Development and Field Validation of a Burst Localization Methodology

B. Farley[1]; S. R. Mounce[2]; and J. B. Boxall[3]

Abstract: Reducing water loss through bursts is a major challenge throughout the developed and developing world. Currently burst lifetimes are often long because awareness and location of them is time- and labor-intensive. Advances that can reduce these periods will lead to improved leakage performance, customer service, and reduce resource wastage. In water-distribution systems the sensitivity of a pressure instrument to change, including burst events, is greatly influenced by its own location and that of the event within the network. A method is described here that utilizes hydraulic-model simulations to determine the sensitivity of potential pressure-instrument locations by sequentially applying leaks to all potential burst locations. The simulation results are used to populate a Jacobian matrix, quantifying the different sensitivities. This matrix may then be searched to identify different instrument locations to achieve required goals: maximising overall sensitivity to all potential events or selective sensitivity to events in different network areas. It is shown here that by searching this matrix to optimize such selective sensitivity, while minimising instrument numbers, it is possible to provide useful burst-localization information. Results are presented from field studies that demonstrate the practical application of the method, showing that current standard network models can provide sufficiently accurate quantification of differential sensitivities and that, once combined with event-detection techniques for data analysis, events can effectively be localized using a small number of instruments. **DOI: 10.1061/(ASCE)WR.1943-5452.0000290.**

CE Database subject headings: Water distribution systems; Water management; Leakage; Hydraulic models; Field tests; Data analysis.

Author keywords: Water distribution systems; Water management; Leakage; Hydraulic models; Field tests; Data analysis.

Introduction

According to EIRIS (2011), the world is facing climate, energy, and food crises. However, these cannot be fully discussed without understanding the impact of water scarcity. It is estimated that two thirds of the world's population will live in water-scarce areas by 2025 (EIRIS 2011). Global demand for water is forecast to outstrip supply by 40% by 2030 due to factors such as population growth and climate change (POST 2011). Therefore it is critical to ensure that water resources are managed carefully and in particular that losses from pipe networks are tackled. Losses can occur from many sources; one of these is leakage or bursts arising from breaks or fractures in water distribution systems (WDSs). Globally, the level of leakage varies tremendously; in the UK, it is estimated that leakage from WDSs accounts for 25–30% of the total water supply. Water is widely considered as abundant in the UK, however low rainfall in 2011 has lead to concerns about crops and the potential for drought (Environment Agency 2011). In recent years, information and communication technologies (ICT), water system-simulation and modeling-optimization technologies, and improved leakage control have all progressed to enable water engineers to effectively tackle and reduce water loss (Wu et al. 2011), however more is urgently needed.

This paper presents a methodology for locating low numbers of pressure instruments in a WDS to effectively detect and localize leak/burst events. This is achieved through optimization of the location of additional instrumentation, in combination with the existing instruments, to subdivide a system into smaller detection zones. The work utilizes current UK industry-standard hydraulic models and is demonstrated for WDSs of differing size and complexity. Results are presented from field tests using hydrant flushing to simulate leak/burst events in real distribution systems. This field validation made use of an automated data analysis detection system to identify events within time-series data (Mounce et al. 2010a).

[1]Univ. of Sheffield, Pennine Water Group, Dept. of Civil and Structural Engineering, Sheffield S1 3JD, UK.

[2]Research Fellow, Univ. of Sheffield, Pennine Water Group, Dept. of Civil and Structural Engineering, Pennine Water Group, Sheffield S1 3JD, UK (corresponding author). E-mail: S.R.Mounce@sheffield.ac.uk

[3]Professor of Water Infrastructure Engineering, Univ. of Sheffield, Pennine Water Group, Dept. of Civil and Structural Engineering, Sheffield S1 3JD, UK.

Note. This manuscript was submitted on December 9, 2011; approved on May 16, 2012; published online on May 21, 2012. Discussion period open until April 1, 2014; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Water Resources Planning and Management***, Vol. 139, No. 6, November 1, 2013. © ASCE, ISSN 0733-9496/2013/6-604-613/$25.00.

Background

The distribution of potable water to consumers in the developed world is accomplished via a complex network of pipes. The complexity of WDSs varies tremendously from area to area. In the UK, and increasingly in other parts of the world (Brothers 2003), WDS are subdivided into district meter areas or distribution management areas (DMAs). To measure and assess the performance of these, WDS instrumentation is installed measuring flow and pressure at certain locations. These flow and pressure instruments are generally located at predetermined positions within each WDS. In the UK, flow and pressure instruments are typically installed at the inlet (and any outlet) to each DMA and an additional pressure instrument (referred to as the DG2) is installed at the point of highest elevation or another critical point in the DMA. The highest elevation is selected to comply with regulations regarding minimum pressure levels in the WDS.

Understanding complex WDSs has historically been hard because of the difficulty of collecting accurate data. Manual data collection meant that data was analyzed as infrequently as every two months. However, in recent years, developments in measuring and recording data have made data collection easier and the use of telemetry, like general packet radio service (GPRS), has allowed for data collected from instrumentation to be accessed quickly, with data now available in near-real time, often at fifteen minute sampling intervals. These advances, together with the availability and ease of installation of pressure instrumentation, at any hydrant or other tapping point, provide significant potential for improving the understanding and management of WDSs. It is proposed that, by developing an understanding of the sensitivity of pressure responses across WDSs, potentially important information in regard to system performance could be gained from increasing the number of pressure instrumentation devices permanently installed. However, there are significant capital and operational expenditure costs, maintenance requirements, and information technology (IT) issues associated with any instrumentation deployment, hence their number and locations need to be optimized with demonstrable benefits accruing.

To identify the optimal number and locations of instruments for any particular application, such as detection of water-quality events (Berry et al. 2006), leakage-hotspot identification (Wu and Sage 2008), or hydraulic-model calibration (Bush and Uber 1998), it is necessary to solve a complex optimization problem. The problem is complex because there are many possibilities in terms of potential instrument locations and because the optimal number of instruments will differ from WDS to WDS. Optimal instrument placement using hydraulic simulation software has been studied for different purposes in WDSs. Through the application of hydraulic mathematical models, simulations can be run in batches and multiple different events can be investigated. However, multiple simulations create a large amount of data that requires analysis. Once this data is analyzed, the problem of where to situate instruments for different purposes can be solved. Work along these lines by Bush and Uber (1998) and Kapelan et al. (2003) demonstrated methods by which the optimal location of instrument(s) for calibration purposes can be found. Another field in which optimal instrument/sensor placement has been widely studied is for the early detection of contamination events in WDSs (Berry et al. 2006; Janke et al. 2006; Watson et al. 2010). Despite trying to solve different problems, the principles behind these approaches are similar, namely utilising hydraulic models and running multiple simulations of different circumstances. Once simulations have been run, it is important to search the resultant data in an efficient and well-thought-out fashion, to find the optimal location(s). A commonly used search approach in the field is the genetic algorithm (GA).

GAs are a search procedure based on the mechanics of natural selection and natural genetics (Goldberg 1989). They are highly parallel, mathematical algorithms that transform a set (population) of mathematical objects (typically strings of ones and zeros referred to as genes) into a new population. They work by combining survival of the fittest for individual genes; these are then passed on to the next generation. As the successful (fittest) genes *breed* over generations they quickly converge to optimal solutions after examining only a small fraction of the search space. Mutations and crossover are also included in generations to ensure that a string of genes that may help provide an optimal solution are not lost too early. GAs and other evolutionary algorithms have been successfully applied to many complex engineering optimization problems and extensively for water resources engineering and management (Nicklow et al. 2010). They have been widely applied to water distributions for calibration (Kapelan et al. 2003; Kapelan and Savic 2009), existing leakage detection (Wu et al. 2010), and contamination-event detection (Ostfeld and Salomons 2004). These applications have often led to resultant algorithmic advances.

Research has been conducted in the application of multiple hydraulic simulation and GA search approaches to burst-event detection. A methodology for optimal placement of pressure instruments for improved detection was first proposed in Farley et al. (2008), and fully presented with field validation in a real water distribution system in Farley et al. (2010a). Perez et al. (2009) presented a similar method for identifying burst events, however the method was reliant on heavily instrumenting networks with more than fifteen sensors and has not been tested on in a real WDS with simulated or real events. Romano et al. (2011) has used pressure instruments to localize leak/burst events using an ordinary cokriging technique (an interpolation technique utilizing a cross-correlated secondary variable to reduce the variance of the estimation error) and other geostatistical approaches to successfully locate a series of flushing events. Thirteen pressure instruments were deployed in a single DMA, and again this is far from normal practice for real WDSs. Installation of the number of sensors in each WDS required by approaches such as Perez et al. (2009) and Romano et al. (2011) is not currently practical from a cost, management, and IT perspective for water companies.

Method

The method presented here builds on and develops work by Farley et al. (2008, 2010a, b), to provide a technique which is able to both detect and localize burst events within WDSs.

WDSs usually have both flow and pressure instruments installed in them, but the instrument behavior and approach to data collection are different. Flow data are averaged and then aggregated, leading to the data being smoothed (Mounce et al. 2012). Pressure values are instantaneous values; as a result, some of the subtle variations in pressure may be missed. Flow measurements are usually taken at the inlet, and are sensitive to all downstream changes. Pressure measurements, however, are sensitive to changes in headloss along prescribed upstream flow routes only. They are therefore most sensitive to changes along certain routes and generally most sensitive to change local to the instrument's position. Pressure instrumentation has been used for the present study (as opposed to flow), because pressure instruments are significantly cheaper and can readily be installed in any WDS using any readily available tapping point such as fire hydrants or wash outs, without the need for new fittings, excavations, or decommissioning of pipes.

Flow data has been effectively used for event detection, being more sensitive to leak/burst events than pressure data (Mounce et al. 2011). It is hypothesized that this detection via flow data can be augmented by pressure data, confirming detection and, if positioned intelligently, allowing location information to be inferred. This hypothesis is based on the differential and local sensitivity of the pressure instrumentation. Fig. 1 shows conceptually how, for an extremely simple ideal network, differential sensitivity of instrumentation could be used to provide both detection and location information.

To utilize this approach, it is necessary to determine likely instrument behavior at different locations with sufficient accuracy to identify differential sensitivity in real, complex networks. The work by Farley et al. (2008, 2010a, b) utilized a methodology that produces such a sensitivity matrix. This was achieved by sequentially

J. Water Resour. Plann. Manage., 2013, 139(6): 604-613

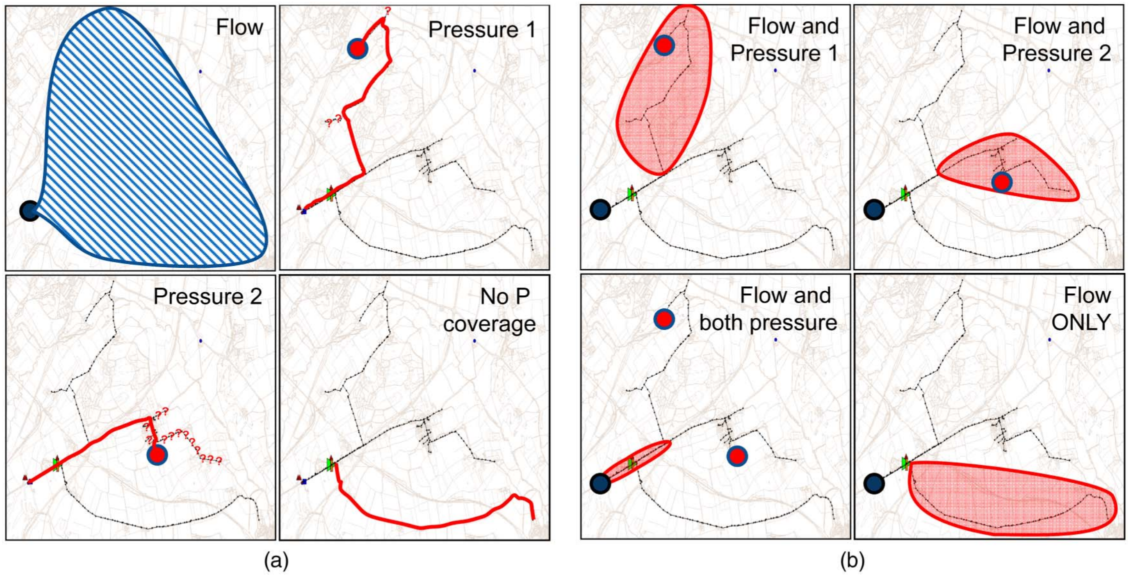

Fig. 1. (a) Simplified differential response of instrument locations; (b) differential detection resulting from selected instrument locations

modeling leak/burst events at all nodes in a model and simulating the pressure response at all possible instrumentation points. The main benefit of building on this work is that it has been subject to extensive validation, including fieldwork using flushing to simulate burst events.

The major challenge is then to search this matrix to maximize overall sensitivity, to minimize the amount of instruments to be added to a given network (ideally complementing existing instrumentation), and to provide a maximum number and even size of detection zones. The search methods previously used by Farley et al. (2008, 2010b) focused on detection rather than localization. Hence the methodological development presented here is for an approach to search the matrix to provide localization information, requiring the integration of a GA search approach to improve efficiency. Whilst the approach has been developed for application to WDSs with a DMA configuration, there is no reason why it could not be applied to different systems (e.g., whole networks or trunk mains). The hydraulic models used by the method are typical UK industry-standard models supplied by a water company and no additional calibration having being conducted.

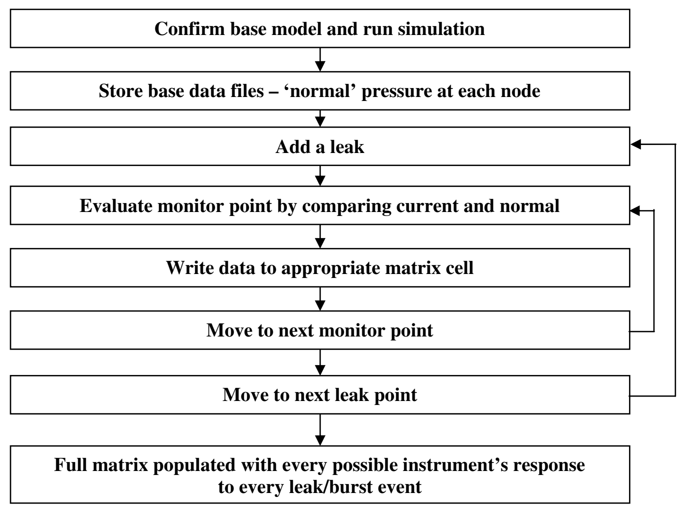

Fig. 2. Flowchart of methodology (adapted from Farley et al. 2008, © ASCE)

Assembling/Producing the Jacobian Sensitivity Matrix

The steps in the process of generating the Jacobian sensitivity matrix via hydraulic model simulation are illustrated in a flow chart in Fig. 2.

New leak/burst events were simulated at every node (representing every possible leak/burst event location), and the change in pressure analyzed using

$$\chi^2 = \sum \frac{(P_{lc} - P_n)^2}{P_n} \tag{1}$$

where χ^2 = chi squared value; P_{lc} = pressure under leak conditions; and P_n = pressure under normal conditions. The pressure under normal conditions is the system modeled with no new leaks present. The χ^2 method provides a good test of sensitivity, as it compares the change in pressure from the system under normal conditions to when a leak/burst event has occurred. As it is normalized by the pressure under normal conditions (P_n), this ensures that events that occur at high pressure are not determined as overly sensitive. The χ^2 values are calculated for every instrument response, summated for a 24-h period, and used to populate the Jacobian sensitivity matrix. In Farley et al. (2010b) it was shown that the dependence of the method on the diameter of the leak was minimal, and that results from a standard-size leak were transferable. However, it should be noted that this requires a pressure-dependent leak function and not simply the addition of a standard demand.

Searching the Jacobian Matrix

To extract relevant/useful data from the Jacobian matrix it is important to develop a search technique that is able to efficiently search large amounts of data. The first stage of this is to define whether leak/burst events would be detected by possible pressure-instrument locations. This is achieved by applying a detection threshold to the data in the sensitivity matrix. The detection threshold is then used to create a binary matrix populated with 1 (indicating detection) and 0 (indicating no detection).

J. Water Resour. Plann. Manage., 2013, 139(6): 604-613

The threshold used to determine detection/nondetection is derived from the hydraulic simulation results; there will be a degree of uncertainty in this value, as water distribution models are only representations of the real system (how close to reality is dependent on model build quality and calibration accuracy) and therefore subject to uncertainty. To limit the impact of model uncertainty and to try to ensure model error does not play a role in detection/nondetection, a model-specific threshold is used. The threshold is selected as the average sensitivity (taken from the sensitivity matrix) resulting from the model in question. Assuming that the model has a degree of uncertainty, a sensitivity value that is close to the threshold response is more uncertain than an instrument which has no (or very minimal) or strong response. Therefore, it is preferable to select instruments with very small or very large sensitivities to changes in pressure, rather than a response close to the threshold. Consequently an uncertainty band was applied to either side of the threshold; the aim of this band was to penalize instruments that have a response close to the mean. The size of the band was explored to establish sensitivity; percentage uncertainties were explored as different percentage values (± 5, 10, 20, 30, and 40%). Following extensive testing with a variety of DMA models, 10% was selected as the penalty function sufficiently accounting for model uncertainty, but not producing excessive penalty zones. Thus the matrix that is searched is no longer binary and is populated with 1 (indicating detection), -1 (indicating no detection), and 0 indicating a response in the uncertain or penalty zone.

Once detection and nondetection (or uncertainty) has been estimated, a method is required by which the location of an event can be inferred. Table 1 shows how theoretically four distinct zones can be identified with only two pressure instruments. This is based upon using two pressure instruments installed within the DMA and utilizes the flow meter installed at the inlet (Zone D detected only at the inlet flow instrument), which is consistent with Fig. 1.

An important aim of the search is to find a combination of locations which subdivide the DMA into evenly sized zones. If all the zones are equal, then the search areas are of equal size. Size may be quantified by number of nodes, not necessarily capturing a geographical area but rather some hybrid of geographical area and network complexity. This effectively provides a useful measure for the time it would take to pinpoint a leak within a given zone, and it is actually this that should be equalized. The target zone size for each DMA is calculated

$$Tz = \text{Nodes}_{\text{Total}}/Z \quad (2)$$

where Tz = target zone size; $\text{Nodes}_{\text{Total}}$ = total number of nodes in the DMA; and Z = number of possible zones (this is dependent on the number of instruments, n). Theoretically the number of instruments per DMA is limited only by the number of potential instrument locations. Therefore, by increasing the number of instruments per DMA, the number of zones increases, as $Z = 2^n$. For a DMA of 100 nodes with two instruments, $Tz = 100/2^2 = 25$. The target zone size is used in the *scoring* of possible instrument combinations. How close each zone is to the target will define how well the combination divides the DMA. Therefore for the example DMA (of 100 nodes), four zones of 25 nodes would represent a perfect division.

Table 1. Combinations of Responses for Two Instruments

Instrument	Zone A	Zone B	Zone C	Zone D
1	Detection	Detection	Nondetection	Nondetection
2	Nondetection	Detection	Detection	Nondetection

To search the Jacobian sensitivity matrix an objective or fitness function was developed, to find the optimal combination of instruments which subdivides the DMA into the most evenly sized zones. The consequence of applying an uncertainty band is that it acts to create an additional zone, a fifth zone in Table 1. This zone will be populated by responses that occur in the penalty zone for one or more of the pressure instruments. It is particularly desirable to keep this zone as small as possible; therefore a multiplier was applied to ensure that it is less favorable for a solution to have a large penalty zone. The fitness function equation used is shown in Eq. (3). The decision variable for the GA was the locations of instruments

$$FF = \sum_{i=1}^{Z} \sqrt{(N_i - T_Z)^2} + P \times 1.25 \quad (3)$$

where P = number of event locations that are in the uncertainty band; and N_i = total number of nodes (events) detected in zone i. Both N_i and P are evaluated by interrogating the binary (with uncertainty zone) Jacobian matrix for each set of instrument locations selected by the GA.

To solve Eq. (3), a genetic algorithm search approach was used. The software developed for this application utilized the *MATLAB* Genetic Algorithm toolbox (Chipperfield et al. 1994). A function (the objective or fitness function) was written as a. *m* file which is then optimized (minimized) by the GA. Note that the number of instruments is not explicitly optimized as part of the fitness function.

By increasing the number of instruments within a DMA the number of zones that the DMA is divisible into also theoretically increases as $Z = 2^n$. However, in practice it is extremely difficult to identify multiple (say 6) instrument locations that would allow for subdivision of a DMA into the theoretically possible number of (in this case 64) zones. There is a diminishing return on the number, size, and usefulness of the zones established by adding instruments; the nature of this trade off is unique to every network. Experience has shown that larger DMAs (typically ≥800 nodes) can usually accommodate the installation of three additional instruments but provide five to seven useful zones rather than the theoretically possible eight. As the number of instruments and subdivisions increases, the more reliant the method becomes on high model-accuracy which can often be suspect in reality. Other constraints on heavily instrumenting DMAs are the cost, maintenance, and IT overheads. Assigning multiple, say greater than three additional, pressure instruments to every DMA in the system would be expensive and may not reduce the search time taken by leakage technicians to pinpoint leak/burst events, particularly when considering the inherent travel time and related operations; therefore this is unlikely to be practical.

Application and Validation

The application and validation section is divided into two sections:

- Firstly the approach was applied to 14 DMA models to test application and explore how evenly these networks were subdivided; and
- Additional pressure instruments were then deployed in a real-life WDS in the UK, based on application of the method, and bursts simulated by opening fire hydrants to validate the DMA sub-division approach.

J. Water Resour. Plann. Manage., 2013, 139(6): 604-613

Ideal Application

A set of 14 DMAs was analyzed by the search technique (ignoring the current instrumentation) to show the application for a range of DMAs with the method evaluated for subdivision of the DMA into smaller burst event detection zones. The industrial partner advised restricting the addition of instruments to two as a practical limit since (potentially) four zones offer a pragmatic, cost effective solution with present technologies and practices. The method is extendable so that, for example, three instruments could render (potentially) eight subdivisions.

The characteristics of the DMA play an important role in the subdivision of the DMA. Generally, it is easier for larger DMAs to be divided into four zones, as a result of the size. When the method was applied to some of the smaller DMAs it was not possible to achieve four zones. However, if the DMA is smaller, then the search area is smaller still, and as a result it will not influence the search time (to find the leak/burst event) significantly. The number of zones for each of the 14 DMAs in the pilot is presented in Table 2, together with summary information to provide an impression of the range of DMA sizes and characteristics.

Table 2 shows that for most of the DMAs used in this pilot it is not possible to subdivide them effectively into four suitably evenly sized zones. The characteristics of the DMA influence the number of zones it is possible to subdivide. Generally a DMA which is smaller in size divides in to a smaller numbers of zones. Table 2 also expresses the fitness function as a fraction of the number of nodes; this has been included to offer a comparison between two DMAs to assess the quality of the subdivision, with a low value indicating a good subdivision.

Subdivision of smaller DMAs into two zones can be as beneficial as dividing a larger DMA into 4 zones. For example dividing a DMA which is 50% of the size of another into half the number of zones evenly produces the same size zones. Therefore the time taken to locate a leak/burst event will be similar. Consequently being unable to subdivide the DMA into four zones for smaller DMAs is not crucial. A potential benefit for the water utility company is that smaller DMAs will need fewer instruments, which will reduce instrumentation costs.

Practical Constrained Application

The water company that participated in this test wished to keep their existing instrumentation at the current locations within the DMA, as they provide important information about system performance for which continuous records are required. Using the method developed, it is possible to include this instrumentation already situated in the DMA, as well as all other possible instrumentation points. The matrix search was hence adapted to run searches with and without utilizing the existing instrumentation locations as fixed points.

In addition to the strategy described for the ideal application, three additional search strategies that include the current instrumentation already in the DMA to varying degrees were applied to the 14 DMAs as follows:

1. The optimal combination of two instruments determined by applying the search technique defined above, ignoring the current instrumentation (results presented in Table 2);
2. The current instrumentation only (i.e., the pressure instruments at inlet and at the point of highest elevation) with no additional instrumentation;
3. One of the current industry instruments and one optimally placed instrument; and
4. One of the current industry instruments (i.e., the pressure instrument at the inlet or the point of highest elevation) and two optimally placed instruments.

Strategy 1 effectively provides the *benchmark* against which to judge the *best* solution from Strategy 2, 3, or 4. A selection of four DMAs are presented below to illustrate the results obtained from the application of these strategies; these are the four DMAs that were then used in the live field trials.

By supplementing the DG2 (critical instrument) with an additional instrument, DMA A is divided into two distinct zones, with a small penalty zone (as shown in Fig. 3) following Strategy 3. The fitness function and fraction obtained are identical to those achieved by Strategy 1, showing that the current DG2 location is actually only sensitive to a defined area of the DMA rather than the majority of the DMA as might be the aspiration for a DG2 as a crucial instrument. For DMA B, one additional instrument (using Strategy 3) provides a good division of the DMA into two large zones and one very small one. However, in this case the DG2 did not provide any value to the subdivision of the DMA and the fitness function and fraction were worse than those reported in Table 2 for Strategy 1. This drop in fitness function versus the small zone obtained by the application of Strategy 4 was deemed of insufficient benefit when discussed with the water utility company. For DMA C, the optimal combination was selected as being obtained by using Strategy 4. This is one of the few situations were the current DG2 instrument location provides some benefit for leak subdivision, but the drop in fitness function and fraction when only adding one instrument is significant. For this DMA using three instruments (one DG2 and two optimally placed) the fitness function and fraction are actually improved enough over Strategy 1 to provide sufficient benefit for installation (as evaluated qualitatively by the water company personnel). For DMA D, the addition of a single optimally located instrument (using Strategy 3) enables division into two distinct zones, obtaining the same fitness function and fraction as the benchmark Strategy 1. This is despite the DG2 instrument location not contributing a zone of detection. This shows that the Strategy 1 solution can actually be achieved with only one instrument.

The division of the DMAs shown in Fig. 3 offer some interesting insight into the strategies. Strategy 2 is generally very poor for both overall sensitivity, as found in previous work (Farley et al. 2008), and shows little effective subdivision, as expected. The division of DMAs A and C offers a lower fitness function

Table 2. Number of Subdivision Areas Achieved for DMAs Used in the Pilot, Based on Using the Optimal Combination

DMA	A	B	C	D	E	F	G	H	I	J	K	L	M	N
Number of zones	3	3	3	2	3	3	4	3	3	3	3	2	3	4
Fitness function	404	353	534	368	771	351	321	722	568	187	386	367	178	147
Fitness function as a fraction of number of nodes	0.81	0.79	0.77	1.00	1.11	0.94	0.32	0.95	0.52	0.72	1.14	1.12	0.87	0.55
Total number of nodes	493	448	696	368	693	373	995	760	1091	260	338	328	204	265
Total length of pipe (km)	16.5	23.5	24	30.2	30	17.8	36	27	20	9	14	12.7	6.3	11.5
Min/max pressure (m)	20/71	17/163	27/64	22/80	25/76	6/71	30/76	4/87	15/71	18/98	20/133	21/71	16/76	15/100

J. Water Resour. Plann. Manage., 2013, 139(6): 604-613

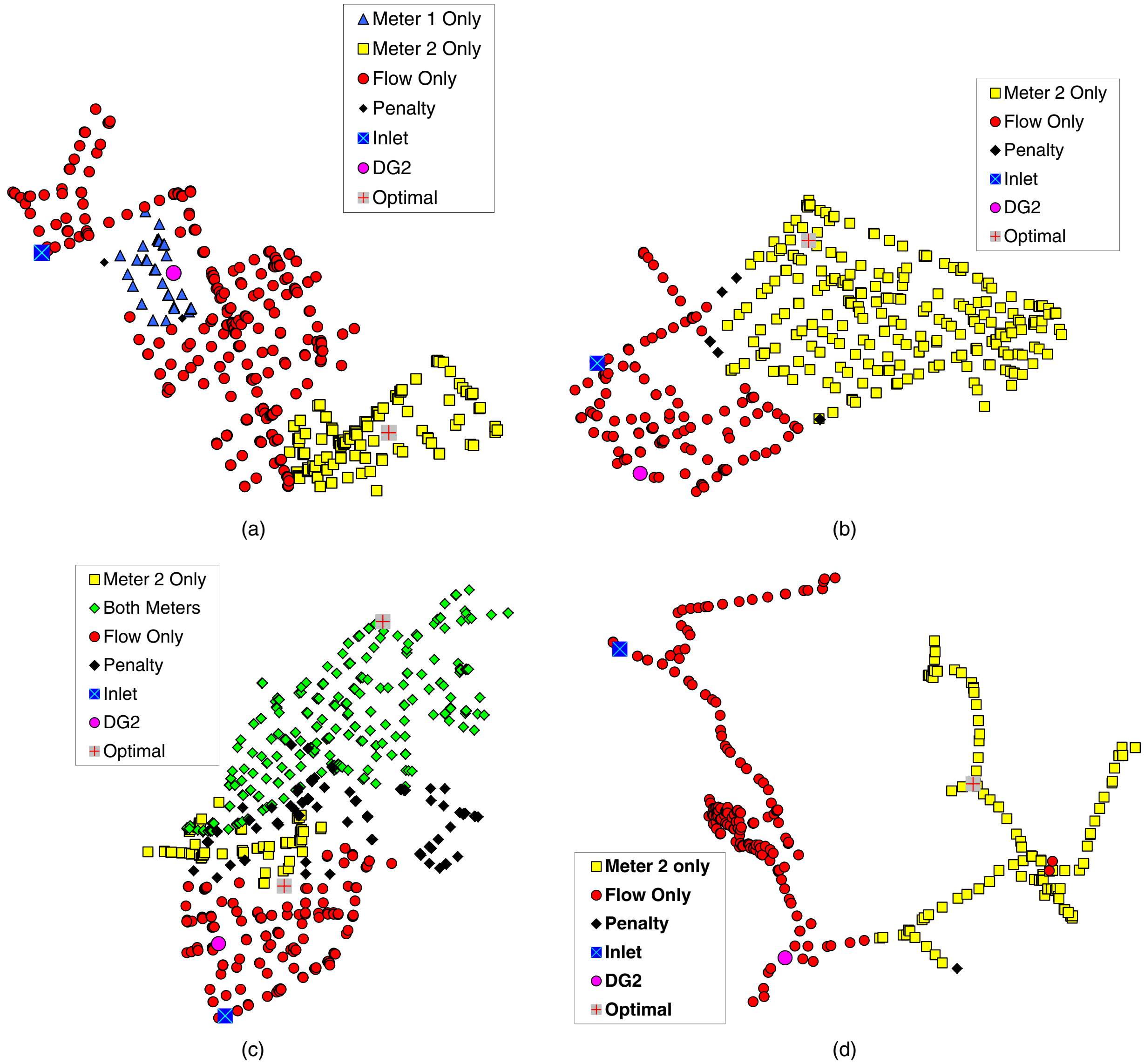

Fig. 3. Division of sample DMAs: (a) A; (b) B; (c) C; (d) D

as a fraction of the total nodes score than the division of DMAs B and D. This is because in both the former DMAs the DG2 instrument is used, however it does not contribute (in terms of creating a detection zone). As a result, the fitness function relative to the number of nodes is higher (1.0 in both cases B and D, compared to 0.81 for A and 0.62 for C). DMA A benefited from having its DG2 instrument situated at a sensitive point and was therefore able to offer some subdivision of the DMA. However, the DG2 instrument is not always situated in a sensitive location and some benefit (in terms of leak/burst-event detection) may be achieved by moving it (Farley et al. 2010b). In general the location of the DG2 point has been shown to be ineffective for leak/burst-event location and detection, however it can provide other useful WDS data. Strategy 3 has been used for both DMAs B and D, however the DG2 point contributes very little in terms of detection and location in these particular DMAs. Strategy 3 is a viable strategy, however it generally depends on the DG2 being situated in a sensitive area.

Field validation

Once the instrumentation had been installed in the DMAs, it was important to test the methodology to see if the actual location of the real leak/burst events was obtained. Previous work by Farley et al. (2008, 2010b) and Mounce et al. (2010b) have used hydrant flushing to simulate the effect of a leak/burst events within a DMA to provide certainty of test conditions and a reasonable timeframe of event. A mixture of pre-determined and blind hydrant flushings was used to evaluate the methodology in this paper. In all, eight events were conducted; these events were created by a water company technician opening a hydrant at a location within one of the four DMAs (see Fig. 4). Once the hydrant was open, a volume of water was allowed to flow from the hydrant (the flow from the hydrant flushing was unknown). The first set of flushing locations were determined by the research team and therefore placed in known locations. The second set of flushing tests were conducted solely by the water company, therefore the location of the flush was not disclosed until after the detection/nondetection and location had been evaluated. The field trial events ran for a period of at least 12 h, with some running for up to 24 h. The following events were conducted in the four DMAs:

- DMA A—one blind flushing was conducted;
- DMA B—one blind flushing was conducted;
- DMA C—two blind flushing tests and three events specified by the research team were conducted; and

J. Water Resour. Plann. Manage., 2013, 139(6): 604-613

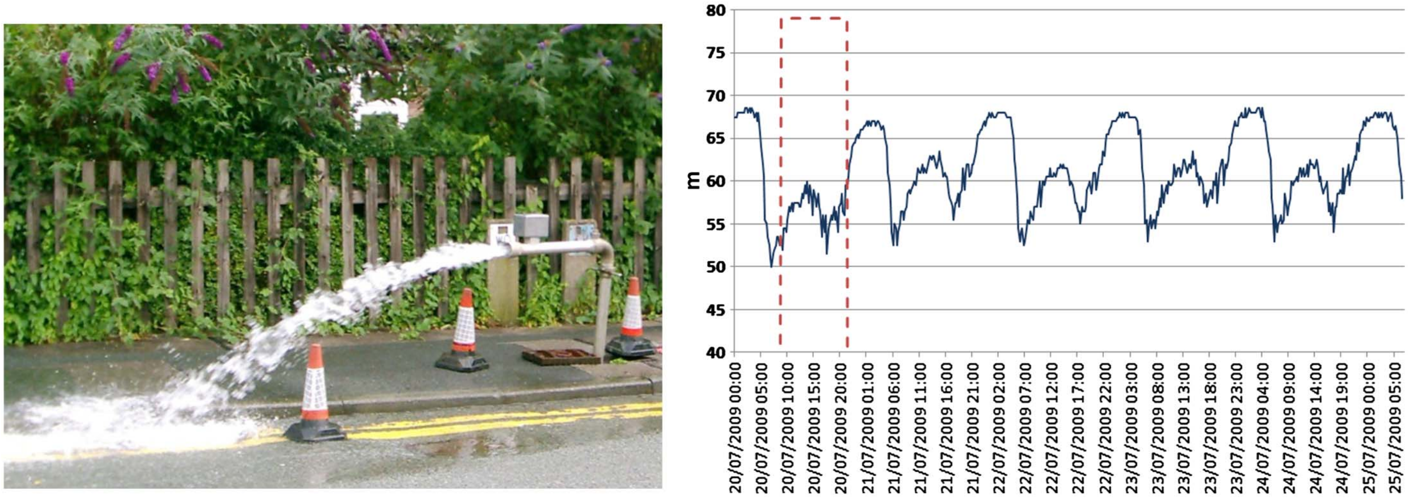

Fig. 4. Hydrant flushing to simulate burst and detected pressure drop in online data

- DMA D—one blind flushing was conducted.

Regular automated data analysis allows the identification of new leaks as they occur, including smaller events not displaying obvious surface signs of their presence. Such data analysis can be as simple as flat-line alarm levels, or use automated profiling for alarm limits. A more sophisticated way it can be achieved is through intelligent *smart alarms*. Recent developments in the field of computational intelligence variously called soft computing, machine learning, or data-driven modeling are helping to solve various problems in the water resources domain. Evora and Coulibaly (2009) presented a review of recent advances in artificial neural-network modeling of remote-sensing applications in hydrology. In order for the optimal siting methodology to be assessed, some form of automated, online system for analyzing the flow and pressure data was required, i.e., to determine when a leak/burst event had occurred within a DMA and evaluate any change at the pressure instrumentation. Mounce et al. (2010a) describe an online system pilot implemented with a UK water company using an artificial neural network (ANN) and fuzzy inference system (FIS) system for detection of leaks/bursts at the DMA level. This event detection system is not reliant on any special hardware or network configuration and produces intelligent alerts. The automated analysis system is data driven, starting from the logger units which initiate calls to the telemetry software every thirty minutes, GPRS signal permitting. The system is designed to provide detection of bursts and leaks as they occur but not existing leaks or background leakage. The system provides *sensitive* detection of abnormal flow and pressure events and, due to the ANN/FIS system developed, provides a confidence estimate, in the form of a percentage, of how unusual the event is together with an estimate of the burst flow rate that can be very effectively used to prioritize events and response. This system was operating on the flows and pressures logged in the case study area for the three month period of the pilot (January–March 2010) which included the blind flushing described above, hence online data was used as verification for event detection.

Results

Where the ANN/FIS system was operating and data sources were available, the near-real-time system detected all events providing a robust confirmation of its ability to detect abnormal flow level. A more in depth investigation of both flow and pressure alerts now follows to assess how well the model methodology for division of DMAs into zones performed.

Table 3 shows that for the three events conducted in DMA C (Events 1, 2, and 3) all successfully correctly identified the zone in which the leak/burst event had occurred. There was full agreement between the model-analyzed response and the ANN/FIS system determined response for all instruments. Table 3 also shows that no event in this DMA was detected by the DG2 instrument, this suggests that this instrument is not optimally placed to detect low pressure events.

Blind Testing Field Validation

The five blind tests of the method were conducted in four DMAs in two phases in a one month period and provided a robust test, as the location and size of the event was dependent on the technician at the water company. The aim was to make the simulated leak/burst events as realistic as possible.

The subdivision of DMA A was previously shown in Fig. 3. A blind flush occurring in DMA A is now used to illustrate results obtained. The DMA is divided into three zones, shown in Fig. 5, and events that occur in the yellow area should be detected by the

Table 3. Comparison of the Model-Analyzed and ANN/FIS System Detection for the Nonblind Test in DMA C

		Detection on flow at inlet		Detection on pressure at inlet		Detection on DG2		Detection on optimal 1		Detection on optimal 2		
DMA	Event	Model	AI	Model	AI	Model	AI	Model	AI	Model	AI	Correct location
C	1	Yes	Yes	No	No	No	No	Yes	Yes	Yes	Yes	Yes
C	2	Yes	Yes	No	No	No	No	Yes	Yes	Yes	Yes	Yes
C	3	Yes	Yes	No	No	No	No	Yes	Yes	Yes	Yes	Yes

Note: AI = ANN/FIS, short for artificial intelligence.

J. Water Resour. Plann. Manage., 2013, 139(6): 604-613

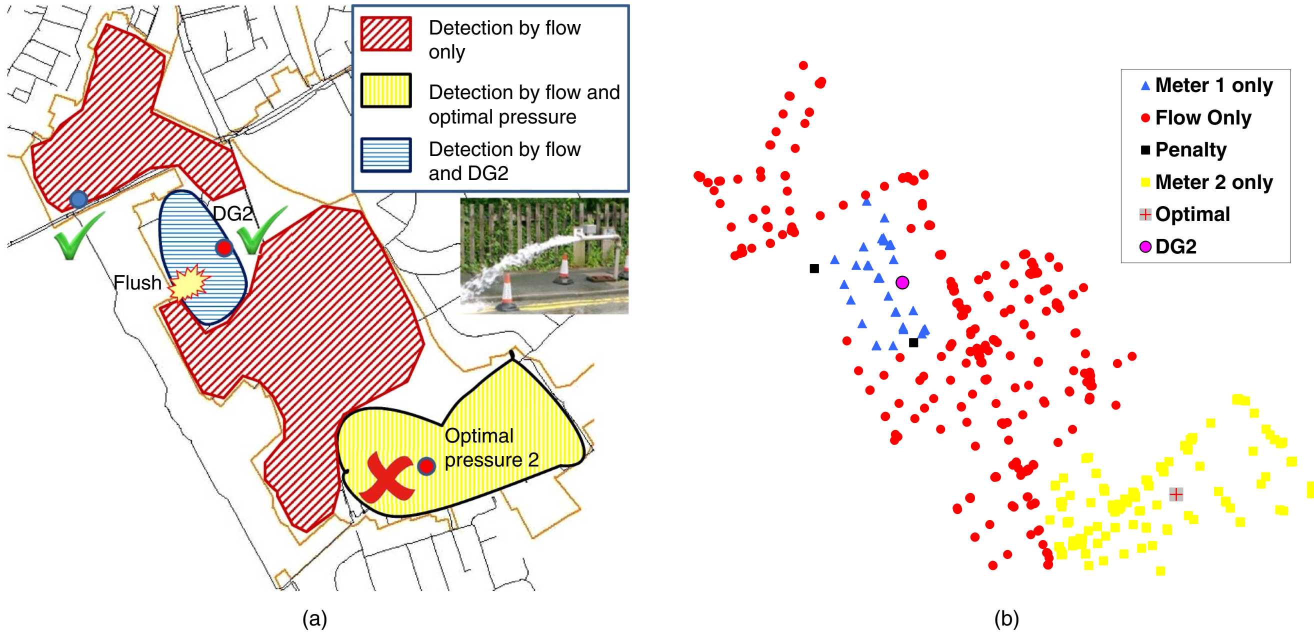

Fig. 5. Example of how the method can be used to subdivide a DMA, using DMA A: (a) response of instruments to events and detection type by area; (b) model-predicted response zones

optimal pressure instrument. Events that occur in the blue zone will be detected by the second optimal/DG2 instrument. Events in the rest of the DMA will only be detected by the flow meter at the inlet. The blind test was conducted in the blue zone (in Fig. 5); as a result it should be detected by the optimal/DG2 instrument. The response of the instruments is shown in Fig. 5(a).

Fig. 5(a) shows that the event can be successfully located in the DMA since there is complete agreement between the ANN/FIS detection and model analyzed detection. In this case, automated alerts for the flow sensor (estimated flush size 6.9 L/s) and DG2 pressure instrument/sensor were generated. The optimal pressure instrument has not responded, so this rules out its reporting zone as a potential location for the event. The flow meter has responded as has the optimal/DG2 instrument. Therefore the event occurred in the small zone close to the DG2 instrument from the optimal division of DMA using two pressure instruments. The correct zone of the leak/burst event was identified. Table 4 summarizes all five blind tests.

Table 4 shows that there was complete agreement between the model-analyzed detection and ANN/FIS detection for Event 4 in DMA A, Event 5 in DMA B, and Event 7 in DMA C. In the case of Event 5, the event was conducted close to the inlet. Throughout the use of the modeling methodology adopted such events have proved difficult to detect and, as a result, the zone close to the inlet are generally flow-only zones. In the case of Event 6, three of the instruments used in this field work have agreement between the model-analyzed and actual response for DMA C. However for the methodology to be successful, all instruments need to agree. Consequently the location of this event was incorrectly attributed to the wrong zone in this DMA. For Event 8, the correct zone was not identified. As predicted in Table 4, the DG2 location did not detect the engineered leak/burst event. This shows that the DG2 location is not the most sensitive for detection of leak/burst events.

Discussion of Flushing Test Results

The results from the flushing tests conducted in the five DMAs were generally positive. There was total agreement in the researcher-defined test conducted in DMA C (Table 3), with the correct zone being identified each time. Three out of the five blind tests produced exact agreement between the model-analyzed and actual response, and therefore the correct zone of the leak/burst event could be identified. In the other two events, factors beyond the control of this methodology led to the incorrect zone being identified. There was a large difference between the modeled and actual normal (nonevent) pressures in DMA D; this reflects common issues with reliability of models. Tests in DMA D were further

Table 4. Comparison of Model Predicted and ANN/FIS Detection for Blind Test Events

		Detection on flow at inlet		Detection on pressure at inlet		Detection on DG2		Detection on optimal 1		Detection on optimal 2		
DMA	Event	Model	AI	Model	AI	Model	AI	Model	AI	Model	AI	Correct location
A	4	Yes	Yes	No	No	Yes	Yes	No	No	N/A	N/A	Yes
B	5	Yes	Yes	No	No	No	No	No	No	N/A	N/A	Yes
C	6	Yes	Yes	No	No	No	No	Yes	No	Yes	No	No
C	7	Yes	Yes	No	No	No	No	No	No	No	No	Yes
D	8	Yes	ND	No	ND	No	No	No	Yes	N/A	N/A	No

Note: ND indicates no data was available for this instrument at this time.

J. Water Resour. Plann. Manage., 2013, 139(6): 604-613

hampered, as two instruments were not working during the period of the blind tests. One instrument was the flow meter and therefore it was not possible to determine the size of the event. For Event 7, conducted in DMA C, the pressure instruments failed to detect this event, when the model predicted they should do so. This is likely to be down to the small size of the leak/burst event, thus making the pressure change difficult to determine. Farley et al. (2010b) showed that smaller leak/burst events (typically less that 1.5 L/s) are more difficult to detect.

The instrumentation was installed in the DMA for a considerably long period of time (approximately 3 months). During this period, the ANN/FIS online system was analysing pressure and flow at all instrument points (where instrumentation operation and communications allowed). In this period a number of ghost detections were recorded at pressure instruments. Ghost detections are more common when using pressure time-series values for detection, as pressure fluctuates much more than the flow in the system. If the pressure instruments were to be used solely to detect leak/burst events, this would be of concern and potentially lead to the revision of the detection bands (parameter settings) of the analysis system. During the period of the pilot only 5% of ghost alerts were encountered for flow analysis compared to 38% for pressure analysis demonstrating that the flow signal is much more reliable for event detection [see also Mounce et al. (2011)]. However, when flow is used first to determine detection, the pressure instruments can then potentially be used to determine location. For six of the eight events, the correct zone in which the leak/burst event had occurred was identified when both flow and pressure instruments were used.

The first set of known events were all conducted in areas where a response from the optimally placed instrument would occur, and as a result were successfully located. As the locations of the blind tests were not preselected, for research purposes, a number of the events were conducted close to the inlet at low flow rates. From a water company's perspective this is ideal as it will cause a minimum impact on the pressures in the system. However it meant that the majority of events were conducted in flow-only detection zones. Therefore the pressure instruments' ability to detect was not comprehensively tested. Only two of the blind events were conducted in zones that would be expected to achieve pressure detection. Of these two events, one was successful. The failure to detect the other event can be attributed to the low flow rate of the event, which did not cause a large change in pressure at the optimally placed pressure instrument/sensors as anticipated. The hydrant flushing was useful to allow for further testing of the methodology in a real world environment and to develop an understanding of how data analysis systems and optimal methodology (for detection and location) can operate together.

Conclusions

1. A method, based on GA optimization of a Jacobian sensitivity matrix derived from hydraulic model simulation results, that is able to detect and reduce the search time for finding a leak/burst event(s) by subdividing a DMA has been presented with the following features:
   - Practicality, having been developed and tested with input from the water industry;
   - Requires only a low number of instruments;
   - Utilizes current industry-standard hydraulic models with a pressure-dependent leak function and with previous validation having demonstrated independence of leak-size diameter; and
   - Allows integration with an appropriate event-detection system (an automated ANN/FIS system was used in this study), enabling the potential for further development as an online sub-DMA location system.
2. The methodology has been successfully applied to a number of DMAs as part of a verification and validation scheme in a real WDS in the UK using pressure and flow data.
   - Verification was conducted on 14 DMA models illustrating sub-division for a range of DMA characteristics; and
   - Validation was carried out with a total of eight events (hydrant openings) conducted in four DMAs. Three of these were in known locations (selected by the research team) and five in blind locations (selected by a water technician).
     a. Six events were correctly localized to sub-DMA areas of the DMA;
     b. One event was inaccurate due to an error in the model; and
     c. One event was inaccurately predicted, this is likely to be due to the small size of the event.
3. A field-verified and validated methodology has been presented in this work offering a practicable method to greatly reducing the time taken for leak/burst events to be located. Such advances are essential to the urgent need for water loss reduction.

Acknowledgments

This work was supported by the EPSRC (grant EP/E003192/1) and industrial collaborators. The authors would like to thank Yorkshire Water Services for their assistance with models and data. In addition, the authors wish to acknowledge 7Technologies for provision and support of the hydraulic modeling software. Particular thanks are due to Mr. R. Patel, Mr. L. Soady, and Dr. J. Machell for their help and support in conducting field work.

References

Berry, J., Hart, W. E., Phillips, C. A., Uber, J. G., and Watson, J. P. (2006). "Sensor placement in municipal water networks with temporal integer programming models." *J. Water Resour. Plann. Manag.*, 132(4), 218–224.

Brothers, K. J. (2003). "A practical approach to water loss reduction." *Water*, 21, 54–55.

Bush, C. A., and Uber, J. G. (1998). "Sampling design methods for water distribution model calibration." *J. Water Resour. Plann. Manag.*, 124(6), 334–344.

Chipperfield, A. J., Fleming, P. J., and Fonseca, C. M. (1994). "Genetic algorithm tools for control systems engineering." *Proc., Adaptive Computing in Engineering Design and Control*, Plymouth Engineering Design Centre, Univ. of Plymouth, Plymouth, UK, 128–133.

EIRIS. (2011). "A drought in your portfolio: Are global companies responding to water scarcity?" *EIRIS water risk report*, London, ⟨http://www.eiris.org/files/research%20publications/EIRISWaterRiskReport2011.pdf⟩.

Environment Agency. (2011). "Environment Agency News, UK." ⟨http://www.environment-agency.gov.uk/news/130645.aspx⟩ (Mar. 7, 2011).

Evora, N. D., and Coulibaly, P. (2009). "Recent advances in data-driven modeling of remote sensing applications in hydrology." *J. Hydroinform.*, 11(3–4), 194–201.

Farley, B., Boxall, J. B., and Mounce, S. R. (2008). "Optimal locations of pressure meter for burst detection." *Proc., 10th Water Distribution System Analysis Symp.*, ASCE, Reston, VA, 1–11.

Farley, B., Mounce, S. R., and Boxall, J. B. (2010a). "Field validation of 'optimal' instrumentation methodology for burst/leak detection and location." *Proc., 11th Water Distribution System Analysis Symp.*, ASCE, Reston, VA, 1093–1102.

Farley, B., Mounce, S. R., and Boxall, J. B. (2010b). "Field testing of an optimal sensor placement methodology in an urban water distribution network for event detection." *Urban Water J.*, 7(6), 345–356.

Goldberg, D. E. (1989). *Genetic algorithms in search optimization and machine learning*, Addison-Wesley, Boston.

Janke, R., Murray, R., Uber, J., and Taxon, T. (2006). "Comparison of physical sampling and real-time monitoring strategies for designing a contamination warning system in a drinking water distribution system." *J. Water Resour. Plann. Manag.*, 132(4), 310–313.

Kapelan, Z., and Savic, D. (2009). "Editorial: Special issue on WDS model calibration." *Urban Water J.*, 6(1), 1–2.

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

MATLAB R2010a [Computer software]. Mathworks, Natick, MA.

Mounce, S. R., Boxall, J., and Machell, J. (2010a). "Development and verification of an online artificial intelligence system for burst detection in water distribution systems." *J. Water Resour. Plann. Manag.*, 136(3), 309–318.

Mounce, S. R., Farley, B., Mounce, R. B., and Boxall, J. B. (2010b). "Field testing of optimal sensor placement and data analysis methodologies for burst detection and location in an urban water network." *Proc., 9th Int. Conf. on Hydroinformatics*, Vol. 2, Chemical Industry Press, Tianjing, China, 1258–1265.

Mounce, S. R., Mounce, R. B., and Boxall, J. (2012). "Identifying sampling interval for event detection in water distribution networks." *J. Water Resour. Plann. Manag.*, 138(2), 187–191.

Mounce, S. R., Mounce, R. B., and Boxall, J. B. (2011). "Novelty detection for time series data analysis in water distribution systems using support vector machines." *J. HydroInform.*, 13(4), 672–686.

Nicklow, J., et al. (2010). "State of the art for genetic algorithms and beyond in water resources planning and management." *J. Water Resour. Plann. Manag.*, 136(4), 412–432.

Ostfeld, A., and Salomons, E. (2004). "A stochastic early warning detection system model for drinking water distribution systems security." *Proc., EWRI ASCE Conference*, ASCE, Reston, VA.

Parliamentary Office of Science, and Technology (POST). (2011). "Water in production and products." *POSTnote 385*, The Parliamentary Office of Science and Technology, London.

Perez, R., Puig, V., Pascual, J., Peralta, A., Landeros, E., and Jordanas, L. (2009). "Pressure sensor distribution for leak detection in Barcelona water distribution network." *Water Sci. Technol.: Water Supply*, 9(6), 716–721.

Romano, M., Savic, D., and Kapelan, Z. (2011). "Geostatistical techniques for approximate location of bursts in water distribution systems." *Proc. Urban Water Management: Challenges and Opportunities*, D. Savic, Z. Kapelan, and D. Butler, eds., Univ. of Exeter, Exeter, UK, 695–700.

Watson, J. P., Hart, W. E., Woodruff, D. L., and Murray, R. (2010). "Formulating and analyzing multi-stage sensor placement problems." *Proc., 11th Water Distribution System Analysis Symp.*, ASCE, Reston, VA, 347–354.

Wu, Z. Y., et al. (2011). *Water loss reduction*, Z. Y. Wu, ed., Bentley Systems, Dublin, Ireland.

Wu, Z. Y., and Sage, P. (2008). "Water loss detection via genetic algorithm optimization-based model calibration." *Water Distribution Systems Analysis Symp. 2006*, ASCE, Reston, VA, 1–11.

Wu, Z. Y., Sage, P., and Turtle, D. (2010). "Pressure-dependent leak detection model and its application to a district water system." *J. Water Resour. Plann. Manag.*, 136(1), 116–128.

J. Water Resour. Plann. Manage., 2013, 139(6): 604-613