



Selecting the best location of water quality sensors in water distribution networks by considering the importance of nodes and contaminations using NSGA-III (case study: Zahedan water distribution network, Iran)

Siroos Harif¹ · Gholamreza Azizyan¹ · Mohsen Dehghani Darmian² · Mohammad Givehchi¹

Received: 29 August 2022 / Accepted: 18 February 2023 / Published online: 28 February 2023
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract

One of the most effective ways to minimize polluted water consumption is to arrange quality sensors properly in the water distribution networks (WDNs). In this study, the NSGA-III algorithm is developed to improve the optimal locations of sensors by balancing four conflicting objectives: (1) detection likelihood, (2) expected detection time, (3) detection redundancy, and (4) the affected nodes before detection. The research procedure proposed the dynamic variations of chlorine between defined upper and lower bounds, which were determined utilizing the Monte Carlo simulation model. For selecting a contamination matrix with the same characteristics and effects of all possible events, a heuristic method was applied. The coefficients of importance are introduced in this study for the assessment of contamination events and network nodes. The Pareto fronts for each of the two sets of conflicting objectives were computed for benchmark and real water distribution networks using the proposed simulation–optimization approach. Results indicated that sensors should be installed downstream of the network to maximize sensor detection likelihood; however, this increases detection time. For the benchmark network, maximum and minimum detection likelihoods were calculated as 92.8% and 61.1%, respectively, which corresponded to the worst detection time of 11.58 min and the best detection time of 5.06 min. So, the position of sensors regarding the two objective functions conflicts with each other. Also, the sensitivity analysis related to the number of sensors illustrated that the Pareto fronts became a more efficient tool when the number of sensors increased. The best pollution detection likelihood in the real water network increased by 18.93% and 24.66% by incrementing the number of sensors from 5 to 10 and 5 to 15, respectively. Moreover, adding more than 10 sensors to the benchmark network and more than 15 to the real system will provide little additional detection likelihood.

Keywords Chlorine concentration · Contaminant detection · Contamination of important nodes · NSGA-III · Water distribution networks · Sensor placement strategy

Responsible Editor: Xianliang Yi

✉ Gholamreza Azizyan
g.azizyan@eng.usb.ac.ir

Siroos Harif
siroos.harif@pgs.usb.ac.ir

Mohsen Dehghani Darmian
mohsen.dehghanidarmian@gmail.com

Mohammad Givehchi
m_givehchi@yahoo.com

¹ Civil Engineering Department, University of Sistan and Baluchestan, Zahedan, Iran

² Chair of Engineering Hydrology and Water Management, Technical University of Darmstadt, Darmstadt, Germany

Introduction

The urban water distribution system (WDS) is one of an indispensable part of public infrastructure around the globe because of its relation to economic and social activities (Nafi et al. 2018; Qian et al. 2021). WDS is a complex network consisting of a collection of nodes, pumps, reservoirs, valves, storage tanks, treatment facilities, and numerous links (Adedoja et al. 2018a; de Winter et al. 2019; Nilsson et al. 2005; Palleti et al. 2016, 2018; Preis and Ostfeld 2008). The primary function of WDS is to deliver fresh water at the standard quantity and quality levels from sources such as storage tanks, rivers, lakes, and reservoirs to commercial,

industrial and residential users (de Winter et al. 2019; Mazumder et al. 2019; Priya et al. 2018).

Due to WDS's large size, complex configurations, and many access points, they are highly vulnerable to different contamination events (radioactive, chemical, or biological contaminants) that can be accidental, resulting from cross-connections, pipe bursts, wastewater backflow, intrusion of contaminated water due to pressure reduction, polluted source, or deliberate (e.g., terrorist attacks) (Costa et al. 2013; De Sanctis et al. 2010; Di Nardo et al. 2014; Huang and McBean 2009; Ostfeld et al. 2008).

A remarkable number of pollution incidents have been recorded in recent decades. Examples of accidental pollutants include Milwaukee, USA (Mac Kenzie et al. 1994; Corso et al. 2003), Nokia, Finland (Laine et al. 2011), The United States (Adedoja et al. 2018a), Lanzhou, China (Chinadaily 2014), Walkerton, Ontario, Canada (Hrudey et al. 2003), Tuscany, Italy (Nuvolone et al. 2021), West Virginia (Rosen et al. 2014; Whelton et al. 2015; Cooper 2014), Ohio (Wilson 2014), and Strasbourg, France (Deshayes et al. 2001) and recently two accidental events of waterborne diseases due to bacterial infection are recorded in California and Norway (Qiu et al. 2020). Also, in recent years, intentional contaminations have been reported, The United States is the most famous example of this, especially after 9/11 (Ostfeld et al. 2008). Other cases include Moscow (Lambrou et al. 2015), Scotland (Gavriel et al. 1998), the United States, and Japan (Yokoyama 2007) events.

Unfortunately, many of the intentional and accidental pollution events recorded in human history were not identified until people consumed the contaminated water for a long time, which has caused economic and public health losses and adverse social effects (Zheng et al. 2018). Consumers can detect some contaminants, such as chlorine, by noticing changes in the color and odor of water. But what happens if a contaminant enters the network that has no color or odor for easy detection by consumers? What happens if a hazardous substance enters the network accidentally or negligently from areas where chemicals are used for agriculture, or toxic waste is stored? Or even worse, what happens if someone deliberately injects a harmful pollutant into the water distribution network? Studies have confirmed that attacks on water distribution networks are actual, as they have occurred previously and might happen again (Adedoja et al. 2018b). Thus, to prevent significant disasters, it is necessary to identify and warn about contamination events within the WDS, protect water quality against accidental and intentional contamination events, and manage possible pollution events. The first step toward achieving this goal is improving the physical security of the system (Preis and Ostfeld 2008). The second step is utilizing monitoring sensors, a contamination warning system (CWS), also known as an early warning system (EWS), to monitor water quality (Grayman et al. 2016).

The third step is to develop effective response plans to neutralize or remove or minimize the impact of contaminants entering the network (Harif et al. 2020, 2022). It is possible to gain the maximum level of water safety by monitoring all nodes in the system, systematically (Adedoja et al. 2018b; Yazdi. 2018). But it is not feasible to install sensors at every node of the network due to the high cost, budget constraints, and maintenance issues (Adedoja et al. 2018a, b). To cope with these restrictions, only a limited number of these sensors must be installed at specific vital or critical places in the water network. Therefore, to achieve efficient monitoring, it is necessary to develop optimal water quality sensor placement (WQSP) strategies for WDSs.

Metaheuristic algorithms were extensively utilized in the optimization of different types of water systems (Dehghani Darmian et al. 2018; Khodabandeh et al. 2021). Sensor optimization investigation in the network started in the 1990s and has continued up to now. The issue based on the number of objective functions is classified as a single objective optimization problem (Al-Zahrani and Moeid 2001; Berry et al. 2006, 2009; Cheifetz et al. 2015; Comboul and Ghanem 2013; harif et al. 2021; Hu et al. 2017; Kessler et al. 1998; Ohar et al. 2015; Ostfeld and Salomons 2004, 2005; Propato 2006; Propato and Piller 2006; Schwartz et al. 2014a, 2014b; Shastri and Diwekar 2006; Woo et al. 2001) and multiple objective optimization problems (Afshar and Miri Khombi 2015; Bazargan-Lari 2014; de Winter et al. 2019; Dorini et al. 2008; Eliades and Polycarpou 2006; Gueli 2008; He et al. 2018; Huang et al. 2008; Khorshidi et al. 2018, 2019; McKenna et al. 2006; Naserizade et al. 2018; Nazempour et al. 2018; Ostfeld et al. 2008; Ostfeld and Salomons 2008; Preis and Ostfeld 2008; Wu and Walski 2006).

Among the single-objective methods, Lee and Deininger (1992) defined the total demands of nodes monitored by sensors as demand coverage (DC). They utilized mixed-integer linear programming (MILP) to optimally locate quality sensors in the network to maximize DC. Kumar et al. (1997) developed the method used by Lee and Deininger. They calculated the residual chlorine concentration at each node and estimated the DC values based on it. Kessler et al. (1998) defined the level of service as the maximum volume of contaminated water that is in a concentration more than the minimum danger level and consumed before the contamination is detected. They presented an algorithm to determine the optimal location of sensors based on the level of service. Kumar et al. (1999) defined time of detection as the time that elapsed from the entry of the contaminant into the network to the time detected by the sensors. Al-Zahrani and Moeid (2001, 2003) modified the demand coverage using a genetic algorithm (GA). Ostfeld and Salomons (2004) used a genetic algorithm to find the optimal layout of sensors based on the level of service for accidental or deliberate events. Ostfeld and Salomons (2005) developed their previous work

by considering uncertainties regarding the demands and pollution events. Comboul and Ghanem (2013) described the uncertainty analysis for the best sensor location to maximize the percentage of contamination detection. Schwartz et al. (2014a, 2014b) employed genetic algorithm to determine the optimal location of quality sensors to minimize the exposed consumers. Cheifetz et al. (2015) introduced a greedy incremental sensor placement method to be utilized for the optimization of quality sensors in a large real-world water distribution network. Rathi and Gupta (2016) considered maximizing the detection likelihood and water demand coverage as objective functions to determine the optimal location of sensors in water networks. They first normalized two objectives and then weighted them in a linear single-objective function. Hu et al. (2017) investigated the sensor placement in a large water network, utilizing a Spark-based genetic algorithm. In their study, the objective was to minimize the impact of contamination. Comparing the results of the proposed model with the experimental results illustrated that the proposed model has a good performance.

Simultaneous, In the context of multi-objective optimization, McKenna et al. (2006) examined the ability of sensors to minimize the detection time of events, the population exposed to pollution, and the contamination extent within the network by considering the number and detection limits for sensors. They illustrated that the discovery of contamination events is reliant on the detection limits. Preis and Ostfeld (2008) utilized a multi-objective model for designing quality sensors to maximize two objectives of detection likelihood and sensor detection redundancy and minimize detection time. Dorini et al. (2008) employed the noisy cross-entropy (nCE) algorithm for solving the optimization problem to minimize the detection time, the expected demand for polluted water, the affected population, and the maximization of detection likelihood. Aral et al. (2010) to optimally locate the sensors in the network, four different objective functions: detection time, affected population, consumed contaminated water, and detection likelihood using a genetic algorithm considered. Bazargan-Lari (2014) proposed a methodology to minimize four conflict objectives, including detection time, undetected contamination events, affected population before detection, and the number of sensors using NSGA II. Schwartz et al. (2014a, 2014b) investigated the effect of changes in chlorine concentration, pH, and alkalinity to identify the pollutant entering the distribution network. They injected two organophosphates into Net 3 of EPANET. The first is the pesticide chlorpyrifos (CP), and the second is the potent insecticide parathion (PA). The result illustrates the change in total free chlorine is the value employed as the main indication of an external intrusion and the changes in alkalinity and pH are used to confirm this indication in order to reduce the probability of false positive alarms Afshar and Miri Khombi (2015) presented two mathematical models for

the best layout of sensors as the dynamic double-use benefit model (DDUBM) and the static double-use benefit model (SDUBM) that provide a tradeoff between demand coverage and consumption of polluted water. They tested the validity of their proposed models utilizing two example problems with multi-objective ant colony optimization (ACO) algorithm. Du et al. (2015) proposed an algorithm to determine the activation time of wireless sensor nodes in each timeslot to prolong network lifetime with guaranteed monitoring quality in water distribution networks. Antunes and Dolores (2016) employed the NSGA-II algorithm to determine the optimal locations for a set of sensors by considering four objective functions detection likelihood, detection time, consumption of contaminated water, and affected consumers before detection. Nazempour et al. (2018) presented a new model applying complex network theory by considering that a WDS is a compound system. Two objective functions of maximizing the coverage of sensors and water demand coverage are studied by the authors. Naserizade et al. (2018) presented a new model based on conditional value at risk (CVaR) for best sensor placement in water systems. Four objective functions using the NSGA-II algorithm are developed to minimize detection time, affected population, cost, and the percentage of undetected events. Khorshidi et al. (2019) used a decision support framework based on the game theory by considering the two goals of minimizing the detection time and the sensor cost for the optimal locations of quality sensors. Ponti et al. (2021) proposed a novel evolutionary algorithm to estimate and analyze the optimal solutions of Pareto for sensor placement problems. Evaluation of the results of the new algorithm with the NSGA-II algorithm on a water distribution system through applying two objective functions demonstrated an improvement of this algorithm compared with NSGA-II, especially for low iteration counts. Xu et al. (2022) developed a new sensor placement approach, integrating multi-objective optimization and a reduced order model, to minimize costs and maximize monitoring performance. Recently, Shahmirnoori et al. (2022) utilized the particle swarm optimization algorithm to determine the optimal location of the fixed and mobile sensors by considering three objective functions detection likelihood, detection time, and consumption of contaminated water. The results show that increasing mobile sensors from 1 unit to 3 or 5 units, decreased the detection time by 14% and 22%, increased the detection likelihood by 102% and 159%, and reduced the contaminated water consumption by 52% and 70%, respectively.

A review of previous studies on the CWS problem shows some simplifying assumptions as below: in many previous studies, sensors were designed for a fixed demand pattern. While in a real water network, owing to many consumers, the demand is not constant and fluctuates over time. Fluctuations in demand may cause significant variations in pipe flows and

change in contaminant propagation across the water network. Also, in most of the research, the pollutant entering the network has been considered conservative, which can affect sensor's location. The conservative assumption of pollutant may lead to an inaccurate impact approximation and spread of the pollutant because a contaminant entering the network may decrease its concentration over time or produce a toxic substance with its products across the water network. For example, the results of Klosterman's investigation exhibited that in the case of employing adsorption models (reaction with the wall), the concentration of bulk arsenate and consumer exposure is significantly lower than the conservative assumption (Klosterman et al. 2014). Therefore, the optimal location of the sensors will be changed. Moreover, the type of pollutant detection sensor in previous studies has not been determined, and it has been assumed that the sensor in the network detects the smallest concentration of conservative pollutants using a small and predetermined threshold in the network. Unfortunately, this assumption is unrealistic. Due to budget constraints, it is impossible to install sensors that detect a specific type of contamination with the smallest concentration fluctuations, as each network is exposed to different contaminants. It has been confirmed that chlorine has a significant reaction to both microbial and chemical contamination. Thus, in the present study, for a more realistic simulation, chlorine sensors have been utilized to detect contaminants entering the distribution network. In previous studies, researchers have used a fixed boundary of chlorine concentration (minimum permitted levels of chlorine) to identify contamination (Shahmirnoori et al. 2022). Therefore, some contaminant events may not be detected in time and may take a longer time to detect. But, in the proposed approach, chlorine concentration is compared to certain computed lower and upper periodic bounds for each node and at various times. Bounds are set to record variations in chlorine concentration by several Monte Carlo simulations considering uncertainty in demand patterns and roughness coefficients. Using chlorine sensors can be an inexpensive solution because they are currently used for monitoring water quality at many water stations.

In most previous research, sensors have been designed for a series of random or specific scenarios with a limited number of contaminants. In other words, researchers have considered limitations for injection node, starting time, duration, and mass rate of pollutants entering to network. Then, based on these limitations, they designed the optimal location of the sensors. In a real network with a complex topology, contamination events can occur at any node and at any time with any mass rate and duration time. Therefore, the number of possible events is uncountable, and considering a limited and specific number of pollutions is unrealistic. Considering all the different possible events makes the design of sensors impossible. So, a pollution matrix is employed in this study by utilizing

a genetic algorithm which represents the most representative contaminants entering the network. Unlike previous research, the importance of pollutants in the contamination matrix and the importance of nodes in the water network are not considered the same in this investigation. In other words, pollutants that have a considerable impact on consumers and nodes that could harm more consumers if infected are of greater importance. So, with the presentation of two new coefficients, the sensors have been designed such that important contaminants are identified earlier than other contaminants, and important nodes are not infected. Finally, a multi-objective optimization approach is applied to determine the optimal location of the sensors in the WDN to maximize sensor detection likelihood and redundancy, as well as minimize sensor expected detection time, and percentage of affected nodes.

Material and methods

EPANET

In this research, for simulating the chlorine concentration and the effect of a contamination event distributed through the water network, the EPANET 2.0 and EPANET-MSX software were utilized. EPANET is a simulator performing an extended period simulation of hydraulic and water quality behavior (simulation of residual chlorine) within a drinking water distribution system. In the quality analysis, the concentration is calculated at each node and time step (Rossman 2000). The quality analysis in EPANET are advection, diffusion, and reactions that are expressed as follows:

$$\frac{\partial c_i}{\partial t} = -u_i \frac{\partial c_i}{\partial x} \pm R(c_i) \quad (1)$$

In this equation, c_i , u_i , and $R(c_i)$ are the concentration of the contaminant in the i th pipe at time t and location x (mg/L), the velocity of water in the i th pipe (m/s), and the term for the reaction rate including the wall and bulk reactions, respectively.

EPANET-MSX

EPANET's limitation in simulating only a single chemical species makes it unsuitable for simulating the transport and decay of chlorine in response to specific pollutants. An extended version of the original EPANET is developed to model the transport and the decay of multi-chemical species. This software is called EPANET-MSX (multi-species extension) (Shang et al. 2008). In the quality analysis, the concentration is calculated at each node and time step. The quality analysis in EPANET-MSX is expressed as follows:

$$\frac{d[A]}{dt} = -k_1[A][D] - k_3[A][C] \quad (2)$$

$$\frac{d[D]}{dt} = -k_2[A][D] \quad (3)$$

$$\frac{d[C]}{dt} = -k_4[A][C] \quad (4)$$

In these equations, A , D , and C represent chlorine, dissolved organic carbon (DOC), and contaminant concentration, respectively. k_1 and k_2 are reaction rate coefficients between chlorine and DOC. k_3 and k_4 are reaction rate coefficients between chlorine and contaminant.

In the present study arsenic, a cheap and available toxic heavy metal was selected as the contaminant. In the water network, arsenic reacts with chlorine and produces arsenate, then arsenate adsorption onto the exposed iron on the pipe wall surface occurs.

The EPANET and EPANET-MSX include Programmers Toolkit, a dynamic link library, that allows users to customize the EPANET and EPANET-MSX computing engine according to their requirements (Rossman 2000). The output of this can be used as the input to MATLAB. Then, this software is utilized for optimizing the location of sensors in distribution networks. The methods of evolutionary computation, such as genetic algorithms are beneficial and powerful tools in optimization problems (Hashemi Monfared et al. 2017, 2023; Dehghani Darmian et al. 2020), and solving searches for their unbiased nature allow them to perform very well in situations with little domain knowledge (Gong et al. 2014, 2015). The NSGA-III, an improved version of GA and NSGA-II (Srinivas and Deb 1994), is one of the most efficient algorithms for many-objective evolutionary algorithms.

NSGA-III algorithm

In this study, the objective functions conflict with each other; hence, no single answer can be obtained that optimizes all functions simultaneously. Therefore, multi-objective optimization or many-objective optimization methods should be used. Usually, multi-objective optimization problems (MOP) are defined for problems having two or three objectives, while optimization problems with four or more objectives are categorized as many-objective optimization problems (MaOP) (Gu and Wang 2020).

During the past decades, evolutionary algorithms (EAs) have been developed to solve multi- and many-objective optimization problems. In this research, the evolutionary optimization algorithm known as the non-dominated sorting genetic algorithm (NSGA-III) has been used (Deb and Jain 2013). NSGA-III is designed for multi- and many-objective

optimizations. This algorithm has demonstrated its efficiency in solving optimization problems such as economic/environmental dispatch problems (Bhesdadiya et al. 2016), hydro-thermal-wind scheduling (Yuan et al. 2015), engineering design problems (Gaur et al. 2017), industrial symbiosis system (Cao et al. 2020), rush order insertion rescheduling (He et al. 2020), feature selection (Zhu et al. 2017), biomedical search engines (Gupta et al. 2021), and air quality in buildings, (Martínez-Comesaña et al. 2022). The basic framework of NSGA-III is like the NSGA-II algorithm, but the method of selection in NSGA-III differs from the original NSGA-II (Deb and Jain 2013). Figures 1 and S1 illustrate the flowchart of NSGA-III algorithm and the selection mechanism.

Objective functions

This research aims to optimize the locations of quality sensors in the network as part of the battle of the water sensors networks (BWSN) utilizing the NSGA-III algorithm. The assumptions of the research are given below:

1. Contamination events can occur at any time of the day.
2. Each contamination event can occur at one point in the network.
3. Pollution reacts with other species like residual chlorine in the water network.
4. All nodes except booster nodes can be possible locations for the monitoring station.
5. The response time of the system is assumed to be 60 and 90 min for small and large networks, respectively. In the other words, when any of the sensors detect contamination, after 60 or 90 min, the network is closed and during this time the monitoring station will record data.
6. Monitoring stations record concentrations every 10 min.
7. The sensors can record concentration fluctuations greater than 0.01 mg/L without errors.

Four specific objectives are considered for the optimal locations of water quality sensors in the network: 1—sensor detection likelihood (f_1); 2—sensor expected detection time (f_2); 3—sensor detection redundancy (f_3), and 4—percentage of affected consumer nodes (f_4).

Sensor detection likelihood (f_1)

In a specific layout of sensors, the probability of detection is described as below:

$$f_1 = \frac{1}{TS} \sum_{i=1}^{TS} d_i \quad (5)$$

In this equation, TS denotes the total number of contamination scenarios. d_i equal to one. If the difference between the

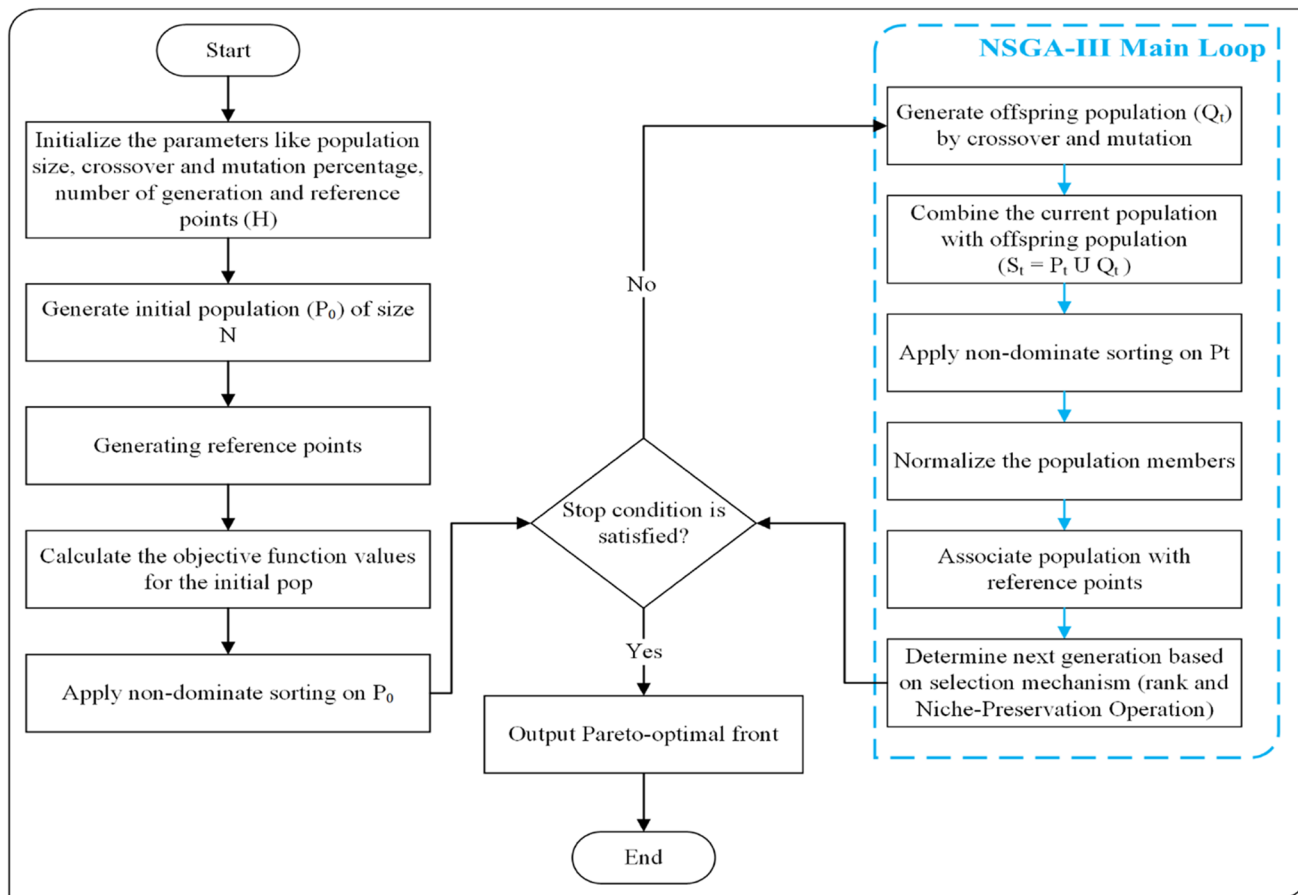


Fig. 1 Flowchart of NSGA-III algorithm

chlorine concentration and the lower bound of the allowable chlorine concentration in the candidate location of the sensor at three consecutive simulation time steps is more than 0.01 (the least assumed accuracy for the sensors), and $d_i=0$ otherwise.

It should be noted that in this study, unlike previous studies, to reduce the false positive detection by sensors, it is assumed that if the sensor detects the presence of contamination for at least three consecutive time steps, contamination will be considered detected contaminants. The optimal location of the sensors is where f_1 is maximized.

Sensor expected detection time (f_2)

For each pollution incident, the elapsed time from the start of the contamination event, to the first identified presence of contaminant by a sensor is defined as the time of detection by a sensor. t_i is the time of the first detection by i th sensor. The time of detection (t_d) is the minimum detection time among all sensors present in this design.

$$td = \min t_i \tag{6}$$

The sensor expected detection time is computed by

$$E(t_d) = \frac{1}{\sum_{i=1}^{TS} d_i} \sum_{i=1}^{TS} t d_i \times I[t_d(i, t)] \tag{7}$$

where $E(t_d)$ represents the mathematical expectation of the minimum detection time t_d . As previously stated in this study, the importance of contaminants in the contamination matrix is not considered the same. The contamination which is infected more nodes should be detected quickly than the others. This concept is considered in the present study for determining the optimal location of the sensors using modified sensor expected detection time as follows:

$$f_2 = \text{Modified } E(t_d) = \frac{1}{\sum_{i=1}^{TS} d_i} \sum_{i=1}^{TS} t d_i \times I[t_d(i, t) \times IF_i] \tag{8}$$

$$I[td(i, t)] = \begin{cases} 1 & td(i, t) > 0 \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

$$IF = \frac{1}{N} \sum_{i=1}^N PN_i \tag{10}$$

where N is the total number of nodes in the network, $PN_i = 1$. If the difference between the chlorine concentration and the lower bound of the allowable chlorine concentration in the node at three consecutive simulation time steps is more than 0.01 (the minimum assumed accuracy for the sensors), e.g., the node is polluted and $PN_i = 0$ otherwise. $I[t_d(i,t)] = 1$, if the contamination is identified by sensors and equals zero otherwise. IF is the impact factor. Utilizing this coefficient, the optimal location of the sensors is selected in such a way that the contaminants that infect more nodes of the network are detected earlier.

Sensor detection redundancy (f_3)

Developing of sensors that detect the contamination in water distribution networks in real time is still ongoing research, although there will be uncertainty in sensor detections. Then, to avoid false positive sensor detections and to increase the reliability of sensor detections, it is necessary to increase redundancy among sensors. A triply redundant measure that three sensors were required at least to detect the presence of contaminant concentration at a maximum time of 30 min between the first and third detections is considered. For each event, the redundancy of a sensor network design is equal to 1 if all the following conditions are true. Also, if at least one of them is false, the redundancy is equal to 0.

$$\begin{aligned} 1. & |t_1 - t_2| \leq 30\text{min} \\ 2. & |t_1 - t_3| \leq 30\text{min} \\ 3. & |t_2 - t_3| \leq 30\text{min} \end{aligned} \tag{11}$$

Where $t_1, t_2,$ and t_3 are the detection times by sensor 1, sensor 2, and sensor 3 respectively.

The redundancy (f_3) for a sensor network design is

$$f_3 = \frac{1}{\sum_{i=1}^{TS} di} \sum_{r=1}^{TS} R_r \tag{12}$$

$$R_r = \begin{cases} 1 & |t_1 - t_3| \leq 30\text{min} \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

Percentage of affected consumer nodes (f_4)

In previous research, the percentage of affected nodes is defined as follows:

$$PAF = \left(\frac{1}{N} \sum_{i=1}^N PN_i \right) \times 100 \tag{14}$$

where $PN_i = 1$ if the node is polluted; otherwise, $PN_i = 0$, and N denotes the total number of nodes in the water distribution

network. As previously stated in this study, the importance of contaminants and nodes in the contamination matrix are not considered the same. So, optimal sensor placement in the network should prevent infection of nodes with high demands.

For each water distribution network, several nodes have no demand. In other words, they do not have a consumer. These nodes are considered inactive nodes (internal nodes), and nodes with demand are considered active nodes. Therefore, in this section, the percentage of nodes that have demand and have been infected before the system response has been calculated. This objective function should be minimized and could be expressed as follows:

$$f_4 = \text{ModifiedPAF} = \frac{1}{N} \left(\sum_{i=1}^N PN_i \times D_i \times IF_i \right) \times 100 \tag{15}$$

$$IF = \frac{\sum_{t=0}^{t=td+tr} D(i,t)}{\sum_{i=1}^N \sum_{t=0}^{t=td+tr} D(i,t)} \tag{16}$$

D_i is the demand of nodes, it equals one if the node demand is greater than zero; otherwise, it is equal to zero. $D(i,t), t_r,$ and IF are the demand of node i at time t , the response time, and the impact factor respectively. According to this factor, the nodes that have more demands should be protected by sensors before being infected.

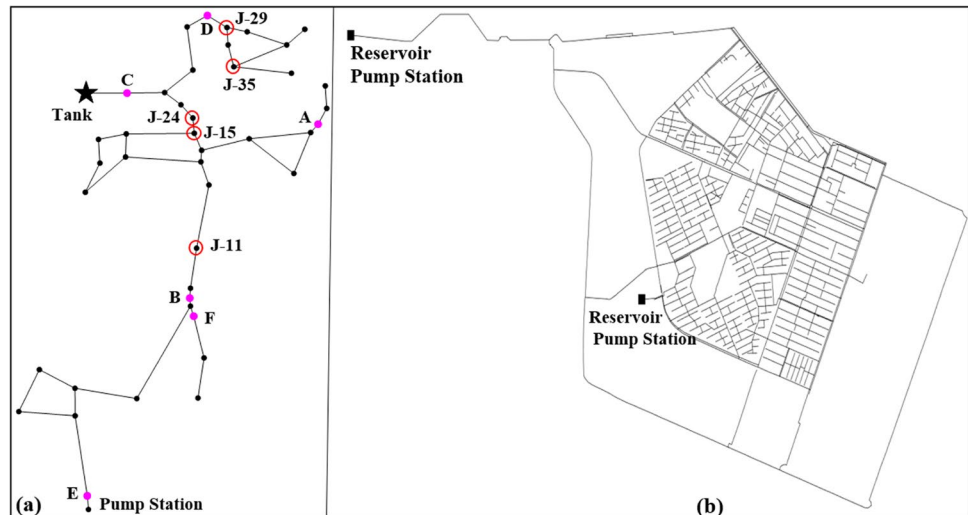
Case studies

In this study, to design water quality sensors by applying the proposed approach, two different networks with different complexity and characteristics are considered: a small benchmark network (44 nodes) and a real water network of large size (916 nodes). Various analyses were conducted for every network, including a base run, contamination simulations, and sensitivity analyses.

Case study 1

Various researchers have used the southern region of Central Connecticut water distribution network to test and validate water quality models numerous times. Munavalli and Kumar (2003) modified the reaction coefficient (i.e., bulk and wall coefficients) and the location of booster stations of this network. In this study, the modified network of Munavalli and Kumar, shown in Fig. 2a, was employed. This network consists of 37 nodes (32 consumer and 5 internal nodes), 48 pipes linking the nodes, one elevated storage tank, one pump station, and six booster stations, represented by letters A to F. Time steps for water quality and hydraulic simulations were 10 min, and the demand flow patterns were 24 h.

Fig. 2 (a) Benchmark WDN.
(b) Real WDN



Case study 2

The proposed approach was applied to a real water distribution network in southeastern Iran (Zahedan city) shown in Fig. 2b. This network is more complex and larger than case study 1. The topology of this network consists of 916 nodes (804 consumer and 112 internal nodes), 1025 pipes linking the nodes, two pump stations, two reservoirs, and ten booster stations. Time steps for water quality and hydraulic simulations were 10 min, and the demand flow patterns were 24 h.

Contamination matrix

To evaluate the fitness function of the sensor network detection likelihood (f_1), the sensor network detection redundancy (f_3), and to evaluate the cost function of sensor expected detection time (f_2), and percentage of affected consumer nodes (f_4), a contamination matrix should be produced. To construct a contamination matrix, assumptions about the number, starting time, mass rate, duration of injection, and location of contaminations must be considered. In this study, the following assumptions for networks are considered:

1. Number of injections: one node.
2. Starting time of injection: randomly selected from the beginning to the end of the simulation.
3. The mass rate of injection: randomly selected between 0.03 and 0.12 gr/min and between 0.05 and 0.2 gr/min for benchmark and real network, respectively.
4. Duration of injection: randomly selected between 30 and 240 and between 10 and 100 min for benchmark and real network, respectively.
5. Location of injection: randomly selected from all nodes, excluding dead ends.

The contamination matrix in this research for benchmark network has 26,112 contamination events, which is obtained as follows: injecting at 34 nodes (the number of dead-end nodes is 4, which is subtracted from the total number of network nodes as well as the tank is added), every 30 min for 24 h, with four mass injection rates of 0.03, 0.06, 0.09, and 0.12 gr/min, at four injection durations of 40, 80, 150, and 220 min (i.e., $34 \times 48 \times 4 \times 4 = 26,112$ events). Also, the contamination matrix for the real network has 463,872 contamination events, injecting at 604 nodes (the number of dead-ends and booster nodes is 312, which is subtracted from the total number of network nodes) every 30 min for 24 h, with four mass injection rates of 0.05, 0.1, 0.15, and 0.2 gr/min, at four injection durations of 10, 20, 60, and 90 min (i.e., $604 \times 48 \times 4 \times 4 = 463,872$ events).

Construction of reduced contamination matrix

Contamination events can occur at any node and at any time with any mass rate and duration times. Therefore, as the size of the system increases, the number of possible contamination events increases and is uncountable. To deal with this problem, Preis and Ostfeld (2008) presented a heuristic process that utilized a small sample of contaminants representing the total possible contaminants. Instead of employing the entire contamination matrix, they used a sampling approach to select the most representative contaminations and obtained similar results to those achieved utilizing the full matrix. They developed the following formulas to create a set of contaminants that included a reduced contamination matrix:

$$\text{Minimize: } \sum_{i=1}^5 |AS_i - ANil + | \sigma Si - \sigma Nil \quad (17)$$

$$\text{Subject to : } qj > 0 \quad j = 1, 2, \dots, N \quad (18)$$

where AS_i and σSi are average and standard deviation values of the geographical x coordinate of a sampled contamination events set. AN_i and σNi are average and standard deviation values of the geographical x coordinate of the water distribution system nodes. q_j is discharge flow from node j , and N is the total number of system nodes. In these equations, $i=2$ refers to the geographical y coordinate; $i=3$ represents the injection mass rate; $i=4$ expresses the injection starting time, and $i=5$ is the injection duration time.

Equation (18) states that when searching for the sample, network’s endpoints should not be used because if the infection enters the network in these nodes, the infection will not spread in the network.

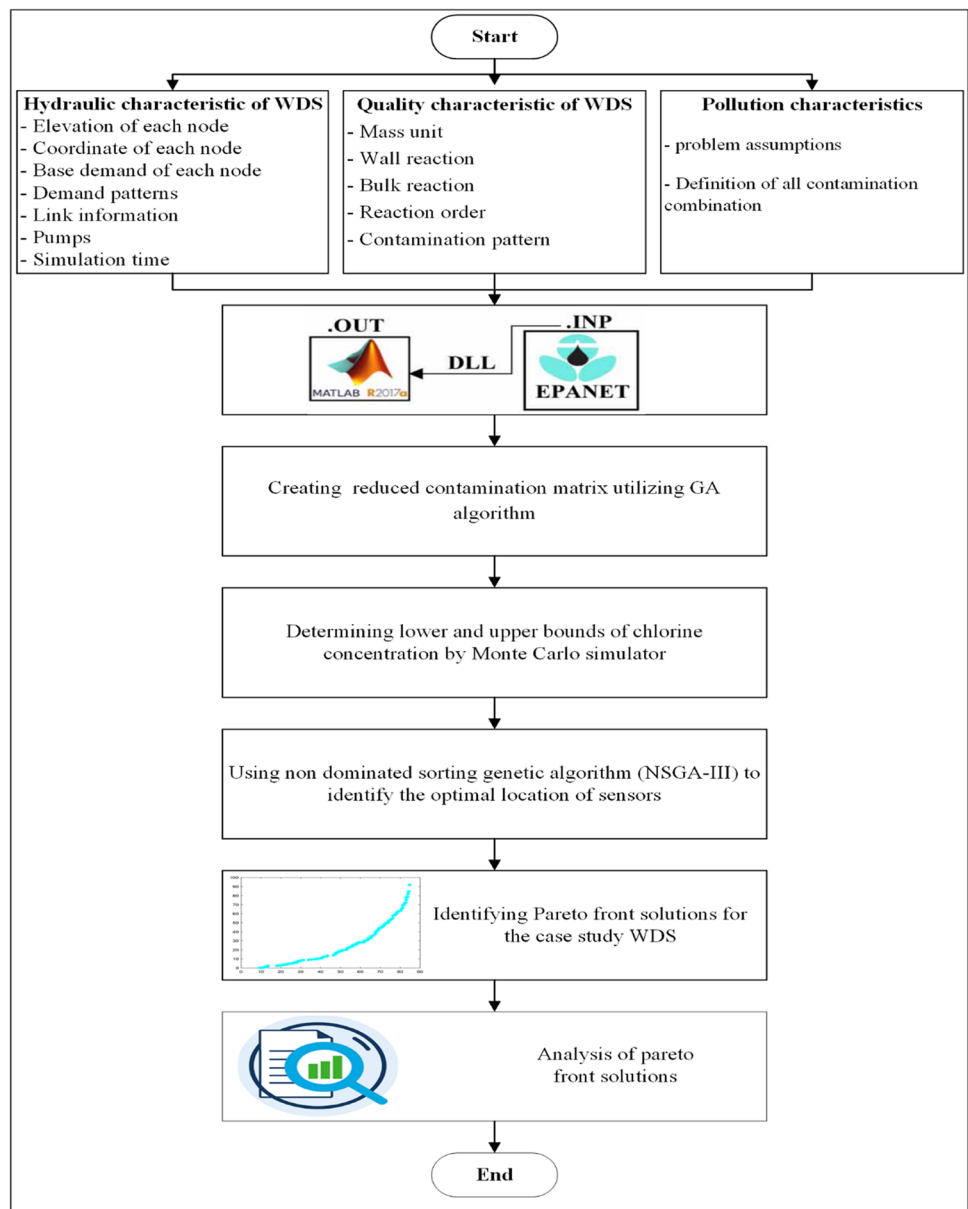
By solving the optimization problem presented in Eqs. (17) and (18), we come to a sample of pollution, which

represents the total pollution of the pollution matrix. This optimization problem is solved using a genetic algorithm. To create a reduced matrix by Eqs. (17) and (18), 1000 contamination events were selected. These reduced matrices were utilized as contamination events that are injected into the two networks (case studies 1 and 2), and optimization is performed based on these contamination events. This study’s general outline is presented in Fig. 3.

Contamination detection with chlorine boundaries

As previously mentioned, in this study, contamination event detection is based on chlorine concentration changes. Utilizing chlorine concentration sensors is more realistic and cost-effective. The chlorine concentration in the water network

Fig. 3 Simulation–optimization flowchart of the research



must first be determined for all nodes at different times under normal conditions (without contaminants in the network). In general, when chlorine is injected into the distribution network, after a period, the chlorine concentration reaches a steady condition in the network. Figure 4 illustrates the changes in chlorine concentration at node J-11 (see Fig. 2a) during a 10-day simulation of the benchmark network. As can be seen, the chlorine concentration is in a steady state from the fifth day. Therefore, pollutants can enter the network from the beginning to the end of the sixth day. Also, the effects of contaminant entry into the network up to 24 h after injection have been evaluated to locate the sensors optimally. Therefore, the simulation period time for network 1 was considered 7 days with 10 min time steps and a response time of 60 min.

In a real network, there are uncertainties in various parameters, such as demand patterns and roughness of pipes. Fluctuations in demand and roughness of pipes may cause significant variations in hydraulic (pipe flows) and water quality (change in contaminant propagation) across the water network. To make the simulation conditions realistic, an uncertainty of $\pm 20\%$ for the demand patterns and $\pm 10\%$ for the pipe roughness coefficient concerning normal values is considered. Then, due to these uncertainties, fluctuations in chlorine concentration were recorded for all nodes at different times employing the Monte Carlo simulator, and the

upper and lower bounds of chlorine concentration for each node were determined. By injecting a contaminant into the network, the sensors detect the presence of a contaminant if the chlorine concentration in the network exceeds the defined boundaries for more than three-time steps of simulation (30 min).

To explain how the proposed method detects contamination events. We simulated node J-35 due to the entry of pollutants into network 1. Figure 5a shows changes in chlorine concentration at node J-35 (see Fig. 2a) during normal conditions. As can be seen, the chlorine concentration is between the upper and lower boundaries. But Fig. 5b demonstrates variations in chlorine concentration when contamination enters the network (injection at node J-29, starting at 20:00 of the 6th simulated day (140 h after simulation started) for 100 min with a mass of 0.12 gr/min). The result indicates that the concentration of chlorine in the time of 149 to 151 h due to the reaction with the pollutant is less than the lower bound, which means that the contamination enters the distribution network, and then sensors can detect it.

It should be noted that if a fixed boundary is set for detection, some contaminant events may not be detected in time and may take a longer time to detect. To understand more, Fig. 6a and b illustrates changes in chlorine concentration at node J-24 during the normal condition and a contamination event, respectively (injection at node J-15, starting at 04:40

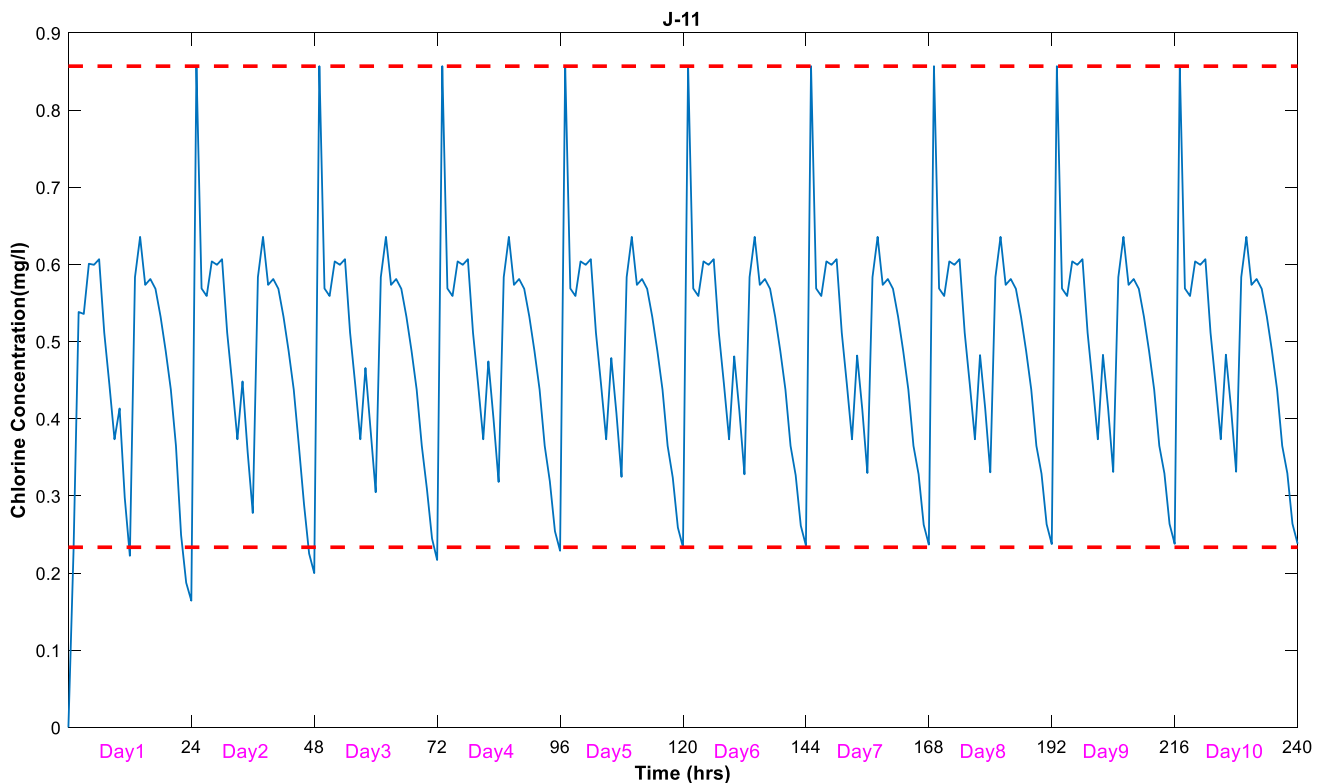


Fig. 4 Chlorine concentration at node J-11

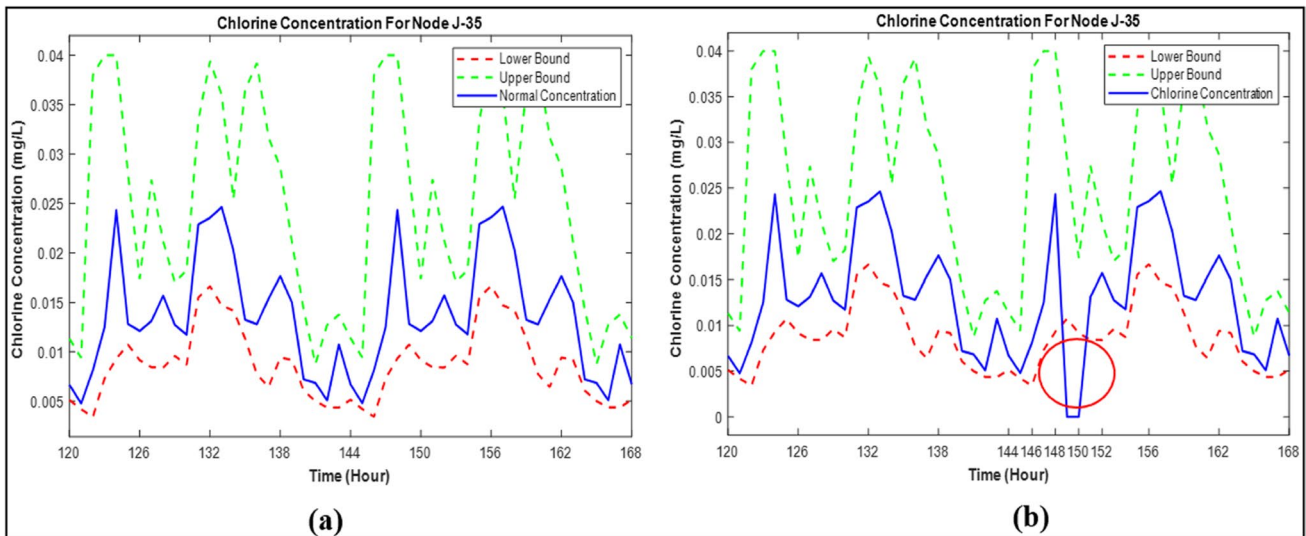


Fig. 5 Chlorine concentration at node J-35 during (a) normal conditions and (b) a contamination event

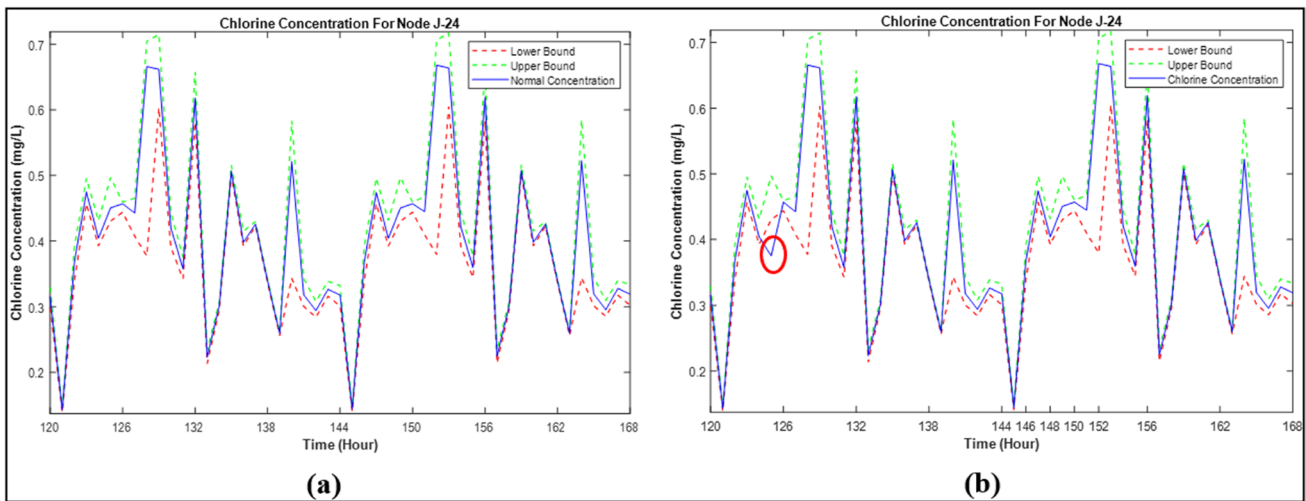


Fig. 6 Chlorine concentration at node J-24 during (a) normal conditions and (b) a contamination event

of the 6th simulated day for 30 min with a mass of 0.03 gr/min). If we consider a fixed boundary for chlorine concentration, such as 0.15 mg/L, the sensor installed in node 24 cannot detect the contamination. In summary, the larger the contaminant concentration, the faster the detection time because of more changes in chlorine concentration.

Results and discussion

The NSGA-III algorithm was applied to determine the optimal location of sensors by considering maximizing the sensor detection likelihood (f_1) and sensor detection

redundancy (f_3), as well as minimizing the modified sensor expected detection time (f_2), modified percentage of affected consumer nodes (f_4), and the number of sensors. The objective functions conflict with each other; hence, no single answer can be found that optimizes all objective functions simultaneously, and the optimal solutions will be a set of Pareto fronts rather than one specific solution.

Five sensors have been placed, and sensitivity analysis has been done for two networks. Various two-dimensional diagrams have been interpreted as four-dimensional Pareto because the description is complex,

according to Hu et al. (2018). Genetic algorithm parameters in this study are as follows crossover probability is 0.75, and mutation probability is 0.1. The number of populations was 1000, and the maximum number of iterations was set to 300. There were two stop conditions for the NSGA-III algorithm: either reaching 300 iterations or finding no new non-dominated solution after 20 successive generations.

Benchmark WDN

Base run results

Figures 7, 8, and 9 show base run results for placing five sensors in water network 1. In Fig. 7a, an optimal Pareto front for maximizing sensor detection likelihood (f_1) versus minimizing modified sensor expected detection time (f_2) is summarized. The interaction between the two objective

Fig. 7 (a) Optimal Pareto front for detection likelihood versus sensor expected detection time, and (b) selected locations of the sensors according to Pareto front

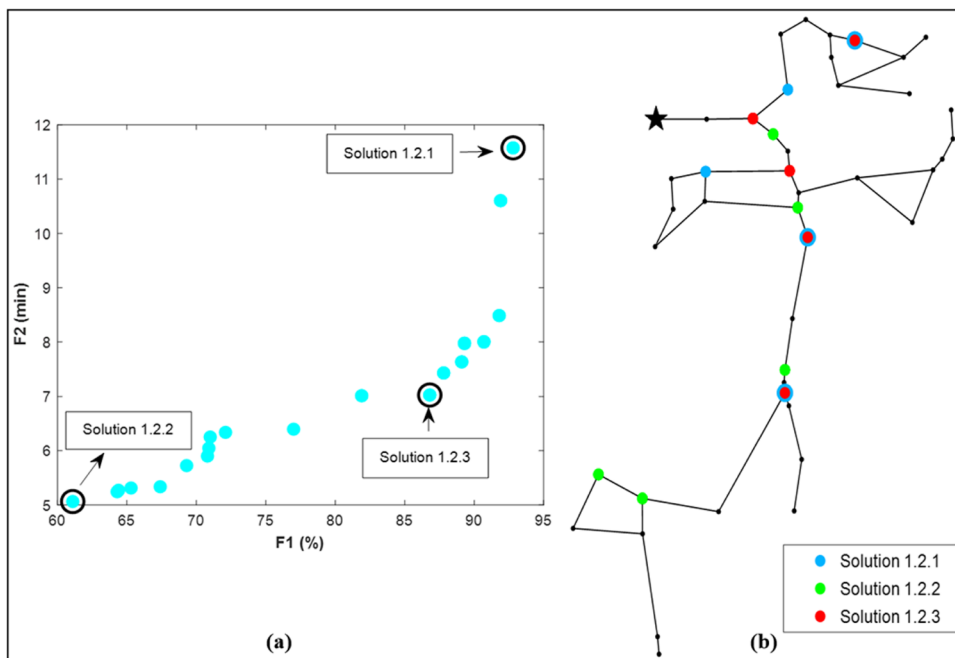


Fig. 8 (a) Optimal Pareto front for detection likelihood versus sensor redundancy, and (b) selected locations of the sensors according to Pareto front

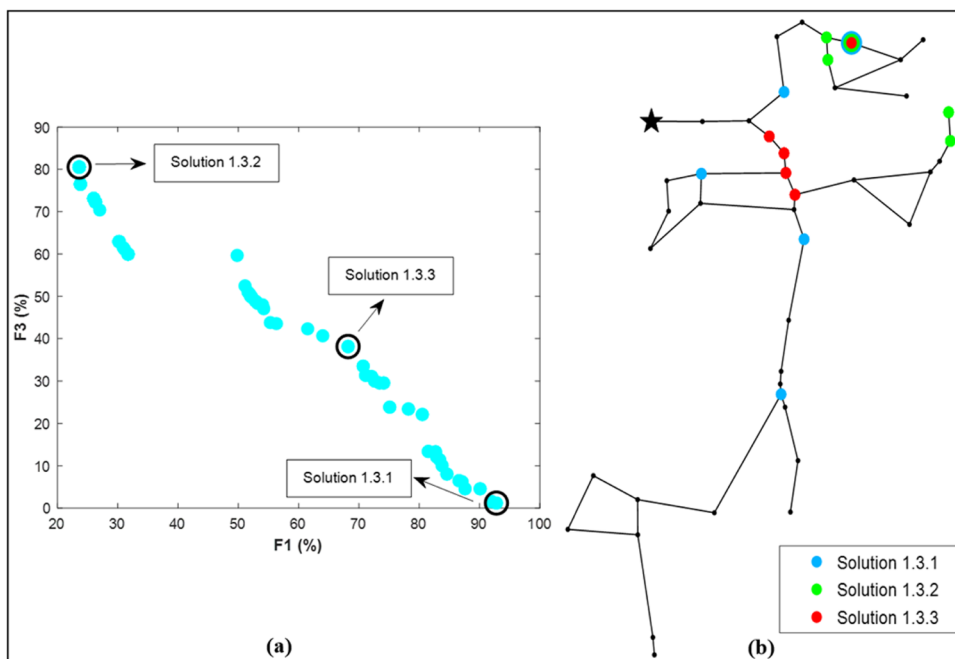
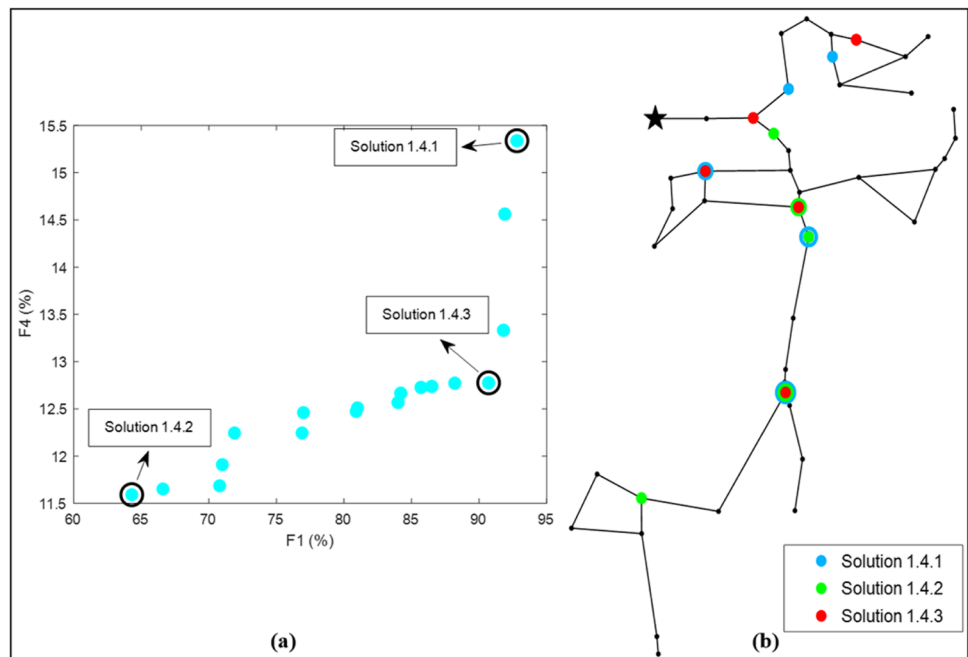


Fig. 9 (a) Optimal Pareto front for detection likelihood versus percentage of active polluted nodes, and (b) selected locations of the sensors according to Pareto front



functions is such that to maximize sensor detection likelihood in the network, sensors should be installed in the downstream nodes of the network; Installing sensors downstream of the network increases detection time. On the other hand, in the second objective function, sensors should be installed in nodes close to the input point of infection to minimize the detection time of pollutants, especially pollutants that have a more significant impact on the consumers, by using the introduced importance coefficient; So, the layout of sensors according to the two objective functions conflict with each other.

Figure 7b illustrates the selected locations of the sensors in the network for the three solutions from the Pareto front. To place the sensors in the network, three different solutions (locations) were selected based on the Pareto front points. The first solution has the best answer for objective function 1 and the worst for objective function 2. The second solution has the best answer for objective function 2 and the worst for objective function 1. Finally, the third solution is where both objective functions have an intermediate state. The sensor locations according to solution 1.2.1 indicate the best detection likelihood equal to 92.8% but the worst solution for the modified detection time (11.58 min). Solution 1.2.2 shows the location of the sensors according to the minimum contamination—modified detection time (5.06 min), but the worst detection likelihood is 61.1%. In solution 1.2.3, the method presented by Young (1993) is used to select the optimal point of the Pareto front for locating the sensors in the distribution network. According to Young's method, the optimal point has a detection likelihood of 86.8% and a modified contamination detection time of 7.03 min.

As expected, the optimal location of the sensors according to solution 1.2.1 is the downstream points of the network, and based on solution 1.2.2, the points are close to the injection areas.

In Fig. 8a, an optimal Pareto front for maximizing sensor detection likelihood (f_1) versus maximizing sensor detection redundancy (f_3) is summarized. The interaction between the two objective functions is such that to achieve the maximum detection likelihood, the sensors must be spread in the downstream nodes of the network to cover more nodes, while in objective function 3, to achieve maximum redundancy, the sensors must have a short distance from each other. Therefore, the location of the sensors in two objective functions conflicts with each other.

Figure 8b describes the selected locations of the sensors in the network for the three solutions from the Pareto front. Solution 1.3.1 indicates that the best percentage of detection likelihood is 92.8%, but the worst value of redundancy is 1.19%. Solution 1.3.2 shows the location of the sensors according to the best redundancy value of 80.51% but the worst detection likelihood (23.6%). In solution 1.3.3, the method presented by Young (1993) is used to select the optimal point of the Pareto front for locating the sensors in the distribution network. According to Young's method, the optimal point has a detection likelihood of 68.2% and a redundancy value of 38.12%.

Figure 9a depicts an optimal Pareto front for maximizing sensor detection likelihood (f_1) versus minimizing the modified percentage of affected active nodes (f_4). To achieve the maximum detection likelihood, the sensors must be spread across the downstream nodes of the network, which cause

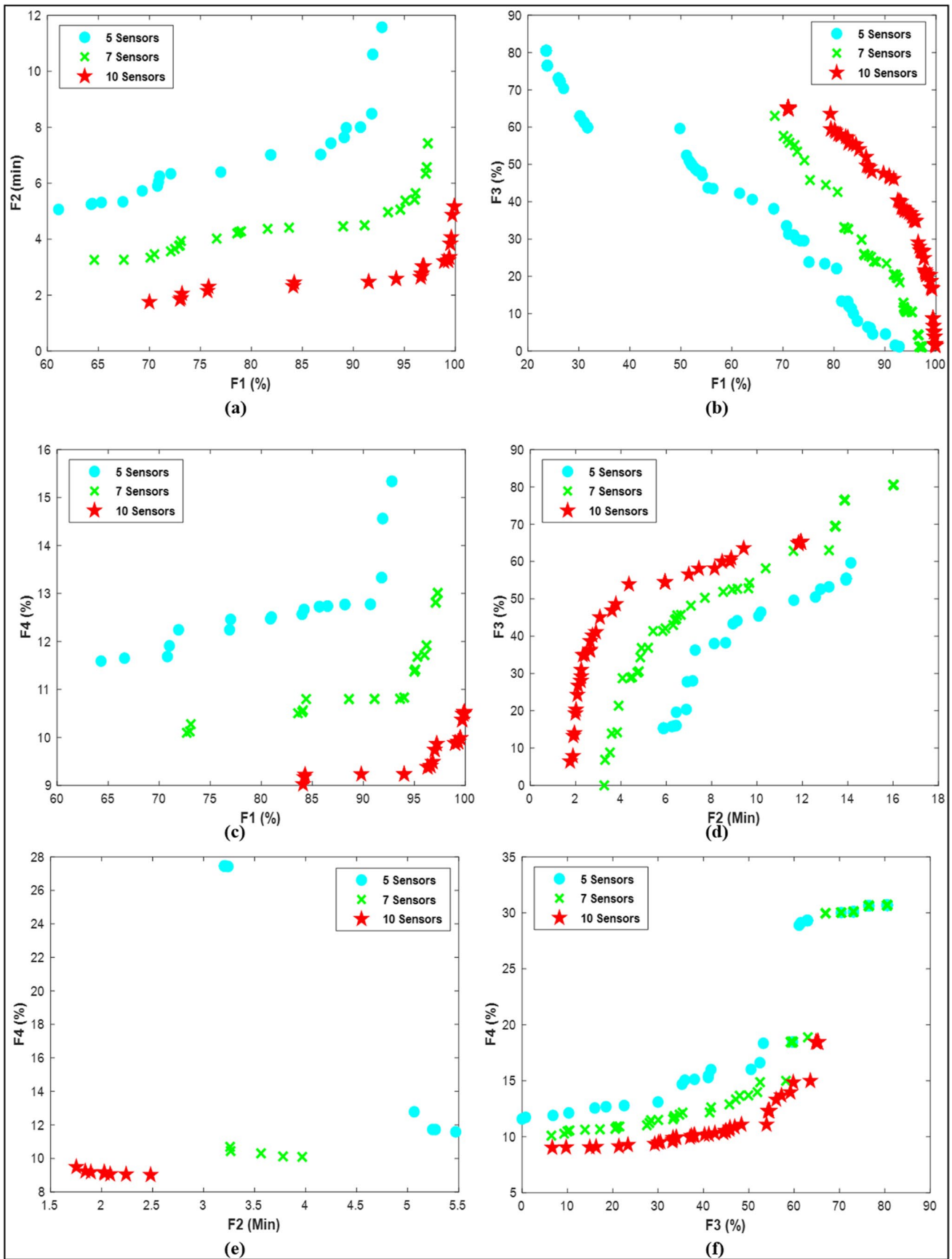


Fig. 10 Comparing the optimal Pareto fronts for 5 sensors with 7 and 10 sensors utilized in the base and sensitivity analysis runs for different objective functions of benchmark network

to infect many network nodes. However, in objective function 4, to achieve the minimum number of infected nodes, especially nodes that have more demands, by using the introduced importance coefficient, sensors must be installed in locations close to where the pollution enters the network. Therefore, the optimal location of the sensors in the two objective functions conflicts with each other.

Figure 9b illustrates the selected locations of the sensors in the network for the three solutions from the Pareto front. The location of the sensors according to solution 1.4.1 indicates the best percentage of detection likelihood equal to 92.8% but shows the worst modified percentage of infected nodes (15.34%). Solution 1.4.2 shows the location of the sensors according to the minimum (the best) modified percentage of infected nodes equal to 11.59%, but the percentage of detection is the lowest (64.3%). In solution 1.4.3, according to Young's method, the optimal point has a detection likelihood equal to 90.7% and the modified percentage of affected nodes of 12.78%.

Optimal Pareto fronts and selected sensor locations for objective functions f_2 , f_3 , and f_4 (Eqs. (8), (12), and (15)) were obtained for three solutions. These results are illustrated in figures S2 to S4 of the supplementary file.

Sensitivity analysis results

This section investigates the effect of the number of sensors utilized in the network on the objective functions. The purpose is to check whether the increase in the number of sensors improves the optimal Pareto fronts. For this purpose, the number of sensors was increased from five to seven and ten sensors. Figure 10a–f compares the optimal Pareto fronts for the five, seven, and ten sensors applied in the base and sensitivity analysis runs. As presented in Fig. 10a–f, the Pareto fronts for 7 sensors dominate the Pareto fronts of 5 sensors, and the Pareto fronts for 10 sensors dominate the Pareto fronts of 5 and 7 sensors. For example, according to Fig. 10a. When employing five sensors, the best detection likelihood is 92.8%, and the worst solution for the modified detection time is 11.58 min. But with increasing the number of sensors to seven, the best detection likelihood by sensors increased by 4.85% and reached 97.3%. Also, the worst solution for the modified detection time decreased by 4.15 min (35.9%) and reached 7.43 min, which has improved the two objective functions. Then, with increasing the number of sensors to ten, the best detection likelihood compared to five and seven sensors increased by 7.65 and 2.67 percent, respectively, and reached 99.9 percent. Also, the worst solution for the modified detection time compared to five and

seven sensors decreased by 55.35 and 30.42 percent and reached 5.17 min.

As observed, with increasing the number of sensors, the solutions of different objective functions have improved compared to the base run (5 sensors). Figure 11 shows the optimal Pareto front for the number of sensors utilized in the distribution network 1 versus the best detection likelihood. Results express that by installing only one or two sensors in the network, 36.4 and 63.2% of pollutants can be detected. As can be seen, Initially, when the number of sensors is small, the detection percentage changes rapidly as the number of sensors increases. But, as the number of sensors increases, the detection percentage changes slower. So increasing the number of sensors to more than 9 sensors in this small system will provide little additional detection likelihood. For example, by increasing the number of sensors from 1 to 2, the detection likelihood increased by 73.6%, but by increasing the number of sensors from 3 to 4, the detection likelihood just increased by 12.7%. The maximum detection likelihood corresponds to 11 sensors, which means that increasing the number of sensors to more than 11 will provide any additional detection likelihood of pollutants entering the network.

Real WDN

To evaluate the effectiveness of the proposed method in large networks, a real water distribution network in southeastern Iran (Zahedan city) is used (Fig. 2b). Figure 12 illustrates the optimal Pareto front for the number of sensors utilized in the distribution network 2 versus the best detection likelihood. Results represent by installing only one or two sensors in the network, 24.5 and 45.2% of pollutants can be detected. As can be seen, the maximum detection likelihood corresponds to 20 sensors, which means that increasing the number of sensors to more than 20 will provide any additional detection likelihood of pollutants entering the network. Also, increasing the number of sensors to more than 15 in this large system will provide little additional detection likelihood. Therefore, in this study, to analyze the sensitivity of the number of sensors, the effect of 5, 10, and 15 sensors on the network has been investigated.

Figures 13a–f and 14 demonstrate optimal Pareto front and sensor locations for placing five sensors in water network 2 according to base run results. Optimal Pareto fronts for the four conflict objective functions are plotted in various two-dimensional diagrams. According to the Fig. 13a, the more the modified detection time, the more the detection percentage because when the detection percentage is maximum, it means that the sensors are installed at the end areas of the network, which causes them to be far from the entry point of the pollutant; thus, detection time increase. Figure 13b exhibits that increasing the detection

Fig. 11 Optimal Pareto front for the number of sensors versus detection likelihood for benchmark network

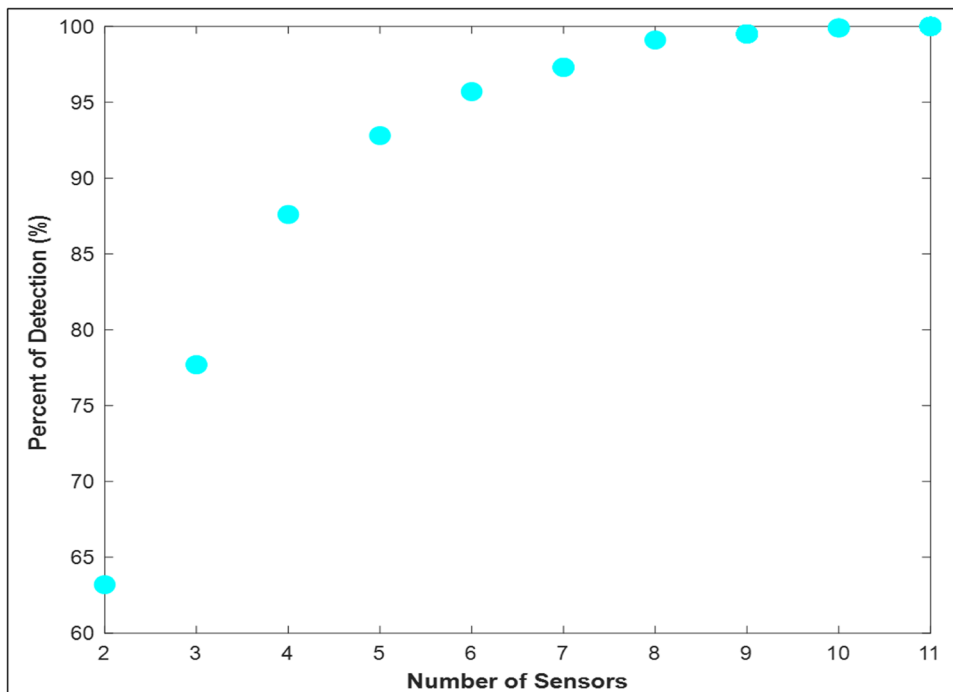
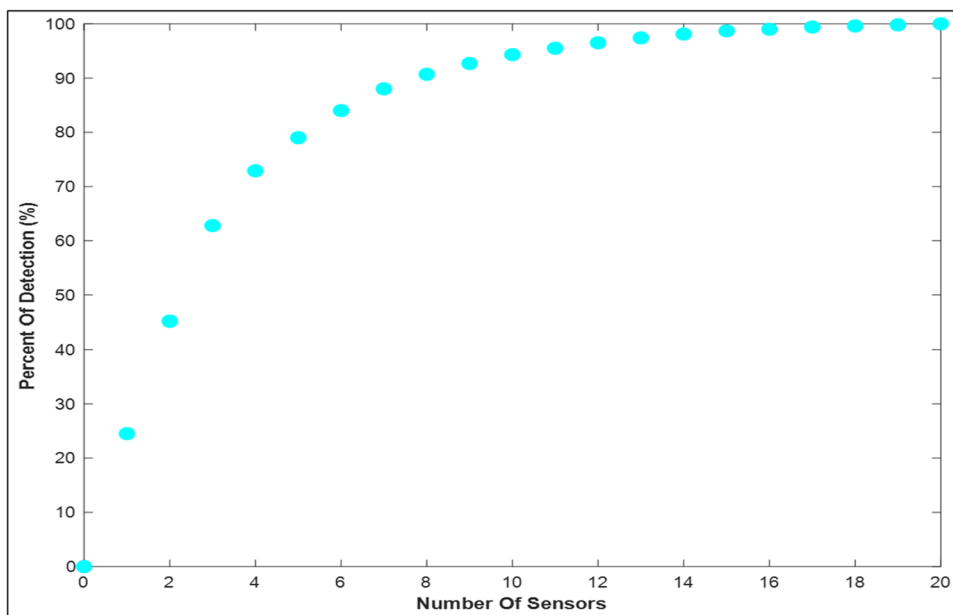


Fig. 12 Optimal Pareto front for the number of sensors versus detection likelihood for real network



percentage reduces redundancy. In order to achieve the maximum detection percentage, the sensors must be installed across the network and far from each other, which reduces redundancy. Also, the more the detection likelihood leads to the more the modified percentage of polluted nodes (Fig. 13c). The results of Fig. 13d represent that with increasing redundancy, the detection time increases. Because the sensors are located close to each other and far from the input nodes. Finally, Fig. 13e and f present

that decreasing the modified percentage of infected nodes decreases the redundancy and increase detection time.

Figure 15 compares the optimal Pareto fronts for five, ten, and fifteen sensors utilized in the base and the sensitivity analysis runs. As confirmed in Fig. 15–f, the Pareto fronts for 10 sensors dominate the Pareto fronts of 5 sensors, and the Pareto fronts for 15 sensors dominate the Pareto fronts of 5 and 10 sensors. For example, in Fig. 15b, when five sensors in the network exists, the best detection likelihood

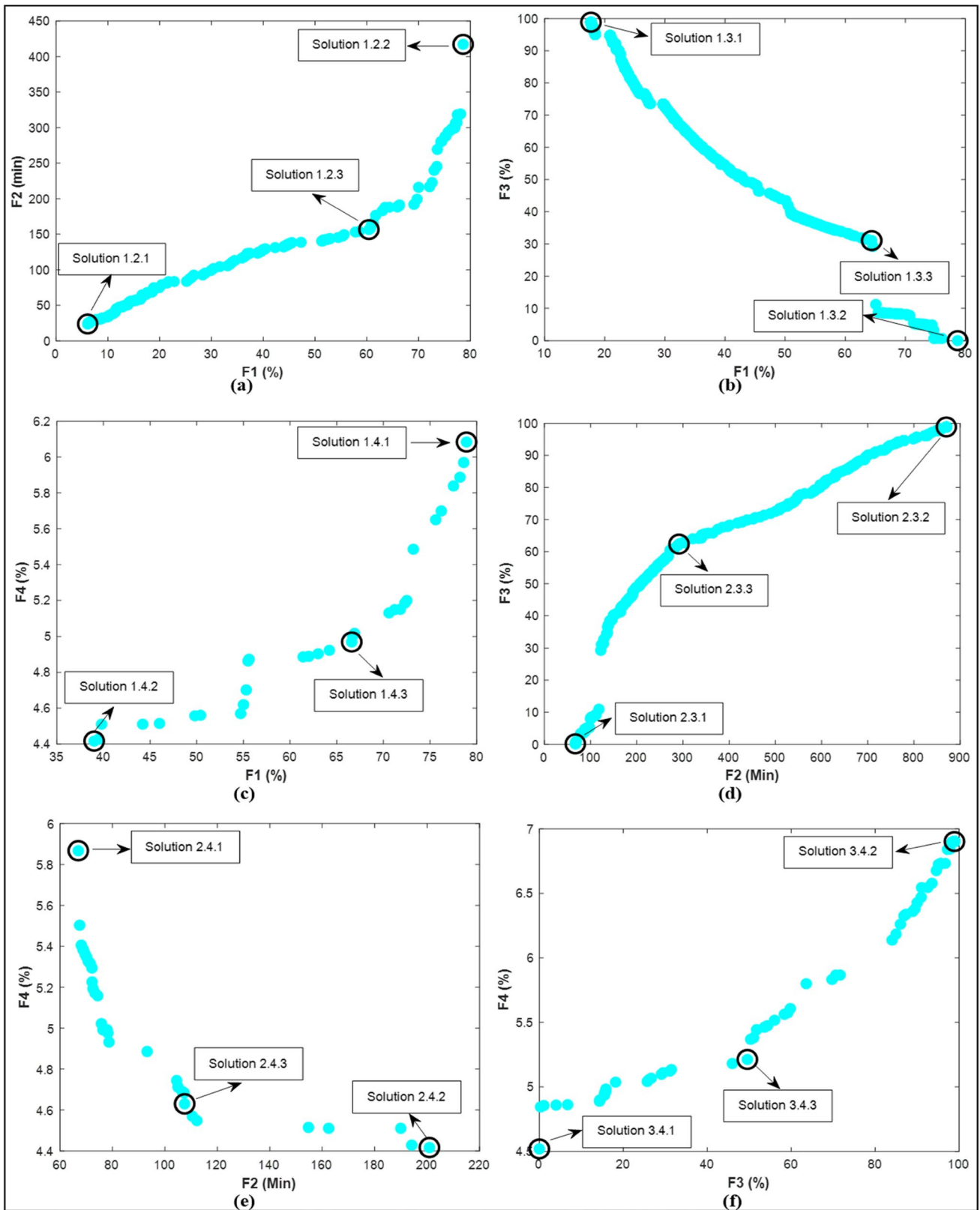


Fig. 13 Optimal Pareto front for a base run of real network for different objective functions

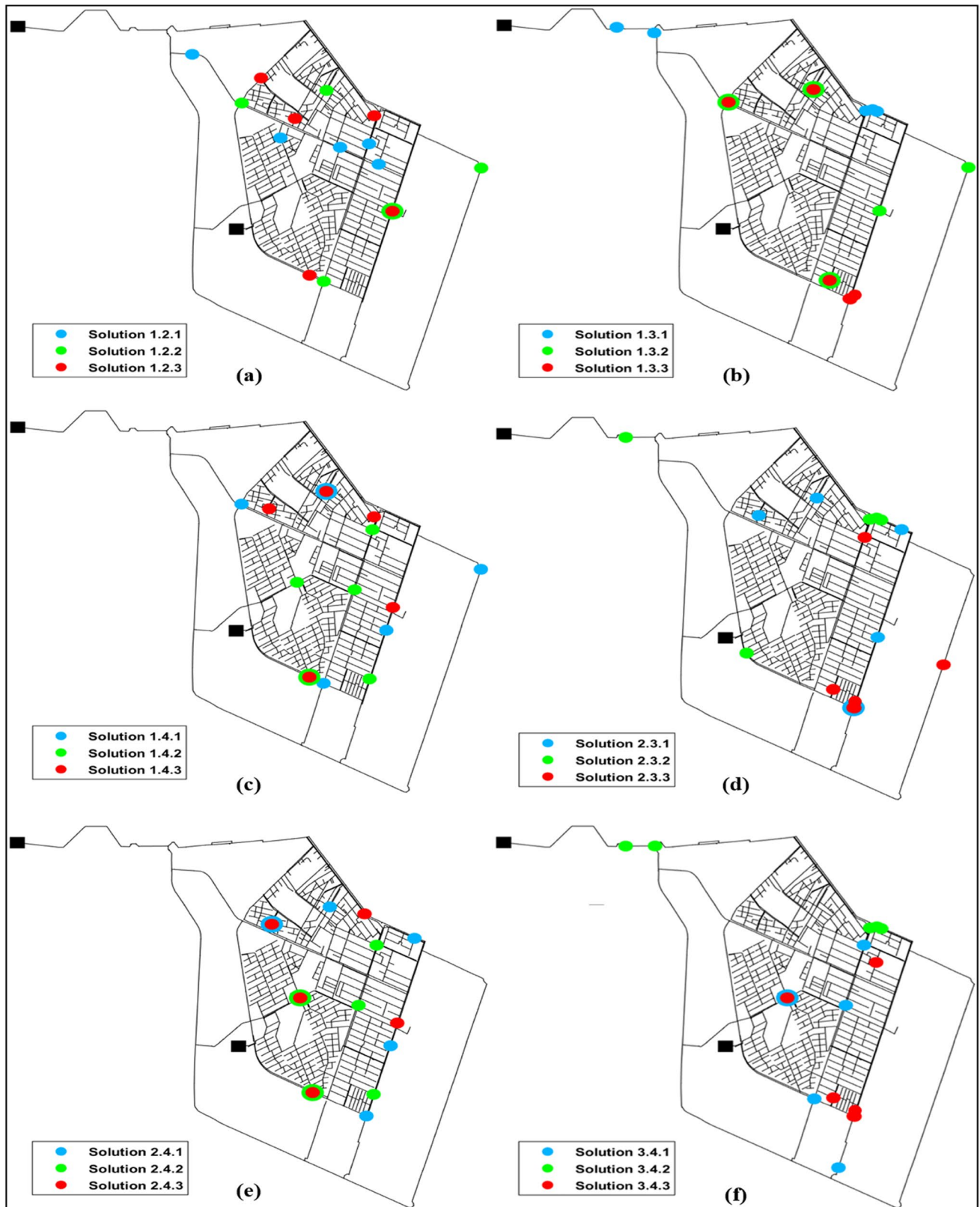


Fig. 14 Selected locations of the sensors according to Pareto fronts for a base run of real network for different objective functions

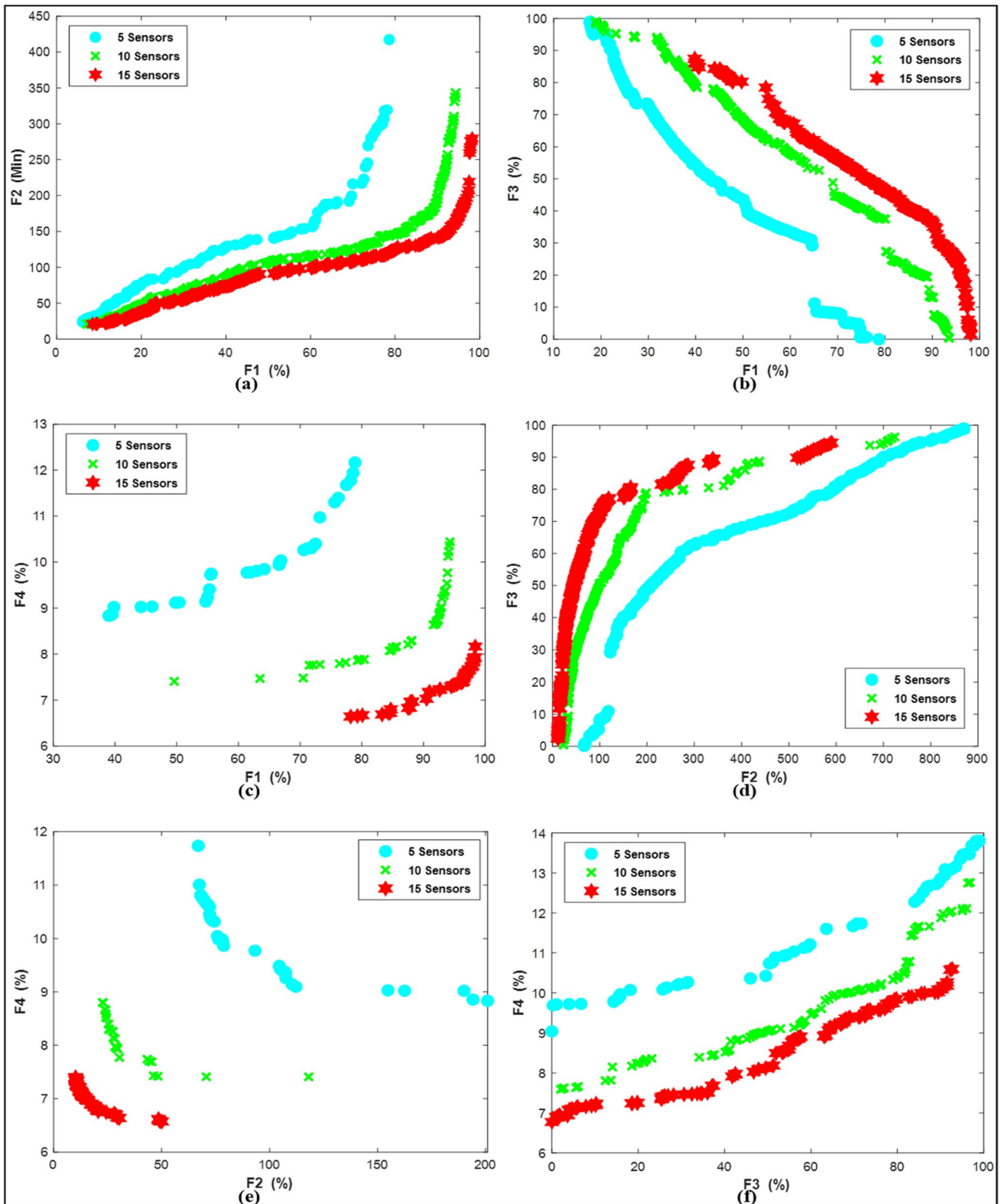


Fig. 15 Comparing the optimal Pareto fronts for 5 sensors with 10 and 15 sensors utilized in the base and sensitivity analysis runs for different objective functions of real network

is equal to 78.7%, and the best solution for redundancy is 98.87 percent. But by increasing the number of sensors to ten, the best detection likelihood by sensors increased by 18.93% and reached 93.6%. Moreover, the best solution for redundancy increased to 98.94%, which has improved the two objective functions. Then, with increasing the number of sensors to fifteen, the best detection likelihood compared to five and ten sensors increased by 24.66 and 4.8 percent, respectively, and reached 98.1 percent. Also, the best solution for the redundancy related to fifteen sensors increased to 99.1 percent.

Conclusions

Water distribution networks (WDNs) are one of the main components of public infrastructure that distribute safe drinking water to billions of customers worldwide. Due to the complexity of their structures and several access points, WDNs are vulnerable to intentional or accidental contamination events. Contaminants entering the water distribution systems are one of the most dangerous events that may occur either deliberately or accidentally. Polluted water can cause sickness or even death among consumers. Therefore, the protection of WDNs is crucial, and monitoring tools need to be improved. This led utilizing of sensors for identifying the pollutants in water distribution systems. This study proposes a multi-objective optimization approach for sensor network design to precisely detect possible pollution events. For the evaluation of the presented approach, benchmark and real water networks have been selected.

Contamination events are a potential risk, which can occur at any node, at any time, with any mass rate, and duration times. The number of contamination events increases with system size. For this reason, a heuristic method was employed for selecting a representative sample of contaminations that had similar characteristics and effects. A contamination matrix based on 1000 pollutant events was created for benchmark and real water networks. The developed contamination event detection procedure is based on the dynamic change of chlorine concentration relative to defined upper and lower bounds. The upper and lower bounds of chlorine concentration for each node were determined utilizing the Monte Carlo simulator. The optimal placement of five sensors in two networks was analyzed utilizing a simulation–optimization approach. EPANET, EPANET-MSX, and NSGA-III were applied as a hydraulic-quality simulator and an optimizer approach, respectively. Sensor locations have been selected based on the following four objectives: 1—sensor detection likelihood, 2—sensor expected detection time, 3—sensor detection redundancy, and 4—the affected nodes before detection. To consider the importance of contamination events and network nodes, importance coefficients

based on the amount of damage caused by contamination events have been introduced. According to the important coefficient 1, the contaminations that cause more damage to the distribution network (affect more nodes) are more important and should be identified quickly. Also, according to the important coefficient 2, nodes with more demands are more important, and sensors must be installed in places to detect contaminations before infecting these nodes. The optimal Pareto fronts were computed for each of the two sets of conflicting objectives. For example, the results of investigating the first objective function versus the second objective function showed that to maximize sensor detection likelihood in the network, sensors should be installed in the downstream nodes of the network. But installing sensors downstream of the network increases detection time. So, the layout of sensors according to the two objective functions conflict with each other. The optimal Pareto front results of these two objective functions for the benchmark network illustrated that the best detection likelihood is equal to 92.8%, which is equivalent to the worst detection time (11.58 min). Also, the best detection time is equal to 5.06 min, which indicates the worst detection likelihood (61.1%). In the next step, the sensitivity analysis related to the number of sensors on different objective functions was investigated so that the number of sensors increased from five to seven and ten for the benchmark system and increased from five to ten and fifteen for a real system. The results illustrated that as the number of sensors increased, the Pareto fronts became more efficient tools. Also, by increasing the number of sensors, the new Pareto front dominates the old Pareto front (the Pareto front of fewer sensors). For example, by increasing the number of sensors from five to ten for the real network, the best detection likelihood increased by 18.93%. Moreover, the best solution for redundancy increased from 98.87 to 98.94%, which has improved the two objective functions. Then, by increasing the number of sensors to fifteen, the best detection likelihood compared to five and ten sensors increased by 24.66 and 4.8 percent, respectively. Also, the best solution for redundancy increased to 99.1 percent. Finally, the optimal Pareto front for the number of sensors versus the best detection likelihood was depicted. For example, by increasing the number of sensors from 1 to 2, the detection likelihood increased by 73.6% for the benchmark network, and by installing only one or two sensors in the real network, 24.5% and 45.2% of pollutants can be detected. Moreover, the results demonstrated that increasing the number of sensors to more than 10 and 15 sensors in benchmark and real systems respectively, will provide little additional detection likelihood. Also, the maximum detection likelihood corresponds to 11 and 20 sensors for benchmark and real water systems. It is suggested to study the failure probability of different pollution detection scenarios as a limitation of this study. In addition, future research could address other

uncertainties, such as imperfect water quality sensors, the importance of nodes based on their applications, and the use of mobile sensors.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11356-023-26075-5>.

Author contribution All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Siroos Harif, Gholamreza Azizyan, Mohsen Dehghani Darmian, and Mohammad Givehchi. The first draft of the manuscript was written by Siroos Harif, and Mohsen Dehghani Darmian. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Availability of data and materials Data and materials are available to be shared upon request.

Declarations

Ethics approval No ethical issue is to be declared in this article.

Consent to participate This is not applicable.

Consent for publication All of the authors have read and approved the paper for publication. We confirmed that it has not been published previously nor is it being considered by any other peer-reviewed journal.

Competing interests The authors declare no competing interests.

References

- Adedoja OS, Hamam Y, Khalaf B, Sadiku R (2018a) A state-of-the-art review of an optimal sensor placement for contaminant warning system in a water distribution network. *Urban Water Journal* 15:985–1000. <https://doi.org/10.1080/1573062X.2019.1597378>
- Adedoja OS, Hamam Y, Khalaf B, Sadiku R (2018b) Towards development of an optimization model to identify contamination source in a water distribution network. *Water* 10:579. <https://doi.org/10.3390/w10050579>
- Afshar A, Miri Khombi S (2015) Multi objective optimization of sensor placement in water distribution networks dual use benefit approach. *Iran Univ Sci Technol* 5:315–331
- Al-Zahrani M, Moeid K (2001) Locating optimum water quality monitoring stations in water distribution system, Bridging the Gap: Meeting the World's Water and Environmental Resources Challenges 1–9. [https://doi.org/10.1061/40569\(2001\)393](https://doi.org/10.1061/40569(2001)393)
- Al-Zahrani MA, Moied K (2003) Optimizing water quality monitoring stations using genetic algorithms. *Arab J Sci Eng* 28:57–75
- Antunes CH, Dolores M (2016) Sensor location in water distribution networks to detect contamination events—a multi objective approach based on NSGA-II, 2016 IEEE Congress on Evolutionary Computation (CEC). IEEE. 1093–1099. doi:<https://doi.org/10.1109/CEC.2016.7743910>
- Aral MM, Guan J, Maslia ML (2010) Optimal design of sensor placement in water distribution networks. *J Water Resour Plan Manag* 136:5. [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0000001](https://doi.org/10.1061/(ASCE)WR.1943-5452.0000001)
- Bazargan-Lari MR (2014) An evidential reasoning approach to optimal monitoring of drinking water distribution systems for detecting deliberate contamination events. *J Clean Prod* 78:1–14. <https://doi.org/10.1016/j.jclepro.2014.04.061>
- Berry J, Hart WE, Phillips CA, Uber JG, Watson J-P (2006) Sensor placement in municipal water networks with temporal integer programming models. *J Water Resour Plan Manag* 132:218–224
- Berry J, Carr RD, Hart WE, Leung VJ, Phillips CA, Watson J-P (2009) Designing contamination warning systems for municipal water networks using imperfect sensors. *J Water Resour Plan Manag* 135:253–263. [https://doi.org/10.1061/\(ASCE\)0733-9496\(2009\)135:4\(253\)](https://doi.org/10.1061/(ASCE)0733-9496(2009)135:4(253))
- Bhesdadiya RH, Trivedi IN, Jangir P, Jangir N, Kumar A (2016) An NSGA-III algorithm for solving multi-objective economic/environmental dispatch problem. *Cogent Engineering* 3:1269383. <https://doi.org/10.1080/23311916.2016.1269383>
- Cao X, Wen Z, Xu J, De Clercq D, Wang Y, Tao Y (2020) Many-objective optimization of technology implementation in the industrial symbiosis system based on a modified NSGA-III. *Journal of Cleaner Production* 245:118810. <https://doi.org/10.1016/j.jclepro.2019.118810>
- Cheifetz N, Sandraz A-C, Féliers C, Gilbert D, Piller O, Lang A (2015) An incremental sensor placement optimization in a large real-world water system. *Procedia Engineering* 119:947–952. <https://doi.org/10.1016/j.proeng.2015.08.977>
- Chinadaily (2014) Lanzhou tap water tainted with benzene. Updated April 11:2014 http://www.chinadaily.com.cn/china/2014-04/11/content_17428825.htm
- Comboul M, Ghanem R (2013) Value of information in the design of resilient water distribution sensor networks. *J Water Resour Plan Manag* 139:449–455. [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0000259](https://doi.org/10.1061/(ASCE)WR.1943-5452.0000259)
- Cooper WJ (2014) Responding to crisis: the West Virginia chemical spill. *ACS Publications* 48:3095–3095. <https://doi.org/10.1021/es500949g>
- Corso PS, Kramer MH, Blair KA, Addiss DG, Davis JP, Haddix AC (2003) Costs of illness in the 1993 waterborne *Cryptosporidium* outbreak, Milwaukee. *Wisconsin Emerging Infectious Diseases* 9:426
- Costa D, Melo L, Martins F (2013) Localization of contamination sources in drinking water distribution systems: a method based on successive positive readings of sensors. *Water Resour Manag* 27:4623–4635. <https://doi.org/10.1007/s11269-013-0431-z>
- De Sanctis AE, Shang F, Uber JG (2010) Real-time identification of possible contamination sources using network backtracking methods. *J Water Resour Plan Manag* 136:444–453. [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0000050](https://doi.org/10.1061/(ASCE)WR.1943-5452.0000050)
- de Winter C, Palleti VR, Worm D, Kooij R (2019) Optimal placement of imperfect water quality sensors in water distribution networks. *Comput Chem Eng* 121:200–211. <https://doi.org/10.1016/j.compchemeng.2018.10.021>
- Deb K, Jain H (2013) An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: solving problems with box constraints. *IEEE Trans Evol Comput* 18:577–601. <https://doi.org/10.1109/TEVC.2013.2281535>
- Dehghani Darmian M, Hashemi Monfared SA, Azizyan G, Snyder SA, Giesy JP (2018) Assessment of tools for protection of quality of water: uncontrollable discharges of pollutants. *Ecotoxicol Environ Saf* 161:190–197. <https://doi.org/10.1016/j.ecoenv.2018.05.087>
- Dehghani Darmian M, Khodabandeh F, Azizyan G, Giesy JP, Hashemi Monfared SA (2020) Analysis of assimilation capacity for conservation of water quality: controllable discharges of pollutants. *Arab J Geosci* 13:1–13. <https://doi.org/10.1007/s12517-020-05907-5>
- Deshayes F, Schmitt M, Ledrans M, Gourier-Frery C, de Valk H (2001) Pollution du reseau d'eau potable a Strasbourg et survenue concomitante de gastro-enterites-mai 2000. *Bulletin épidémiologique hebdomadaire* 2:5–7

- Di Nardo A, Di Natale M, Musmarra D, Santonastaso G, Tzatchkov V, Alcocer-Yamanaka V (2014) A district sectorization for water network protection from intentional contamination. *Procedia Engineering* 70:515–524. <https://doi.org/10.1016/j.proeng.2014.02.057>
- Dorini G, Jonkergouw P, Kapelan Z, Di Pierro F, Khu S, Savic D (2008) An efficient algorithm for sensor placement in water distribution systems. *Water Distribution Systems Analysis Symposium* 2006:1–13. [https://doi.org/10.1061/40941\(247\)101](https://doi.org/10.1061/40941(247)101)
- Du R, Gkatzikis L, Fischione C, Xiao M (2015) Energy efficient sensor activation for water distribution networks based on compressive sensing. *IEEE J Sel Areas Commun* 33:2997–3010. <https://doi.org/10.1109/JSAC.2015.2481199>
- Eliades D, Polycarpou M (2006) Iterative deepening of Pareto solutions in water sensor networks, Proc. 8th Annual Water Distribution Systems Analysis Symp, ASCE, Reston, USA. 1–19. [https://doi.org/10.1061/40941\(247\)114](https://doi.org/10.1061/40941(247)114)
- Gaur A, Talukder AK, Deb K, Tiwari S, Xu S, Jones D (2017) Finding near-optimum and diverse solutions for a large-scale engineering design problem, 2017 IEEE Symposium Series on Computational Intelligence (SSCI). IEEE. 1–8. <https://doi.org/10.1109/SSCI.2017.8285271>
- Gavriel A, Landre J, Lamb A (1998) Incidence of mesophilic *Aeromonas* within a public drinking water supply in north-east Scotland. *J Appl Microbiol* 84:383–392. <https://doi.org/10.1046/j.1365-2672.1998.00354.x>
- Gong W, Cai Z, Liang D (2014) Adaptive ranking mutation operator based differential evolution for constrained optimization. *IEEE Transactions on Cybernetics* 45:716–727. <https://doi.org/10.1109/TCYB.2014.2334692>
- Gong W, Zhou A, Cai Z (2015) A multioperator search strategy based on cheap surrogate models for evolutionary optimization. *IEEE transactions on Evolutionary Computation* 19:746–758. <https://doi.org/10.1109/TEVC.2015.2449293>
- Grayman WM, Murray R, Savic aA, Farmani R, (2016) Redesign of water distribution systems for passive containment of contamination. *J Am Water Works Ass* 108:E381–E391. <https://doi.org/10.5942/jawwa.2016.108.0105>
- Gu Z-M, Wang G-G (2020) Improving NSGA-III algorithms with information feedback models for large-scale many-objective optimization. *Futur Gener Comput Syst* 107:49–69. <https://doi.org/10.1016/j.future.2020.01.048>
- Gueli R (2008) Predator-prey model for discrete sensor placement. *Water Distribution Systems Analysis Symposium* 2006:1–9. [https://doi.org/10.1061/40941\(247\)104](https://doi.org/10.1061/40941(247)104)
- Gupta M, Kumar N, Singh BK, Gupta N (2021) NSGA-III-based deep-learning model for biomedical search engines. *Math ProblEng* 2021. <https://doi.org/10.1155/2021/9935862>
- Harif S, Azizyan G, Givehchi M, Dehghani Darmian M (2021) Determining the optimal location of sensors in the wat. *Iran J Irrig Drain* 15:342–356
- Harif S, Azizyan G, Givehchi M, Dehghani Darmian M (2020) Investigation of the efficiency of dilution flow, detention time and discharge of contaminated water on quality management of water distribution network after pollution incidence (case study: drinking water network of Zahedan). *Iran J Ecohydrol* 7:1033–1045. <https://doi.org/10.22059/IJE.2020.309105.1377>
- Harif S, Azizyan G, Givehchi M, Dehghani Darmian M (2022) Analysis and evaluation of water quality management tools in water distribution networks against pollution entrance in Zahedan water distribution network. *Irrig Water Eng* 12:437–453. <https://doi.org/10.22125/IWE.2022.146421>
- Hashemi Monfared SA, Dehghani Darmian M, Snyder SA, Azizyan G, Pirzadeh B, Azhdary Moghaddam M (2017) Water quality planning in rivers: assimilative capacity and dilution flow. *Bull Environ Contam Toxicol* 99:531–541. <https://doi.org/10.1007/s00128-017-2182-7>
- Hashemi Monfared S, Walsh CL, Curtis TP, Jarvis AP, Darmian MD, Khodabandeh F (2023) New coefficient for water quality modeling in meandering rivers: fatigue factor. *Eco Inform* 101999
- He G, Zhang T, Zheng F, Zhang Q (2018) An efficient multi-objective optimization method for water quality sensor placement within water distribution systems considering contamination probability variations. *Water Res* 143:165–175. <https://doi.org/10.1016/j.watres.2018.06.041>
- He X, Dong S, Zhao N (2020) Research on rush order insertion rescheduling problem under hybrid flow shop based on NSGA-III. *Int J Prod Res* 58:1161–1177. <https://doi.org/10.1080/00207543.2019.1613581>
- Hrudey SE, Payment P, Huck P, Gillham R, Hrudey E (2003) A fatal waterborne disease epidemic in Walkerton, Ontario: comparison with other waterborne outbreaks in the developed world. *Water Sci Technol* 47:7–14. <https://doi.org/10.2166/wst.2003.0146>
- Hu C, Ren G, Liu C, Li M, Jie W (2017) A Spark-based genetic algorithm for sensor placement in large scale drinking water distribution systems. *Clust Comput* 20:1089–1099. <https://doi.org/10.1007/s10586-017-0838-z>
- Hu C, Li M, Zeng D, Guo S (2018) A survey on sensor placement for contamination detection in water distribution systems. *Wireless Netw* 24:647–661. <https://doi.org/10.1007/s11276-016-1358-0>
- Huang JJ, McBean EA (2009) Data mining to identify contaminant event locations in water distribution systems. *J Water Resour Plan Manag* 135:466–474. [https://doi.org/10.1061/\(ASCE\)0733-9496\(2009\)135:6\(466\)](https://doi.org/10.1061/(ASCE)0733-9496(2009)135:6(466))
- Huang JJ, McBean EA, James W (2008) Multi-objective optimization for monitoring sensor placement in water distribution systems. *Water Distribution Systems Analysis Symposium* 2006:1–14. [https://doi.org/10.1061/40941\(247\)113](https://doi.org/10.1061/40941(247)113)
- Kessler A, Ostfeld A, Sinai G (1998) Detecting accidental contaminations in municipal water networks. *J Water Resour Plan Manag* 124:192–198. [https://doi.org/10.1061/\(ASCE\)0733-9496\(1998\)124:4\(192\)](https://doi.org/10.1061/(ASCE)0733-9496(1998)124:4(192))
- Khodabandeh F, Dehghani Darmian M, Azhdary Moghaddam M, Hashemi Monfared SA (2021) Reservoir quality management with CE-QUAL-W2/ANN surrogate model and PSO algorithm (case study: Pishin Dam, Iran). *Arab J Geosci* 14:1–18. <https://doi.org/10.1007/s12517-021-06735-x>
- Khorshidi MS, Nikoo MR, Sadegh M (2018) Optimal and objective placement of sensors in water distribution systems using information theory. *Water Res* 143:218–228. <https://doi.org/10.1016/j.watres.2018.06.050>
- Khorshidi MS, Nikoo MR, Ebrahimi E, Sadegh M (2019) A robust decision support leader-follower framework for design of contamination warning system in water distribution network. *J Clean Prod* 214:666–673. <https://doi.org/10.1016/j.jclepro.2019.01.010>
- Klosterman ST, Uber JG, Murray R, Boccelli DL (2014) Adsorption model for arsenate transport in corroded iron pipes with application to a simulated intrusion in a water distribution network. *J Water Resour Plan Manag* 140:649–657. [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0000353](https://doi.org/10.1061/(ASCE)WR.1943-5452.0000353)
- Kumar A, Kansal M, Arora G (1997) Identification of monitoring stations in water distribution system. *Journal of Environmental Engineering* 123:546–752. [https://doi.org/10.1061/\(ASCE\)0733-9372\(1997\)123:8\(746\)](https://doi.org/10.1061/(ASCE)0733-9372(1997)123:8(746))
- Kumar A, Kansal ML, Arora G, Ostfeld A, Kessler A (1999) Detecting accidental contaminations in municipal water networks. *J Water Resour Plan Manag* 125(5):308–310. [https://doi.org/10.1061/\(ASCE\)0733-9496\(1999\)125:5\(308\)](https://doi.org/10.1061/(ASCE)0733-9496(1999)125:5(308))
- Laine J, Huovinen E, Virtanen M, Snellman M, Lumio J, Ruutu P, Kujansuu E, Vuento R, Pitkänen T, Miettinen I (2011)

- An extensive gastroenteritis outbreak after drinking-water contamination by sewage effluent, Finland. *Epidemiol Infect* 139:1105–1113. <https://doi.org/10.1017/S0950268810002141>
- Lambrou TP, Panayiotou CG, Polycarpou MM (2015) Contamination detection in drinking water distribution systems using sensor networks, 2015 European Control Conference (ECC). IEEE 3298–3303. <https://doi.org/10.1109/ECC.2015.7331043>
- Lee BH, Deininger RA (1992) Optimal locations of monitoring stations in water distribution system. *J Environ Eng* 118:4–16. [https://doi.org/10.1061/\(ASCE\)0733-9372\(1992\)118:1\(4\)](https://doi.org/10.1061/(ASCE)0733-9372(1992)118:1(4))
- Mac Kenzie WR, Hoxie NJ, Proctor ME, Gradus MS, Blair KA, Peterson DE, Kazmierczak JJ, Addiss DG, Fox KR, Rose JB (1994) A massive outbreak in Milwaukee of Cryptosporidium infection transmitted through the public water supply. *N Engl J Med* 331:161–167. <https://doi.org/10.1056/NEJM199407213310304>
- Martínez-Comesaña M, Eguía-Oller P, Martínez-Torres J, Febrero-Garrido L, Granada-Álvarez E (2022) Optimisation of thermal comfort and indoor air quality estimations applied to in-use buildings combining NSGA-III and XGBoost. *Sustain Cities Soc* 80:103723. <https://doi.org/10.1016/j.scs.2022.103723>
- Mazumder RK, Salman AM, Li Y, Yu X (2019) Reliability analysis of water distribution systems using physical probabilistic pipe failure method. *J Water Resour Plann Manage* 145:04018097. [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0001034](https://doi.org/10.1061/(ASCE)WR.1943-5452.0001034)
- McKenna SA, Hart DB, Yarrington L (2006) Impact of sensor detection limits on protecting water distribution systems from contamination events. *J Water Resour Plan Manag* 132:305–309. [https://doi.org/10.1061/\(ASCE\)0733-9496\(2006\)132:4\(305\)](https://doi.org/10.1061/(ASCE)0733-9496(2006)132:4(305))
- Munavalli G, Kumar MM (2003) Optimal scheduling of multiple chlorine sources in water distribution systems. *J Water Resour Plan Manag* 129:493–504. [https://doi.org/10.1061/\(ASCE\)0733-9496\(2003\)129:6\(493\)](https://doi.org/10.1061/(ASCE)0733-9496(2003)129:6(493))
- Nafi A, Crastes E, Sadiq R, Gilbert D, Piller O (2018) Intentional contamination of water distribution networks: developing indicators for sensitivity and vulnerability assessments. *Stoch Env Res Risk Assess* 32:527–544. <https://doi.org/10.1007/s00477-017-1415-y>
- Naserizade SS, Nikoo MR, Montaseri H (2018) A risk-based multi-objective model for optimal placement of sensors in water distribution system. *J Hydrol* 557:147–159. <https://doi.org/10.1016/j.jhydrol.2017.12.028>
- Nazempour R, Monfared MAS, Zio E (2018) A complex network theory approach for optimizing contamination warning sensor location in water distribution networks. *International Journal of Disaster Risk Reduction* 30:225–234. <https://doi.org/10.1016/j.ijdrr.2018.04.029>
- Nilsson KA, Buchberger SG, Clark RM (2005) Simulating exposures to deliberate intrusions into water distribution systems. *J Water Resour Plan Manag* 131:228–236. [https://doi.org/10.1061/\(ASCE\)0733-9496\(2005\)131:3\(228\)](https://doi.org/10.1061/(ASCE)0733-9496(2005)131:3(228))
- Nuvolone D, Petri D, Aprea MC, Bertelloni S, Voller F, Aragona I (2021) Thallium contamination of drinking water: health implications in a residential cohort study in Tuscany (Italy). *Int J Environ Res Public Health* 18:4058. <https://doi.org/10.3390/ijerph18084058>
- Ohar Z, Lahav O, Ostfeld A (2015) Optimal sensor placement for detecting organophosphate intrusions into water distribution systems. *Water Res* 73:193–203. <https://doi.org/10.1016/j.watres.2015.01.024>
- Ostfeld A, Salomons E (2004) Optimal layout of early warning detection stations for water distribution systems security. *J Water Resour Plan Manag* 130:377–385. [https://doi.org/10.1061/\(ASCE\)0733-9496\(2004\)130:5\(377\)](https://doi.org/10.1061/(ASCE)0733-9496(2004)130:5(377))
- Ostfeld A, Salomons E (2005) Securing water distribution systems using online contamination monitoring. *J Water Resour Plan Manag* 131:402–405. [https://doi.org/10.1061/\(ASCE\)0733-9496\(2005\)131:5\(402\)](https://doi.org/10.1061/(ASCE)0733-9496(2005)131:5(402))
- Ostfeld A, Salomons E (2008) Sensor network design proposal for the battle of the water sensor networks (BWSN). *Water Distribution Systems Analysis Symposium 2006*:1–16. [https://doi.org/10.1061/40941\(247\)108](https://doi.org/10.1061/40941(247)108)
- Ostfeld A et al (2008) The battle of the water sensor networks (BWSN): A design challenge for engineers and algorithms. *J Water Resour Plan Manag* 134:556–568. [https://doi.org/10.1061/\(ASCE\)0733-9496\(2008\)134:6\(556\)](https://doi.org/10.1061/(ASCE)0733-9496(2008)134:6(556))
- Palleti VR, Narasimhan S, Rengaswamy R, Teja R, Bhallamudi SM (2016) Sensor network design for contaminant detection and identification in water distribution networks. *Comput Chem Eng* 87:246–256. <https://doi.org/10.1016/j.compchemeng.2015.12.022>
- Palleti VR, Kurian V, Narasimhan S, Rengaswamy R (2018) Actuator network design to mitigate contamination effects in water distribution networks. *Comput Chem Eng* 108:194–205. <https://doi.org/10.1016/j.compchemeng.2017.09.003>
- Ponti A, Candelieri A, Archetti F (2021) A new evolutionary approach to optimal sensor placement in water distribution networks. *Water* 13:1625. <https://doi.org/10.3390/w13121625>
- Preis A, Ostfeld A (2008) Multiobjective contaminant sensor network design for water distribution systems. *J Water Resour Plan Manag* 134:366–377. [https://doi.org/10.1061/\(ASCE\)0733-9496\(2008\)134:4\(366\)](https://doi.org/10.1061/(ASCE)0733-9496(2008)134:4(366))
- Priya SK, Shenbagalakshmi G, Revathi T (2018) Design of smart sensors for real time drinking water quality monitoring and contamination detection in water distributed mains. *Int J Eng Technol* 7:47–51. <https://doi.org/10.14419/ijet.v7i1.1.8921>
- Propato M (2006) Contamination warning in water networks: General mixed-integer linear models for sensor location design. *J Water Resour Plan Manag* 132:225–233. [https://doi.org/10.1061/\(ASCE\)0733-9496\(2006\)132:4\(225\)](https://doi.org/10.1061/(ASCE)0733-9496(2006)132:4(225))
- Propato M, Piller O (2006) Battle of the water sensor networks, 8th annual Water Distribution System Analysis Symposium. University of Cincinnati, Cincinnati, Ohio USA
- Qian K, Jiang J, Ding Y, Yang S-H (2021) DLGEA a deep learning guided evolutionary algorithm for water contamination source identification. *Neural Comput Appl* 33:11889–11903. <https://doi.org/10.1007/s00521-021-05894-y>
- Qiu M, Salomons E, Ostfeld A (2020) A framework for real-time disinfection plan assembling for a contamination event in water distribution systems. *Water research* 174:115625. <https://doi.org/10.1016/j.watres.2020.115625>
- Rathi S, Gupta R (2016) A simple sensor placement approach for regular monitoring and contamination detection in water distribution networks. *KSCE J Civ Eng* 20:597–608. <https://doi.org/10.1007/s12205-015-0024-x>
- Rosen JS, Whelton AJ, McGuire MJ, Clancy JL, Bartrand T, Eaton A, Patterson J, Dourson M, Nance P, Adams C (2014) The crude MCHM chemical spill in Charleston. *W Va Journal-American Water Works Association* 106:65–74. <https://doi.org/10.5942/jawwa.2014.106.0134>
- Rossman LA (2000) EPANET 2: users manual
- Schwartz R, Lahav O, Ostfeld A (2014a) Optimal sensor placement in water distribution systems for injection of chlorpyrifos. *World Environmental and Water Resources Congress 2014*:485–494. <https://doi.org/10.1061/9780784413548.052>
- Schwartz R, Lahav O, Ostfeld A (2014b) Integrated hydraulic and organophosphate pesticide injection simulations for enhancing event detection in water distribution systems. *Water Res* 63:271–284. <https://doi.org/10.1016/j.watres.2014.06.030>
- Shahmirnoori A, Saadatpour M, Rasekh A (2022) Using mobile and fixed sensors for optimal monitoring of water distribution network under dynamic water quality simulations. *Sustain Cities Soc* 82:103875. <https://doi.org/10.1016/j.scs.2022.103875>

- Shang F, Uber JG, Rossman LA, Janke R (2008) EPANET multi-species extension user's manual. Risk Reduction Engineering Laboratory, US Environmental Protection Agency, Cincinnati, Ohio
- Shastri Y, Diwekar U (2006) Sensor placement in water networks: a stochastic programming approach. *J Water Resour Plan Manag* 132:192–203. [https://doi.org/10.1061/\(ASCE\)0733-9496\(2006\)132:3\(192\)](https://doi.org/10.1061/(ASCE)0733-9496(2006)132:3(192))
- Srinivas N, Deb K (1994) Multiobjective optimization using nondominated sorting in genetic algorithms. *Evol Comput* 2:221–248. <https://doi.org/10.1162/evco.1994.2.3.221>
- Whelton AJ, McMillan L, Connell M, Kelley KM, Gill JP, White KD, Gupta R, Dey R, Novy C (2015) Residential tap water contamination following the freedom industries chemical spill: perceptions, water quality, and health impacts. *Environ Sci Technol* 49:813–823. <https://doi.org/10.1021/es5040969>
- Wilson E (2014) Danger from microcystins in Toledo water unclear. *Chem Eng News* 92:9
- Woo H-M, Yoon J-H, Choi D-Y (2001) Optimal monitoring sites based on water quality and quantity in water distribution systems, Bridging the Gap: Meeting the World's Water and Environmental Resources Challenges 1–9
- Wu ZY, Walski T (2006) Multi objective optimization of sensor placement in water distribution systems, 8th Annual Symp. on Water Distribution Systems Analysis. Environmental and Water Resources Institute of ASCE (EWRI of ASCE) New York. [https://doi.org/10.1061/40941\(247\)105](https://doi.org/10.1061/40941(247)105)
- Xu Z, Guo Y, Saleh JH (2022) Multi-objective optimization for sensor placement: an integrated combinatorial approach with reduced order model and Gaussian process. *Measurement* 187:110370. <https://doi.org/10.1016/j.measurement.2021.110370>
- Yazdi J (2018) Water quality monitoring network design for urban drainage systems, an entropy method. *Urban Water Journal* 15:227–233. <https://doi.org/10.1080/1573062X.2018.1424215>
- Yokoyama K (2007) Our recent experiences with sarin poisoning cases in Japan and pesticide users with references to some selected chemicals. *Neurotoxicology* 28:364–373. <https://doi.org/10.1016/j.neuro.2006.04.006>
- Young HP (1993) An evolutionary model of bargaining. *Journal of Economic Theory* 59:145–168. <https://doi.org/10.1006/jeth.1993.1009>
- Yuan X, Tian H, Yuan Y, Huang Y, Ikram RM (2015) An extended NSGA-III for solution multi-objective hydro-thermal-wind scheduling considering wind power cost. *Energy Convers Manage* 96:568–578. <https://doi.org/10.1016/j.enconman.2015.03.009>
- Zheng F, Du J, Diao K, Zhang T, Yu T, Shao Y (2018) Investigating effectiveness of sensor placement strategies in contamination detection within water distribution systems. *J Water Resour Plan Manag* 144:06018003. [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0000919](https://doi.org/10.1061/(ASCE)WR.1943-5452.0000919)
- Zhu Y, Liang J, Chen J, Ming Z (2017) An improved NSGA-III algorithm for feature selection used in intrusion detection. *Knowledge-Based Syst* 116:74–85. <https://doi.org/10.1016/j.knosys.2016.10.030>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.