

The combined usage of the hydraulic model calibration residual and an improved vectorial angle method for solving the BattLeDIM problem

Lefeng Huang¹, Kun Du¹, Miaoting Guan², Qi Wang²

¹Faculty of Civil Engineering and Mechanics, Kunming University of Science and Technology, Kunming 650500, China

²School of Civil and Transportation Engineering, Guangdong University of Technology, Guangzhou 510006, China

¹dukun_cq@foxmail.com

ABSTRACT

We propose an integrated method that combines the model-based and data-driven techniques to deal with the leakage detection and location challenge in the L-Town network. The method consists of five stages, including (1) model decomposition, (2) partition of SCADA data, (3) nodal demand calibration, (4) calibration residual-based leakage detection, and (5) an improved vectorial angle method based leakage localization.

In the first stage, the L-Town network was decomposed into subsystems according to the three DMAs existing in the system (see Figure 1). For each subsystem, the pressure at its inlet can be simulated and controlled through the pressure reduction valve (PRV) installed. As such, the entire L-Town network was divided into three smaller, independent systems with known interactions, which significantly simplified the leakage detection and localization problem. Below, only the Area A is used to demonstrate how the proposed method can solve the BattLeDIM problem.

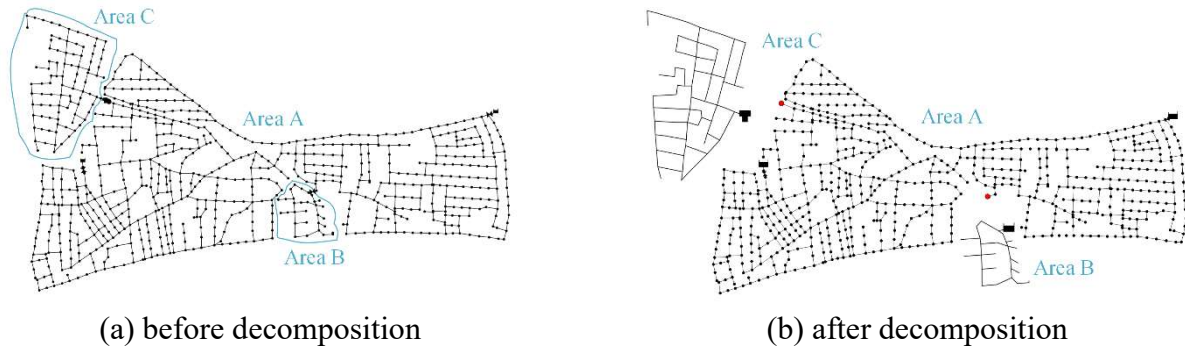


Figure 1. Model decomposition of the L-Town benchmark network (three DMAs denoted as Areas A, B, and C, respectively)

In the second stage, the partition of SCADA data is to alleviate the impact of demand uncertainties on leakage detection and localization accuracy. Nodal demands are usually affected by factors such as weather, urban functional arrangements, and users' consumption profiles, showing cyclical patterns. Such data play a crucial role in the accuracy of a hydraulic model. Therefore, it is necessary to calibrate nodal demands before implementing leakage detection. The base demand at each junction has three categories, including residential, commercial, and industrial, and is repeated weekly according to the Epanet input file.

For each type of other datasets (e.g., pressure and flow), the time series data collected from the utility's SCADA system in 2018 were reorganized into 52 groups (considering only 364 days), each containing a week's operational recordings (see Equation 1). This is to prepare the data feed for nodal demand calibration. Consequently, each original dataset of 104,832 entries was transformed to 52 matrices of 2016-by-31. The row number (i.e., 2016) denotes the amount of data collected in a

¹ Corresponding author: Dr. Kun Du

week at a 5-minute time step (i.e., $60 \div 5 \times 24 \times 7 = 2016$); while the column number (i.e., 31) refers to the number of sensors deployed in Area A, including 29 pressure sensors and two flow meters at the pipes directly connected to 2 reservoirs.

$$[(\mathbf{H}_1, \dots, \mathbf{H}_{29}), (\mathbf{q}_1, \mathbf{q}_2)] \Rightarrow \begin{bmatrix} \begin{bmatrix} H_1^{1,1} & \dots & H_{29}^{1,1} & q_1^{1,1} & q_2^{1,1} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ H_1^{1,2016} & \dots & H_{29}^{1,2016} & q_1^{1,2016} & q_2^{1,2016} \end{bmatrix} & \text{1st week} \\ \begin{bmatrix} H_1^{2,1} & \dots & H_{29}^{2,1} & q_1^{2,1} & q_2^{2,1} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ H_1^{2,2016} & \dots & H_{29}^{2,2016} & q_1^{2,2016} & q_2^{2,2016} \end{bmatrix} & \text{2nd week} \\ \vdots & \\ \begin{bmatrix} H_1^{52,1} & \dots & H_{29}^{52,1} & q_1^{52,1} & q_2^{52,1} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ H_1^{52,2016} & \dots & H_{29}^{52,2016} & q_1^{52,2016} & q_2^{52,2016} \end{bmatrix} & \text{52nd week} \end{bmatrix} \quad (1)$$

Where H_i is a column vector of pressure at node i with the length of 104,832; q_j is a column vector of flow at pipe j with the length of 104,832; $H_i^{n,t}$ is the pressure of node i at time t in the n -th week; $q_j^{n,t}$ is the flow of pipe j at time t in the n -th week.

In the third stage, nodal demands were calibrated based on the prior information by the least square method iteratively. We considered the base demand used in the Epanet input file as empirical values. A single-objective optimization model was established to minimize the $f(\mathbf{Q})$. This converted the under-determined problem (the number of unknown variables is greater than that of equations) to an over-determined one. Specifically, this conversion included two steps. Firstly, we ran the original hydraulic model of the L-Town network and retrieved the averaged nodal demands by accumulating the actual demands of three categories at each time step (see Equation 2). Secondly, the objective function was formulated, as shown in Equation 3.

$$\mathbf{Q}^a = \begin{bmatrix} Q_{1,1}^a & \dots & Q_{1,2016}^a \\ \vdots & \ddots & \vdots \\ Q_{nn,1}^a & \dots & Q_{nn,2016}^a \end{bmatrix} \quad (2)$$

Where Q^a is the empirical value used in the objective function; $Q_{i,t}^a$ is the total demand at node i and time t ; nn is the number of nodes.

$$\min f(\mathbf{Q}^{n,t}) = w_1 \sum_{i=1}^{nH} [H_i^{n,t} - H_i(\mathbf{Q}^{n,t})]^2 + w_2 \sum_{j=1}^{nq} [q_j^{n,t} - q(\mathbf{Q}^{n,t})]^2 + w_3 \sum_{k=1}^{nn} (Q_{k,t}^a - Q_k^{n,t})^2 \quad (3)$$

Where $H_i^{n,t}$ and H_i are the monitored and simulated pressures of node i at time t in the n -th week, respectively; $q_j^{n,t}$ and q_j are the monitored and simulated flows of pipe j at time t in the n -th week, respectively; $Q_{k,t}^a$ and $Q_k^{n,t}$ are the empirical and simulated nodal demands of node k at time t in the n -th week, respectively; w is the weight coefficient (w_1 and w_2 are equal to 1, while w_3 is equal to 0.1); nH is the number of pressure sensors; nq is the number of flow meters.

Equation 3 can also be expressed as the following matrix. The calibrated nodal demands were obtained by solving Equations 5-6 iteratively until the 2-norm of $\Delta \mathbf{Q}_l$ is less than the specified threshold (i.e., 0.01). The Jacobian matrices (i.e., $\mathbf{J}_H(\mathbf{Q}_l^{n,t})$ and $\mathbf{J}_q(\mathbf{Q}_l^{n,t})$) were solved according to [1]. The detail on the least square method can be accessed via [2].

$$\min \mathbf{f}(\mathbf{Q}^{n,t}) = \begin{bmatrix} \mathbf{H}^{n,t} - \mathbf{H}(\mathbf{Q}^{n,t}) \\ \mathbf{q}^{n,t} - \mathbf{q}(\mathbf{Q}^{n,t}) \\ \mathbf{Q}_l^a - \mathbf{Q}^{n,t} \end{bmatrix}^T \mathbf{W} \begin{bmatrix} \mathbf{H}^{n,t} - \mathbf{H}(\mathbf{Q}^{n,t}) \\ \mathbf{q}^{n,t} - \mathbf{q}(\mathbf{Q}^{n,t}) \\ \mathbf{Q}_l^a - \mathbf{Q}^{n,t} \end{bmatrix} \quad (4)$$

$$\Delta \mathbf{Q}_l = \left\{ \begin{bmatrix} \mathbf{J}_H(\mathbf{Q}_l^{n,t}) \\ \mathbf{J}_q(\mathbf{Q}_l^{n,t}) \\ \mathbf{I} \end{bmatrix}^T \mathbf{W} \begin{bmatrix} \mathbf{J}_H(\mathbf{Q}_l^{n,t}) \\ \mathbf{J}_q(\mathbf{Q}_l^{n,t}) \\ \mathbf{I} \end{bmatrix} \right\}^{-1} \begin{bmatrix} \mathbf{J}_H(\mathbf{Q}_l^{n,t}) \\ \mathbf{J}_q(\mathbf{Q}_l^{n,t}) \\ \mathbf{I} \end{bmatrix} \mathbf{W} \begin{bmatrix} \Delta \mathbf{H}_l \\ \Delta \mathbf{q}_l \\ \Delta \mathbf{Q}_l^0 \end{bmatrix} \quad (5)$$

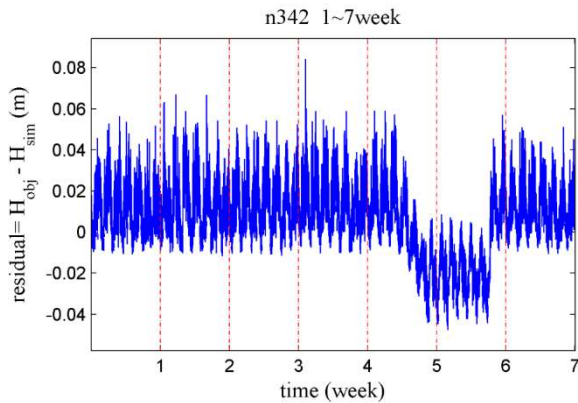
$$\mathbf{Q}_{l+1}^{n,t} = \mathbf{Q}_l^{n,t} + \Delta \mathbf{Q}_l \quad (6)$$

Where $\mathbf{J}_H(\mathbf{Q}_l^{n,t})$ is the Jacobian matrix of nodal pressure to nodal demand; $\mathbf{J}_q(\mathbf{Q}_l^{n,t})$ is the Jacobian matrix of pipe flow to nodal demand; \mathbf{I} is the identity matrix of size nn .

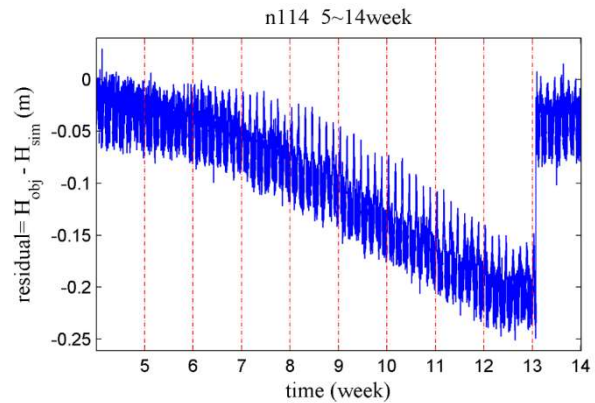
In the fourth stage, the locations of pipe burst were identified based on the calibration residual obtained in the previous stage. The nodal calibration residuals of pressure were recorded at each time step. When a pipe burst, the calibration residuals between monitored and simulated pressures at some sensitive nodes would increase or decrease accordingly. Such fluctuations in calibration residuals were used as signals to judge whether a burst event occurred. During 2018, there were eight burst events reported in Area A of the L-Town network. We managed to identify six of these events with good approximation according to the Euclidean distances between the repaired and identified locations (see Table 1). Some signals that successfully detected the six pipe burst events in Area A are presented in Figures 2(a) to 2(f). As shown in Figure 2(a), the calibration residuals at node n342 drop dramatically from the middle of Week 5 and recover at the end of Week 6, which indicates that a pipe burst was likely to occur during this period.

Table 1. Information of six pipe bursts identified in Area A of the L-Town network in 2018

Event No.	Burst ID	Start Time (Estimated)	End Time	Burst Locations by This Study	Euclidean Distance between the Repaired and Identified Locations (m)
①	p232	02-02 03:00	02-10 09:20	p138	157.52
②	p461	02-19 00:00	04-02 11:40	p452	84.93
③	p628	05-08 00:00	05-29 21:20	p630	38.64
④	p183	08-07 02:30	09-01 17:10	p856	39.79
⑤	p158	10-06 04:35	10-23 13:35	p702	135.18
⑥	p369	10-26 02:05	11-08 20:25	p371	100.74



(a) n342



(b) n114

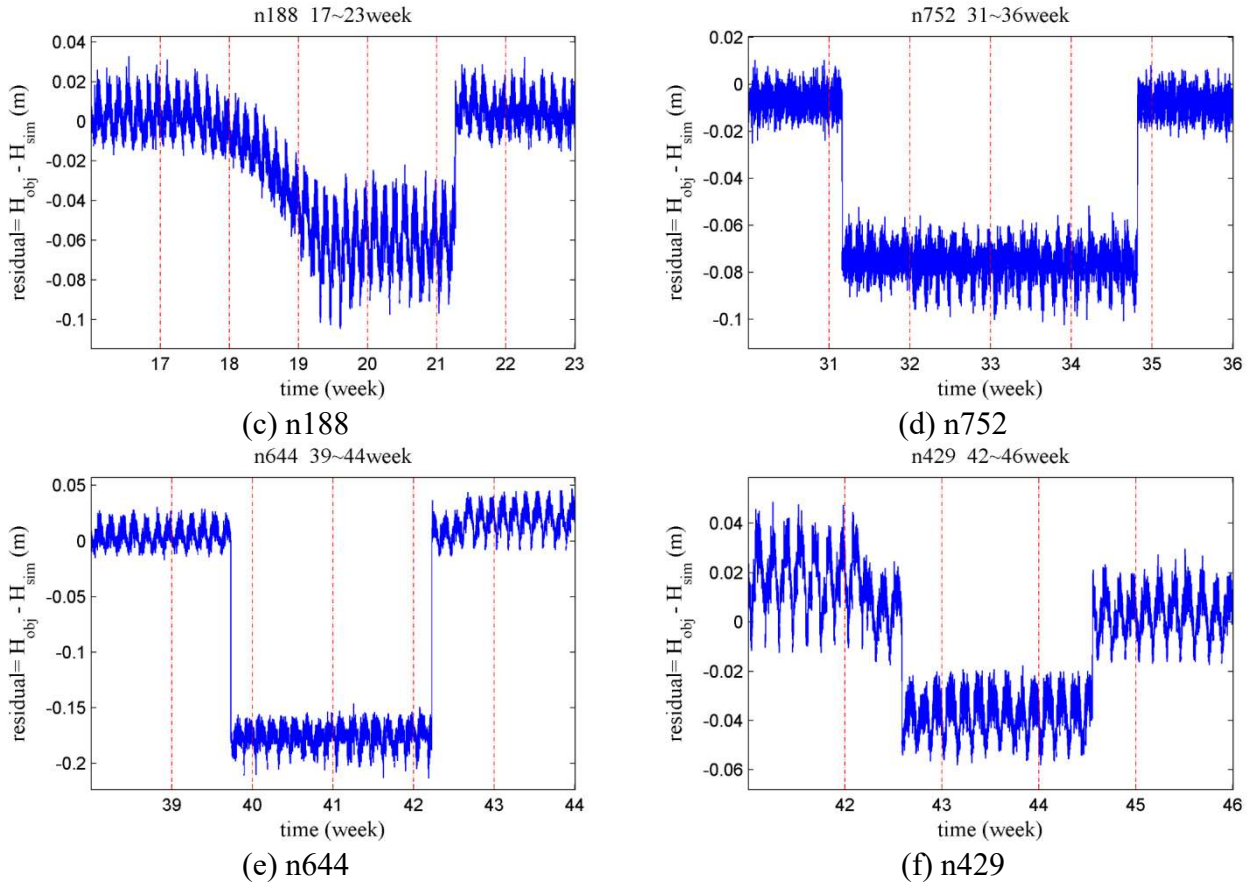


Figure 2. Detection of pipe burst events based on the calibration residuals at different nodes

In the final stage, the locations of pipe burst events were approximated by an improved vectorial angle method proposed for the BattLeDIM problem. The original vectorial angle method was developed in [3]. It was used to locate the pipe burst places by comparing the vector of calibration residuals with the Jacobian matrix of nodal pressures to nodal demands. However, due to the assumption that pipe burst occurs at nodes, this method could only identify the location in a broader range. If a burst happens in a pipeline, the original vectorial angle method may fail to determine an accurate position. Therefore, we modified this method by introducing the Jacobian matrix of nodal pressures to pipe flows. This change extended the original vectorial angle method for dealing with the burst events that occurred in pipelines. The pressure residuals are shown in Equation 7.

$$\mathbf{R} = [\mathbf{r}_1, \dots, \mathbf{r}_t] = \begin{bmatrix} r_{1,1} & \cdots & r_{1,m} \\ \vdots & \ddots & \vdots \\ r_{nH,1} & \cdots & r_{nH,m} \end{bmatrix} = \begin{bmatrix} H_{1,1}^o - H_{1,1}(\mathbf{Q}^{l,1}) & \cdots & H_{1,m}^o - H_{1,m}(\mathbf{Q}^{l,m}) \\ \vdots & \ddots & \vdots \\ H_{nH,1}^o - H_{nH,1}(\mathbf{Q}^{l,1}) & \cdots & H_{nH,m}^o - H_{nH,m}(\mathbf{Q}^{l,m}) \end{bmatrix} \quad (7)$$

Where $r_{i,t}$ is the calibration residual of pressure at node i at time t ; $H_{i,t}^o$ is the observed pressure at node i at time t ; $H_{i,t}(\mathbf{Q}^{l,t})$ is the pressure of node i at time t in the previous week; $\mathbf{Q}^{l,m}$ is the calibrated nodal demand in the previous week when the burst event did not occur; m is the number of samples selected for the identification of a pipe burst event (m was equal to 288 in this study).

$$\mathbf{J}_H^j(\mathbf{Q}^{l,t}) = 0.5[\mathbf{J}_H^{up}(\mathbf{Q}^{l,t}) + \mathbf{J}_H^{down}(\mathbf{Q}^{l,t})] \quad (8)$$

$$\mathbf{J}_H^p(\mathbf{Q}^{l,t}) = [\mathbf{J}_H^1(\mathbf{Q}^{l,t}), \dots, \mathbf{J}_H^{mm}(\mathbf{Q}^{l,t})] \quad (9)$$

Where $\mathbf{J}_H^j(\mathbf{Q}^{l,t})$ is the sensitivity vector of pressure to the flow of pipe j ; $\mathbf{J}_H^{up}(\mathbf{Q}^{l,t})$ and $\mathbf{J}_H^{down}(\mathbf{Q}^{l,t})$ are the Jacobian matrices at the upstream and downstream nodes of pipe j , respectively; $\mathbf{J}_H^p(\mathbf{Q}^{l,t})$ is the nH -by- mm Jacobian matrix of nodal pressure to pipe flow; mm is the number of

pipes (excluding the ones connected to the two reservoirs directly). The detail on how Equations 8 and 9 were obtained can be accessed via [4].

For each time step, the angle between the vector of calibration residuals and each column of the Jacobian matrix was calculated by Equation 10.

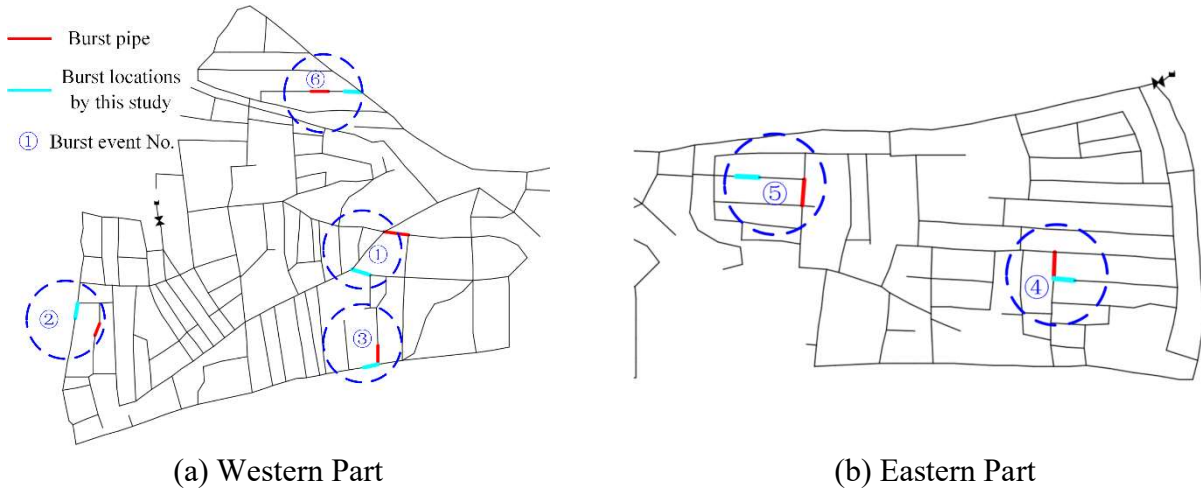
$$\alpha_{j,t} = \arccos \left(\frac{\mathbf{r}_t \cdot \mathbf{J}_H^j(\mathbf{Q}^{l,t})}{|\mathbf{r}_t| \cdot \|\mathbf{J}_H^j(\mathbf{Q}^{l,t})\|} \right) \quad (10)$$

Where $\alpha_{i,t}$ is the angle between the vector of calibration residuals and the vector of pipe j in the Jacobian matrix; \mathbf{r}_t is the vector of the calibration residual at time t ; $\mathbf{J}_H^j(\mathbf{Q}^{l,t})$ is the vector of pipe j in the Jacobian matrix at time t .

For each pipe burst event, we calculated the average angle over 288 time steps (i.e., one day) before a burst location was localized. This is to mitigate the impact of demand uncertainties on the calculation of the vectorial angle (see Equation 11). The pipe with the minimum value of $\bar{\alpha}_j$ was identified as the broken one. The six locations detected as pipe bursts mentioned in the third stage are shown in Figure 8.

$$\bar{\alpha}_j = \frac{\sum_{t=1}^m \alpha_{j,t}}{m} \quad (11)$$

where, $\bar{\alpha}_j$ is the average angle between the vector of calibration residuals and the vector of pipe j in the Jacobian matrix; m is the same as defined previously.



(a) Western Part

(b) Eastern Part

Figure 8. Locations of pipe burst identified by the proposed method

Keywords: Burst detection and location; Prior information; Nodal demand calibration; Vectorial angle method; Water distribution system

ACKNOWLEDGMENT

This work was supported by both the Research on Key Technologies and Applications of Building Safety Risk Prevention and Control, Key R & D Plan of Yunnan Province, P. R. China and National Science Foundation Award (Grant No.: 0836046).

REFERENCES

- [1] Nian-dong, L., Kun, D., Jia-peng, T., & Wei-xin, D. (2017). Analytical Solution of Jacobian Matrices of WDS Models. *Procedia Engineering*, 186, 388–396.
- [2] Kang, D., & Lansey, K. (2009). Real-Time Demand Estimation and Confidence Limit Analysis for Water Distribution Systems. *Journal of Hydraulic Engineering*, 135(10), 825–837.

- [3] Ponce, M. V. C., Castañón, L. E. G., & Cayuela, V. P. (2012). Extended-Horizon Analysis of Pressure Sensitivities for Leak Detection in Water Distribution Networks. *IFAC Proceedings Volumes*, 45(20), 570–575.
- [4] Geng, Z., Hu, X., Han, Y., & Zhong, Y. (2019). A Novel Leakage-Detection Method Based on Sensitivity Matrix of Pipe Flow: Case Study of Water Distribution Systems. *Journal of Water Resources Planning and Management*, 145(2), 04018094.

SUMMARY

The combined usage of the hydraulic model calibration residual and an improved vectorial angle method is presented for the burst detection and diagnosis in the L-Town network. It consists of five stages: (1) model decomposition, (2) partition of SCADA data, (3) nodal demand calibration, (4) calibration residual-based leakage detection, and (5) an improved vectorial angle method based burst localization. Compared with existing methods, the proposed method has the following advantages that make it a robust burst detection and localization approach. First, the bursts are detected based on the calibration residuals of nodal demands, by which the hydraulic model and SCADA data are taken into account simultaneously. Second, the concept of pipe sensitivity vectors is proposed, considering the burst occurs in the middle of pipelines. This sensitivity vector is calculated based on the nodes' sensitivity vectors at both ends of a pipe. It is then used to localize pipe detect the burst pipes accurately through calculating the angle between the calibration residual vector and each pipe sensitivity vector. The burst pipe is the sensitivity vector that presents the smallest angle with the residual vector. We first applied the method mentioned above to the pipe burst events reported in 2018. Then, our estimations regarding the location and start time of pipe bursts in 2019 are yield.