Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering00 (2014) 000–000

Procedia Engineering

www.elsevier.com/locate/procedia

16th Conference on Water Distribution System Analysis, WDSA 2014

Online burst detection in a water distribution system using the Kalman filter and hydraulic modelling

Isaac Okeya[a], Zoran Kapelan[a], Christopher J. Hutton[a], Devina Naga[b]

[a]*Centre for Water Systems, University of Exeter, North Park Road, Exeter,EX4 4QJ, UK*
[b]*United Utilities, Lingley Green Avenue, Great Sankey, Warrington, WA5 3LP*

Abstract

This paper presents a burst detection methodology that utilises distributed real time sensor data in a district metered area using a data assimilation method and a hydraulic model. A sensitivity analysis was applied to evaluate the performance of various burst detection metrics under different conditions and to identify appropriate thresholds for online burst detection using artificial generated burst events. It was found that the best performing burst detection metric is the normalised corrected flow residual. This burst detection metric can be effective to detect bursts in a timely and reliable manner within a district metering area under assumed test conditions.

Peer-review under responsibility of the Organizing Committee of WDSA 2014.

*Keywords:*Burst Detection Metric, District Metred Area, Kalman Filter, Water Distribution System, Water Demand Forecasting Model

1. Introduction

Water is one of the most important resources in both developed and developing countries. Due to the growing concern of rapid climate change and increasing urban population, there is a need to reduce the loss of treated water from Water Distribution System (WDS) to the environment. Water loss is mainly caused by pipe bursts and high leakages due to aging and deteriorating assets of WDS [1] . This water loss damages the environment including nearby infrastructures [1], causes service disruption to the customer and increases unnecessary energy cost and carbon footprint. Therefore, the United Kingdom water utilities aspire to use a combination of cost-effective technologies that can reduce water loss via early detection of the abnormal flows [2].

Many researchers attempt to use a state estimation approach to evaluate the current state of the WDS section for diagnosing and identifying abnormal flows. Such an approach can be affected by erroneous WDS observations, disaggregation of nodal demand and limited number of hydraulic sensors in the WDS due to financial constraints.

Peer-review under responsibility of the Organizing Committee of WDSA 2014.

Detection of abnormal flows with a WDS hydraulic model is mostly based on either the residual analysis or WDS state estimates pattern evaluation [3], [4]. The residual analysis makes use of the difference between the online WDS hydraulic model predictions and the corresponding WDS observations [3]. The WDS state estimates' pattern evaluation compares the existing pattern to the expected or past pattern of flow or pressure data [4].

Artificial Intelligence (AI) based techniques with a hydraulic model such as the combination of Artificial Neural Network (ANN) and fuzzy logic [2] and Support Vector Machine (SVM) [5] to evaluate the WDS state estimates' patterns for abnormal event detection and identification have shown some promising results. However, a large amount of data is required to train either an ANN or a SVM based method to detect abnormal flows. Jung and Lansey (2013) explain how the Extended Kalman filter (EKF) can be used to detect a pipe burst online when an operational has change occurred. Compared to the conventional method of using a combination of statistical analysis and AI tools, the EKF method has the benefit of utilising small quantities of calibration data and lowering the computational time to detect pipe bursts in near real-time [4], [7]. The application of using the EKF with a hydraulic model for burst detection is not reliably established since it is only applied on a trunk main with synthetic flow data. Only two burst detection metrics have been used in the literature to identify abnormal flow events with Kalman filtering procedure, 1) standardised innovation sequence [4] and 2) flow residual between the observed and corrected flow data [3].

The objective of this paper is to compare the performance of the various burst detection metrics relating to the Kalman Filtering procedure. This paper presents a burst detection methodology that makes use of a water demand forecasting model, a WDS hydraulic model and a modified Kalman filter method to detect pipe bursts with distributed noisy flow data. The report is organised as follows: section 2 explains the burst detection methodology, burst simulation, burst detection metrics and its sensitivity analysis; section 3 describes the case study area and the data used; section 4 discusses the results obtained and section 5 summaries the report and provides the concluding remarks.

2. Methodology

The goal of the burst detection methodology is to detect any pipe bursts in near real-time utilising various burst detection metrics by linking distributed noisy flow observations with a recently calibrated WDS hydraulic model that solves the nonlinear hydraulic relationships. The burst detection methodology comprises three major components: Water Demand Forecasting Model (WDFM), a Hydraulic Model, and the Kalman Filter (KF). The detection methodology proceeds as follows:

1. The historic inflows and outflows from the DMA boundary are used to derive the historical DMA demands.
2. The WDFM for the DMA is developed based on the historic DMA demands.
3. The WDFM forecasts demands using the historic demands,
4. The forecasted demands are then used to drive the hydraulic model and derive predicted hydraulic state estimates (i.e. predicted flow rates and pressure heads) corresponding to forecasted demand.
5. The KF is then applied to calculate the corrected demands at the current time step driven by the difference between the predicted hydraulic state estimates and the corresponding WDS observations.
6. The corrected demands are inserted into the hydraulic model to obtain the corrected hydraulic state estimates.
7. The burst detection metrics (see section 2.4) are calculated and converted into a dimensionless burst detection metric value.
8. Compare the dimensionless burst detection metric value against the burst threshold (explained in section 2.5); if and when a burst is detected, a burst alarm is raised.

The methodological steps 3 – 8 are repeated at subsequent time steps for the application of online burst detection in a WDS.

2.1. Water Demand Forecasting Model

The objective of the WDFM is to forecast demands every 15 minutes with a 15 minute lead time. The forecast demand is the function of the historical demands from the previous time steps. These demands are estimated using the DMA boundary flow meters (i.e. historic DMA inflows and outflows). The WDFM is calibrated offline using a historical time series of demands. The WDFM is based on Multi-Linear Regression (MLR) approach which requires associated weight coefficients to determine the WDFM outputs. The cross-correlation between the demand at the time step, *t-n (n =0 and n is the number of time steps)* and demand from previous time steps, *t-n (n = 1, 2, 3*) was performed from times *t-n* (current time step) to *t-n-2016* (3 weeks maximum lag) at the intervals of 15minutes. Based on the cross-correlation results, the WDFM was developed using a combination of demands from time step, *t-1*, *t-2* and *t-672*. The Maximum Likelihood Estimation (MLE) is used to derive the weight coefficients. The WDFM use previous demands at DMA level as input to forecast demand and it can be written as:

$$d_t = w_{t-1} d_{t-1} + w_{t-2} d_{t-2} + w_{t-672} d_{t-672} \tag{1}$$

Where d_t is the forecasted demand at the time step, t; d_{t-1} is the corrected demand from a previous time step (15minutes ago), $d_{t-2}$is the corrected demand from 30 minutes ago; d_{t-672} is the corrected demand from a week ago; w_{t-1}, w_{t-2} and $w_{t-672}$are the associated weight for d_{t-1}, d_{t-2}andd_{t-672} respectively.

2.2. Kalman Filter

The Kalman Filter [5]is a recursive data processing algorithm for the linear stochastic dynamic system. The KF uses the combination of observed hydraulic states and hydraulic model predictions of WDS along with the associated uncertainties to obtain the corrected hydraulic states. The hydraulic model predictions of the WDS are obtained by simulating the hydraulic model based on the forecast demands. The relationship between system observations and hydraulic model predictions can be modelled by:

$$z_t = h\left(d_t\right) + e_t \tag{2}$$

where z_t is the system observations (i.e. flow rates, pressure heads) vector at time step, t; $h(.)$is the nonlinear network hydraulic model of WDS and e_t is the observations error vector.

The residual between the system observations and the hydraulic model predictions are used to correct the forecast demands and attain posterior error covariance matrix:

$$d_t^c = d_t + HK_t\left(z_t - h\left(d_t\right)\right) \tag{3}$$

$$P_t^c = \left(I - K_t\right)P_t \tag{4}$$

where K_t is the Kalman gain matrix at time step, t; H is the observation operator that map from flow space to demand space; I is the identity matrix; P_t is the prior error covariance matrix; d_t^c is the corrected demand at time step, t and P_t^c is the posterior error covariance at time step, *t*.

The Kalman gain is the weight factor based on the prior and observation error covariance matrix:

$$K_t = P_t / \left[P_t + R_t\right] \tag{5}$$

and

$$P_t = \left[P_{t-1}^c + P_t^Q \right] \tag{6}$$

$$P_t^Q = \left[qH\left(h\left(d_t\right)\right)\right]^2 \tag{7}$$

$$R_t = \left[rH\left(z_t\right)\right]^2 \tag{8}$$

where R_t is the observation error covariance matrix; q and r are the average percent error of the hydraulic model predictions and the system observations respectively and P_t^Q is the process error covariance.

2.3. Burst Modelling

The bursts are modelled as pressure-dependent emitter flows [6] located at system nodes:

$$q_{i,t} = C\, p_{i,i}^e \tag{9}$$

where $q_{i,t}$ is the burst flow at node, i at time step, t; C is the emitter discharge coefficient; $p^e_{i,I}$ is the nodal pressure at node, i at time step, t; and e is the emitter pressure exponent.

2.4. Burst Detection Metrics

The following burst detection metrics are used for the burst detection analysis here:

- Flow/pressure residual
- Normalised flow/pressure residual
- Moving Average (MA) of flow/pressure residual

The flow/pressure residual (equation 10) is the difference between the observed flow rate/pressure and predicted flow rate/pressure (from the hydraulic model). The flow/pressure residual and the normalised flow/pressure residual can be calculated as:

$$PQ_t^r = PQ_t^o - PQ_t^c \tag{10}$$

$$nPQ_t^r = PQ_t^r / PQ_t^o \tag{11}$$

where PQ_t^o is the observed flow rate/pressure at time step, t; PQ_t^c is the predicted flow rate at time step, t (via the hydraulic model); PQ_t^r is the current flow residual and nPQ_t^r is the normalised flow residual at the time step, t.

MA (equation 12) of flow/pressure residual is the mean of the previous n flow/pressure residuals and n is the number of time steps selected.. The flow/pressure residual is written as:

$$MA_t = \left[PQ_t^r + PQ_{t-1}^r + \ldots + PQ_{t-(n-1)}^r \right] \Big/ n \tag{12}$$

where PQ_{t-1}^r is the flow/pressure residual from the previous time step, t; MA_t is the moving average of flow/pressure residuals at the time step, t.

All the burst detection metrics' values are converted into a dimensionless burst detection metric value by using the one of the following equations (equation 13 and 14):

$$NMV_{i,i} = \left[MV_{i,i} - MV_i^{avg} \right] \Big/ \left[MV_i^{\max} - MV_i^{avg} \right] \; if \; MV_i > MV_i^{avg} \tag{13}$$

$$NMV_{i,i} = \left[MV_i^{avg} - MV_{i,i} \right] \Big/ \left[MV_i^{avg} - MV_i^{\min} \right] \; if \; MV_i < MV_i^{avg} \tag{14}$$

where $NMV_{i,t}$ is the normalised burst detection metric value at the hydraulic sensor, i at time step, t; $MV_{i,t}$ is the burst detection metric value at the hydraulic sensor, i at time step, t; m is the total number of burst detection metric values used; $MV_i^{\max}$ is the maximum average value of the highest 5% of the burst detection metric values for hydraulic sensor, i; MV_i^{avg} is the average value of all the burst detection metric values for hydraulic sensor, i and $MV_i^{\min}$ is the minimum average value of the lowest 5% of the burst metric values for hydraulic sensor, i.

The averages of the burst detection metric values for the individual hydraulic sensors in step 2 must be obtained offline before the burst method is applied online. Equation 13/14 is used at each time step to obtain the dimensionless burst detection metric value at the methodological step 7 of the burst detection method. The performance of the burst detection metrics depends on the number of detected bursts; false alarms raised and undetected bursts.

2.5. Sensitivity Analysis and Detection Thresholds

In this paper, a sensitivity analysis is performed to develop detection thresholds for different bursts detection metrics. The steps to conduct the sensitivity analysis of the burst detection metric are as follows:

Step 1: The step 3-6 of the burst detection method is repeated at the subsequent time step offline to generate burst detection metric values for various normal conditions due to the uncertainties in the DMA demand and nodal base demand.

Step 2: All the burst detection metric values obtained under various normal conditions are then sorted in ascending order to estimate the burst detection metric values' average and the average of the minimum, maximum burst detection metric values for the individual hydraulic sensors.

Step 3: The burst detection metric value is then normalised into a dimensionless burst detection metric value using the one of the following equations (equation 13 and 14).

Step 4: The average of dimensionless burst detection metric value is estimated and the dimensionless burst detection metric value that are above their respective dimensionless burst detection metric value average is considered as a burst detection metric outlier. However, the average of the flow-based dimensionless burst detection metric is estimated with the exclusion of values that are below 1 to prevent unusual high number of burst detection metric outliers.

*Step 5:*The number of burst detection metric outliers in the last n time steps (n is the number of selected time steps , i.e. 4 time steps (1hour), 5 time steps (1.25hours)) is then estimated and used to suggest the minimum number of dimensionless burst detection metric outliers for the individual previous k time steps for burst detection.

3. Case Study

The case study area is a real-life District Metered Area (DMA) which renamed as DMA001 and it is located in North-West of England (see Figure 1). The DMA001 supplies water to approximately 1145 customers under gravity with an average daily demand of 9l/s. The DMA001 model consists of 305nodes, 238 pipes and 87 valves.

The difference between the real-life inflows and outflows data of the DMA001at intervals of 15 minutes between 16th September 2013 and 24thNovember 2013 are the DMA net flows. The net flow is considered as the DMA demand which is divided by the total base demand to obtain the demand coefficients. The base demand is the allocated demand for the catchment area of the node. The product of the demand coefficients and base demand at the individual nodes represents the amount of water required at the node. The DMA demand data obtained for the time period between 16th September 2013 and 20th October 2013 (5 weeks) are used to develop and calibrate the WDFM offline. The calibrated parameters are the weight coefficients used in the WDFM (see equation 1).

The hydraulic model of the DMA is calibrated offline by using the demand and pressure data between 30th Sept and 13th Oct 2013. There is one pressure sensor within the DMA and it is located at the highest elevation point in the DMA. All the plastic pipes' roughness and the base demand remain unchanged. The hydraulic model calibration parameters (adjusted using the trial and error technique) are: 1) the roughness value of the cast iron and ductile iron pipes; 2) the relative opening (i.e. tau value) of the Throttle Control Valve (TCV) located at the DMA inlet was increased; 3) the demand coefficients at 15minutes intervals were modified so that predicted flow rates match the real-life inlet and outlet flow rates. All these modifications were made to ensure the pattern of the predicted pressure heads closely match the observed pressure heads' pattern. Due to the limited number of flow and pressure observations used for the DMA001 hydraulic model calibration, it is rational to say that the DMA001 hydraulic model is probably not calibrated offline sufficiently well.

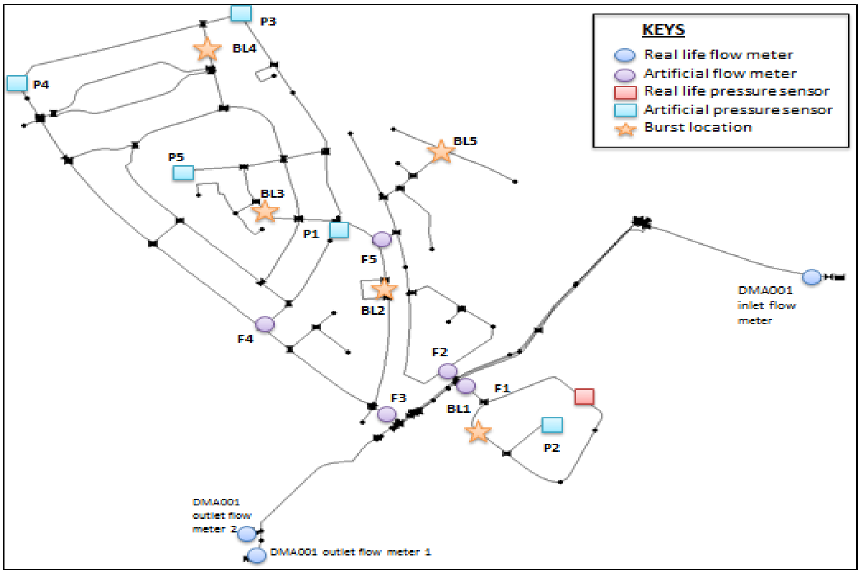

Fig.1: The DMA001 hydraulic model and the location of hydraulic sensors

The demand coefficients between 21st October and 18th November 2013 are inserted into the hydraulic model to obtain synthetic flow and pressure observations within the DMA. The burst detection method was run without methodological step 7 and 8 between 21th October and 17th November 2013 (4 weeks) offline toto obtain burst metric values under various normal noises. . The sensitivity analysis is performed on the burst detection metric value to develop thresholds for the individual bursts detection metrics. The synthetic flow and pressure observations on the 18th November 2013 are used for burst detection analysis.

The bursts are synthetically generated on the 18th Nov 2013 based on equation 12 using a hydraulic simulation model. The emitter exponent of 0.5 in the hydraulic model is applied [7]. Various emitter discharge coefficients have been tested to find the acceptable emitter discharge coefficient to represent the 5 burst flows. The 5 burst flows used in this case study are approximately 5%, 10%, 20%, 30% and 50% of the daily average DMA demand. The bursts are simulated at 3 different time periods (morning peak (6.30am – 10.30am), evening peak (4.30pm – 8.30pm) and night (1.30am – 5.30am)). All the bursts are assumed to last for 4 hours. Bursts are simulated at 5 different burst locations (see Figure 3). The 'perfect' flow/pressure observations obtained from the corresponding hydraulic model outputs were altered by adding random noise and/or systematic errors, as shown in Table 1.

Table 1: The type of flow observation noises

Case No	Type of error	Error
1	Random noise	$N(0, 0.005Q_{i,t})$
2	Random noise	$N(0, 0.01Q_{i,t})$
3	Systematic errors and random noise	$+0.01Q_{i,avg}$ and $N(0, 0.01Q_{i,t})$
4	Systematic errors and random noise	$-0.01Q_{i,avg}$ and $N(0, 0.01Q_{i,t})$

$Q_{i,t}$ represents the flow rate at the flow meter, i at time step, t; and $Q_{i,avg}$ is the average flow rate at flow meter, i at the time step, t.

The location of the 5 synthetic flow meters and pressure sensors (refer to figure 1) are selected based on a modified methodology of optimal sensor placement for event detection [8].

4. Results and Discussion

The performance of the dimensionless burst detection metric's thresholds is tested on the different type of bursts on 18th November 2013. The summary of all the burst detection metrics' performance are shown in table 2 and 3.The MA value of the residuals in this paper is the average of the residual within window size of 8 time steps (2hours). The window size of 8 time steps is selected because the pipe bursts are noticeable with a window size between 4 time steps (1hour) and 16 time steps (4hours). Table 2 and 3 showed the flow residuals-based burst detection metrics seem to detect more bursts when compared to the burst detection metrics based on pressure residuals, probably the pressure-residual burst detection metrics are not sensitive to low-medium burst magnitude (e.g. 5% - 20% of average daily DMA demand). However, the pressure residual- based burst detection metrics tend to have a lower number of false alarms raised compared to the flow residuals-based burst detection metrics due to low sensitivity to demand change

Table 2: The overall burst detection rates

Dimensionless burst detection metric	Detected bursts	Undetected bursts	False alarms
Corrected flow residual	319 (85%)	56 (15%)	19 (5%)
Normalised corrected flow residual	324 (87%)	51 (13%)	17 (5%)
MA of corrected flow residuals	319 (85%)	56 (15%)	21 (6%)
Corrected pressure residual	159 (42%)	216 (58%)	10 (3%)
Normalised corrected pressure residual	181 (48%)	194 (52%)	7 (3%)
MA of corrected pressure residuals	175 (47%)	200 (53%)	7 (3%)

Table 3: The summary of detection rates for different burst magnitudes

Burst magnitude	5% of average daily DMA demand)		10% of average daily DMA demand)		20% of average daily DMA demand)		30% of average daily DMA demand)		50% of average daily DMA demand)	
Dimensionless burst detection metric	detected bursts	undetected bursts	detected bursts	undetected bursts	detected bursts	undetected bursts	detected bursts	undetected bursts	detected bursts	undetected bursts
Corrected flow residual	34 (45%)	41 (55%)	65 (87%)	10 (13%)	65 (87%)	10 (13%)	75 (100%)	0 (0%)	75 (100%)	0 (0%)
Normalised corrected flow residual	34 (45%)	41 (55%)	65 (87%)	10 (13%)	75 (87%)	10 (13%)	75 (100%)	0 (0%)	75 (100%)	0 (0%)
MA of corrected flow residuals	34 (45%)	41 (55%)	65 (87%)	10 (13%)	65 (87%)	10 (13%)	75 (100%)	0 (0%)	75 (100%)	0 (0%)
Corrected pressure residual	0 (0%)	75 (100%)	0 (0%)	75 (100%)	40 (53%)	35 (47%)	56 (75%)	19 (25%)	68 (91%)	7 (9%)
Normalised corrected pressure residual	0 (0%)	75 (100%)	0 (0%)	75 (100%)	40 (53%)	35 (47%)	66 (88%)	9 (12%)	75 (100%)	0 (0%)
MA of corrected pressure residuals	0 (0%)	75 (100%)	0 (0%)	75 (100%)	41 (55%)	34 (45%)	59 (79%)	16 (21%)	60 (80%)	15 (20%)

As it can be seen from Table 3, the detection rate of each burst detection metric improves as the burst magnitude increases and all the flow residual-based burst detection metrics have the same detection rate. However, the MA of corrected flow residual have the highest number of false alarms raised due to the accumulation of high corrected flow residuals caused by the sudden increase in flow observations that are not related to the burst.

Based on the detection rates reported in Table 2, it seems that the normalised corrected flow residual is the best performing burst detection metric because of the highest detection rate and the lowest false alarm rate. Generalising this observation, however, requires future work on additional case studies.

5. Summary, Conclusion Recommendations for Further Work.

This paper has presented a burst detection methodology based on the Kalman filtering of synthetic flow observations and predictions (from the hydraulic model). The KF corrects the forecast demand of the DMA from the WDFM using the flow residual between the synthetic flow observations and model predictions. The proposed sensitivity analysis of burst detection metric is utilised to develop the threshold(s) for the individual dimensionless burst detection metrics.

The methodology makes use of different burst detection metrics to detect pipe bursts online. Each of the burst detection metrics are analysed and compared to assess their capabilities. The best performing burst detection metric is the normalised corrected flow residual. However, the flow residuals-based burst detection metric tends to have also the slightly higher false alarm rates when compared to the pressure residuals-based burst detection metrics.

Further research will analyse different number/location combination of hydraulic sensors comparison to the existing detection method(s) to improve the robustness of the current burst detection methodology. The effectiveness of using different combinations of burst detection metrics will be analysed too. The burst detection methods will be applied on real-life bursts/measurement data. The relationship between the location of the burst location and the hydraulic sensors will also be studied for burst location approximation.

Acknowledgements

The authors are grateful to Engineering and Physical Sciences Research Council (EPSRC) and United Utilities (UU) including Mr D. Clucas, Mr T. Allen and UU hydraulic modelling team for providing the case study data and supporting financially the STREAM EngD project.

References

[1] R. Puust, Z. Kapelan, D. Savic and T. koppel, “A review of methods for leakage managment in pipe networks,” *Urban Water,* vol. 7, pp. 25-45, 2010.

[2] M. Romano, Z. Kapelan and D. Savic, “Geostatistical techniques for approximate location of pipe burst events in water distribution systems,” *Hydroinformatics,* pp. 634-651, 2013.

[3] T. C. Arsene, B. Gabrys and D. Al-Dabass, “Decision Support System for Water Distribution System based on neural networks and graphs theory for leakage detection,” *Expert Systems with Applications,* vol. 39, pp. 13214-13224, 2012.

[4] D. Jung and K. Lansey, “Burst Detection in Water Distribution System using the Extended Kalman Filter,” Bari, Italy, 2013.

[5] J. Mashford, V. Highett, D. De Silva, D. Marney and S. Burn, “An Approach to Leak Detection in Pipe Networks Using Analysis of Monitored Pressure Values by Support Vector Machine,” *Network and System Security,* pp. 534 - 539, 2009.

[6] G. Ye and R. Fenner, “ Kalman filtering of hydraulic measurements for burst detection in water distribution systems,” *Pipeline Systems Engineering and Practice (ASCE),* pp. 14-22, 2011.

[7] C. J. Hutton, Z. Zoran Kapelan, L. Vamvakeridou-Lyroudia and D. Savić, “Dealing with Uncertainty in Water Distribution System Models: A Framework for Real-Time Modeling and Data Assimilation,” *Water Resources Planning and Management,* pp. 169-183, 2014.

[8] R. Kalman, “A New Approach to Linear Filtering and Prediction Problems,” *Journal of Basic Engineering, ASME,* vol. 82, pp. 35-45, 1960.

[9] Z. Y. Wu, T. Walski, R. Mankowski, G. G. R. Herrin and M. Tryby, “Calibrating Water Distribution Models via Genetic Algorithms,” Kansas City, 2002.

[10] M. O. Lambert, “International Report: Water losses management techniques,” Water Supply, 2002.

[11] B. Farley, S. R. mounce and J. B. Boxall, “Field testing of an optimal sensor placement methodology in an urban water distribution network for event detection,” *Urban Water,* vol. 7, no. 6, pp. 345-356, 2010.