Flow Measurement and Instrumentation 80 (2021) 101995

Contents lists available at ScienceDirect

Flow Measurement and Instrumentation

journal homepage: www.elsevier.com/locate/flowmeasinst

# Leakage detection in water networks by a calibration method

Reza Moasheri [a], Mohammadreza Jalili Ghazizadeh [a,*], Mohammadreza Tashayoei [b]

[a] *Faculty of Civil, Water and Environmental Engineering, Shahid Beheshti University, Tehran, Iran*
[b] *Water and Wastewater Company of District 1, Tehran, Iran*

ARTICLE INFO

*Keywords:*
Water supply networks
Leakage detection
Calibration
Imperialist competitive algorithm
Hydraulic model
Field data

ABSTRACT

This paper presented a new method of locating the probable leakage zones in the water distribution networks. The proposed method has consisted of two steps: first, a zone with the most leakages is identified by analysing the results of the pressure metering at different parts of the network. In the second step, the leaky zone is divided into some virtual zones. According to the field pressure or flow metering results in the network, the simultaneous calibration of the nodal demand and the pipes' roughness coefficients are dealt with by the Imperialist Competitive Algorithm (ICA) at different hours of the day. Considering the implementation results at any time, the probability of the leakage in each virtual sub-zone and the roughness coefficient of the pipes are estimated, simultaneously. The proposed method was implemented on two hypothetical networks and a real network. The results showed that this method could prioritize the leaky zones with good accuracy. The proposed method can be used by the water utilities for leak detection in water distribution networks.

## 1. Introduction

Water is a valuable resource for human survival. Despite the abundance of water on Earth, only about 0.5% of all the water on our planet is accessible [1]. Water is delivered to customers by water distribution networks (WDNs) and water transmission mains. WDNs play a significant role for individual and industrial customers [2]. Instead, a portion of water is lost and it does not reach users. Non-Revenue Water (NRW) is expressed by the International Water Association (IWA) as one of the important evaluation indexes of the water systems [1]. This index includes the real losses, apparent losses, and unbilled authorized consumption [3]. Leakage is an important part of real losses in WDNs and causes a lot of energy and water losses [4]. Each year, more than 32 billion $m^3$ treated water is lost through WDNs leakage [5]. It is widely between networks and countries from low as 3–7% in the well-maintained networks to 50% of input in some undeveloped countries [6], and its global average is about 35% [5]. In general, the leakage detection methods could be divided into physical and simulation-field methods [7]. Methods such as thermography, tracer gas, acoustic equipment, video inspection, and ground-penetrating radar [8] are physical leakage detection methods.

Because of the need for expert forces, timing, and costliness of the common leakage detection methods, recent studies have focused on determining the leakage locations or zones by simulation-field methods. One of these methods is the calibration of the hydraulic model of the networks [7].

There are different approaches to calibration, and the most common one is the use of implicit models. In this approach, an objective function is defined as the field data (by the pressure or flow metering) and the simulation data (results of the hydraulic network model) differences. Then, it is usually minimized using an optimization algorithm. In the calibration method, the amount and location of leakages are estimated by considering them as one of the adjusted parameters. The following are some of the studies that have been carried out on this topic.

Di Nardo, Di Natale, Gisonni and Iervolino [9] presented a calibration method to estimate the leakage parameters and the hourly water consumption pattern. In this study, a Genetic Algorithm (GA) was used to optimize the objective function. This method was investigated on a part of the water network of Pozzuoli (located in Italy). The results showed that the amount of the leakage was 18%.

Sousa, Ribeiro, Muranho and Sá Marques [10] proposed a method for detecting leakages in the water supply networks with Simulated Annealing (SA) algorithm and graph theory. In this method, first, using previously measured field data (before leakages), the hydraulic model of the network is calibrated by the SA algorithm. Then, to locate leakages, using new field data of pressure meters (related to the leaky state of the network), the network model nodal demand is calibrated using the SA algorithm. The proposed method was investigated on a water network

* The corresponding author.
*E-mail addresses:* r_moasheri@sbu.ac.ir (R. Moasheri), m_jalili@sbu.ac.ir (M.J. Ghazizadeh), tashayoei@t1ww.tpww.ir (M. Tashayoei).

https://doi.org/10.1016/j.flowmeasinst.2021.101995
Received 24 November 2020; Received in revised form 27 May 2021; Accepted 7 June 2021
Available online 11 June 2021
0955-5986/

for simultaneous one and two leakages. The overall results indicated that this method was satisfactory.

Hajibandeh and Nazif [11] used the multi-objective ant colony algorithm to detect the location and the amount of leakage in WDNs. This method was investigated on the Rossman network [12].

Moasheri and Jalili-Ghazizadeh [7] proposed a method for the simultaneous calibration of the pipes' roughness and the leakages of the hydraulic model of the water distribution networks. In the proposed method, the leakage was considered as the additional demand of the node. After grouping the pipes, to determine the roughness coefficients of each pipe group, using the Imperialist Competitive Algorithm (ICA), a metaheuristic optimization algorithm, the simulated and measured flows and pressures difference were minimized. The output of this method was finding the pipes' roughness coefficients and the probabilistic estimation of the zones suspected to the leakage in the network. The proposed method was investigated on a hypothetical network and a real network. The overall results showed that the method could estimate the roughness coefficients of the pipes with low error, and it was able to prioritize the network leakage zones correctly.

The other related studies are: Wu and Sage [13]; Faghfoor Maghrebi, Hasanzadeh and Yazdani [14]; Sophocleous, Savic, Kapelan, Shen and Sage [15]; Sophocleous, A. Savić, Kapelan and Giustolisi [16]; Shekofteh, jalili Ghazizadeh and Yazdi [17]; Rajeswaran, Narasimhan and Narasimhan [18]; Tabesh, Jamasb and Moeini [19].

Most of the previous methods have a limitation on the number of simultaneous leakages, while it is unknown in the real networks. Also, a

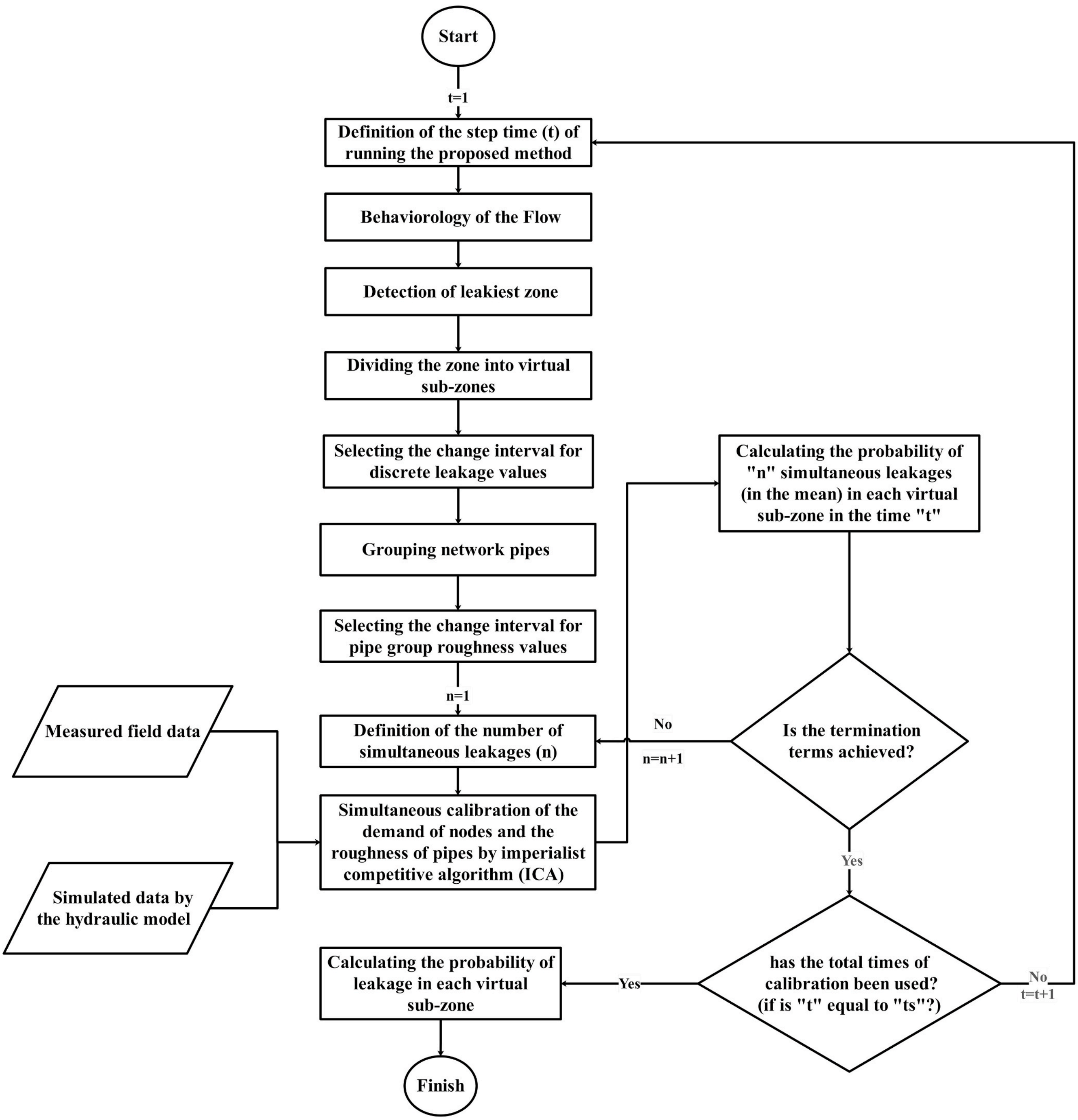


**Fig. 1.** The flowchart of the proposed method.

large part of the proposed method is not investigated on a real network. In such studies, it is assumed that the only parameter that causes the difference between field and simulation data (in the simulation model) is the leakages. However, this assumption is not necessarily correct and there are many uncertainties about the parameters of the hydraulic model in reality.

In this paper, a new method of locating the leakages in a water distribution network has been presented, detecting the leaky zones by analysing the field data and conducting the calibration process. This method has been investigated on the two hypothetical networks and a real network.

## 2. Material and methods

In the present study there are two main steps for the leakage detection of the water networks. In the first step, called the flow Behaviorology step, considering the field pressure metering results from the different parts of the network, the zone with the highest amount is identified. In the second step, to reduce the volume of the leakage detection operation, the leakage zone (reported from the first step) is divided into several virtual sub-zones. Then, the simultaneous calibration of the nodal demand and the pipes' roughness coefficients of the hydraulic network model is carried out by the ICA. Finally, the possibility of the leakage is evaluated in each virtual sub-zone. Unlike most previous studies, this method has no limit on the number of simultaneous leakages of the network, and it probabilistically evaluates the leakage percentage in each zone. While the previous methods require a calibrated hydraulic model of the network, in this method, the uncertainty in the roughness of the network pipes is considered as an important calibration factor. Also, implementing the Behaviorology step reduces the search space in the calibration step leading to an increase in the speed of the calculations and thus the cost reduction, especially in the wide networks.

Fig. 1 depicts the general flowchart of the proposed algorithm.

### 2.1. Behaviorology of the flow

The first step of the proposed method involves the behaviorology of the flow.

The purpose of the behaviorology is to analyze the results of the pressure parameter related to before and after the occurrence of leakages in the network. The leakiest zone is found by placing several pressure meters in the network and dividing the network into some zones. The higher the number of the pressure meters in the network, the smaller the leaky zones would be and this reduces the cost of leakage detection in that zone.

It should be noted that one of the main goals of a water network utility is to locate such a zone in the network. Therefore, this step leads to achieving the objective.

The authors first introduced this method in 2019 ([20]). Its general algorithm in a network with any dimension and any number of the nodes is as follows [20]:

(1) Selecting at least three nodes from a certain branch of the network (the first, internal, and end nodes of that branch) as the field pressure metering locations. Each branch means a set of network nodes connected in a certain direction by a group of pipes.
(2) Calculation of the absolute value of hydraulic grade line slope is done based on the computer simulation results of the network in the leakage-free state between the consecutive points (or nodes) which are selected for the field pressure metering locations. At this step, a matrix called the computational hydraulic grade line slope is formed ($A((r-1)\times 1)$), where r is the total number of points at which the pressure metering would be performed.
(3) Implementing the pressure metering in the field pressure measurement locations.
(4) Calculating the hydraulic line using the field data of the third step in the field pressure metering locations.
(5) Calculating the absolute value of hydraulic grade line slope based on the results of the fourth step; between both consecutive points of the field pressure measurement locations in the network. As a result, a matrix called the field hydraulic grade line slope is formed ($B((r-1)\times 1)$).
(6) Calculating the ratio of the changes of the hydraulic grade line slope in the presence of the leakage(s) state to the absence of leakage (s) state for the pipes in the determined branch (first step), between both of the consecutive locations of the field pressure measurement. In other words, at this step, matrix B is divided by matrix A, and the result would be matrix K.
(7) Analysis of the results based on the following points:
    - In both of the consecutive entries of hydraulic grade line slope ratio matrix (matrix K), where the ratio of the second entry to the first entry is less than 1, the zone corresponding to the first entry of the two entries could be the leaky zone.
    - If the value of the last entry of hydraulic grade line slope ratio matrix (matrix K) is more than one, the network has leakage in the corresponding zone with the last entry of hydraulic grade line slope ratio matrix or a zone outside the defined branch of the network's last node.
    - If the network has leakage, but the hydraulic grade line slope ratio matrix entries have a value equal to one, the network has leakage in the zone outside the defined branch to its first node.

Due to the hydraulic complexity of the looped networks, through the proposed method, the zone of the most leakage(s) in the defined direction (which could be the zone where all the network leakages are in it) could be determined. Thus, the nodes suspected to the leakage (based on the mentioned points), corresponding to the two consecutive entries of the hydraulic grade line slope ratio matrix, in which the change in the value of the second entry relative to the value of the first entry or the change in the value of its last entry (provided that the value is more than one) is maximal compared to one, they could determine the zone of the most leakage(s) in the defined direction.

In this step, the Monte Carlo simulation method, first discussed by Metropolis and Ulam [21], has been used to apply the uncertainty of the error of the measurement devices. Thus, to calculate the hydraulic grade line slope ratio matrix (matrix K), there was a change in the field pressure values (randomly and based on the uncertainty interval) to a certain number (for example, 10,000 times). After calculating the mentioned matrix in each iteration, their average is considered as the final value of the matrix K.

### 2.2. Network calibration

The general method used at this step is the same as the one proposed in the authors' previous study in 2020 ([7]).

In the first stage, the reported zone of the previous step is divided into several sub-virtual zones, according to the geographical location of the network pipes and the measurement field tools. Then, assuming the leakage from the network nodes, the calibration of the node demands and the roughness of the pipes is done by the ICA.

ICA was first introduced by Atashpaz-Gargari and Lucas in 2007 [22]. It is a metaheuristic optimization algorithm which has been used to solve many optimization problems [23–29]. It has shown the best performance in global optimum achievement and convergence rate. More details about the algorithm are found in the reference No [7].

The objective function of the optimization algorithm is defined as Eq. (1). In the calibration process, it is first assumed that there is a single leakage in the network. The calibration process is carried out with this assumption, and the leakage probability of each sub-zone is calculated.

Then, assuming the two simultaneous leakages in the network, the calibration process is conducted. As a weighted average (due to one or two simultaneous leakages), the leakage probability of each sub-zone is calculated. In each iteration, one unit is added to the number of possible simultaneous leakage in the network. Generally, in the tth iteration, the network is calibrated with the assumption of "t" simultaneous leakage, and the leakage probability is estimated in each sub-zone as the weighted average (caused by 1 to "t" simultaneous leakage). This process is continued until changing less than 5% of the leakage probability of each sub-zone in the tth iteration compared to the "t-1″th iteration. In the proposed method, to reduce the dimensions of the search space of ICA, the optimization process and estimation of the values of the adjusted parameters (leakage and roughness coefficient of pipes) are performed in a discrete space. More details of this approach are available in the reference No [7].

$$f(x) = minimize\left(\frac{\sum_{np=1}^{NP}\left(\frac{Pobs_{np}}{\sum Pobs_{np}}\right)\left|\frac{Psim_{np}-Pobs_{np}}{Pobs_{np}}\right| + \sum_{nf=1}^{NF}\left(\frac{Fobs_{nf}}{\sum Fobs_{nf}}\right)\left|\frac{Fsim_{nf}-Fobs_{nf}}{Fobs_{nf}}\right|}{NP+NF}\right) \tag{1}$$

In this Equation, $f(x)$ is the objective function, $np$ is the counter of network pressure meters; $NP$ is the total number of pressure meters; $Psim_{np}$ is the simulated pressure of the $np^{th}$ node; $Pobs_{np}$ is the field measured pressure of the $np^{th}$ node, $nf$ is counter of network flow meters, $NF$ is the total number of flow meters, $Fsim_{nf}$ is the simulated flow of the $nf^{th}$ pipe, $Fobs_{nf}$ is the field measured flow of the $nf^{th}$ pipe.

It should be noted that the Monte Carlo method has been used to apply the uncertainty of the measurement devices as well as the base nodal demands in the hydraulic model of the network. Thus, in the optimization stage with the ICA, a random change (based on the uncertainty interval) was made to a certain number (for example, 10,000 times) in the field pressure or flow values, and their average is assumed to be the field measurements. Similarly, this process has been implemented to apply the uncertainty of the node base demands in the network's software model.

Also, considering the stochastic nature of the ICA, the proposed method is implemented at different times, and according to the results of each time, the probability of the leakage is estimated as a weighted average. The more field information is available at more times; the more reliable the output result would be.

Eq. (2) is proposed, to calculate the probability of the leakage incidence in each virtual sub-zone of the leaky zone (reported from the behaviorology step).

$$P_z = \frac{\sum_{s=1}^{ts} P_{z,s}\ e^{-ave(Cost_s)}}{\sum_{s=1}^{ts} e^{-ave(Cost_s)}} \tag{2}$$

where:

$P_z$: the probability of the leakage in the Zth virtual sub-zone (in the leakiest zone)
S: Counter of the times (s = 1, …, ts)
ts: The total number of times that field data is available
$P_{z,s}$: the probability of the leakage in the Zth virtual sub-zone at the time s
$ave(Cost_s)$: The average values of the objective function of the ICA at the time s

Eq. (3) is used to calculate the probability of the leakage of the Zth virtual sub-zone in each iteration at any time.

$$P_{z,s,t} = \frac{\sum_{T=1}^{t} \frac{(TLZ)_{z,s,T}}{Q_{s,tz}}\ e^{-Cost_{s,T}}}{\sum_{T=1}^{t} e^{-Cost_{s,T}}} \tag{3}$$

Where:

$P_{z,s,t}$: The probability of the leakage of the Zth virtual sub-zone in the tth iteration at the time s
$(TLZ)_{z,s,T}$: The total estimated leakage of the Zth virtual sub-zone in the $T^{th}$ iteration at the time s
$Q_{s,tz}$: sum of The leakage flow of total sub-zones at the time s
$Cost_{s,T}$: The value of the objective function of the ICA in the Tth iteration at the time s

In the proposed method, the grouping of network pipes should be done. For this purpose, the pipes that have the same age and material can be placed in the same group so that during the calibration process, the roughness of all pipes in each group changes equally. This assumption reduces the space search of the optimization algorithm. So, the computation speed would be increased.

To calculate the roughness of each pipe group, and the weighted average of the roughness of the pipe group at the different times, Eq. (4) is used:

$$C_j = \frac{\sum_{s=1}^{ts} C_{j,s}\ e^{-ave(Cost_s)}}{\sum_{s=1}^{ts} e^{-ave(Cost_s)}} \tag{4}$$

where;

$C_j$: The corrected roughness coefficient of the jth pipe group.
$C_{j,s}$: The roughness coefficient of the jth pipe group at the time s.

Eq. (5) is used to calculate the roughness coefficient of the jth pipe group in tth iteration at the time s ($C_{j,s,t}$)

$$C_{j,s,t} = \frac{\sum_{T=1}^{t} C_{j,s,T}\ e^{-Cost_{s,T}}}{\sum_{T=1}^{t} e^{-Cost_{s,T}}} \tag{5}$$

In the implementation step of the calibration method with the ICA, it is necessary to know the total amount of leakage in the search zone of that algorithm (the zone of most leakages). Considering the total number of the pressure meters in the flow behaviorology step, if the number of the networks is divided into the p zone, the amount of the total leakage in the zone with the most leakages is in the $[\frac{Q_{L,s}}{p}, Q_{L,s}]$ interval, where $Q_{L,s}$ is the total amount of the leakage in the network at the time s. In the proposed method, to estimate the amount of the total leakage in the zone, the average value in the mentioned interval based on Eq. (6) has been used.

$$Q_{s,tz} = \frac{p+1}{2p} Q_{L,s} \tag{6}$$

On average or at the time of the customers' minimum night flow, the total amount of the leakage could be estimated in different ways. Since in the proposed method, it is necessary to know the total amount of the leakage at different times (to determine the search range of the optimization algorithm), the FAVAD equation [30] with the known leakage exponent is used to calculate it at those times. The equation is a relationship between network pressure and leakage parameters.

If the method of using the water balance table has been used to determine the total amount of the leakage, Eq. (7) is suggested to determine its corresponding amount at different times (based on the modified form of the FAVAD equation)

$$\frac{Q_{L,s}}{Q_{ave.day}} = \left(\frac{P_{ave.s}}{P_{ave.day}}\right)^N \tag{7}$$

where:

$Q_{ave.day}$: The average daily flow of the total leakage
$P_{ave.s}$: The average pressure of the network at the time s
$P_{ave.day}$: The average daily pressure of the network
$N$: The leakage exponent in the FAVAD equation

To calculate the average pressure of the network at any time, the results of the pressure-metering in the network at the Average Point Height (APH) could be used to calculate the Average Zone Point (AZP) [31] at that time.

Eq. (8) is suggested to calculate the average daily pressure (Pave. day):

$$P_{ave.day} = \frac{\sum_{s=1}^{ts} P_{ave.s}}{ts} \tag{8}$$

It should be noted that in practice, by measuring the pressure at the APH and based on Eq. (7), the total amount of the leakage flow is obtained at any time. In the hypothetical examples, the simulation must be done in the hydraulic model. The simulated pressures should be used as the field pressures, whereas, in simulation, the amount of the leakage flow of the nodes at the different times is unknown. A computer program was written to solve this problem. This program computes the amount of AZP and the leakage flow at different times (based on Eq. (9)) by the iterative solution method.

$$\frac{Q_{i,s}}{Q_{i,ave.day}} = \left(\frac{P_{ave.s}}{P_{ave.day}}\right)^N \tag{9}$$

$Q_{i,s}$: The leakage flow of the ith node at the time s
$Q_{i,ave.day}$: The average daily leakage of the ith node

If the Minimum Night Flow (MNF) method is used to estimate the total amount of the leakage, Eq. (10) is used to calculate the total amount of leakage at any time (based on the modified form of the FAVAD equation). Similarly, in the hypothetical examples, Eq. (11) is used to calculate the nodes' leakage in the software model of the network.

$$\frac{Q_{L,s}}{Q_{MNF}} = \left(\frac{P_{ave.s}}{P_{ave.MNF}}\right)^N \tag{10}$$

$$\frac{Q_{i.s}}{Q_{i,MNF}} = \left(\frac{P_{ave.s}}{P_{ave.MNF}}\right)^N \tag{11}$$

wherein, $Q_{MNF}$, $Q_{i,MNF}$ and $P_{ave.MNF}$ are the total leakage of the network, the leakage of ith **node, and** the average pressure of the network at the time of MNF, respectively.

It should be noted that if the existing network is a total irregular looped type, it is recommended to use the calibration step directly without conducting the behaviorology step. Because, many pressure meters are needed to analyze each branch in the behaviorology step, and it might be challenging to determine the total flow of the leakage of that branch (which should be clear in the calibration step). In this case, instead of Eq. (2) and (3), Eq. (12) and (13) are recommended, respectively.

$$P_Z = \frac{\sum_{s=1}^{ts} P_{Z,s}\, e^{-ave(Cost_s)}}{\sum_{s=1}^{ts} e^{-ave(Cost_s)}} \tag{12}$$

$$P_{Z,s,t} = \frac{\sum_{T=1}^{t} \frac{(TLZ)_{Z,s,T}}{Q_s}\, e^{-Cost_{s,T}}}{\sum_{T=1}^{t} e^{-Cost_{s,T}}} \tag{13}$$

where:

$P_Z$: The probability of the leakage in the Zth virtual zone
$P_{Z,s}$: The probability of the leakage in the Zth virtual zone at the time of s
$P_{Z,s,t}$: The probability of leakage in the Zth virtual zone in the tth iteration at the time s
$(TLZ)_{Z,s,T}$: The estimated total leakage of the Zth virtual zone in the tth iteration at the time s

In this study, The EPANET programmer's toolkit has been used for hydraulic analysis of the network. The proposed method has been developed in Matlab R2017R software.

## 3. Results and discussion

In this part, the proposed method is examined in the two case studies. The first case study includes model networks, while the second one includes a real network. In these networks, the relevant parameters of the proposed method are presented in Table 1.

### 3.1. Case study one

This case study includes two hypothetical networks, which are described in what follows.

#### 3.1.1. The first network

This network (Fig. 2) which the authors have studied in their previous research ([7]), includes 100 pipes, 76 nodes, 3 pressure reducing valves, a tank, and a reservoir. In the steady-state analysis, the total inlet flow and junction demands were 80.88 L/s, 38.55 L/s, respectively, and the total length of pipes of the network was 26.4 km. Also, the Hazen-Williams coefficient of all the pipes was 130. Due to the Hazen-Williams coefficient's simple estimation, the Hazen-Williams equation is more widely used than the other head loss equations for the engineering problems (such as Darcy-Weisbach and Manning equations).

According to the material and age, it is assumed that the network pipes could be divided into 5 groups. In Fig. 2, the first group has been shown with blue, the second group with purple, the third with black, the fourth with red, and the fifth with orange.

To implement the proposed method, the results of the three different times, including midnight, 6 a.m., and 6 p.m., have been used in this network with the demand pattern coefficients of 0.5, 1, and 1.6, respectively. In the Monte Carlo method, the error of measuring devices and the uncertainty in the base demand of the nodes in the computer model of the network have been considered equal to 10% interval of their real values and the number of iterations equal to 10,000 times.

As mentioned before, to calculate the average pressure of a network, the pressure must be measured at the APH. One way to estimate the APH is to divide a network into several elevation zones and calculate the average elevation of that network by weight (according to the number of customers in each zone).

In this network, assuming the equal distribution of the subscribers, the APH is obtained from the average height of the nodes. So, in the mentioned network, the pressure should be measured at the height level of 124.60 m. A yellow star in Fig. 2 indicates its location.

In the behaviorology step, nodes 3, 5, 7, and 8 have been considered as the nodes with pressure meters. Note that in the previous study [7], this network had been examined in the case of installing ten flow meters in the network.

As it was previously explained, one of the objectives of combining

**Table 1**
The used parameters of the proposed algorithm for the case studies.

| Parameter | Value |
|---|---|
| Number of imperialists | 10 |
| Number of colonies | 40 |
| Maximum number of iterations (in ICA) | 100 |
| $\alpha$ | 1 |
| $\beta$ | 2 |
| $\Delta q$ (lit /s) | 0.1 |
| $\Delta C$ | 1 |
| $\zeta$ | 0.1 |
| $nC$ | 13 |

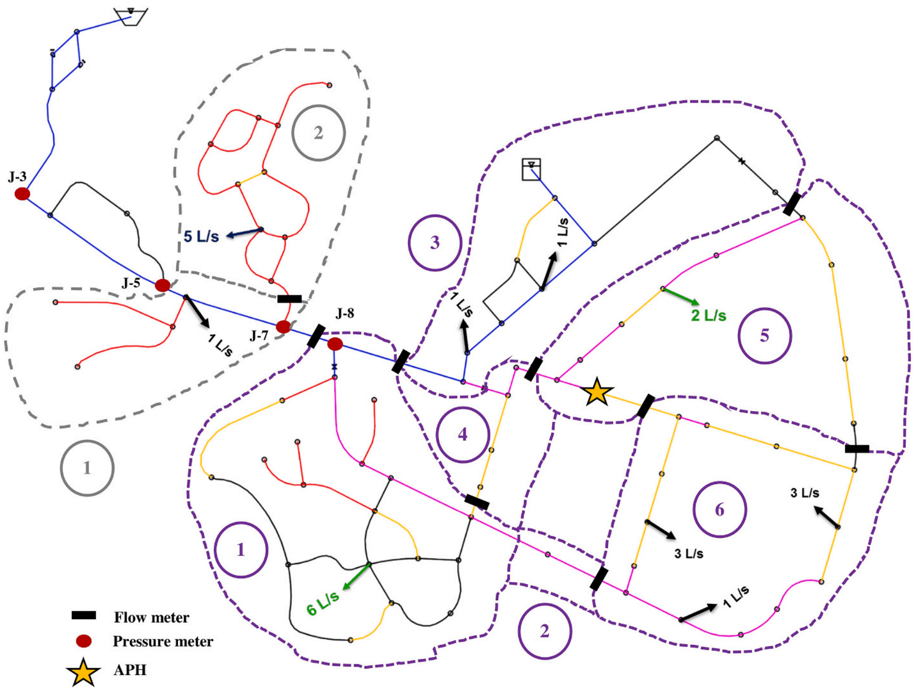


**Fig. 2.** The first network of the first case study.

the behaviorology methodology with the calibration algorithm is reducing the number of field measurement tools required during the calibration step. Also, if it becomes clear that based on the network pressure meters' installation location, the most leakages in this network are before the J-8 node. This result eliminates the need to install the pressure or flow meters in the post-J-8 zone to create the virtual sub-zones. This is economically significant.

The proposed method has been studied in the three different scenarios in this network, including the one, two, and six simultaneous leakages. The hydraulic model of the network was used to simulate each scenario and to obtain the field data.

- Studying a leakage

To determine the capability of this method for detecting a leakage, it has been assumed that a leakage in the network has occurred at an average daily flow of 5 L/s. This leakage has been marked in Fig. 2 with the blue arrow at the mentioned node.

The leakage of this node was evaluated at midnight (12), 6 a.m. and, 6 p.m. as 10.04, 5.00, and 3.11 L/s, respectively (based on Eq. (9)).

The results of the behaviorology method have been presented in the first row of Table 2. Based on this analysis, because only the ratio of the third to the second entry of the hydraulic grade line slope matrix is less than one, the nodes corresponding to the second entry of the matrix (i.e., the zone of the network between J-5 and J-7) have the most leakage amounts (in this example, the leaky zone). Considering the location of the flow meter installed in this zone (which has been shown with the black rectangle in Fig. 2), it has been divided into the two virtual sub-zones, whose zones have been shown in the grey dashed lines in Fig. 2.

By implementing the calibration algorithm for this zone, the probability of the leakage in each sub-zone and the Hazen-Williams roughness coefficient of each pipe group have been estimated, and the results have been presented in Tables 3 and 4, respectively. Based on the results of the probability of the leakage in each sub-zone (Table 3), the algorithm correctly estimated the probability (and consequently the real percentage) of the leakage in each sub-zone. The results in Table 4 show the mean difference of 4.6% of the estimated Hazen-Williams coefficient of each tube group with its actual value.

- Studying the two simultaneous leakages

The leakages with the daily average flow of 6 and 2 L/s at the two different nodes have been considered to study the two simultaneous leakages in the network. They are marked with the green arrow in Fig. 2.

After implementing the proposed method, the results of the behaviorology step have been presented in the second row of Table 2.

Based on the second row of Table 2, at midnight (12) and 6 a.m., only the last entry of the hydraulic grade line slope ratio matrix was more than one. So, according to the process of analysis described in the flow behaviorology step (section 3-3), the most leakages (here all of them) were in the zone corresponding to the last entry of the matrix up to the zone outside of the defined direction (the zone between J-3 and J-8); i.e., part of the network that is located after the node J-7.

Also, based on the results of the mentioned matrix at 6 p.m., the ratio of the third entry to the second entry was less than 1. However, its last entry was more than 1. Based on what was mentioned in section 2.1, because the difference of 1 from the last entry of this matrix (1–1.0014 = 0.0014) was more than the difference of the second entry from the third entry (1.0016–1.0014 = 0.0002), the zone after node J-7 showed the zone of the most leakages. Also, in reality, all the leakages have been located in the zone after the node J-7, so the algorithm has correctly identified their zones.

**Table 2**
The results of the hydraulic grade line slope ratio matrix (K) for the first network of the case study one.

| | Initial and end nodes of hydraulic grade line | 3 & 5 | 5 & 7 | 7 & 8 |
|---|---|---|---|---|
| First scenario | Midnight | 1.0028 | 1.0031 | 0.8955 |
| | 6 a.m. | 1.0009 | 1.001 | 0.9462 |
| | 6 p.m. | 1.0011 | 1.0012 | 0.9663 |
| Second scenario | Midnight | 1.0011 | 1.0013 | 1.0014 |
| | 6 a.m. | 1.0000 | 1.0004 | 1.0022 |
| | 6 p.m. | 1.0013 | 1.0016 | 1.0014 |
| Third scenario | Midnight | 1.0004 | 1.0006 | 1.0007 |
| | 6 a.m. | 1.0007 | 1.0009 | 1.0009 |
| | 6 p.m. | 1.0006 | 1.0008 | 1.0009 |

**Table 3**
Results of the leakage probability of different virtual sub-zones for the first network of the case study one.

| | Virtual sub-zone | Probability of leakage | | | | Real leakage percentage |
|---|---|---|---|---|---|---|
| | | Midnight | 6 a.m. | 6 p.m. | Weighted average | |
| First scenario | **1** | **100** | **100** | **100** | **100** | **100** |
| | 2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Second scenario | **1** | **84.7** | **71.8** | **73.6** | **76.7** | **75.0** |
| | 2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | 3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | 4 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | **5** | **15.3** | **28.2** | **26.4** | **23.3** | **25.0** |
| | 6 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Third scenario | 1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | 2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | **3** | **0.0** | **1.0** | **1.0** | **0.6** | **22.2** |
| | 4 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | 5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | **6** | **100** | **99.0** | **99.0** | **99.4** | **77.8** |

**Table 4**
Results of estimated roughness of each pipe group for the first network of the case study one.

| Pipe group number | Scenario | | |
|---|---|---|---|
| | First | Second | Third |
| 1 | 127 (2.3%) | 125 (3.8%) | 119 (8.5%) |
| 2 | 124 (4.6%) | 126 (3.1%) | 119 (8.5%) |
| 3 | 121 (6.9%) | 122 (6.2%) | 124 (4.6%) |
| 4 | 125 (3.8%) | 121 (6.9%) | 127 (2.3%) |
| 5 | 123 (5.4%) | 122 (6.2%) | 127 (2.3%) |
| Averaged error (%) | 4.6 | 5.2 | 5.2 |

Considering 8 flow meters in the network, the zone reported from the previous step has been divided into 6 virtual sub-zones (those zones were shown in Fig. 2 with a purple dashed line). After implementing the calibration method, the results of estimating the probability of the leakage in each sub-zone as well as the roughness coefficient of each pipe group have been presented in Tables 3 and 4, respectively. As shown in Table 3, the algorithm estimated the probability of the leakage (and therefore its value) in each zone, with less than 5% error. Another point that could be deduced from this table is that at the peak hour (6 p.m.), the estimated leakage values for each sub-zone had less error than the real value of the leakage. The reduction of the network pressure during the peak hours led to the reduction of the leakage of each node. As a result, the amount of the leakage of the whole network was decreased based on the leakage-pressure relationship. Due to the reduction, the search space of the optimization algorithm to locate the leakages in the network was decreased.

As a result, it was expected that, the obtained results would have less difference from their real value in this case. Although, due to the nature of the optimization algorithms, this argument may not always be correct.

In the third column of Table 4, it could be seen that the estimated values of the roughness coefficient of each pipe group differed about 5.2% from its real value.

- Studying the six simultaneous leakages

For this purpose, a scenario including the six simultaneous leakages with the average daily flow of 3, 3, 1, 1, 1, and 1 L/s was considered. These leaky nodes have been marked with the black arrow in Fig. 2.

The results of conducting the behaviorology method of the flow hydraulic grade line have been presented in the third row of Table 2. Based on the hydraulic grade line slope matrix results, there were no two consecutive entries of the mentioned matrix whose ratio of the second to the first entry was less than one. However, because the last entry of the hydraulic grade line slope matrix was more than 1, the leakiest zone corresponded to the last entry of the mentioned matrix up to the existing zone after the defined direction. In other words, the mentioned zone was related to the part of the network that was located after the J-7 node. Also, in reality, 90% of all the leakages were in this part of the network.

Then, considering the 8 flow meters (as in the previous scenario) in the sub-zones of the reported zone (from the behaviorology step), and as a result of dividing the network into 6 virtual sub-zones, the calibration method was conducted. Its results have been presented in the third row of Table 3.

Accordingly, although the probability values and consequently the estimated amounts of the leakage in the sub-zones that have leakage are very different from their real amounts, however, the priority of the probability of the leakage in each sub-zone has been correctly reported. This may be because, with an increase in the number of simultaneous leakages, the total number of the possible scenarios would be increased. As a result, the ability of the metaheuristic algorithm would be decreased.

The results of the roughness coefficient estimation of each pipe group have also been presented in the fourth column of Table 4. As it could be seen, the reported values have less than 9% error of the actual values.

It was observed that with the increase in the number of simultaneous leakages in the network, the difference between the real roughness coefficients and the estimated values was increased. This could also be related to the decrease in the ability of the metaheuristic algorithm due to the increase in the number of possible leakage scenarios in the network.

### 3.1.2. The second studied network

In this section, the proposed method has been implemented on the "Poulakis" network [32]. In this network (Fig. 3), the base demand of all the nodes was equal to 50 L/s, all the pipes were made out of the ductile cast iron, their Hazen-Williams coefficient in the hydraulic model was equal to 130, and the hydraulic grade of the tank has been considered equal to 100 m.

All nodes were also located at the zero height line. In this network, it was assumed that the field extraction has been done three times of the day, including midnight, 8 a.m., and 4 p.m., with the demand pattern coefficients of 0.5, 1.1, and 1.05, respectively.

In the behaviorology analysis step of the slope of the hydraulic level line of the flow, three nodes including 13, 16, and 18 have been considered as the nodes equipped with the pressure meter. The parameters of the calibration algorithm were like Table 1. The error of the measuring devices and the uncertainty in the base demand of the nodes have been applied by the Monte Carlo method (like the first studied network).

Given that this network has been located at the zero height line, the APH was in the middle of pipe No. 25. It was assumed that for some reasons (such as the lack of a hydrant in that location), it was impossible

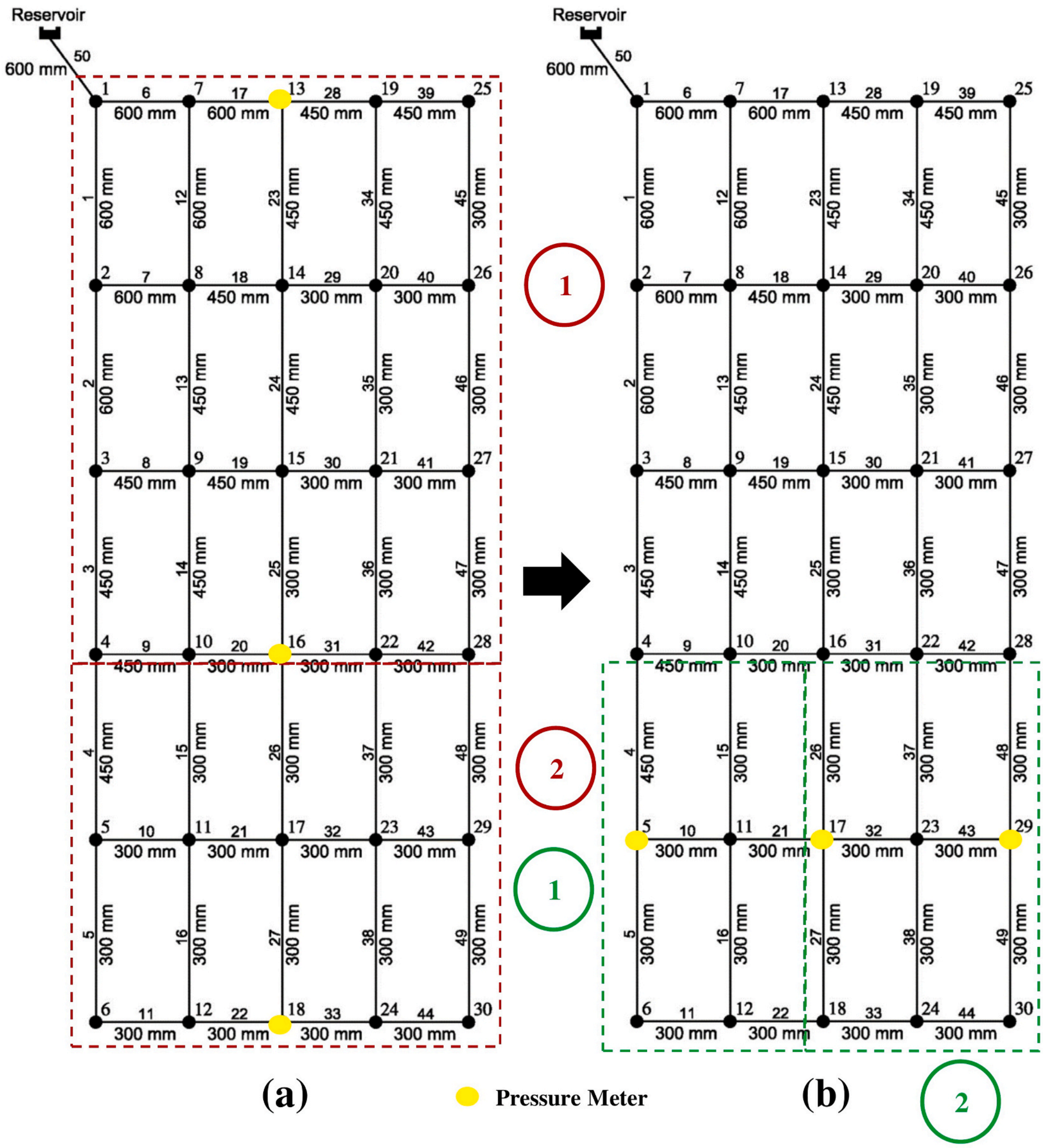


**Fig. 3.** The second network of the first case study.

to measure the pressure at this point, and instead, the pressure was measured at node 16. Also, considering the material of the pipes of this network (Ductile cast iron), the leakage exponent in the FAVAD equation was considered equal to 0.5.

Furthermore, considering the age of the pipes, all of them have been divided into 5 groups. Thus, the pipes No. 1 to 10 were in the first group, No. 11 to 20 in the second group, No. 21 to 30 in the third group, No. 31 to 40 in the fourth group, and No. 41 to 50 in the fifth group.

This network has been investigated in the states of one, three, and six simultaneous leakages, and here only the results of the scenario of the six simultaneous leakages have been presented. It should be noted that the expected results have also been achieved in the first and second scenarios.

To simulate the six simultaneous leakages, in the hydraulic model of the network in nodes No. 5 (with the flow of 15 L/s), No. 30 (with the flow of 10 L/s), No. 12 (with the flow of 15 L/s), No.10 (with the flow of 10 L/s), No. 9 (with the flow of 5 L/s) and 20 (with the flow of 5 L/s) were added as the average daily leakages. Then, the leakage flow of the mentioned nodes was estimated by Eq. (9) at three times (12 midnights, 8 a.m., and 4 p.m.). The summary of these calculations has been presented in Table 5.

**Table 5**
Leakage flow (l/s) of the six simultaneous leakages for the second network of the case study one.

| Leaky node | Time | | |
|---|---|---|---|
| | Midnight | 8 a.m. | 4 p.m. |
| 10 | 14.4 | 6.16 | 7.40 |
| 12 | 21.59 | 9.24 | 11.11 |
| 20 | 7.20 | 3.08 | 3.70 |
| 30 | 14.40 | 6.16 | 7.40 |
| 5 | 21.59 | 9.24 | 11.11 |
| 9 | 7.20 | 3.08 | 3.70 |

Assuming that in the first step, the three nodes with numbers 13, 16, and 18 were equipped with pressure meters (marked with a yellow circle in Fig. 3), the network was divided into two zones (Fig. 3-a). Through conducting the behaviorology method at the three mentioned times, the summary of the results has been presented in Table 6. Based on the analysis of the hydraulic grade line slope ratio matrix (matrix K), the rectangular zone whose four vertices were formed by nodes 4, 28, 6 and 30, showed the leakiest zone in the network. As it could be seen, the sum

**Table 6**
The hydraulic grade line slope ratio matrix (K) for the second network of the case study one.

| Initial and end nodes of the hydraulic grade line | 13&16 | 16& 18 |
|---|---|---|
| Midnight | 1.309 | 1.480 |
| 8 a.m. | 1.057 | 1.086 |
| 4 p.m. | 1.071 | 1.108 |

of the leakage flow in this zone (about 67%) is higher than in the other zone.

During the calibration step, this zone was divided into two virtual sub-zones by installing the three pressure meters at nodes 5, 17, and 29 (Fig. 3-b). In Table 7 the result of the calculations to estimate the probability of the leakage in each sub-zone has been presented. Based on these results, the sub-zone with the higher percentage of the leakage has been identified with a higher probability than the other sub-zone, which indicated the correct function of the proposed method.

The results of the roughness coefficients of each pipe group showed that the mean values estimated with its real value have less than 4% error.

### 3.2. Case study two

In this case study, the proposed method was implemented on the real network (shown in Fig. 4), which was located in Iran. The primary hydraulic model of the network was not calibrated. In this network, the three pressure meters were located on H-1, H-2, and H-3 hydrants (shown with bold black diamonds in Fig. 4).

Based on the material and age of the pipes, the network pipes were divided into two groups, the first and second groups were shown with grey and red colors in Fig. 4, respectively.

In this case study, two leakage scenarios were investigated. In the first scenario, by opening the H-2 fire hydrant and in the second scenario, by opening the H-3 fire hydrant, the leakages with values of 8 L/s and 1.5 L/s were created, respectively. In this network, there was only the field data of one time (ts = 1). According to the number and arrangement of the existing pressure meters, the whole network was initially divided into two zones (one zone is between the H-2, H-1 zone, and another one is between H-3, H-2).

Through conducting the behaviorology method, in the first scenario, the zone between H-2 and H-1 and, in the second scenario, the zone between H-2 and H-3 was reported. Therefore, the leaky zones were identified correctly. Subsequently, each of these zones was divided into some virtual sub-zones, shown in Fig. 4, in the first scenario with a blue dashed line and the second scenario with a green dashed line.

Then, by conducting the calibration method, the leakage percentage of each virtual sub-zone in both scenarios was precisely estimated. In the previous work of the authors ([7]), this network has been examined, but the value of the estimated leakage in the first scenario was 11% different from the actual value; however, here the value of the estimated leakage was equal to the actual value. Most probably, the reason could be the reduction of the search space of the optimization algorithm in the calibration step due to the use of the behaviorology step. This main advantage of the proposed method is very important, especially in the wide water networks.

## 4. Conclusion

In this paper, a new algorithm is proposed to locate the leakage zones in the water supply networks. In the first step, the leakiest zone is identified by comparison between the measured and simulated pressure results. In the second step, the reported zone is divided into several virtual sub-zones. Then, by the simultaneous calibration of the nodal demand and the pipes' roughness coefficients, the probability of leakage in each sub-zone is estimated. The main objectives of using the proposed method are to reduce number of the pressure and flow meters required to create each separate zone as well as reducing the search space in the calibration method (especially for large networks). In the proposed method, the error of the measurement devices and the uncertainty in the base nodal demands are considered by the Monte Carlo method. Also, due to the stochastic nature of the used optimization algorithm, estimation of the probability of the leakage at the different zones of a water supply network is provided. This algorithm was successfully implemented on the two hypothetical networks for different simultaneous leakages as well as on a real network for the two different leakage scenarios. The results showed that the proposed method is independent of the number of simultaneous leakages. Reducing the number of field measurement devices which is one of the results of the proposed method, can reduce the executive and the operational costs.

**Table 7**
Results of the leakage probability of each area for the second network of the case study one.

| Virtual sub-zone | Probability of leakage | | | | Real leakage percentage |
|---|---|---|---|---|---|
| | Midnight | 8 a.m. | 4 p.m. | Weighted average | |
| **1** | **99.90** | **97.33** | **99.13** | **98.78** | **75.00** |
| 2 | 0.10 | 2.67 | 0.87 | 1.22 | 25.00 |

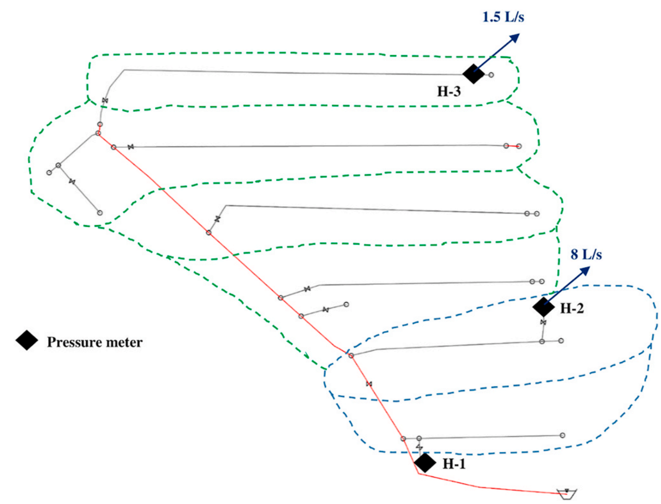


**Fig. 4.** The network of the second case study.

## Author statement

**Reza Moasheri:** Methodology, Software, Writing - Original Draft, Formal analysis, Investigation. **Mohammadreza Jalili Ghazizadeh:** Conceptualization, Validation, Writing - Review & Editing, Data Curation. **Mohammadreza Tashayoei:** Resources, Writing - Review & Editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

[1] Z.Y. Wu, M. Farley, D. Turtle, Z. Kapelan, J. Boxall, S.R. Mounce, S. Dahasahasra, M. Mulay, Y. Kleiner, Water Loss Reduction, Bentley Institute Press, USA, 2011.
[2] P. Przystałka, Performance optimization of a leak detection scheme for water distribution networks, IFAC-PapersOnLine 51 (24) (2018) 914–921.

[3] C. van den Berg, Drivers of non-revenue water: a cross-national analysis, Util. Pol. 36 (2015) 71–78.

[4] S.A. Sadr-Al-Sadati, M. Jalili Ghazizadeh, The experimental and numerical study of water leakage from High-Density, Polym. Test. 74 (2019) 274–280.

[5] M. Farley, G. Wyeth, Z.B.M. Ghazali, A. Istandar, S. Singh, The Manager's Non-revenue Water Handbook: a Guide to Understanding Water Losses, Ranhill Utilities Berhad and the United States Agency for International, Bangkok, 2008.

[6] R. Puust, Z. Kapelan, D.A. Savic, T. Koppel, A review of methods for leakage management in pipe networks, Urban Water J. 7 (1) (2010) 25–45.

[7] R. Moasheri, M. Jalili-Ghazizadeh, Locating of probabilistic leakage areas in water distribution networks by a calibration method using the imperialist competitive algorithm, Water Resour. Manag. 34 (1) (2020) 35–49.

[8] D.I.C. Covas, H.M. Ramos, Leakage Detection Single Pipelines Using Pressure Wave Behaviour, Water Industry System: Modelling and Optimisation Application, Baldock, Hertfordshire, England, 1999, pp. 287–299.

[9] A. Di Nardo, M. Di Natale, C. Gisonni, M. Iervolino, A genetic algorithm for demand pattern and leakage estimation in a water distribution network, J. Water Supply Res. Technol. - Aqua 64 (1) (2015) 35–46.

[10] J. Sousa, L. Ribeiro, J. Muranho, A. Sá Marques, Locating leaks in water distribution networks with simulated annealing and graph theory, Procedia Eng. 119 (2015) 63–71.

[11] E. Hajibandeh, S. Nazif, Pressure zoning approach for leak detection in water distribution systems based on a multi objective ant colony optimization, Water Resour. Manag. 32 (7) (2018) 2287–2300.

[12] L.A. Rossman, EPANET Users Manual, Version 1.1, Cincinnati: US Environmental Protection Agency1994.

[13] Z.Y. Wu, P. Sage, Water loss detection via genetic algorithm optimization-based model calibration, in: 8th Annual Water Distribution Systems Analysis Symposium, Cincinnati, Ohio, USA, 2006.

[14] M. Faghfoor Maghrebi, Y. Hasanzadeh, S. Yazdani, Calibration of water supply systems based on ant colony optimization (in Persian), J. Water and Wastewater; Ab va Fazilab 24 (1) (2013) 101–111 (in persian).

[15] S. Sophocleous, D. Savic, Z. Kapelan, Y. Shen, P. Sage, Advances in water mains network modelling for improved operations, in: 13th Computer Control for Water Industry Conference, CCWI 2015, Elsevier, 2015, pp. 593–602.

[16] S. Sophocleous, D.A. Savić, Z. Kapelan, O. Giustolisi, A two-stage calibration for detection of leakage hotspots in a real water distribution network, in: XVIII International Conference on Water Distribution Systems Analysis, WDSA2016, Elsevier, 2017, pp. 168–176.

[17] M.R. Shekofteh, M. jalili Ghazizadeh, J. Yazdi, Finding the leakage zones in district metered areas (DMAs) of water distribution networks (in Persian), in: 2nd iran Water and Wastewater Science Engineering Congress and National Conference on Demand Supply of Drinking Water and Sanitation, Iran, Isfahan, 2018.

[18] A. Rajeswaran, S. Narasimhan, S. Narasimhan, A graph partitioning algorithm for leak detection in water distribution networks, Comput. Chem. Eng. 108 (2018) 11–23.

[19] M. Tabesh, M. Jamasb, R. Moeini, Calibration of water distribution hydraulic models: A comparison between pressure dependent and demand driven analyses, Urban Water Journal 8 (2) (2011) 93–102.

[20] R. Moasheri, M. Jalili-Ghazizadeh, A Proposed Method for Locating of Leaky Areas in Water Supply Networks Based on Flow Hydraulic Grade Line, AUT Journal of Civil Engineering, 2019.

[21] N. Metropolis, S. Ulam, The Monte Carlo method, Am. Statis. Assoc. 44 (247) (1949) 335–341.

[22] E. Atashpaz-Gargari, C. Lucas, Imperialist Competitive Algorithm: an Algorithm for Optimization Inspired by Imperialistic Competition, IEEE Congress on Evolutionary Computation, Singapore, 2007, pp. 4661–4667.

[23] Z. Pan, D. Lei, Q. Zhang, A new imperialist competitive algorithm for multiobjective low carbon parallel machines scheduling, Math. Probl Eng. 2018 (2018) 5914360.

[24] X. Ji, Q. Gao, F. Yin, H. Guo, An efficient imperialist competitive algorithm for solving the QFD decision problem, Math. Probl Eng. 2016 (2016) 2601561.

[25] M. Abdollahi, A. Isazadeh, D. Abdollahi, Imperialist competitive algorithm for solving systems of nonlinear equations, Comput. Math. Appl. 65 (12) (2013) 1894–1908.

[26] A.R. Kashani, A.H. Gandomi, M. Mousavi, Imperialistic Competitive Algorithm: a metaheuristic algorithm for locating the critical slip surface in 2-Dimensional soil slopes, Geosci. Front. 7 (1) (2016) 83–89.

[27] H. Moayedi, M. Gör, L. Kok Foong, M. Bahiraei, Imperialist competitive algorithm hybridized with multilayer perceptron to predict the load-settlement of square footing on layered soils, Measurement 172 (2021) 108837.

[28] M. Khanali, A. Akram, J. Behzadi, F. Mostashari-Rad, Z. Saber, K.-w. Chau, A. Nabavi-Pelesaraei, Multi-objective optimization of energy use and environmental emissions for walnut production using imperialist competitive algorithm, Appl. Energy 284 (2021) 116342.

[29] M. Rezaiee-Pajand, J. Mohebi Najm Abad, A. Karimipour, A. Rezaiee-Pajand, Propose new implement models to determine the compressive, tensile and flexural strengths of recycled coarse aggregate concrete via imperialist competitive algorithm, J. Build. Eng. 40 (2021) 102337.

[30] J. May, Pressure dependent leakage, World Water Environ. Eng. 17 (8) (1994) 10.

[31] P.K. Amoatey, R. Minke, H. Steinmetz, Leakage estimation in developing country water networks based on water balance, minimum night flow and component analysis methods, Water Pract. Technol. 13 (1) (2018) 96–105.

[32] Z. Poulakis, D. Valougeorgis, C. Papadimitriou, Leakage detection in water pipe networks using a Bayesian probabilistic framework, Probabilist. Eng. Mech. 18 (4) (2003) 315–327.