Optimal meter placement for pipe burst detection in water distribution systems

Medhanie Hagos, Donghwi Jung and Kevin E. Lansey

ABSTRACT

Pipe bursts in water distribution systems (WDS) must be rapidly detected to minimize the loss of system functionality and recovery time. Pipe burst is the most common failure in WDS. It results in water loss out of the system, increased head losses, and low pressure at the customers' taps. Therefore, effective and efficient detection of pipe bursts can improve system resilience. To this end, this study proposes an optimal meter placement model to identify meter locations that maximize detection effectiveness for a given number of meters and type of meter. The linear programming model is demonstrated on a modified Austin EPANET hydraulic network. Receiver operating characteristic (ROC) curves for alternative pressure and flow meters are applied to investigate the relationship between the level of available information and pipe burst detection effectiveness. The optimal sensor locations were distinctly different depending on the type of meter and the objective to be considered. The ROC curves for alternative pressure and pipe flow meters showed that pipe flow meters are vulnerable to false alarms, and that using many pipe flow meters could detect all pipe bursts. Pressure meters could detect up to 82% of the burst events.

Key words | detection, pipe bursts, receiver operating characteristic (ROC), water distribution system, Western Electric Company (WEC) rules

Medhanie Hagos
Kevin E. Lansey
Department of Civil Engineering and Engineering Mechanics,
University of Arizona,
Tucson,
AZ 85732,
USA

Donghwi Jung (corresponding author)
Research Center for Disaster Prevention Science and Technology,
Korea University,
Seoul 136-713,
Korea
E-mail: *donghwiku@gmail.com*

NOMENCLATURE

a	emitter pressure exponent
C	emitter discharge coefficient
$\mathbf{D}$	detection matrix
D_{ij}	an element of $\mathbf{D}$ having a value of 1 or 0, indicating whether the ith burst event would be or would not be detected, respectively, by the jth meter
i	index for burst events
j	index for meters
m	total number of nodes or pipes
N_d	number of burst events that were detected
N_{tb}	total number of burst events
$p_{t,i}$	pressure at node i at time t
$q_{t,i}$	burst flow at node i at time t
t_{di}	time (h) of detection of the ith detected burst event
t_{bi}	time (h) of occurrence of the ith detected burst event
$\mathbf{X}$	configuration vector
X_j	an element of $\mathbf{X}$ indicating whether there is a meter at the jth location (i.e., node or pipe) or not
X_{max}	number of available meters
ADT	average detection time
CUSUM	cumulative sum
CL	control limit
DP	detection probability
EWMA	exponentially weighted moving average
MCUSUM	multivariate CUSUM
MEWMA	multivariate EWMA
ROC	receiver operating characteristic
RF	rate of false alarm

doi: 10.2166/hydro.2016.170

SCADA	supervisory control and data acquisition
SPC	statistical process control
WL	warning limit
WEC	Western Electric Company

INTRODUCTION

Pipe burst refers to the rupture of a pipe due to pipe deterioration, excessive pressure, or ground shifts caused by temperature changes or earthquakes. Water loss through the rupture increases the total flow entering the system, overall pipe flow rates, and head losses through pipes, which in turn, lowers the pressure at the customers' taps. Morrison (2004) reported that the time for detection and location of a 1.1 lps burst was approximately 5 days. Degradation of system functionality (i.e., performance level) persists until the broken pipe is fully repaired. Therefore, early detection of pipe bursts helps improve system resilience by minimizing water losses and system impact. In addition, it saves the energy cost spent in lifting the additional quantity of water.

Over the last two decades, several methodologies have been developed to detect pipe bursts. Most approaches assume the availability of information on measured pipe flows and/or system pressures from supervisory control and data acquisition (SCADA) systems, and include artificial neural networks (Mounce *et al.* 2002, 2003, 2010; Mounce & Machell 2006), state estimation (Andersen & Powell 2000; Ye & Fenner 2011, 2013), Bayesian approach (Poulakis *et al.* 2003), and time series modeling (Quevedo *et al.* 2010). Constructing a district metered area (DMA)-structured network was considered the most common alternative for detecting unreported leakages and small bursts (Morrison 2004; Misiunas *et al.* 2005; Sturm & Thornton 2005; Alkasseh *et al.* 2013).

Statistical process control (SPC) methods are most widely applied in the detection of pipe bursts (Misiunas *et al.* 2006; Romano *et al.* 2010, 2014; Palau *et al.* 2012; Jung *et al.* 2015). SPC methods apply statistical theory to the system output parameters (i.e., the measurable quality parameters, also known as quality characteristics) to identify non-random patterns that may be caused by bursts.

Romano *et al.* (2010) applied the Western Electric Company (WEC) rules (Western Electric Company 1958) to the pressure data from a 13 m long pipeline deployed in a DMA to identify pipe bursts generated by opening hydrants. Normal pressure data were used to compute time-varying mean and standard deviation that are the bases of Shewhart control chart and the resulting warning and control limits (CLs). Violating any of the four rules in the WEC rules caused an alarm. In a follow-up paper, Romano *et al.* (2014) employed the WEC rules in a real-time event detection model that included preprocessing of pipe flow/pressure measurements, forecasting pipe flows/pressures, detecting pipe bursts/other events, and estimating the probability of the events. The WEC rules are fixed as identifiers, thus making its application straightforward.

Jung *et al.* (2015) compared the performance of six SPC methods for pipe burst detection with respect to their detection effectiveness and efficiency: three univariate methods (the WEC rules, the cumulative sum (CUSUM), and the exponentially weighted moving average (EWMA) methods) and three multivariate methods (Hotelling T-squared, multivariate CUSUM (MCUSUM), and multivariate EWMA (MEWMA) methods). Hypothetical pipe flow and pressure measurements were generated at predefined locations in a modified Austin network (Brion & Mays 1991). The EWMA had the best detection effectiveness and efficiency among the six SPC methods.

Optimal sensor network design to detect bursts

In addition to being affected by the detection method, pipe bursts' detectability is affected also by the network structure and the available information. During design, the network structure is largely determined by land use, topography, and water demand distribution; pipe burst detection is not a primary concern at this stage. The amount of information on pipe burst detection in a measurement is a function of the number and types of meters and their locations, as well as the network structure and burst location. As such, measurements at some locations can include more signals of an anomaly than measurements at other locations.

To improve pipe burst detectability, the number of meters within the system can be increased. However, the number of meters that can be installed is normally limited

by the network's structure, and more importantly, by the available budget. Therefore, finding an optimal sensor network design under budget constraints is a challenging task. Sensor network design methodology must identify: (1) the optimal number of meters, (2) the optimal combination of meters (pressure and pipe flow meters), and (3) the locations of the meters, and incorporate an optimal methodology for pipe burst detection.

In the water distribution systems (WDS) domain, a few studies have proposed sensor placement methodologies for pipe burst detection. Farley *et al.* (2010) have developed a sensor placement methodology to determine the most sensitive locations of pressure meters to detect pipe bursts at different locations. The approach was based on the Jacobian sensitivity matrix of pressure change as a function of burst rate at a given location. The Jacobian matrix was converted to a binary matrix where '1' indicates that the pressure change is less than an arbitrary chosen threshold value, and '0' indicates otherwise. The number of burst events detected was maximized using the generated binary matrix. In a follow-up paper, they proposed a pipe burst localization method using the Jacobian matrix (Farley *et al.* 2013). Perez *et al.* (2009) proposed a sensor placement method that determines the location of the smallest number of pressure meters for burst detection and localization. An enumeration technique was used to determine the threshold value for the Jacobian-binary matrix conversion. Wu & Song (2012) developed a pressure meter location method that maximizes the number of unique leakage events detected for a given number of candidate locations. Analogous to Farley *et al.* (2013) the Jacobian matrix of the pressure change was converted to a binary matrix, using the sensor instrument's measurement error as the threshold value. Huang *et al.* (2012) developed a clustering-based pressure meter placement method for pipe burst detection, by which the optimal number of meters was determined. A few studies considered a different criterion (e.g., accuracy of state estimation) for determining sensor locations (Kang & Lansey 2010).

While most previous studies were based on the Jacobian matrix of pressure change as a function of burst rate, an arbitrary threshold value was applied to convert the Jacobian matrix to a binary matrix. However, a more efficient pipe burst detection methodology (e.g., SPC method) should be applied to the raw field measurements to develop the binary matrix. In addition, most of the previous sensor placement studies focused on pressure meters only, without considering the potential detection effectiveness of pipe flow meters. Further, they neglected the importance of false alarms in determining sensor locations.

In this study, we develop and solve an optimal WDS meter placement model to identify the optimal meter locations that maximize detection effectiveness for a given number of meters and type of meter. An SPC method is used to populate the detection matrix (i.e., binary matrix) in the model. The linear programming-based model is demonstrated on a modified Austin network (Brion & Mays 1991). First, the pipe burst detection characteristics are investigated based on a single meter's performance. Then, the optimal sensor networks for pressure and flow meters are independently determined and compared. Finally, the trade-off relationship between the number of available meters and pipe burst detection effectiveness (with respect to not only detection probability (DP), but also false alarm frequency) is investigated by plotting the latter against the former.

METHODOLOGY

Detectability measures

Detectability of a burst detection method involves a combination of the method's detection effectiveness and detection efficiency. Detection effectiveness is related to how well burst events are detected and false alarms in natural random patterns are avoided. Therefore, it is measured by the *DP* and the rate of false alarm (*RF*). *DP* is the proportion of burst events that were detected (N_d) among the total number of burst events (N_{tb}):

$$DP = \frac{N_d}{N_{tb}} \times 100 \tag{1}$$

RF is the proportion of natural random events in which a false alarm is issued among the total number of natural events. Therefore, *DP* is related to the false negative (type II error) rate, while *RF* indicates false positive (type I

error) rate. The receiver operating characteristic (ROC) curve (Egan 1975), a graphical plot that illustrates the performance of a binary classifier system, is created by plotting the true positive rate against the false positive rate at various threshold settings. Therefore, the ROC curve for a sensor network for pipe burst detection displays plots of *DP* and *RF* for different number of meters to explore the trade-off between the two metrics.

Another measure of burst detection sensor network quality is detection efficiency, which evaluates how fast a burst event is detected. The average detection time (*ADT*, h), i.e., the averaged value of the time taken for detection, is used here as the detection efficiency indicator:

$$ADT = \frac{\sum_{i=1}^{N_d}(t_{di} - t_{bi})}{N_d} \tag{2}$$

where t_{di} is the time (h) of detection of the ith detected burst event, and t_{bi} is the time (h) of occurrence of the ith detected burst event. Normally, t_{bi} of a burst is not known in real life, so this statistic is tested using synthetic data. Given the same detection effectiveness (i.e., the same *DP* and *RF*), a sensor network with a shorter *ADT* is more favorable. As the detection time is infinite for a non-detected pipe burst, water utilities generally place more importance on the effectiveness measures than the *ADT*. In summary, to improve resilience, it is desirable to have a high *DP*, low *RF*, and low *ADT*.

Binary integer programming model for optimal meter location

The optimal meter location model is posed as a binary integer programming problem that maximizes *DP* for a given number of meters (X_{max}).

$$\text{Maximize } DP = \frac{\left(\sum_{i=1}^{N_{tb}} \alpha_i\right)}{N} \tag{3}$$

subject to

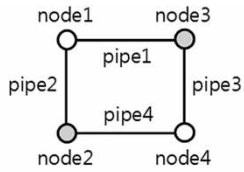
$$\alpha_i < \sum_{j=1}^{m}(D_{ij} \times X_j) \text{ for all } i \tag{4}$$

$$\alpha_i = 0 \text{ or } 1 \tag{5}$$

$$\sum_{j=1}^{m} X_j \le X_{max} \tag{6}$$

where m is the total number of potential meter locations (i.e., total number of nodes or pipes). Prior to optimizing the system, a detection matrix (**D**) is generated. The rows and columns of **D** correspond to potential burst events and potential meter locations, respectively. The elements of **D** (i.e., D_{ij}) are 1 or 0, which indicate whether the ith burst event would be or would not be detected, respectively, by the jth meter. The development of **D** is described in detail below. The problem decisions are also binary and are denoted by the configuration vector (**X**). If $X_j = 1$, a meter is placed at location j (zero for no meter). For a given burst event i, the product of D_{ij} and X_j in Equation (4) equals 1, if a meter is installed at location j ($X_j = 1$) and the burst can be detected by the meter at that location ($D_{ij} = 1$).

Figure 1 shows an example of a network with pressure meters installed at nodes 2 and 3. Therefore, the configuration vector has the element value of 1 for nodes 2 and 3, and 0 for the other nodes. The detection matrix shows the detection of four pipe bursts that occurred in the network, with a meter installed at each node. For example, event 3 can be detected only by the pressure meter at node 3, while event 4 can be detected by the meters at nodes 2 and 4.



D_{ij}		node1	node2	node3	node4
Event	1	1	0	1	0
	2	0	1	0	0
	3	0	0	1	0
	4	0	1	0	1

Detection matrix

node	X_j
1	0
2	1
3	1
4	0

Configuration vector

Figure 1 | The detection matrix and configuration vector of a network (example). Pressure sensors are installed at gray-filled nodes (nodes 2 and 3).

The problem's objective (Equation (3)) is to maximize *DP* that is computed from the intermediate binary problem variable α_i. If one or more meters would detect burst *i*, then α_i is set equal to 1 (Equations (4) and (5)). If all possible sets of meters are considered, the search space of the problem will be 2^m. Here, the number of meters is included as a constraint (Equation (6)): (1) to reduce the search space, (2) to consider a field situation where the possibility of placing meters in the system is limited, and (3) to exclude the need for system-specific meter cost data.

Cost can be the main decision driver in this problem; the other drivers are meter types' access to power or SCADA data lines, and the location relative to these lines. For example, installing flow meters in a WDS is more expensive than installing pressure meters, because they require excavation, installation of a valve and a meter, backfilling, and pavement work (Walski 2012). On the other hand, pressure meters can often be installed inexpensively on fire hydrants. The flow meter is also more expensive than a pressure meter. To minimize cost, the above problem could be modified by deleting Equation (6) and changing the objective to the cost of installed meters and introducing a budget constraint. Equation (3) would become a constraint with a lower bound on the acceptable *DP*. Similarly, the false alarm rate can also be bounded by adding a false alarm matrix similar to **D,** but corresponding to a set of non-burst conditions that some meter locations identify as bursts. An alternative approach is to minimize *RF*; this also requires a false detection matrix ($\mathbf{D}_f$) analogous to **D**. The rows and columns of $\mathbf{D}_f$ correspond to natural random events and potential meter locations, respectively. As in the case of **D**, the elements of $\mathbf{D}_f$ ($D_{f,ij}$) are 1 or 0 indicating whether a false alarm would be or would not be raised, respectively, on the *i*th natural event by the *j*th=meter.

Overview of model development

The flow chart in Figure 2 shows the process followed to formulate the detection matrix **D**. For a given WDS, a period of random demands is simulated to determine the mean and variance of nodal pressures and pipe flow rates throughout a diurnal cycle. Demands are based on a historical record or a predicted range of variability. These statistics are the basis for the Shewhart control charts for each node or pipe. Shewhart control charts and the WEC detection rules form the foundation for this application, but other detection tools can be substituted, if desired.

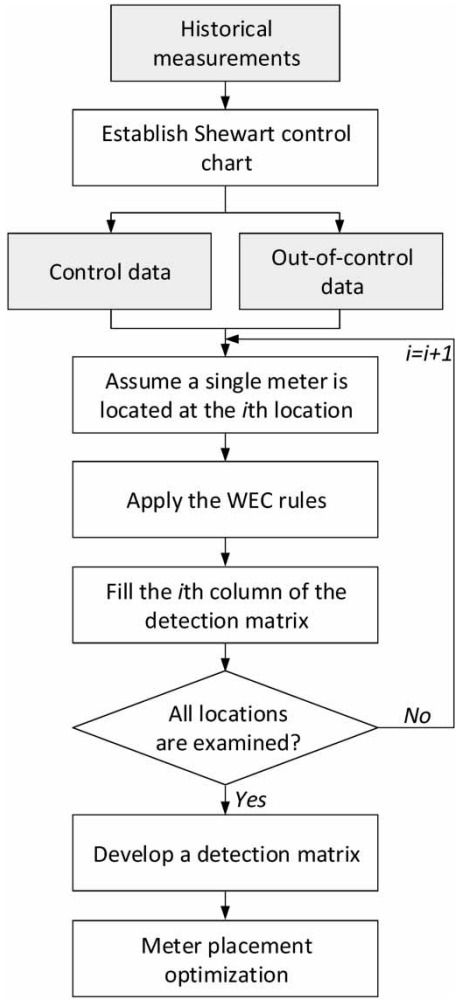


Figure 2 | Flowchart for the proposed optimal meter placement approach.

Shewhart control charts plot the mean value (centerline) of the quality characteristics, thresholds on the expected upper and lower ranges of the variable, and warning limits (WLs) that are multiples of the standard deviation on each side of the mean value (Figure 3). The chart implicitly assumes that the quality characteristic is normally distributed (Surendran *et al.* 2005; Jung *et al.* 2015).

To assess a sensor network's effectiveness, control measurements are generated considering randomness in nodal demands only, while out-of-control measurements are produced considering both random pipe bursts and

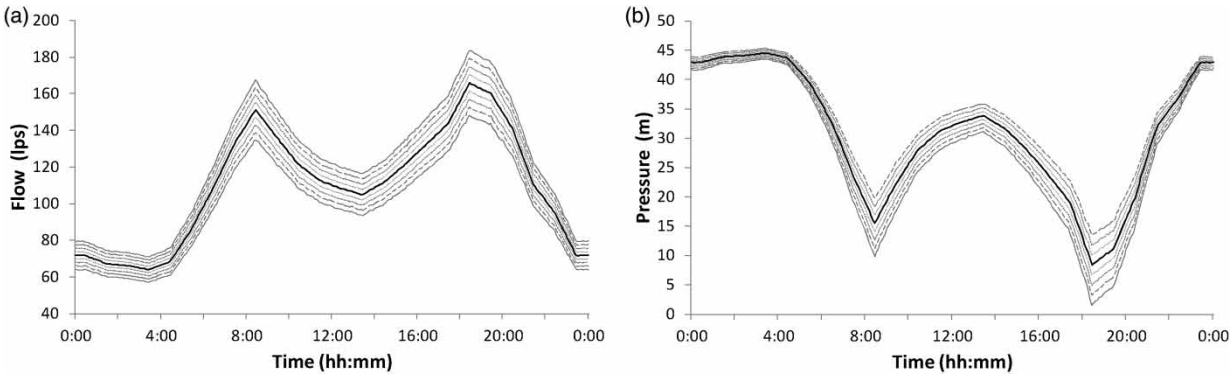


Figure 3 | Representative Shewhart control charts (24 h) for (a) flow and (b) pressure at a representative location.

random demands. The latter is used to develop a detection matrix, and the former is used to form a false alarm matrix. For both conditions, the WEC rules described below are applied to detect the outliers. With the detection matrix **D**, the optimization problem (Equations (3)–(6)) can be formed and solved. The optimal locations to limit the frequency of false alarms are also determined. Therefore, the objective function is to minimize *RF* by replacing D_{ij} with $D_{f,ij}$ in Equation (4).

Burst simulation

Given the system state under normal operating conditions, the impacts of pipe failures are assessed to determine if they cause the system to go out of control at one or more locations. The pipe burst/leakage flow rate is a function of the pipe pressure at the burst location. A higher pressure causes more water loss out of the system. In this study, a pipe burst is modeled as an emitter in EPANET solver (Rossman 2000). The emitter flow is represented as a power function of nodal pressure as:

$$q_{t,i} = Cp_{t,i}^{a} \tag{7}$$

where $q_{t,i}$ is the burst flow at node i at time t, C is the emitter discharge coefficient, $p_{t,i}$ is the pressure at node i at time t, and a is the emitter pressure exponent. Lambert (2001) conducted an experimental study to investigate the accuracy of the power function model represented by Equation (7), and provided guidelines on using the exponent a based on the pipe material and the level of burst/leakage. He has suggested that an exponent of 0.5 be used to simulate large detectable leakages in metal pipes, while an exponent of 1.0 (indicating a linear relationship between discharge and pressure) can be used if no information on the pipe material is available. Fuchs-Hanusch *et al.* (2015) summarized and reported the exponent values derived from field and experimental studies, which cover various pipe materials and failure modes (e.g., round hole and longitudinal cracks). While a wide parameter range was observed for the exponent, the most commonly identified range was around 0.5 (May 1994; Greyvenstein & van Zyl 2007; Walski *et al.* 2009). To simulate pipe bursts of various magnitudes, the emitter discharge coefficient C is considered as a random variable.

WEC rules

The WEC rules (1958) are a set of decision rules based on the Shewhart control chart for detecting non-random patterns in measured data (Montgomery 2009). Romano *et al.* (2010) modified the original WEC rules for WDS pipe burst detection and for identifying a non-random pattern, if the measured quality characteristic satisfies any of the following rules.

Rule 1. Any single measurement is beyond the $\pm 4\sigma$ CL.
Rule 2. Two out of three consecutive measurements are beyond the $\pm 3\sigma$ WLs.
Rule 3. Four out of five consecutive measurements are beyond the $\pm 2\sigma$ WLs.
Rule 4. Eight consecutive measurements are beyond the $\pm 1\sigma$ WLs.

The term $\pm n\sigma$ refers to the warning/CL that is plus or minus n standard deviations from the mean. The WEC rules are applied to one side of the centerline at a time. In other words, a measurement on one side, immediately followed by a measurement on the other side of the centerline outside the WLs will not be taken into account as a non-random pattern. The WEC rules consider, at most, the eight most recent past measurements, having a potential to consider series of previous measurements in decision. Most of the previous detection methods used in sensor placement methodologies took into account a single/current measurement (Farley *et al.* 2010; Wu & Song 2012).

The WEC rules are easy to implement with no parameters to be estimated. Jung *et al.* (2015) confirmed that the WEC rules outperformed multivariate SPC methods with respect to *DP*. A *DP* of over 76% was obtained by applying the WEC rules when measurements of five meters are provided in a loop-dominated system. The *DP* increased up to 93% when pressure measurements of five meters are used for applying the WEC rules. On the other hand, application of the WEC rules results in relatively high false alarm rates when multiple meters are used. Therefore, determining the location of meters that yields minimum false alarm is also important while applying the WEC rules.

Detection and false alarm matrix

The detection matrix (**D**) has binary elements in which the ith row lists all the pressure/pipe flow meters that detect the ith pipe burst as per the WEC rules; its jth column lists all the pipe bursts detected by the jth pressure/pipe flow meter. The following steps are followed to populate **D** (Figure 2).

Step 1. Generate a large number of pipe bursts with random magnitudes and locations.

Step 2. Generate 100 2-day traces of pressure/pipe flow measurements considering random demands and the pipe bursts generated in Step 1.

Step 3. Assuming that the jth pressure/pipe flow meter is installed in the network, determine if the ith burst is detected by examining the post-burst pressure/pipe flow measurements as per the WEC rules.

Step 4. Set the element at the jth column's ith row as 1 if the ith pipe burst is detected by the jth meter and 0 otherwise.

Step 5. Repeat Step 4 for all pipe bursts ($i = 1, \ldots, N$).

Step 6. Repeat Steps 3–5 for all pressure/pipe flow meters ($j = 1, \ldots, m$).

To identify the optimal meter locations that yield minimum false alarm rates, a false alarm matrix can be developed in a similar manner, but constructed for random demand conditions without pipe bursts. The false alarm matrix has binary elements in which the ith row lists all pressure/pipe flow meters that raise a false alarm from the measurements resulting from the ith series of natural variability.

APPLICATION NETWORK

The optimal meter placement model was applied to a modified Austin network (Brion & Mays 1991). The modified network is a branch-dominated transmission system with several looped subareas. It consists of 125 nodes, a fixed head reservoir, 90 pipes, and seven pumps. The pipe sizes range from 152 mm (6 in) to 1,829 mm (72 in). Three out of the seven pumping units were consistently operated.

Following the flowchart shown in Figure 2, three time series sets were generated. First, a normal 2,000-day trace of the nodal pressures and pipe flow rates at a 5 min step was generated using the network's hydraulic model, and this was used to calculate the time-varying mean and standard deviation of the quality characteristics. Normal random nodal demands were inputted to the models. The nodal demands were assumed to be spatially uncorrelated and each one was assumed to have a coefficient of variation of 0.1 based on the study by Kang & Lansey (2009). The demands have a strong temporal correlation because of the diurnal pattern (Figure 4).

The other data sets were (1) a control sample and (2) the out-of-control set used to construct the false alarm and detection matrices, respectively. The control sample considers only natural demand randomness. The out-of-control data set includes demand randomness and pipe bursts. In total, 100 burst events and 100 natural random events were considered. The random burst characteristics

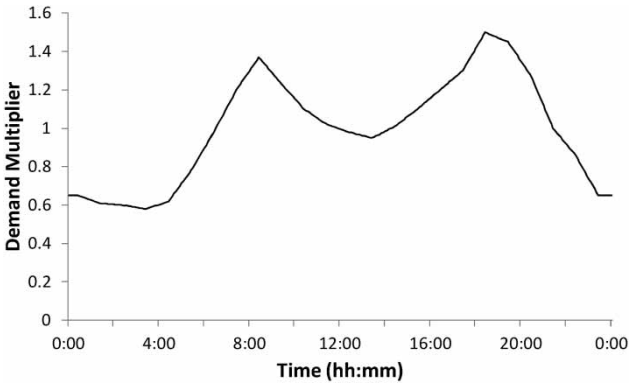


Figure 4 | Diurnal demand pattern repeated twice for 24 h.

are the burst location, initiation time, and burst magnitude. The burst magnitude is defined by the emitter coefficient C that was assumed to be uniformly distributed over the range of 1 to 50. A leakage exponent a of 0.5 is used for all the bursts. The resulting burst magnitudes ranged from 0.1 to 3.3% of the total system mean demand (726 lps). The quality characteristics' control and out-of-control measurements from a location were provided at a time for evaluation as per the WEC rules to populate the false alarm and detection matrices, respectively.

Shewhart control charts with a centerline, three WLs, and a single CL were constructed from the historical data set generated (Figure 3) for each potential meter location (i.e., all nodes and pipes). The mean nodal pressures and pipe flow rates vary over time because of the diurnal demand pattern. While the variances of pipe flow rates are relatively constant through the day, the nodal pressures have high uncertainty during peak demand hours, resulting in wider CLs. Low variations in nodal pressure were observed at night and early morning because of the low demand and low variability during these periods. Control and out-of-control measurements were evaluated as per the WEC rules, using the Shewhart control chart for the desired meter location, to construct the false alarm and detection matrices, respectively. The results for a pressure meter installed at node 15 and representing 15 burst events are summarized in Table 1.

A number of assumptions and simplifications were made in this hypothetical case study: (1) the hydraulic

Table 1 | Detection information using pressure meter at node 15

Burst events								
Event ID	**Location (node ID)**	**Burst time index***	**Detection time index**	**Time to detect (h)**	**Emitter coefficient**	**Burst flow rate (lps)**	**Detection (Yes/No)**	**Detected by rule**
1	89	307	314	0.6	29	16.6	Yes	4
2	77	345			8	4.0	No	
3	26	522			13	6.5	No	
4	22	504	910	33.8	15	7.5	Yes	1
5	35	56	57	0.1	41	22.2	Yes	1
6	76	291	313	1.8	27	14.4	Yes	1
7	108	378	540	13.5	32	18.3	Yes	1
8	99	444	848	33.7	9	5.2	Yes	1
9	48	92			13	7.3	No	
10	59	92	281	15.8	29	15.5	Yes	4
11	94	516	563	3.9	48	27.4	Yes	2
12	5	320	330	0.8	45	20.1	Yes	1
13	5	58	586	44.0	23	10.3	Yes	3
14	34	316	320	0.3	38	20.5	Yes	2
15	61	406			10	5.3	No	

*Burst time index refers to the index of time, starting from a value of 0 at midnight and increasing by 1 every 5 min.

model perfectly represents the real system; (2) the pipe roughness coefficients and other system parameters are known with certainty; (3) a burst that is not detected within 48 h after its occurrence is considered as a non-detected event; (4) a natural random event that does not cause a false alarm within 48 h is considered as a non-false alarming event; (5) only one burst occurs in a span of 48 h; (6) the measurement error of a meter is negligible; (7) measurements are available every 5 min without any missing measurement; (8) only one meter is installed per node/pipe; and (9) the pressure and pipe flow rates are normally distributed around their mean centerlines in the Shewhart control chart, as confirmed in the study by Jung *et al.* (2015). In order to compare alternative pressure/pipe flow meters' configurations, detectability measures (Equations (1) and (2)) are calculated.

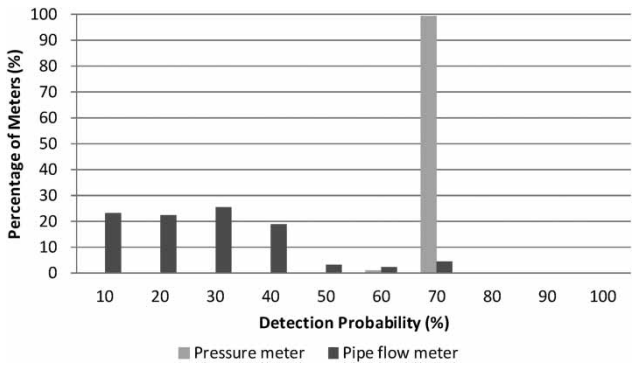


Figure 5 | Histogram of detection probability (*DP*) by the meters.

APPLICATION RESULTS

Impact of pipe burst characteristics on detectability

A histogram of *DP* for pressure and flow meters is given in Figure 5. Nearly all the pressure meters have a *DP* of approximately 70%, while the pipe flow meters generally have relatively low *DP*s. The few locations with low *DP*s are in the fringes of the network. In many cases, common events are detected by all the pressure meters, as shown in Figure 6(a) (heat plot of detections on the vertical axis and meters on the horizontal axis). Each flow meter detects different events (Figure 6(b)). A black spot indicates that a burst is detected by the given meter, so a dark row shows detection by many meters. In this case, 63% of the bursts are detected by all the pressure meters (Figure 7). The

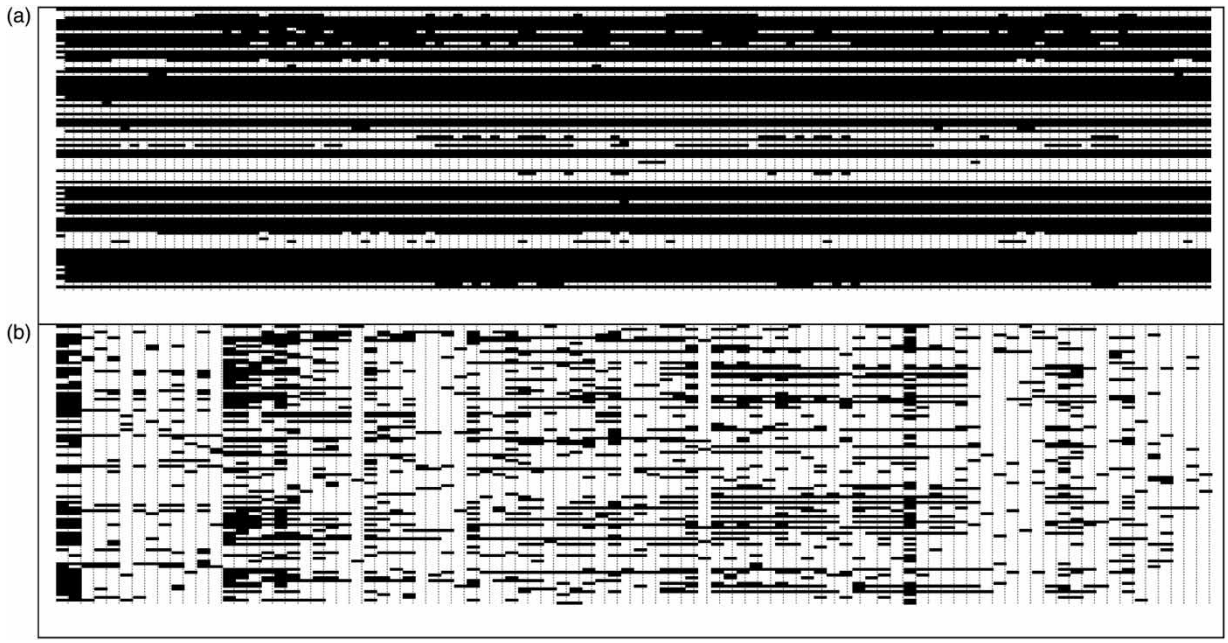


Figure 6 | Detection matrix of (a) pressure meters (100 events (vertical) by 125 nodes (horizontal)) and (b) flow meters (100 events by 90 pipes) with black fill at the cell (*i*, *j*) indicating that the event *i* was detected by meter *j*.

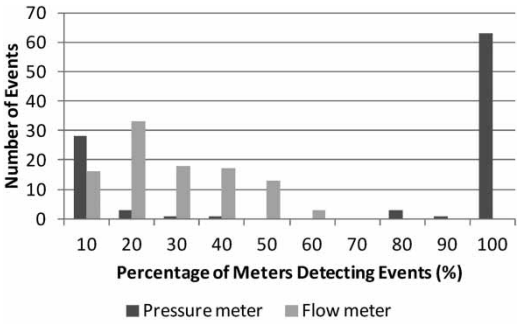


Figure 7 | Histogram of percentage of meters by number of events.

remaining 37% are identified by much smaller subsets; this result suggests the need for a study of the optimal sensor placement problem. However, burst detection by meters at many locations raises questions on the value in locating a burst after a detection using pressure meters that is beyond the scope of this study.

The sensitivity of flow meters to pipe bursts is dependent on their location (Figures 5 and 8). The maximum flow meter *DP* of 66% was for the pipe connecting the source to the pump station. The other flow meters with *DP* higher than 60% also have high flow rates (the pipe just downstream of the pump station, the transmission pipe closest to the pump station, and a distribution pipe within the southwestern loop). This result is reinforced by the plot of *DP* versus flow rate (Figure 9), which suggests that transmission lines are the best locations for flow meters. In practice, therefore, flow meters are installed on each inlet pipe (i.e., transmission pipe) in a DMA-structured network (Alkasseh *et al.* 2013). It is important to note that while many pressures meters can detect a large set of the same bursts, each flow meter detects fewer events that are likely to be more localized (Figure 6(b)). As might be expected, the ratio of the number of meters that detect a pipe burst to the total number of meters (i.e., detection rate) is related to the burst magnitude (Figure 10). All the bursts greater than 1.2% of the total demand were detected by all the pressure meters, while only some of the flow meters identified the bursts.

The *RFs* are small for most pressure and flow meters in this system. A few flow meter locations did have high *RFs* (e.g., a distribution pipe within the southeast loop has an *RF* of 50%) (Figure 11).

Demand and time of day are important factors affecting pipe burst detection. Most bursts were detected at night and early morning because the low demand and pressure/flow variability during these periods makes it easier to detect out-of-control signals in the pressure measurements (Figures 3(b) and 12). Therefore, in a DMA-structured network, the minimum night flow measured at the inlet pipe is generally

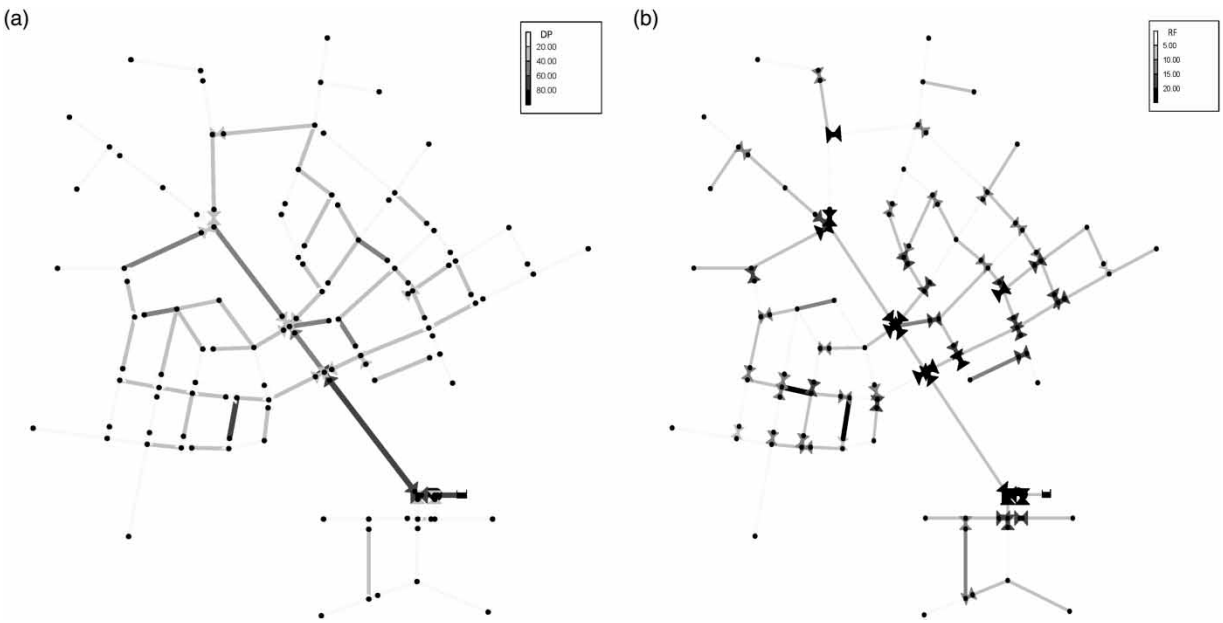


Figure 8 | Spatial distribution of (a) *DP* and (b) *RF* with flow meters. The node size and link thickness are proportional to *DP* or *RF*.

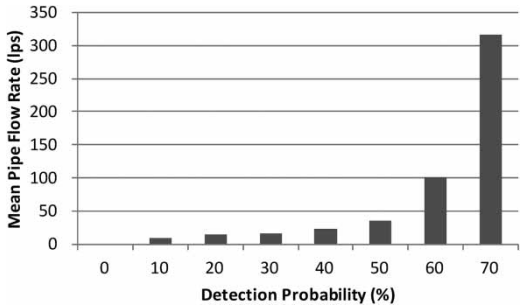


Figure 9 | Histogram of flow meter *DP* versus mean pipe flow rate.

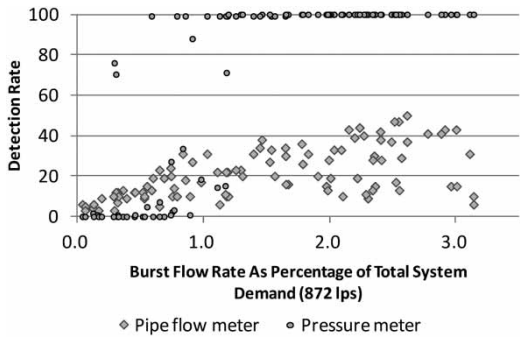


Figure 10 | Detection rate as a function of burst flow rate for flow and pressure meters.

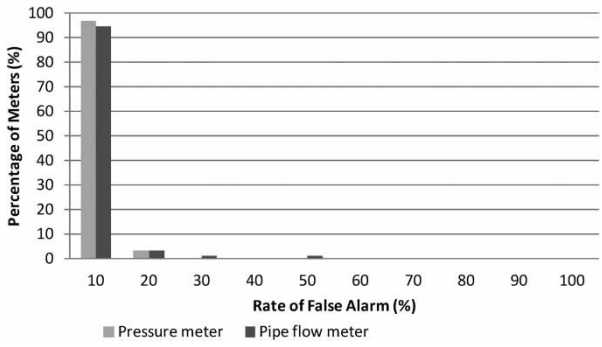


Figure 11 | Histogram of rate of false alarm (*RF*) by the meters.

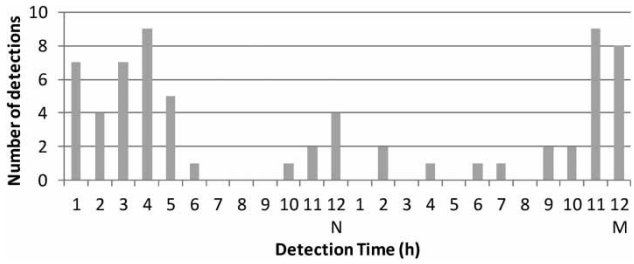


Figure 12 | Histogram of detected events by the detection time for a pressure meter at node 15.

used to detect leakages/pipe bursts in a DMA. As the customer demand is at its lowest at night, the leakage/pipe burst component is at its highest percentage of the inlet flow at night (Sturm & Thornton 2005; Alkasseh *et al.* 2013).

Figure 13 shows the number of detected events for each of the triggering WEC rules for two representative meters. Most of the bursts detected by the pipe flow meter were identified by Rule 1 that flags an event when any single measurement falls beyond the ±4σ CL; that is, a sudden large discrepancy from the historical mean causes an alarm. On the other hand, pressure meter alarms were more often triggered by consistent small pressure shifts. Fifty-seven percent of the events detected by the pressure meters were identified by Rules 3 and 4 that required four out of five consecutive measurements to fall beyond the ±2σ WLs, and eight consecutive measurements to fall beyond the ±1σ WLs, respectively. Therefore, detectability may be improved if an SPC method that takes into account a long measurement history is applied to the pressure measurements (Jung *et al.* 2015).

Optimal meter placements

The meter location problem (Equations (3)–(6)) was solved for the Austin system for pressure meters with X_{max} ranging from 1 to 125. For pipe flow meters, a maximum of 90 pressure meter locations was available. The *RF* for the best *DP* set was computed. In addition to optimizing the *DP*, meter locations that minimized false alarm were also determined. The three results were combined in the ROC curves. A hundred random pipe bursts events were considered for the optimization. The detection and false alarm matrices had dimensions of 100 by 125 binary matrices for pressure

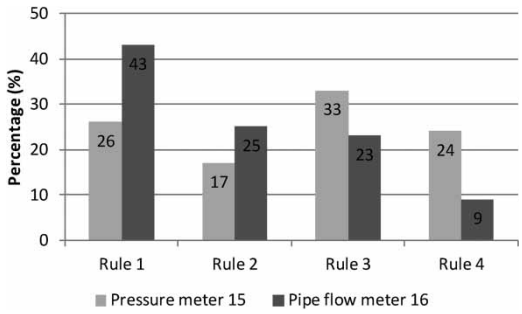


Figure 13 | Histogram of events detected by the individual WEC rules that triggered an alarm.

meters, and 100 by 90 for flow meters. The binary integer programming problem was solved by the general reduced gradient (Abadie 1970) nonlinear solver in Microsoft Excel 2007. The solver was implemented on a 3.4 GHZ Intel Core i7 Quad processor with 8 GB of memory, running Window 7. A CPU time of approximately 0.27 s was required to find an optimal solution of the problem with a predefined number of available meters.

Figures 14 and 15 show the optimal pressure and flow meter locations, respectively, for X_{max} ranging from 1 to 4. While the optimal flow meter locations were near the source (Figure 15), the best sites for pressure meters were in the distal network locations (Figure 14). For example, when only a single meter was permitted, a pressure meter was installed at node 98 within the northeast loop with a *DP* of 68% (Figure 14), while a flow meter was installed in pipe 2 with a *DP* of 63% (Figure 15). Although node 98 is an optimal location for the pressure meter, many other nodes in this system provide almost equivalent coverage. In the case of two allowable pressure meters, the meters were located in the north end of the network at nodes 58 and 65. On the other hand, one of the two flow meters was located at the source connecting pipe, while the other meter was installed within the southwestern loop. Similarly,

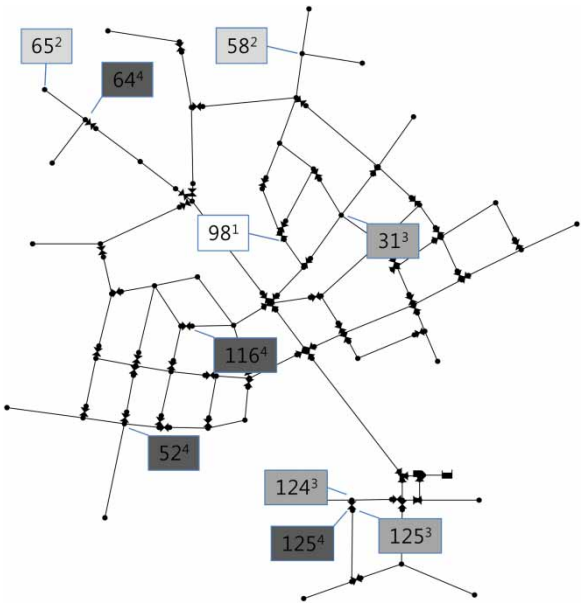


Figure 14 | Optimal pressure meter locations when one to four meters are installed and *DP* is maximized. The value in each box is the node number, with the superscript denoting the allowable number of meters in the sensor network.

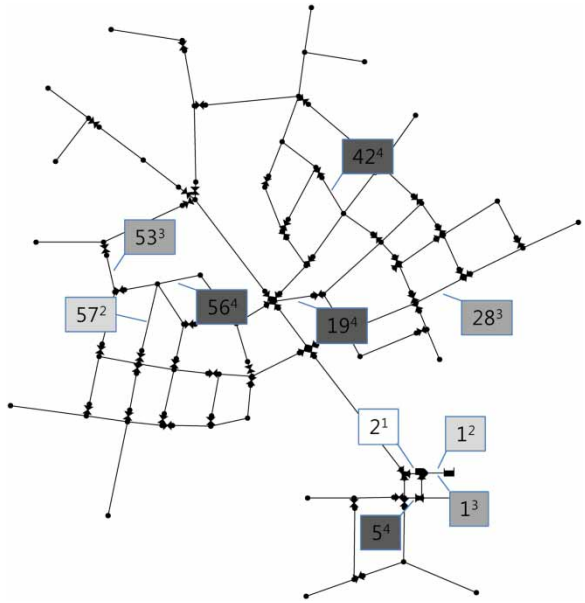


Figure 15 | Optimal flow meter locations when one to four meters are installed and *DP* is maximized. The value in each box is the pipe number, with the superscript denoting the allowable number of meters in the sensor network.

one of the four flow meters was located near the source (Figure 15). It may be noted that several optimal or near-optimal solutions are available. As more meters were used, the *ADT* decreased sharply.

The differences in the optimal meter layout are because of the hydraulic characteristics of pressure and pipe flows. The pressure meter location having the most information for pipe burst detection is at the end of the network, because the pressure will drop when any nodal demand in the system is increased because of higher upstream head losses. On the other hand, the best flow meter location is in the pipes near the source, because pipe flows are affected most by the changes in flow downstream of the meter location.

Figures 16 and 17 show the meter locations that minimize the *RF*. In contrast to the best *DP* locations, pressure meters were located near the source, and flow meters were installed in pipes in the network extremities. For example, a single flow meter's location that yields the lowest false alarm rate is pipe 91 which is connected to a dead end node. The pipe flows at the local pipes provide little information to detect a burst. This result demonstrates that maximizing the *DP* and minimizing the *RF* are contradictory goals, and the best *DP* locations are not likely to be the minimum *RF* locations. In practice, an alternative is to

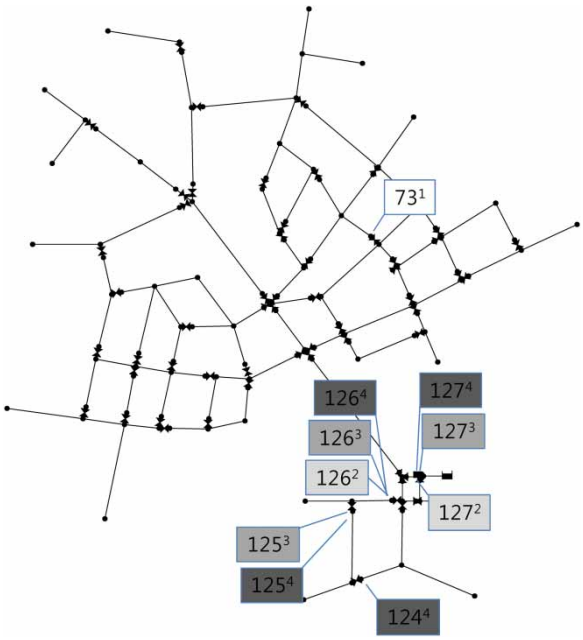


Figure 16 | Optimal pressure meter locations when one to four pressure meters are installed and *RF* is minimized.

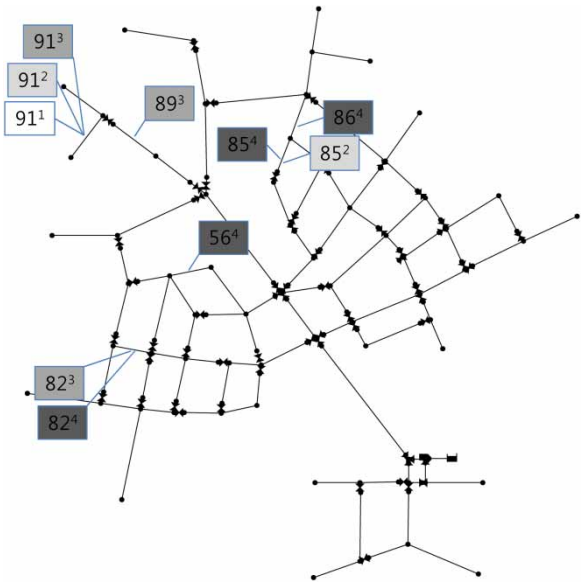


Figure 17 | Optimal flow meter locations when one to four meters are installed and *RF* is minimized.

install sensors at both the best *DP* and minimum *RF* locations, and determine an alarm considering the decisions made from the two locations.

The ROC curves provide information on the trade-off between *DP* and *RF*. As the number of meters in the sensor network increases, both *DP* and *RF* increase. Figures 18(a) and 18(b) show the ROC curves for the pressure and flow meters, respectively. The x-axes represent the percentage of nodes (Figure 18(a)) or pipes (Figure 18(b)) with the installed meters, and the y-axes represent the pipe burst detection and false alarm rates, respectively, in the two figures. For a given number of meters, the optimal meter placement model identified their best locations in the network. Pressure and flow meter sensor networks were determined independently. The *RF* computed for the optimal *DP* sensor networks and the minimum *RF* for the same number of meters are also plotted.

Both meter types show a rapid increase in *DP* when the sensor network has only a few meters. For example, the five optimally located pressure meter network (meter installation level of 4%) has a *DP* of 72%, while the best five flow meter network (meter installation level of 5.6%) has a DP of 87%. When 10% of the pipes are monitored with flow meters, the *DP* is nearly 100%, and the *DP* is virtually 100% with 37 flow meters (meter installation level of 30%) (Figure 18(b)). The highest *DP* for pressure meters is approximately 82%. Although the *DP* is higher for flow meters, the false alarm rate is also high. The *RF* for the optimal pressure meters is consistent and lies between 10 and 20%. On the other hand, increasing the number of flow meters increases the *RF* until an *RF* of approximately 96% is reached.

SUMMARY AND CONCLUSIONS

Burst detectability is related to the magnitude of the burst and proximity to a sensor. Given the ability to install meters at many locations, it is difficult to identify the best sensor locations using only engineering judgment. Hence, a general optimal meter placement model to maximize the detectability of pipe bursts has been formulated as a binary integer optimization problem, and solved for a real network with hypothetical demands. The model designs a sensor network that attempts to detect a set of randomly generated bursts.

A multi-step process is required to formulate the meter placement problem, including: (1) developing Shewhart control charts for each possible flow and pressure meter

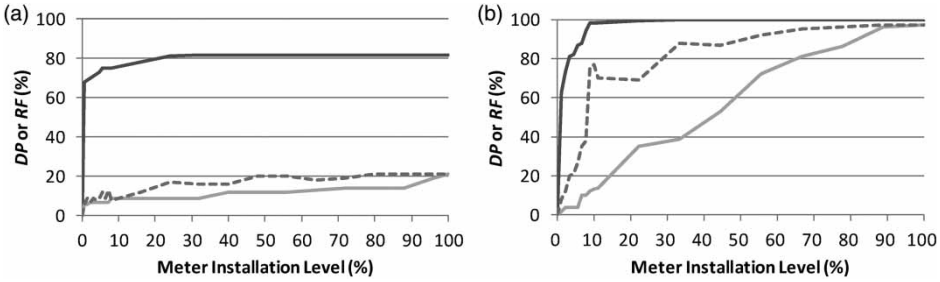


Figure 18 | ROC curves for (a) pressure meters and (b) flow meters (*DP* (solid black line), *RF* for the best *DP* meter sets (gray dashed line), and minimum false alarm rate (gray solid line)). Note that the total number of pressure meters and flow meters (i.e., meter installation level is 100%) are 125 and 90, respectively.

location using the network's hydraulic model and random nodal demands; (2) generating control and out-of-control events to test the sensor network's ability to detect anomalies; (3) applying the WEC rules to produce false alarm and detection matrices for the test events; and (4) solving the binary integer optimization problem. Two objective functions, to maximize the *DP* and minimize the *RF*, were posed while constraining the allowable number of meters. The ROC curves were evaluated for the resulting designs, and the sensitivities of *DP* and *RF* were investigated with respect to the level of available information.

Prior to solving the optimization problem, the results from the first three process steps were studied to better understand the burst detection monitoring. Pipe burst detection characteristics were investigated based on a single meter's performance. Most pressure meters had a *DP* of approximately 70%, which is much higher than the overall *DP* of the flow meters. On the other hand, the *RF* was less than 10% regardless of the type of meter. Many pressure meters can detect a large set of the same bursts, while each flow meter detects fewer events that are likely to be more localized. Finally, most pipe bursts were detected at night and early morning because of low demand variability during these periods.

The optimal sensor networks were distinctly different depending on whether pressure or flow meters were installed. While maximizing the *DP*, the pressure meters were located at distal locations relative to the water source, while the pipe flow meters were installed near the source. The meter locations that yield minimum false alarm were also identified. The results were opposite to those corresponding to the best *DP* locations. The optimum location for pressure meters is near the source, while the optimum location is near the end of the network for pipe flow meters. It appears that the downstream nodes are impacted more by bursts due to higher head losses that accumulate through the network during a burst. Conversely, the flow at the upstream pipes is more sensitive to the downstream anomalies. The ROC curves for the alternative pressure and pipe flow meters showed that using many pipe flow meters could detect all the pipe bursts when the maximum *DP* with the pressure meters was 82%. As flow meters are vulnerable to false alarms, using pressure meters is likely to be more valuable for pipe burst detection in the modified Austin network.

The proposed sensor placement model employed the WEC rules which are the easiest to implement, as there are no parameters to be estimated. However, other SPC methods with higher detection effectiveness, such as CUSUM and EWMA can replace the WEC rules for improving detectability, although the system-specific parameters should be estimated by paying high computational costs.

This study has several limitations that future research must address. First, as in most sensor network designs, the layout is based on a limited set of conditions. Here, we have defined a set of bursts; the final design may be biased by this set, and some study should be conducted to determine how large a set is needed to converge to a consistent design. Multiple simultaneous bursts can be generated and provided for burst detection. Second, this study determined the optimal locations for pressure and flow meters independently. The optimal combinations of pressure and flow meters and their locations can be determined with a small extension of the posed model. Along this line, meter costs and a budget constraint can be included in the optimization model, instead of only limiting the number of meters. Third,

as confirmed from this study, a trade-off exists between *DP* and *RF*. Multi-objective optimal meter placement problems can be formulated to minimize total meter cost and *RF* and maximize *DP*. This study and the previous studies focused on burst occurrence detection. However, since all the pressure meters detect the same set of bursts, they provide little information on where the burst is located. To provide a truly effective monitoring system, it is necessary to simultaneously consider the possibility of detecting and locating bursts, while minimizing false alarms and detection time. Finally, the proposed approach can be applied to detect real-life bursts by including additional preprocessing steps such as raw measurement filtering.

ACKNOWLEDGEMENTS

This material is based in part upon the work supported by the National Science Foundation under Grant No. 0835930. The opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

Abadie, J. 1970 Application of the GRG method to optimal control problems. In: *Integer and Nonlinear Programming* (J. Abadie, ed.). North-Holland Publishing Company, Amsterdam, North Holland, The Netherlands, pp. 191–211.

Alkasseh, J., Adlan, M., Abustan, I., Aziz, H. & Hanif, A. 2013 Applying minimum night flow to estimate water loss using statistical modeling: a case study in Kinta Valley, Malaysia. *Water Resour. Manage.* **27** (5), 1439–1455.

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

Brion, L. & Mays, L. 1991 Methodology for optimal operation of pumping stations in water distribution systems. *J. Hydraul. Eng.* **117** (11), 1551–1569.

Egan, J. P. 1975 *Signal Detection Theory and ROC Analysis*. Series in Cognition and Perception, Academic Press, New York, USA.

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

Farley, B., Mounce, S. & Boxall, J. 2013 Development and field validation of a burst localization methodology. *J. Water Resour. Plann. Manage.* **139** (6), 604–613.

Fuchs-Hanusch, D., Steffelbauer, D., Gunther, M. & Muschalla, D. 2015 Systematic material and crack type specific pipe burst outflow simulations by means of EPANET2. *Urban Water J.* **13** (2), 108–118.

Greyvenstein, B. & van Zyl, J. E. 2007 An experimental investigation into the pressure–leakage relationship of some failed water pipes. *J. Water Supply Res. Technol.–AQUA* **56** (2), 117–124. doi:10.2166/aqua.2007.065.

Huang, H., Tao, T. & Xin, K. 2012 Optimal pressure meters placement for bursts detection based on SOM. In: *Proceedings of the 14th Water Distribution Systems Analysis Conference*, Adelaide, Australia.

Jung, D., Kang, D., Liu, J. & Lansey, K. 2015 Improving the rapidity of responses to pipe burst in water distribution systems: a comparison of statistical process control methods. *J. Hydroinform.* **17** (2), 307–328.

Kang, D. & Lansey, K. 2009 Real-time demand estimation and confidence limit analysis for water distribution systems. *J. Hydraul. Eng.* **135** (10), 825–837.

Kang, D. & Lansey, K. 2010 Optimal meter placement for water distribution system state estimation. *J. Water Resour. Plann. Manage.* **136** (3), 337–347.

Lambert, A. 2001 What do we know about pressure-leakage relationships in distribution systems. In: *Proceedings of IWA Conference on Systems Approach to Leakage Control and Water Distribution System Management*, Brno, Czech Republic.

May, J. 1994 Pressure dependent leakage. World Water and Environmental Engineering, October, pp. 1–10.

Misiunas, D., Lambert, M., Simpson, A. & Olsson, G. 2005 Burst detection and location in water distribution networks. *Water Sci. Technol. Water Supply* **5** (3–4), 71–80.

Misiunas, D., Vítkovský, J., Olsson, G., Lambert, M. & Simpson, A. 2006 Failure monitoring in water distribution networks. *Water Sci. Technol.* **53** (4–5), 503–511.

Montgomery, D. 2009 *Statistical Quality Control: A Modern Introduction*, 6th edn. John Wiley & Sons, New York, USA.

Morrison, J. 2004 Managing leakage by district metered areas: a practical approach. *Water21*, 44–46.

Mounce, S. R. & Machell, J. 2006 Burst detection using hydraulic data from water distribution systems with artificial neural networks. *Urban Water J.* **3** (1), 21–31.

Mounce, S., Day, A., Wood, A., Khan, A., Widdop, P. & Machell, J. 2002 A neural network approach to burst detection. *Water Sci. Technol.* **45** (4–5), 237–246.

Mounce, S., Khan, A., Wood, A., Day, A., Widdop, P. & Machell, J. 2003 Sensor-fusion of hydraulic data for burst detection and location in a treated water distribution systems. *Information Fusion* **4**, 217–229.

Mounce, S. R., Boxall, J. B. & Machell, J. 2010 Development and verification of an online artificial intelligence system for detection of bursts and other abnormal flows. *J. Water Resour. Plann. Manage.* **136** (3), 309–318.

Palau, C. V., Arregui, F. J. & Carlos, M. 2012 Burst detection in water networks using principal component analysis. *J. Water Resour. Plann. Manage.* **138** (3), 47–54.

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

Poulakis, Z., Valougeorgis, D. & Papadimitriou, C. 2003 Leakage detection in water pipe networks using a Bayesian probabilistic framework. *Probabilist. Eng. Mech.* **18**, 315–327.

Quevedo, J., Puig, V., Cembrano, G., Blanch, J., Aguilar, J., Saporta, D., Benito, G., Hedo, M. & Molina, A. 2010 Validation and reconstruction of flow meter data in the Barcelona water distribution network. *Control Eng. Pract.* **18** (6), 640–651.

Romano, M., Kapelan, Z. & Savic, D. A. 2010 Real-time leak detection in water distribution systems. In: *Proceedings of Water Distribution System Analysis 2010*, Tucson, AZ, USA.

Romano, M., Kapelan, Z. & Savic, D. A. 2014 Automated detection of pipe bursts and other events in water distribution systems. *J. Water Resour. Plann. Manage.* **140** (4), 457–467.

Rossman, L. 2000 EPANet2 User's Manual, US Environmental Protection Agency, Washington, DC, USA.

Sturm, R. & Thornton, J. 2005 Proactive leakage management using district metered areas (DMA) and pressure management – is it applicable in North America? In: *Proceedings of the Leakage IWA Conference*, Halifax, Canada.

Surendran, S., Tanyimboh, T. & Tabesh, M. 2005 Peaking demand factor-based reliability analysis of water distribution systems. *Adv. Eng. Softw.* **36**, 789–796.

Walski, T. 2012 Planning-level capital cost estimates for pumping. *J. Water Resour. Plann. Manage.* **138** (3), 307–310.

Walski, T., Whitman, B., Baron, M. & Gerloff, F. 2009 Pressure vs. flow relationship for pipe leaks. In: *Proceedings of the World Environmental and Water Resources Congress 2009*, American Society of Civil Engineers. Kansas City, MO, USA, 1–10, doi:10.1061/41036(342)10.

Western Electric Company 1958 *Statistical Quality Control Handbook*, 2nd edn. AT&T Technologies, Indianapolis, IN, USA.

Wu, Z. & Song, Y. 2012 Optimizing pressure logger placement for leakage detection and model calibration. In: *Proceedings of the 14th Water Distribution Systems Analysis Conference*, Adelaide, Australia.

Ye, G. & Fenner, R. A. 2011 Kalman filtering of hydraulic measurements for burst detection in water distribution systems. *J. Pipeline Syst. Eng. Pract.* **2** (1), 14–22.

Ye, G. & Fenner, R. 2013 Weighted least squares with EM Algorithm for burst detection in UK water distribution systems. *J. Water Resour. Plann. Manage.* **140** (4), 417–424.

First received 14 August 2015; accepted in revised form 4 January 2016. Available online 18 February 2016