. 2006 Sensor placement in water networks: a stochastic programming approach. *J. Water Resour. Plann. Manage.* **132**, 192–203. doi:10.1061/(ASCE) 0733-9496(2006)132:3(192).

van Zyl, J. E. & Cassa, A. M. 2013 Modeling elastically deforming leaks in water distribution pipes. *J. Hydraul. Eng.-ASCE* **140**, 182–189. doi:10.1061/(ASCE)hy.1943-7900.0000813.

van Zyl, J. E. & Clayton, C. 2007 The effect of pressure on leakage in water distribution systems. *Proc. Inst. Civil Eng.: Water Manage.* **160**, 109–114.