Water Supply

© 2022 The Authors

Water Supply Vol 22 No 4, 4558 doi: 10.2166/ws.2022.117

Leakage detection in water distribution networks using space search reduction and local fitness in steady-state and EPS (case study: water distribution network of Birjand)

Ali Nasirian* and Marzieh Ahrari

Department of civil engineering, Engineering faculty, University of Birjand, Birjand, Iran
*Corresponding author. E-mail: a.nasirian@birjand.ac.ir

ABSTRACT

Using hydraulic model calibration of a water distribution network is one of the methods that can reduce leak detection costs by identifying areas with high leakage. In the present paper, a new factor called local fitness is used in combination with the search space reduction method to obtain an optimal answer. In this method, a fitness function is calculated for each zone. For each zone, proximity of the pressure and flow rates obtained from the model to the corresponding observational data indicates that the values selected for the nodes of that zone are close to the correct values. In the next steps, using a search space reduction method and performing the optimization process with the local fitness function, the chance of selecting those values is increased. In this study, leakage was assigned to nodes with emitter coefficient. This method was used on the water distribution network of Birjand. Three factors of background leakage, leakage hotspot and apparent losses were considered as unknown parameters. Results of the present method were compared with the results of a genetic algorithm (GA) and the corresponding exact values. Based the results, the present method showed better results in terms of convergence speed and accuracy.

Key words: calibration, Epanet2, leak detection, optimization, search space reduction method

HIGHLIGHT

- This model detects leakage more accurate

1. INTRODUCTION

As essential as water is to life, it is also considered an economic commodity due to the cost of supply, treatment and transmission (Soares *et al.* 2010). For this reason, non-revenue water (NRW) is important for water experts both economically and as being vital to life (Cassidy *et al.* 2020). NRW includes several parts. Much of this water is wasted due to the leaks in the water distribution network (WDN). This part of the NRW is related to WDN pressure. Apparent losses (AL) constitute another important part of the NRW, but it has not been measured for reasons such as faulty flow meters and unauthorized consumption. Fluctuations in this section are related to the consumption patterns. It should be mentioned that there are other parts such as firefighting consumption and WDN washing, but their amount is less than the previous two parts and therefore excluded from the present study.

One way to reduce NRW in WDNs is to identify the location and the amount of leakage. Using field methods for identifying the location of leakage in the entire WDN is time consuming and not economical. Before using field methods, it is better to first identify the leakage zone by simulating and calibrating the hydraulic model of the WDN, and then determine the exact location of the leak using leak detection devices. In these methods, the WDN leakage is examined by WDN modeling and analysis. These methods are valued by researchers because of being more economical and time-saving. Researchers have focused on using steady and transient simulation of WDS to detect leakage in WDS.

Regarding transient-based leak detection methods the review by Colombo *et al.* (2009) is also useful. Duan *et al.* (2020) wrote a review on leak detection in transient-based model simulation. Also Che *et al.* (2021) have considered transient wave-based methods for pipe anomaly detection. Transient-based methods for leak detection have been successfully used in relatively complex systems: Meniconi *et al.* (2015, 2021) show transient-based leak detection in a WDS and in a transmission main, respectively; while Pan *et al.* (2021) show the results of transient-based leak detection in a bit complex

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (http://creativecommons.org/licenses/by/4.0/).

Downloaded from http://iwa.silverchair.com/ws/article-pdf/22/4/4558/1105189/ws022044558.pdf
by guest

system made of polymeric pipes, paving the way to the extension of his technique to WDS. Although transient-based methods have the advantages of finding the exact location of leaks, there are limitations in their use in large-scale complex WDS. In addition, unlike the steady-state method, which is worked with pressure and flow measurements harvested for network monitoring, the transient method requires additional and more specific field data.

Pudar & Liggett (1992) first used the inverse steady-state analysis to detect leaks in the pipe network. Almandoz *et al.* (2005) calibrated the WDN in a steady state to detect leakage. They examined leakage as being related to pressure and consumption based on consumption pattern. Zheng Yi Wu *et al.* (2010) calibrated and detected leakage on a hypothetical WDN and an actual WDN based on a genetic algorithm. They considered leakage to be pressure dependent and as an emitter coefficient in the model nodes. Results of their research showed that by detecting leaks, calibration accuracy of the hydraulic model is improved. By integrating the genetic algorithm method and an innovative method called step-by-step elimination algorithm, Nasirian *et al.* (2013) identified leaky nodes in two hypothetical WDNs. Their research showed that after several steps of search space reduction, leakage position can be detected with high accuracy. Eryiğit (2019) used an artificial immune system optimization algorithm to calibrate and determine the amount and area of water losses in the WDN. In his research, he considered water losses as unauthorized consumption in nodes and leaks in pipes. This method was able to detect the amount and area of water losses in the nodes and pipes of a hypothetical WDN only through flow observations. Sophocleous *et al.* (2017) used a two-step method to calibrate a real WDN model for determining the leak points. A genetic algorithm was used to solve the optimization problem. Optimization parameters in their study were valve condition, improvement of pipe roughness and leakage correction in nodes. They also used the search space reduction method to speed up convergence. Using a non-sorting genetic algorithm and based on search space reduction, they identified leak points in a real WDN. Using calibration of consumption by the ant algorithm, Hajibandeh & Nazif (2018) presented a new method for determining the amount and location of leakage in WDNs. Their method was evaluated on two hypothetical WDNs. Results showed that the efficiency of the proposed method depends on the number of barometric points. Also, as the number of leak nodes increases, the leak detection process becomes more difficult and more complicated. Amoatey *et al.* (2018) used the Monte Carlo method to detect WDN leaks. Using the orifice leakage equation, leakage was considered proportional to the nodal pressures. Zhang *et al.* (2018) proposed a numerical method for calibration of pipe roughness coefficients and nodal demands. The calibration objective function is formulated as a least-square equation, followed by the derivation of iterative formulas. To reduce the calibration dimensions, the pipes and nodes are grouped according to their physical properties. The proposed approach is calibrated both pipe roughness coefficients and nodal demands, acceptably. Moasheri & Jalili-Ghazizadeh (2020) calibrated the demands and roughness coefficient of the pipes by the imperialist competitive algorithm (ICA). Their results showed that this method can probably prioritize leaky areas without limitation in the number of simultaneous leaks. Huang *et al.* (2020) developed a multistage technique for burst leak localization based on smart demand metering and valve operations. The results indicate that the proposed method is effective and efficient to detect burst leak. Li *et al.* (2021) proposed a new model-based approach for WDS leak localization. The method is specified by (1) expanded the dominant sensor sequence for candidate leak nodes; (2) applying multisets of the measurements; (3) ranking leakage zones and nodes.

In the present study, leak detection is investigated based on modeling and calibration of WDN. For this purpose, leakage is determined by adjusting the emitter coefficient and apparent losses as well as assigning an incremental coefficient to the consumption. WDN analysis was performed for steady state and EPS data. In this study, the method of search space reduction was used in combination with a new parameter called local fitness. In this method, WDN nodes are grouped in several zones. At each step of the analysis, some answers are generated by the random search method and their fitness is calculated. Local fitness function is then calculated for each zone and in each answer. By selecting and integrating answers that have superior local fitness in a zone, the leakage range is determined. Next, by limiting the search space, the steps are repeated to get the best answers. In this research, in addition to determining the amount and location of leakage, another coefficient is calculated that determines apparent losses in the entire WDN.

In the present study, a simple random search method was applied to evaluate the efficiency of the proposed method. Simultaneous calibration of roughness and condition of valves was omitted. Obviously, in the proposed method, other optimization methods can be used instead of random search. The present method can also be used in experiments that simultaneously calibrate factors of pipe roughness and condition of the valves.

The proposed optimization program was coded in MATLAB software and by connecting MATLAB and Epanet2. software, hydraulic analysis of the WDN was performed in Epanet2. This method was tested on zone D of Birjand WDN. Leakage hotspot, background leakage and, to some extent, apparent losses are investigated simultaneously in the WDN.

Downloaded from http://iwa.silverchair.com/ws/article-pdf/22/4/4558/1105189/ws022044558.pdf
by guest

2. METHODOLOGY

In order to calibrate the WDN, the difference between the observed and calculated values of pressure and flow at selected points should be minimized. This is done by setting some WDN modeling parameters that are erroneously determined. In the present study, the main parameters to be set are apparent losses and leakage. Accordingly, WDN is calibrated by adjusting the values of emitter coefficient and the apparent loss coefficient at the nodes.

2.1. Leak detection by WDN calibration method

The difference between the amount of water consumed and recorded by flow meters and the amount of water entering the WDN is called non-revenue water. Total non-revenue water consists of several components; each having a different nature. Some of these components, such as unauthorized consumption and flow meter error, change during the day and night based on the consumption patterns. Some other components, such as pipe breakage and leakage from junctions, are a function of water pressure in the WDN. In this study, non-revenue water is divided into two categories: apparent losses and real losses (leaks). Once the recorded consumption in the nodes is known, the nodal demand is corrected by estimating the emitter coefficient in the nodes as well as estimating the apparent loss. Therefore, the demand in the node j is calculated with:

$$\boldsymbol{D_j}(t) = \boldsymbol{Pt}(t)(1 + C) \times \boldsymbol{DB}_j + K_j \times (\boldsymbol{P_j}(t))^n \tag{1}$$

In this equation, j is the number of the node, $\boldsymbol{D_j}(t)$ is water withdrawal from node j at time t, $\boldsymbol{DB}_j$ is the base consumption at node j and $\boldsymbol{Pt}(t)$ is the consumption pattern at time t. the unit of demands in this article is liter per minute. C is the apparent loss coefficient, which is considered as a percentage of the base demand. Kj is the emitter coefficient in node j and is calculated in terms of liters per minute divided by the pressure to the exponent of n $((l/\min)/((m_{H_2O})^n))$. $\boldsymbol{P_j}(t)$ is the pressure at node j at time t and n is the pressure's exponent. Pressure's exponent depends on the type of pipes and pressure changes in the WDN and can be determined using theoretical relationships, laboratory data or field measurements. FAVAD[1] theory was used to determine the pressure's exponent (Karamouz *et al.* 2005). According to the WDN zoning and grouping of nodes, emitter coefficient for each zone of the WDN is considered as an unknown parameter and all nodes in a specified group are given the same emitter coefficient based on:

$$K_j = \frac{K_N}{Nj} \tag{2}$$

Here, K_j is the emitter coefficient for node j, K_N is the emitter coefficient for zone N, and Nj is the number of nodes in zone N. Therefore, the unknown vector is defined based on:

$$\mathbf{X} = (K_N, C) \qquad N = 1, 2, , \ldots, NG \tag{3}$$

that X is set parameters of the model, NG is the number of zones and K_N is the emitter coefficient for zone N. The fitness function which should be minimized is represented by:

$$F(X) = \frac{\sum_{h=1}^{NH} (Hs_h - Ho_h)^2 + \sum_{f=1}^{NF} (Fs_f - Fo_f)^2}{NH + NF} \tag{4}$$

In Equation (4), $F(X)$ is the fitness function which indicates the proportion of the measured values and the values simulated by the model. Hs_h and Ho_h are respectively calculated and measured hydraulic level for node h. Fs_f and Fo_f show the calculated and measured flows in pipe f, respectively; and NH and NF are the number of pressures and flows measured in the WDN, respectively. Problem constraints are also shown in Equations (5) and (6):

$$K_{sum} - t \le \sum_{N=1}^{NG} K_N \le K_{sum} + t \tag{5}$$

[1] Fixed and Variable Area Discharge

Downloaded from http://iwa.silverchair.com/ws/article-pdf/22/4/4558/1105189/ws022044558.pdf
by guest

Equation (5) limits the sum of the selected coefficients for the WDN to the value of this parameter for the whole WDN with a tolerance value. In this equation, K_{sum} is the emitter coefficient for the WDN, which is obtained from the analysis of non-revenue water, and t is the acceptable tolerance based on the accuracy of measurements:

$$C_{min} \leq C \leq C_{max} \tag{6}$$

C_{min} and C_{max} are the minimum and maximum limits of apparent loss coefficient, respectively. The ratio of apparent losses to total non-revenue water is different for each WDN. If a WDN has much leakage, the share of this parameter in the total non-revenue water decreases, and if there is a lot of unauthorized consumption in a WDN and the WDN has a small amount of leakage, the apparent loss coefficient increases. To determine these parameters in a real WDN, first the minimum and maximum limit of the apparent loss coefficient is assumed with a certain step interval. If the obtained C is in the middle of the range, the considered interval is selected correctly and can be narrowed. If the best fitness is obtained by selecting a C at the beginning or end of the range, it should be extended in the direction of the value obtained with the best fitness. To evaluate the accuracy of the results, a function based on the calculation of the root mean square error (RMSE) is considered as Equation (7). It should be noted that the calculation of this parameter is not possible in real conditions and can be calculated only if the amount of actual leakage is known:

$$E(d) = \sqrt{\frac{\sum_{i=1}^{NG} (Ko_i + Ks_i)^2}{NG}} \tag{7}$$

In this relation, $E(d)$ is the error index of the answers, Ko_N is the actual emitter coefficient, and Ks_N is the simulated emitter coefficient and NG is the zone number.

2.2. Local fitness function

In the proposed optimization algorithm, first some answers are generated randomly in a discrete space by considering the constraints of the problem. The hydraulic analysis of each answer is then performed with a dynamic relationship between the Epanet2. and MATLAB software. Then the fitness function is calculated for each answer. It is unlikely for one answer to regulate all the pressures and flows across the WDN, but a partially good answer may well regulate the pressures and flows in one zone of the WDN. Based on this fact, the local fitness method is introduced and used for the first time in this research.

In this method, the WDN is divided into different zones and according to Equation (8), a function called local fitness is calculated for each zone per each answer. The local fitness function is defined for each zone in proportion to the pressure and observational flows in that zone:

$$FL_{n,i} = \frac{\sum_{h=1}^{ZH} (Hs_{n,i}^{h} - Ho_{n,i}^{h})^2 + \sum_{l=1}^{ZL} (Fs_{n,i}^{l} - Fo_{n,i}^{l})^2}{ZH + ZL} \tag{8}$$

In this relation, $FL_{n,i}$ is the local fitness function of zone n for answer i. $Hs_{n,i}^{h}$ and $Ho_{n,i}^{h}$ are respectively calculated and measured hydraulic level for node h in zone n and answer i, $Fs_{n,i}^{l}$ and $Fo_{n,i}^{l}$ are respectively measured and calculated flow for pipe l in zone n and answer i, ZH and ZL are number of pressures and flows measured in zone n. In order to give more chances to the answers with good fitness in the WDN, the local fitness function is modified according to:

$$FLm_{n,i} = \frac{FL_{n,i} \times F(X)_i}{\overline{F(X)}} \tag{9}$$

In this relation, $FLm_{n,i}$ is modified fitness function of zone n for answer i, $F(X)_i$ is WDN fitness for answer i and $\overline{F(X)}$ is mean fitness throughout the WDN for population of a generation.

Downloaded from http://iwa.silverchair.com/ws/article-pdf/22/4/4558/1105189/ws022044558.pdf
by guest

The K_N search interval to start the program is defined based on the total WDN leakage provided in:

$$\text{for g} = 1 \quad 0 \le K_N \le K_{sum} \tag{10}$$

In this regard, g is the generation number. After calculating the local fitness in each zone, the search interval for K_N is modified. For each zone, an interval is calculated based on the answers corresponding to the best modified local fitness, in which the correct answer is more likely. This interval is presented in:

$$Kmin_N^{g-1} \le K_N \le Kmax_N^{g-1} \text{for} \quad g > 1 \tag{11}$$

This relation shows the search interval in generation g for zone N. In this equation, $Kmin_N^{g-1}$and $Kmax_N^{g-1}$ are respectively the minimum and maximum K_N for the top answers of the $g-1$ generation. The top answers used in zone N include the top 20% of the $g-1$ generation in terms of modified local fitness index. In producing the second generation onwards, the constraint of Equation (11) is considered for 80% of the population. To eliminate the risk of getting stuck in the local minimum, 20% of the population is also selected with the constraint of Equation (10). Therefore, the proposed optimization method, by introducing a reduced search interval, increases the probability of selecting emitter coefficients in this interval in order to create a new population. To maintain the superior answers of the previous stages, in each stage, between the two new and previous generations, the answers with superior fitness are selected according the required number of population and are used to produce the next generation.

2.3. Solution methods

In this study, leak detection was performed on zone D of Birjand WDN. All the information used, including consumption, flow meter, pipe and node properties, and the amount of non-revenue water are considered based on actual information. Knowing the volume of non-revenue water, the amount of apparent loss, the volume of background leakage and leakage hotspots and the location of leakage hotspots were hypothetically considered. This analysis is used to generate pressure and flow observations as well as to validate leak detection results. The results obtained from space search reduction and local fitness (SSR-LF) methods were compared with their real values as well as non-sorting GA (Bentley Systems 2007; Wu *et al.* 2010; Wu & Simpson 2001).

3. RESULTS AND DISCUSSION

3.1. Introducing the studied WDN

Birjand city is located in South Khorasan province in the east of Iran. In this research, zone D of Birjand WDN is studied. This WDN has 1,140 nodes and 1,191 pipes. Investigation of events in Birjand WDN shows that the ratio of the number of leakage orifices with fixed cross-section to the number of leakage orifices with variable cross-section (R_{FVL}) is equal to 0.5. According to the FAVAD and RFVL diagrams, the range of pressure's exponent changes is obtained from 0.9 to 1.2. Also, the ratio (P_1/P_0) ,which indicates different pressure conditions to the average WDN pressure, is equal to 1 for the average flow condition. According to the diagram, the value of pressure's exponent in this WDN is equal to 1.15. Accordingly, the value of emitter coefficient in Equation (1), based on liter per minute divided by the water pressure in meters to the power of 1.15 is calculated ($(l/min/(m_{H_2O})^n)$) and for the sake of brevity in the continuation of the article, its unit is not mentioned.

Figure 1 shows the division of zone D of Birjand WDN to 6 zones. In dividing the network into zones, several factors were considered. (1) Each zone should be connected the others by minimum number of pipes. (2) The WDS was built in different time. The oldest regions of the WDS are placed in zone 2 and 3. The pipes in this region are 30–50 years old. The WDS in districts 1, 4 and 6 are 20–30 years old. The newest part of the network is zone 5 with less than 15 years old. (3) An attempt was made to make the area of all areas almost the same.

Due to the fact that large leakages were assumed in areas 1, 4 and 6, in the second step, these areas were divided into three sub-zones. It should be noted that the present method relies on observational pressures and the refinement of search areas depends on the number of pressure observations. There must be at least one pressure gauge in each sub-zone for the local fitness parameter to be calculated.

The area of the study site is about 5.7 km^2 and each zone and sub-zone is about 1 km^2 and 0.3 km^2, respectively.

Downloaded from http://iwa.silverchair.com/ws/article-pdf/22/4/4558/1105189/ws022044558.pdf
by guest

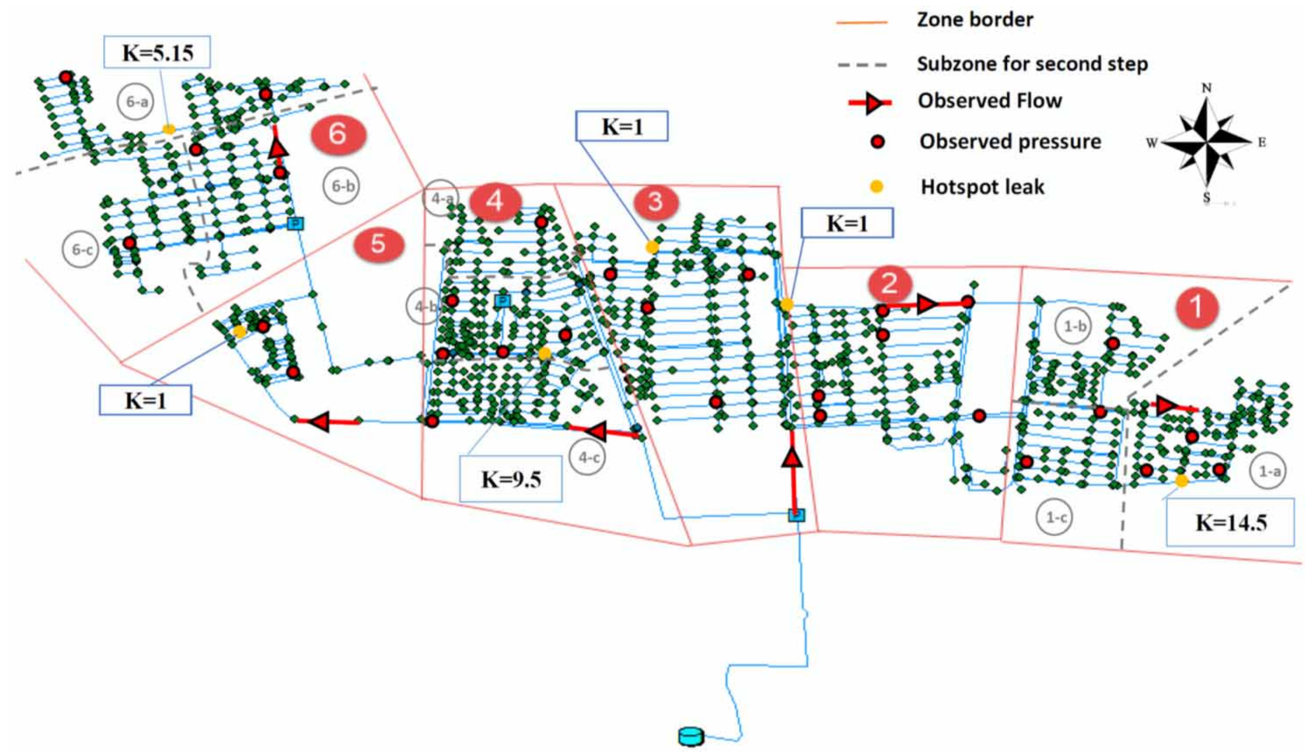

Figure 1 | Division of zone D of Birjand WDN.

Barometric and flow metering points of different zones are also specified. In this WDN, 29 barometric points and six flow metering locations are considered.

According to flow meter measurements, the average consumption of customer is 6,939 liters per minute (115.65 liters per second). Also, based on the measurement of the inflow to the WDN, the average inflow is 11,760 liters per minute (196 liters per second). As a result, non-revenue water in this WDN is equal to 4,821 liters per minute (80.35 liters per second). Having a selective approach, 40% of the non-revenue water is considered as background leakage, which is equal to 1,928.4 liters per minute. According to the relationship between leakage and pressure, the value of 26.25 is obtained as the emitter coefficient for the background leakage in the entire WDN. This value is assigned to all WDN nodes equally. It is also assumed that leakage hotspots of 1,800 liters per minute occur in different parts of the WDN and are concentrated at some points. Thus, the emitter coefficient for leakage hotspots based on the relationship between leakage and pressure for the entire WDN is 32.15. This value is assigned to nodes 269, 53 and 1,049, to hypothetical leak nodes with values of 14.5, 9.5 and 5.15, respectively, and nodes 793, 1,071 and 51, with the value of 1. Due to the amount of non-revenue water and the values estimated for leakage hotspots and background leakages, the amount of 1,092.6 liters per minute remains for apparent losses. Apparent losses are considered as a percentage of water consumption, which is equivalent to 15.7% of the base consumption. Figure 1 shows leaky points in different zones of the WDN. The figure shows that there are leakage hotspots in zones 1, 4 and 6. Table 1 shows the number of nodes in each zone and the values of emitter coefficient for background leakage and leakage hotspot, separately in each zone. Due to the same nature of background leakage and leakage hotspot, values of their emitter coefficient are added together and presented in column 5. It should be noted that since the analysis are performed in a discrete space, these values are rounded in column 6. Numbers in this column are the best values that can be selected by the optimization program.

It should be noted that the accuracy of the results obtained for calibration-based leak detection for a WDS depends on: (1) Quantity and quality of field measurements. (2) The effect of leakage on pressure drops in observational nodes. If the leaks are located near fixed pressure points such as tanks or pressure relief valves, it will be difficult to find. If the position of the leaky node and the pressure sensors are such that the hydraulic drop due to the leak does not have much effect on the pressure of the pressure observation point, it will be difficult to find leaks. (3) WDS characteristics (pipes diameter and nodal demand) and leakage value also affect the quality of results. In the present study, different leakage positions and pressure sensor's locations were investigated and the difference in results was not significant.

Downloaded from http://iwa.silverchair.com/ws/article-pdf/22/4/4558/1105189/ws022044558.pdf
by guest

Table 1 | Actual emitter coefficients in the zones and sub-zones

Zone	Number of junctions	K_N in zones: Leakages hotspot	Background leak	Sum of leakages	Sum of leakages in discrete space
1	237	14.5	6.59	21.09	21
1-a	105	14.5	2.92	17.42	17
1-b	61	0	1.69	1.69	2
1-c	71	0	1.98	1.98	2
2	125	1	3.37	4.37	4
3	206	1	5.53	6.53	7
4	272	9.5	5.15	14.65	15
4-a	59	0	1.12	1.12	1
4-b	102	9.5	1.93	11.43	11
4-c	111	0	2.10	2.1	2
5	72	1	1.21	2.21	2
6	228	5.15	4.4	9.55	10
6-a	97	5.15	1.87	7.02	7
6-b	78	0	1.51	1.51	2
6-c	53	0	1.02	1.02	1

3.2. Leak detection in steady state and apparent loss coefficient of 0.157

First, the WDN was analyzed by SSR-LF method. Emitter coefficient for the whole WDN is equal to 58.4. To consider the constraints of Equation (5), which limits the sum of emitter coefficients, the tolerance value of 2 is assumed. In the analysis, the numbers of population and generation are considered equal to 50 and 5, respectively. At the beginning of the program, the value of K_N was considered between 0 and 58 with intervals of 1 for all zones. According to the method of search space reduction, in the next generations, for each zone, the search interval will be modified based on the local fitness method. According to the constraints of Equation (5), 362 repetitions are performed to obtain all 250 desired answers for five generations. With 362 repetitions, the best fitness is 0.654.

Then, with the previous conditions, the WDN is analyzed by GA. In GA, the value of the emitter coefficient in each node is considered as an unknown parameter. Also, according to the zoning and assigned leaks, the emitter coefficient was considered between 0 to 0.2, with intervals of 0.01. In order to increase the accuracy of calibration, the inflow to the WDN was also introduced to the program as an additional observation. The program is run with a maximum number of repetitions (100,000 times), with a minimum fitness of 0.817. Table 2 shows the results of the analysis.

The absolute optimum in optimization is obtained with a fitness value of zero ($F(X) = 0$). However, due to the zoning of the nodes and the search interval being discrete, it is not possible to choose the absolute optimal answer for the optimization programs in this problem. The first row in Table 2 shows the expected results when choosing the closest answer to the optimal point. In this case, the value of $E(d)$, which indicates proximity of the leakage coefficients to the optimal answer, has reached a small value of 0.35, which is the lowest achievable answer. Also, according to Table 2, fitness function $F(X)$ is equal to 1.03. In other words, the best answer in terms of leakage detection is obtained for a fitness value of 1.03. According to Table 2, the amount of fitness in local fitness optimization method and GA are equal to 0.654 and 0.817, respectively. Comparing the obtained fitness of the two methods with the amount of fitness obtained with absolute optimal, it can be said that both methods are able to adjust the pressure in the nodes and the flow in the pipes with a high degree of accuracy. Therefore, both methods are effective. Also, by comparing the values of $F(X)$ in the two methods, it is observed that the local fitness optimization method with few numbers of repetitions, was able to be more successful in optimization; on the other hand, by comparing $E(d)$, it is clear that the results have more convergence with real leakage values.

Also, by examining the results in different zones, it is observed that using the SSR-LF method, leaks are identified in all zones. In other words, the proposed method has better identified background leakages. This can be seen both by examining

Downloaded from http://iwa.silverchair.com/ws/article-pdf/22/4/4558/1105189/ws022044558.pdf
by guest

Table 2 | Results of WDN calibration analysis

Answers	$F(X)$	$E(d)$	K_N in zones 1	2	3	4	5	6
Discretized K_N	1.03	0.35	21	4	7	15	2	10
1	0.654	2.00	23	2	5	14	2	13
2	0.719	2.11	21	2	6	14	3	14
3	0.727	2.62	21	3	9	11	2	14
4	0.732	3.05	26	0	4	15	2	12
5	0.772	3.39	25	0	4	18	0	13
6	0786	3.59	26	1	2	15	1	14
7	0.789	2.60	23	1	7	11	2	13
8	0.807	2.87	23	6	2	13	2	14
9	0.808	1.96	24	3	4	13	1	11
10	0.819	2.22	25	2	4	15	2	11
AVR. Of 10 best	–	2.15	23.2	2.4	4.7	14	1.5	12.9
GA	0.817	4.43	28.44	0	2.06	16.32	0	13.68

answer 1 and considering the average of the 10 answers. But in terms of quantity, the background leakage was not well detected.

Another point to be made is that achieving a good fitness does not guarantee good leakage detection, and there may be several answers that regulate pressure and flow at the measured points. To further examine this issue, 10 answers of the local fitness method that had the best fitness are also presented in Table 2. The degree of fitness of these 10 analyses is obtained from 0.654 to 0.819. Note that *E(d)* of these answers is greater than the optimal answer and in fact, although they have better fitness, they are not better answers. Also, by examining column *E(d)*, it is seen that answer 9 is the best among 10 answers. Examining several answers with good fitness, as well as considering the best answer, can help to validate the results. If all good answers detect the same leaking zones, the results can be trusted. But if the answers show scattered, different zones, it is very likely that none of the answers are correct.

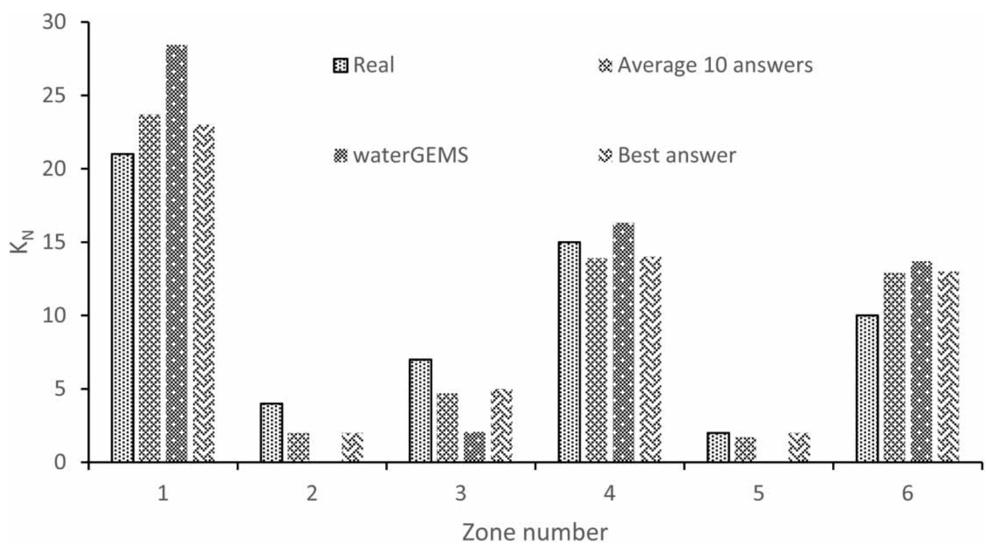

Figure 2 | The real and calculated value of the emitter coefficient in different zones.

Downloaded from http://iwa.silverchair.com/ws/article-pdf/22/4/4558/1105189/ws022044558.pdf
by guest

Figure 2 shows the real leakage coefficient values, the best answer and the average of the top 10 answers of SSR-LF and GA methods. It can be seen that in both methods, zones 1, 4 and 6 are correctly identified as leaking zones. They also show that SSR-LF results are better than GA.

3.3. Calculating apparent loss coefficient (C) using SSR-LF

In this stage of calibration, the same values of the previous stage are considered for the number of population and generation, as well as all the constraints of optimization. In order to calculate the apparent loss coefficient, first the calibration program with high and low limits of 0.1 and 0.3 and step length of 0.05 was analyzed. For each of the selected values for apparent losses, a set of answers is obtained with their corresponding fitness. Results are presented in Figure 3. It should be noted that the amount of fitness presented in Figure 3 is obtained from the average of the top 10 fitness. In this figure, the best fitness is 0.69, with $C = 0.15$.

By running the program in the range of 0.15 to 0.16, with an apparent loss coefficient of 0.155, a fitness value of 0.66 is obtained, which is very close to the actual value of 0.157.

3.4. Leak detection in simulation mode in extended time period (EPS) and known apparent loss coefficient equal to 0.157

At this stage, the WDN is analyzed in EPS. In this analysis, the regulated parameters are K_N values for the zones. For calibration information, pressure and flow values at the measuring points are used at 4-hour intervals (4 am, 8 am, 12 pm, 4 pm, 8 pm and 12 am).

Due to the fact that the WDN inflow fluctuates during 24 hours a day, the WDN consumption pattern should be determined during 24 hours. To calculate the WDN consumption pattern coefficient, the measured inflow on the selected day is used. Figure 4 shows these data. For this purpose, the consumption pattern coefficient during 24 hours is determined in

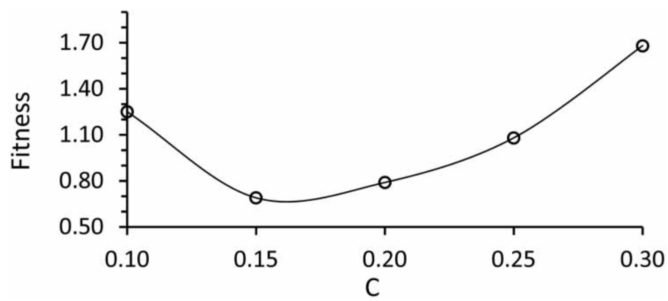

Figure 3 | Results from the selection of apparent loss coefficient.

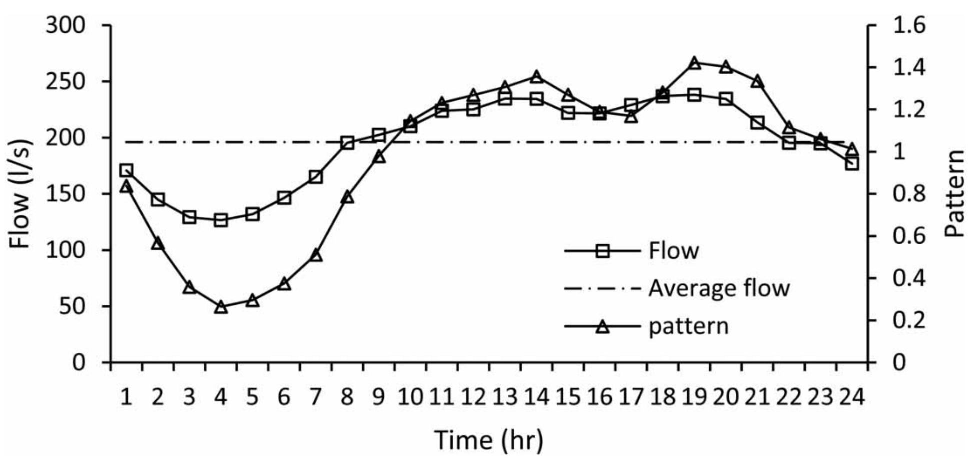

Figure 4 | Changes in inflow and hourly water consumption coefficients.

Downloaded from http://iwa.silverchair.com/ws/article-pdf/22/4/4558/1105189/ws022044558.pdf
by guest

such a way that the amount of outflow from the tank per hour corresponds to the diagram in Figure 4. Since the emitter coefficient and apparent losses are constant during 24 hours, this can be achieved by changing the consumption pattern coefficient. The given consumption pattern is also shown in Figure 4.

The local fitness optimization program (SSR-LF), similar to the previous state, was run with five generations and a population of 50 per generation. 1,806 repetitions were performed to find the optimal answer. The best answer has a fitness value of 0.649. Table 3 shows the10 best results.

Further calibration was performed in an EPS on GA. The maximum and minimum limits of the emitter coefficient were considered for each group of nodes and other regulatory parameters were considered the same as the previous stage. The GA is performed with 100,000 repetitions and the best fitness obtained in this mode is equal to 0.678. Results are presented in the last row of Table 3.

Table 3 | Results of the analysis obtained from calibration of the WDN with six zones in the extended time period mode

			K_N in zones					
Answers	$F(X)$	**E(d)**	**1**	**2**	**3**	**4**	**5**	**6**
Real	1.018	0.35	21	4	7	15	2	10
1	0.649	1.95	18	5	7	16	0	12
2	0.720	3.15	26	2	2	16	1	12
3	0.733	2.81	21	1	8	11	3	14
4	0.765	3.41	26	2	2	13	0	13
5	0.786	3.13	24	7	1	16	0	12
6	0.837	1.83	23	1	5	15	3	11
7	0.840	2.51	18	6	9	13	0	13
8	0.842	4.26	29	1	2	13	2	13
9	0.862	2.71	26	1	4	16	2	9
10	0.864	2.74	26	3	3	14	2	12
Avr.	–	1.89	23.7	2.9	4.3	14.3	1.3	12.1
GA	0.678	2.18	21.33	2.5	6.18	16.32	0	13.68

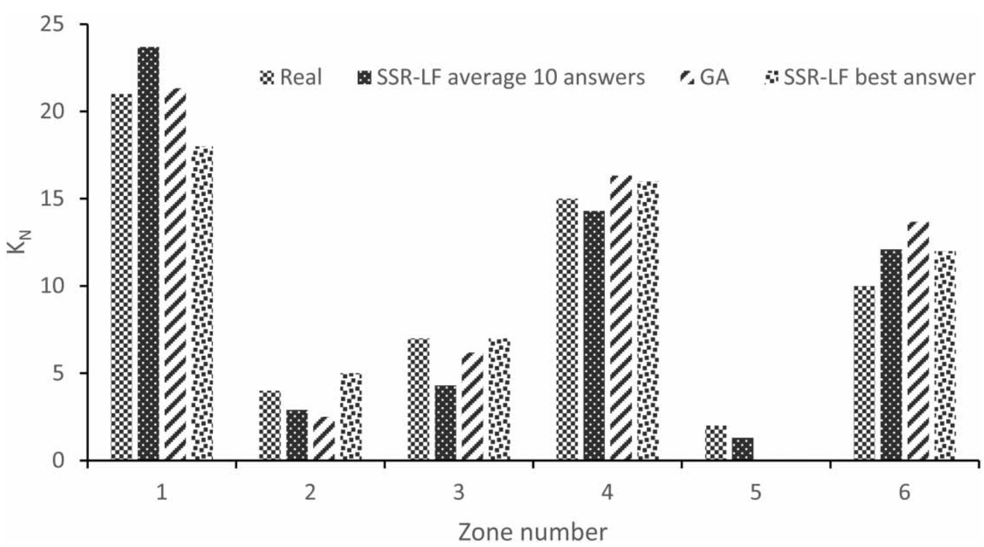

Figure 5 | Emitter coefficient for precise mode, local fitness and GA.

Downloaded from http://iwa.silverchair.com/ws/article-pdf/22/4/4558/1105189/ws022044558.pdf
by guest

Table 4 | Results of the analysis obtained from calibration of the WDN with 12 zones

Answers	*F(X)*	E(d)	K_N in zones 1-a	1-b	1-c	2	3	4-a	4-b	4-c	5	6-a	6-b	6-c
Real	0.818	0.314	17	2	2	4	7	1	11	2	2	7	2	1
1	0.239	2.73	13	5	0	4	11	5	6	2	3	8	0	0
2	0.383	2.53	14	3	5	2	6	4	6	5	1	9	3	0
3	0.389	2.43	12	4	2	5	8	1	8	7	3	9	0	1
4	0.408	2.46	11	2	2	9	8	1	9	2	3	9	0	0
5	0.420	3.12	13	1	11	5	5	1	8	4	2	7	0	1
6	0.446	4.04	12	0	14	2	5	1	9	1	3	9	3	0
7	0.459	3.17	12	4	1	9	4	2	7	7	1	9	0	4
8	0.490	2.14	12	3	5	6	7	1	8	1	1	9	4	1
9	0.496	2.43	14	4	7	3	4	5	10	1	1	9	1	1
10	0.517	1.68	15	3	0	7	9	0	10	4	0	8	1	2
Avr.	0.425	2.67	12.8	2.9	4.7	5.2	6.7	2.1	8.1	3.4	1.8	8.6	1.2	1
GA	0.262	4.94	13.65	1.83	2.13	11.25	2.06	12.39	2.04	0	1.44	7.76	0	3.12

The degree of fitness obtained from the top 10 answers is in the range of 0.649 to 0.864. In all answers, zones 1, 4 and 6 are correctly identified as leaky zones. Figure 5 shows the emitter coefficient in each zone compared to their actual values. Comparing the *F(x)* and *E(d)* obtained in steady state with EPS, it can be concluded that, in both methods, the EPS data provide better results.

3.5. Leak detection in steady state and apparent loss coefficient of 0.157 and 12 zones

According to the results obtained in the previous sections, in order to more accurately identify the location of large leaks, the areas in which the location of large leaks have been identified are divided into smaller areas. Figure 1 show the sub-zones in the WDS and Table 1 present the sub-zones characteristics.

Calibration was analyzed in the second stage with similar to the first stage with populations of 50 and 5 generations. SSR-LF program was answered after 365 repetitions. The results of the 10 answers that had the best fit are presented in Table 4. The fitness of 10 analyzes is from 0.239 to 0.517. Leak detection was also performed in GA with 10,000 replications and the best fitness was 0.262.

Table 4 shows that the SSR-LF was able to identify high leakage in all three areas, while the GA in one area does not have the appropriate accuracy. The value of E(d) also indicates that all 10 answers presented in SSR-LF are more accurate than GA result.

4. CONCLUSION

Leak detection is one of the most important issues related to WDNs. One of the main objectives of this research was to improve the existing leak detection methods based on calibration, which was achieved by improving the optimization methods. In this research, using an algorithm based on search space reduction and local fitness (SSR-LF), we studied calibration and leak detection of zone D of Birjand WDN. Two types of leakage including background leakage and hotspot leakage were considered for the WDN. Also, part of non-revenue water was assumed as apparent losses and was considered as another unknown parameter. To investigate the method, three leakage hotspots with different values were considered simultaneously. Simulations were performed in both steady-state and EPS and pressure-related leakage analysis was done. The results showed that, using local fitness in each zone, the proposed algorithm can identify leaky zones with a high accuracy. All three leakage hotspots were correctly identified and the amount of estimated leakage was highly accurate. Also, the proposed method was able to detect the background leakage in all zones and the amount of detected leakage was close to the actual values. Comparing the proposed optimization method with GA, it was observed that the proposed algorithm reached better answers with few numbers of repetitions. Comparison of the answers obtained to identify leakage hotspots and background

Downloaded from http://iwa.silverchair.com/ws/article-pdf/22/4/4558/1105189/ws022044558.pdf
by guest

leakages shows the superiority of the proposed method. The applied method was also able to calculate the apparent losses with very high accuracy, as it was another objective of this research. By correctly identifying and separating this amount from the total volume of non-revenue water, better results can be achieved in detecting leakage. The results showed that leak detection with EPS model has better results than the steady state. Moreover, it is shown that using multiple answers with relatively good fitness, instead of focusing on the best answer, can help validate the results.

ACKNOWLEDGEMENTS

The managers and experts of South Khorasan Water and Wastewater Company, especially Mr Moradi, who cooperated in the implementation of this research, are appreciated.

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

REFERENCES

Almandoz, J., Cabrera, E., Arregui, F., Cabrera, E. & Cobacho, R. 2005 Leakage assessment through water distribution network simulation. *Journal of Water Resources Planning and Management* **131** (6), 458–466. doi:10.1061/(ASCE)0733-9496(2005)131:6(458).

Amoatey, P. K., Bárdossy, A. & Steinmetz, H. 2018 Inverse optimization based detection of leaks from simulated pressure in water networks, part 1: analysis for a single leak. *Journal of Water Management Modeling* **26**, 1–10. doi:10.14796/JWMM.C460.

Bentley Systems, I. 2007 In: *Watergems v8 xm Users' Manual Bentley Systems Incorporated*. Bentley WaterGEMS V8 XM Edition User's Guide; Haestad Methods Solution Center: Watertown, CT, USA, 2007.

Cassidy, J., Barbosa, B., Damião, M., Ramalho, P., Ganhão, A., Santos, A. & Feliciano, J. 2020 Taking water efficiency to the next level: digital tools to reduce non-revenue water. *Journal of Hydroinformatics*. doi:10.2166/hydro.2020.072.

Che, T.-C., Duan, H.-F. & Lee, P. J. 2021 Transient wave-based methods for anomaly detection in fluid pipes: a review. *Mechanical Systems and Signal Processing* **160**, 107874. https://doi.org/10.1016/j.ymssp.2021.107874.

Colombo, A. F., Lee, P. & Karney, B. W. 2009 A selective literature review of transient-based leak detection methods. *Journal of Hydro-Environment Research* **2** (4), 212–227.

Duan, H.-F., Pan, B., Wang, M., Chen, L., Zheng, F. & Zhang, Y. 2020 State-of-the-art review on the transient flow modeling and utilization for urban water supply system (UWSS) management. *Journal of Water Supply: Research and Technology-Aqua* **69** (8), 858–893. doi:10.2166/aqua.2020.048.

Eryiğit, M. 2019 Water loss detection in water distribution networks by using modified clonalg. *Journal of Water Supply: Research and Technology-Aqua* **68** (4), 253–263. doi:10.2166/aqua.2019.139.

Hajibandeh, E. & Nazif, S. 2018 Pressure zoning approach for leak detection in water distribution systems based on a multi objective ant colony optimization. *Water Resources Management* **32** (7), 2287–2300. doi:10.1007/s11269-018-1929-1.

Huang, Y., Zheng, F., Kapelan, Z., Savic, D., Duan, H.-F. & Qingzhou, Z. 2020 Efficient leak localization in water distribution systems using multistage optimal valve operations and smart demand metering. *Water Resources Research* **56**. doi:10.1029/2020WR028285.

Karamouz, M., Tabesh, M., Nazif, S. & Moridi, A. 2005 Estimation of hydraulic pressure in water networks using artificial neural networks and fuzzy logic. *Journal of Water and Wastewater; Ab va Fazilab (in Persian)* **16** (1), 3–14.

Li, X., Chu, S., Zhang, T., Yu, T. & Shao, Y. 2021 Leakage localization using pressure sensors and spatial clustering in water distribution systems. *Water Supply*. doi:10.2166/ws.2021.219.

Meniconi, S., Brunone, B., Ferrante, M., Capponi, C., Carrettini, C. A., Chiesa, C., Segalini, D. & Lanfranchi, E. A. 2015 Anomaly pre-localization in distribution-transmission mains by pump trip: preliminary field tests in the Milan pipe system. *Journal of Hydroinformatics* **17** (3), 377–389.

Meniconi, S., Capponi, C., Frisinghelli, M. & Brunone, B. 2021 Leak detection in a real transmission main through transient tests: deeds and misdeeds. *Water Resources Research* **57** (3), e2020WR027838.

Moasheri, R. & Jalili-Ghazizadeh, M. 2020 Locating of probabilistic leakage areas in water distribution networks by a calibration method using the imperialist competitive algorithm. *Water Resources Management* **34** (1), 35–49. doi:10.1007/s11269-019-02388-4.

Nasirian, A., Maghrebi, M. F. & Yazdani, S. 2013 Leakage detection in water distribution network based on a new heuristic genetic algorithm model. *Journal of Water Resource and Protection* **05** (03), 10. doi:10.4236/jwarp.2013.53030.

Pan, B., Capponi, C., Meniconi, S., Brunone, B. & Duan, H. F. 2021 Efficient leak detection in single and branched polymeric pipeline systems by transient wave analysis. *Mechanical Systems and Signal Processing* **162**, 108084.

Pudar, R. S. & Liggett, J. A. 1992 Leaks in pipe networks. *Journal of Hydraulic Engineering* **118** (7), 1031–1046. doi:10.1061/(ASCE)0733-9429(1992)118:7(1031).

Soares, A. K., Covas, D. I. C. & Reis, L. F. R. 2010 Leak detection by inverse transient analysis in an experimental PVC pipe system. *Journal of Hydroinformatics* **13** (2), 153–166. doi:10.2166/hydro.2010.012.

Downloaded from http://iwa.silverchair.com/ws/article-pdf/22/4/4558/1105189/ws022044558.pdf
by guest

Sophocleous, S., Savić, D. A., Kapelan, Z. & Giustolisi, O. 2017 A two-stage calibration for detection of leakage hotspots in a real water distribution network. *Procedia Engineering* **186**, 168–176. https://doi.org/10.1016/j.proeng.2017.03.223.

Wu, Z. Y. & Simpson, A. R. 2001 Competent genetic-evolutionary optimization of water distribution systems. *Journal of Computing in Civil Engineering* **15** (2), 89–101. doi:10.1061/(ASCE)0887-3801(2001)15:2(89).

Wu, Z. Y., Sage, P. & Turtle, D. 2010 Pressure-dependent leak detection model and its application to a district water system. *Journal of Water Resources Planning and Management* **136** (1), 116–128. doi:10.1061/(ASCE)0733-9496(2010)136:1(116).

Zhang, Q., Zheng, F., Duan, H.-F., Jia, Y., Zhang, T. & Guo, X. 2018 Efficient numerical approach for simultaneous calibration of pipe roughness coefficients and nodal demands for water distribution systems. *Journal of Water Resources Planning and Management* **144** (10), 04018063. doi:10.1061/(ASCE)WR.1943-5452.0000986.

First received 25 September 2021; accepted in revised form 18 February 2022. Available online 2 March 2022

Downloaded from http://iwa.silverchair.com/ws/article-pdf/22/4/4558/1105189/ws022044558.pdf
by guest