DOI: 10.4408/IJEGE.2018-01.S-15

IMPACT OF OBJECTIVE FUNCTION SELECTION ON OPTIMAL PLACEMENT OF SENSORS IN WATER DISTRIBUTION NETWORKS

Silvia TINELLI, Enrico CREACO & Carlo CIAPONI

University of Pavia - Dipartimento di Ingegneria Civile e Architettura - Via Ferrata, 3 - 27100 Pavia, Italy
Corresponding author: silvia.tinelli01@universitadipavia.it

EXTENDED ABSTRACT

Nell'ambito della filiera di produzione e di erogazione dell'acqua potabile, la rete di distribuzione idrica rappresenta l'elemento più complesso da analizzare e gestire per quanto riguarda la sicurezza qualitativa dell'acqua consegnata all'utenza, a causa della grande dispersione dei punti di potenziale contaminazione. Per questo motivo, soprattutto nei Paesi più sensibili al problema degli attacchi terroristici, sono stati da tempo avviati programmi di ricerca finalizzati alla messa a punto di sistemi di monitoraggio continuo e di allarme precoce (*Early Warning System* - EWS) basati su sensori, da installare in punti opportunamente scelti della rete, in grado di rilevare in tempi rapidi gli eventi di contaminazione.

La definizione del numero e della localizzazione delle stazioni di monitoraggio (più semplicemente indicate come "sensori") da inserire nella rete di distribuzione idrica rappresenta un aspetto cruciale della progettazione di un EWS. Il contaminante, che può essere immesso, accidentalmente o intenzionalmente, in un qualsiasi punto della rete di distribuzione, si propaga prevalentemente in relazione ai flussi idraulici che si determinano in base alla domanda idrica. Se nella rete sono installati molti sensori, la probabilità che la contaminazione venga rilevata in tempi brevi è elevata; viceversa, se i sensori sono pochi, o ubicati in punti poco significativi della rete, il contaminante potrebbe essere rilevato dopo molto tempo dall'istante di immissione, o potrebbe addirittura non essere rilevato se il flusso che lo veicola non transita in un punto monitorato. L'efficacia del sistema EWS dipende quindi dal numero e dalla localizzazione dei sensori. Per un prefissato numero di sensori, necessariamente limitato per ragioni di costo, la migliore localizzazione è quella che ne massimizza l'efficacia, ovvero la capacità del sistema di ridurre l'impatto degli incidenti di contaminazione sulla salute pubblica.

Tale criterio, che contempla il minimo impatto sulla salute pubblica, deve essere tradotto in *funzioni obiettivo* esprimibili in termini quantitativi, così che possano essere introdotte nei modelli di ottimizzazione (Hart & Murray, 2010).

Ad esempio, una possibile funzione obiettivo esprimibile in termini quantitativi corrisponde alla probabilità (da massimizzare) che il flusso contaminato transiti per un punto monitorato della rete (probabilità di rilevamento). In alternativa, la funzione obiettivo può essere espressa attraverso grandezze (da minimizzare) quali, ad esempio, il tempo intercorrente fra la contaminazione e il suo rilevamento, il numero degli abitanti che in questo tempo sono raggiunti dal contaminante, il numero degli abitanti che ricevono una concentrazione di contaminante superiore ad una determinata soglia, il quantitativo di acqua contaminata erogata, la percentuale degli eventi di contaminazione non rilevati.

Nel presente articolo questo problema è affrontato esaminando come la scelta della funzione obiettivo influenzi il risultato finale. A tal fine, sono state esaminate e fra loro comparate due diverse impostazioni, entrambe basate sull'impiego del numero dei sensori come prima funzione obiettivo (da minimizzare). Le due impostazioni si differenziano invece per la seconda funzione obiettivo che è stata assunta rispettivamente corrispondente alla probabilità di rilevamento (da massimizzare) e all'entità della popolazione raggiunta dal contaminante (da minimizzare). I risultati delle ottimizzazioni e le rivalutazioni delle soluzioni ottimali in termini di alcuni indicatori dell'efficacia del sistema di monitoraggio mostrano che la prima impostazione (F.O.= probabilità di rilevamento) produce soluzioni più efficaci per quanto riguarda la probabilità di rilevamento e il grado di ridondanza del sistema di monitoraggio. Per contro, la seconda impostazione (F.O.= n° utenti contaminati) produce soluzioni più efficaci con riferimento alla riduzione dell'entità della popolazione raggiunta dalla contaminazione e del tempo intercorrente fra l'inizio della contaminazione e il suo rilevamento. La scelta fra le due impostazioni va fatta quindi tenendo conto, anche in relazione alla situazione specifica e agli interventi programmati in caso di allarme, se sia preferibile privilegiare la sicurezza del rilevamento o la sua tempestività.

L'articolo evidenzia anche che le due differenti impostazioni danno origine a localizzazioni dei sensori sensibilmente diverse fra loro. Infatti, mentre la prima impostazione tende a localizzare i sensori nell'area in cui converge la maggior parte dei flussi di acqua, la seconda produce una distribuzione più diffusa su tutta la rete.

ABSTRACT

This paper investigates how the choice of the effectiveness-related objective function affects the results of the optimal placement of water quality sensors in water distribution networks (WDNs). The methodology adopted is based on bi-objective optimization, with the number of installed sensors as first objective function to be minimized. As for the choice of the second objective function, representative of the effectiveness of the monitoring system to react to a set of potential contamination events, two variants of optimization were considered: variant 1 - event detection likelihood to be maximized; variant 2 - average contaminated population to be minimized. The analysis of the results of the optimizations, which were re-evaluated in terms of four different system effectiveness indicators, proved that neither optimization variant is numerically superior. The choice of the objective function also impacts the physical placement of the sensors, with locations at maximum distance from the source(s) and more scattered over the layout for the two variants respectively.

Keywords: *sensors; water distribution networks, water pollution, network design*

INTRODUCTION

Water distribution networks (WDNs) may experience various kinds of water quality problems during their useful life, such as those related to terroristic contamination attacks or to the accidental ingress of materials and pollutants during maintenance works, resulting from a lack of a hydraulic seal. Furthermore, contaminants may also enter WDNs because of the presence of other pressurized water supplies, such as the irrigation supply, at user connections, in the absence of suitable and effective hydraulic disconnections.

To successfully detect potential contamination events and to minimize their impact, WDNs can be equipped with water quality monitoring systems. These systems include sensors installed at strategic locations, selected in such a way as to guarantee early warning and reduced impact (Walski *et alii*, 2003).

In the last two decades, numerous optimization methodologies were set-up for monitoring system design, in which the optimal placement of sensors in WDNs need to be determined (e.g., the single objective methodologies of Lee & Deininger, 1992; Kumar *et alii*, 1997; Kessler *et alii*, 1998; Ostfeld & Salomons, 2004, 2005; Berry *et alii*, 2006, 2009; Propato, 2006; Shastri & Diwekar, 2006; Cheifetz *et alii*, 2015 and the multi-objective methodologies of McKenna *et alii*, 2006; Preis & Ostfeld, 2008; Tinelli *et alii*, 2017). In the context of multi-objective optimization, numerous objective functions were adopted to characterize the monitoring system (e.g., Ostfeld *et alii*, 2008; Presis & Ostfeld, 2008), including number of installed sensors, as a surrogate for the cost, and event detection time, contaminated population and sensor redundancy, as surrogates for system effectiveness. Though all these variables could be simultaneously considered in the same optimization framework, optimization techniques lose resolution effectiveness as the number of objective functions grows (Creaco *et alii*, 2016). Furthermore, there is no consensus amongst researchers on the number and type of performance objectives to be considered and several other issues related to sensor location problem (Rathi & Gupta, 2014). Therefore, while considering a total number of two conflicting objectives, associated with the cost and effectiveness of the monitoring system respectively, the issue arises of which single pair of objective functions can give the best results in the context of the optimal placement of sensors in WDNs. This work aims to explore this issue.

In the following sections, first the methodology based on the bi-objective optimization is presented, followed by the applications, where first the case study is described and then the results are reported. The paper end with the conclusions.

METHODOLOGY

A bi-objective optimization is used to search for the optimal locations of water quality sensors in the network. Two different variants are developed, both adopting, for the first objective function f_1, the number of installed sensors, as a surrogate for the total cost of the monitoring system. The difference between the two variants lies in the choice of the second objective function f_2, which accounts for the performance of the monitoring system.

In detail, the former variant considers the detection likelihood, which is the probability of events being detected by at least one of the installed sensors. This function to be maximized inside the optimization is calculated as follows:

$$\textit{Detection likelihood} = 1/S \sum_{r=1}^{S} d_r \tag{1}$$

where S is the total number of potential contamination events considered in the analysis. Variable d_r is equal to 1 if at least one sensor detects the r-th contamination event; otherwise, it is equal to 0.

The second variant, instead, uses the average population contaminated before the first detection of the generic event. This function to be minimized inside the optimization is expressed as follows:

$$\textit{Population} = 1/S \sum_{r=1}^{S} pop_r \tag{2}$$

where pop_r accounts for the inhabitants served by the contaminated nodes, till one of the sensors installed in the network detects the r-th contamination event. It is then assumed that a warning is given to interrupt the network service in a reaction time interval $\Delta treact$ =0 after the first detection. It is worth remarking that, when the r-th event is not detected, pop_r includes all the nodes crossed by the contamination till the whole contaminant mass leaves the network.

For each solution considered inside the optimization process, variables d_r and pop_r can be assessed through simple manipulations on matrices that have been calculated *una tantum* before the beginning of the optimization, starting from water quality simulations of the network (Tinelli *et alii*, 2017). These simulations are carried out through Epanet under the assumption of conservative contaminant.

The optimization problem is solved through the Non-dominated Sorting Genetic Algorithm II (NSGA-II) (Deb *et alii*, 2002), known for its efficiency in solving complex multi-objective optimization problems. In NSGAII, the number of genes in the n_{in} individuals is equal to the number of network nodes that can be fitted with a sensor. The possible values for each gene are 0 and 1, representing absence and presence of the sensor in the associated node, respectively. The operation of NSGAII can be summarized in the following steps:

0 - random generation of n_{in} initial individuals;
1 - selection of parent individuals;
2 - generation of offspring individuals starting from parent individuals through crossover and mutation processes;
3 - combination of parent and offspring individuals;
4 - application of fitness criteria to select n_{in} individuals in the combined group.

It has to be remarked that step 4 is essential to keep the number of individuals constantly equal to n_{in} over the course of the NSGAII run.

The sequence of Steps 1-4 is a generation and is repeated for a pre-fixed number of times. After the maximum number of generations has been reached, the n_{in} individuals obtained in Step 4 constitute the final solution of the NSGAII run. A certain number (n_{iter}) of iterations of n_{par} NSGAII parallel runs (Tinelli *et alii*, 2017) can be carried out to improve the robustness of the end solutions, which are expected to be close to the global optima.

As the objectives clearly compete against each other, the output of the optimization consists of a set of trade-off solutions, that is the Pareto front.

Various criteria can be used by the decision maker to select the ultimate solution, such as a constraint in f_1 or f_2. Otherwise, the knee-point in the Pareto front can be identified, in which an increase in the cost of the monitoring systems f_1 is no longer paid back by a significant benefit in terms of f_2.

APPLICATIONS

Case-study

The case-study considered in this work is a network in northern Italy (Guidorzi *et alii*, 2009; Creaco & Franchini, 2012), made up of 536 demanding nodes, 825 pipes and 2 reservoirs (layout in Figure 1).

The procedure (Tinelli *et alii*, 2017) was used to sample the representative features, in terms of location, mass rate, duration and starting time of the S contamination events to be considered in the analysis. As a result, all 536 demanding nodes of the WDS were considered possible injection locations. Single values of mass rate and injection duration, equal to 200 g/min and 60 min, respectively, were considered following the assumption that contamination events should be massive. Only one representative starting time was accounted for, that is 8:00 a.m., because preliminary analyses showed the network to have a single operating condition (i.e., no flow inversion at any pipes). The overall number S of contamination events was then equal to 536.

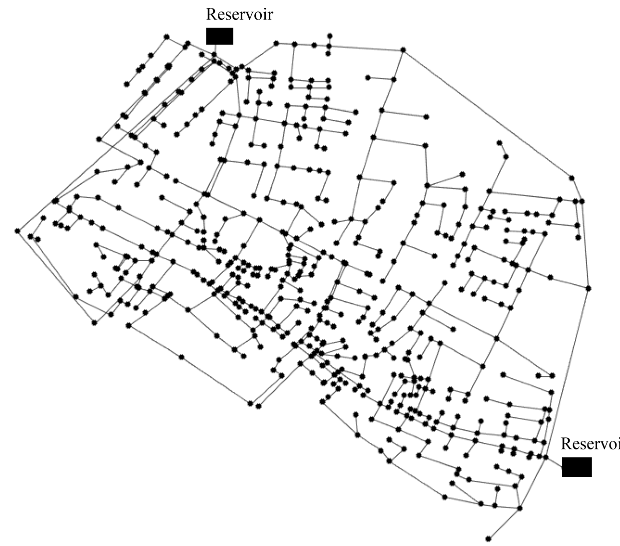

Fig. 1 - Network layout

The NSGAII settings were chosen based on the results of preliminary simulations unreported here, which enabled obtaining a trade-off between accuracy of the results and computational overhead. In detail, n_{in} and the maximum number of generations were both set at 500. Furthermore, n_{par} and n_{iter} were both set at 5.

Results

The graphs in Figure 2 show the Pareto fronts of optimal trade-off solutions in in the two variants of optimization.

In graph a), associate with the first variant, a monotonous trend of $f_2(f_1)$ is shown, in which a significant benefit in terms of detection likelihood (f_2) is obtained as the number of installed sensors (f_1) increases up to about 10, which is close to the knee of the front. A further increase in f_1 does not yield significant benefits. Compared to graph a), the main difference of graph b) lies in the monotonous decreasing trend of the contaminated population f_2. The position of the knee of the front in the results of the second variant is also close to f_1=10.

To compare thoroughly the solutions obtained in the two variants of optimization, these solutions were re-evaluated in terms of four effectiveness indicators for the water quality monitoring system. Besides the detection likelihood and

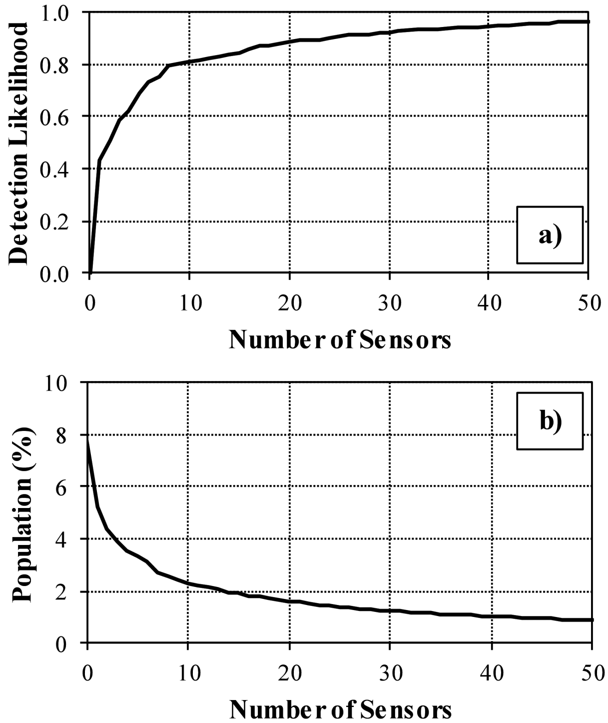

Fig. 2 - Pareto fronts obtained in the first a) and second b) variant of optimization

contaminated populations, evaluated over the whole group of *S* events through equations 1 and 2 respectively, the detection time and the sensor redundancy were adopted as a benchmark. These two additional indicators, instead, were assessed over the sub-group of detected events, that is the events that are detected by at least one sensors. In detail, the detection time is the average time elapsing between the contamination start and the time instant when the first sensor is reached. The redundancy, instead, is defined as the average number of sensors (including the first) that detect the contamination within 30 minutes from the first detection, which contributes to the safety of the monitoring systems.

The graphs in Figure 3 show the curves of re-evaluated solutions plotted against the number of installed sensors.

Looking at the solutions of the first variant of optimization, the curve in graph a) coincides with the Pareto front in Figure 2a and then features a monotonous increasing trend. The trend of the curves in the other graphs is not strictly monotonous since the contaminated population (graph b), the detection time (graph c) and the sensor redundancy (graph d) were not objective functions in the first variant of optimization. In fact, optimal solutions are usually sub-optimal when re-evaluated in terms of different indicators from the objective functions used in the optimization.

Analogously, looking at the solutions of the second variant of optimization, the curve in graph b) coincides with the Pareto front in Figure 2b and then features a monotonous decreasing

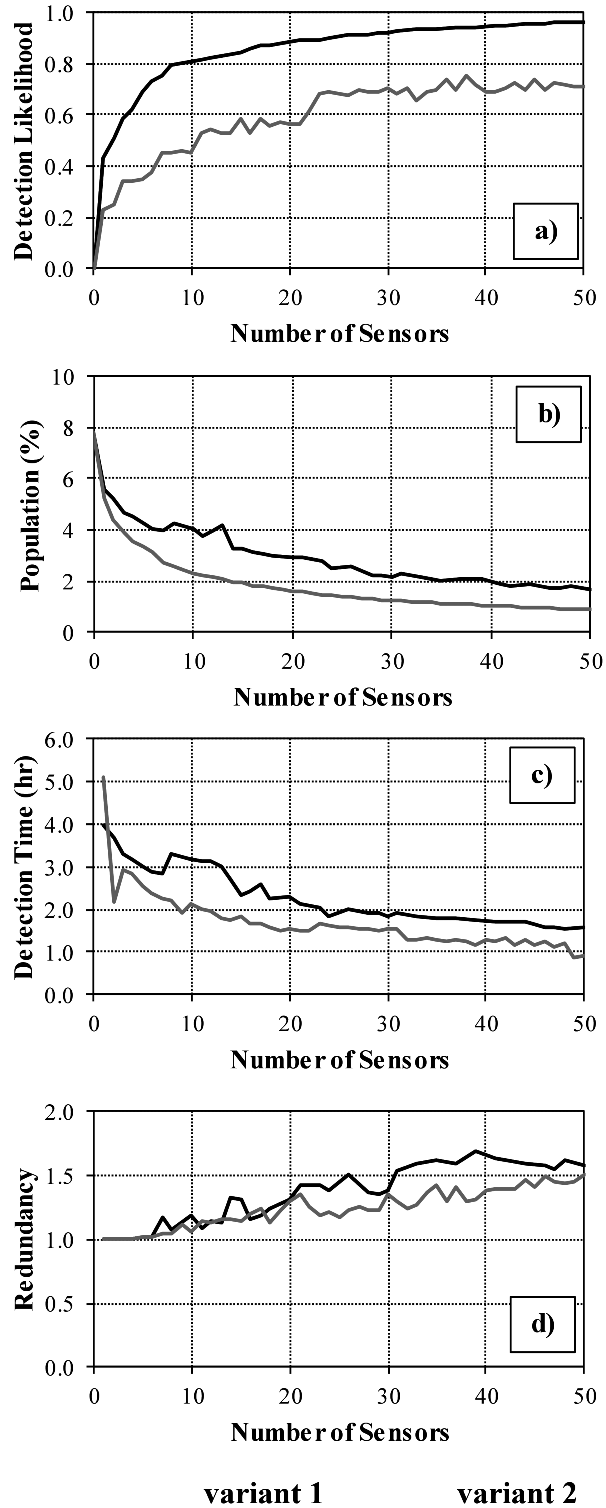

Fig. 3 - Solutions obtained in the two variants of optimization, re-evaluated in terms of a) detection likelihood, b) contaminated population, c) detection time and d) sensor redundancy

	First Variant Detection Likelihood	Second Variant Contaminated Population
Detection Likelihood	1.00	-0.96
Contaminated Population	-0.94	1.00
Detection Time	-0.90	-0.84
Sensor Redundancy	0.83	0.84

Tab. 1 - Correlation coefficient between the objective functions used in the two variants of optimization and the four effectiveness indicators

trend. The trend of the curves in the other graphs is not strictly monotonous since the detection likelihood (graph a), the detection time (graph c) and the sensor redundancy (graph d) were not objective functions in the second variant of optimization. However, as Table 1 shows, the four effectiveness indicators are always strongly intercorrelated in both variants of optimization.

Overall, the analysis of the results in Figure 3 shows that neither variant of optimization is superior. In fact, the first variant yields solutions that perform better in terms of detection likelihood and sensor redundancy, both positive indicators of the effectiveness of the monitoring system (black line above grey line in graphs a and d). The second variant, instead, produces better performing solutions in terms of contaminated population and detection time, both inverse indicators of the effectiveness of the monitoring system (grey line below black line in graphs b and c). However, by leaning on graphs such as those in Figure 3, water utility managers can choose the ultimate solution for in-situ installation based on their budget (which impacts the number of installed sensors), on the effectiveness indicator they prefer and on the degree of effectiveness they aim to reach in terms of the various indicators. As an example, the solution obtained through variant 1 with 10 sensors has a detection likelihood of 0.81, a contaminated population of 4%, a detection time of 3.2 hr and a redundancy of 1.2. The solution obtained through variant 2 with 10 sensors, instead, features almost halved detection likelihood (0.45) and contaminated population (2.3%), a lower detection time (2.1 hr) and a similar sensor redundancy (1.1).

Another criterion that can be adopted for the choice concerns the location of the sensors in the various optimal solutions. As an example, Figure 4 enables analysis and comparison of the results of the two variants of optimization, in terms of optimal placement of 10 sensors. Figure 4a shows that the placement obtained in the first variant is made up of sensors located in the intermediate area of the network, that is at the maximum hydraulic distance from either reservoir. This happens because most water paths outgoing from the reservoirs converge to this area. Therefore, the placement of sensors in this area is essential for maximizing the event detection likelihood. In the second variant, sensors are more scattered over the whole layout at gradually increasing distance from the reservoirs (Fig. 4b), to guarantee early warning and therefore reduced impact in terms of contaminated population.

CONCLUSIONS

In this paper, the problem of optimal placement of water quality sensors was tackled by making use of the bi-objective optimization. Two different variants of optimization were considered. Both variants featured the total number of sensors as first objective function to minimize, as a surrogate for the cost of the monitoring system. The two variants differed, instead, in the second objection function, which was the likelihood of contamination event detection (to maximize) and the contaminated population (to minimize) for the former and latter variant, respectively. Optimizations were carried out through NSGAII. The results of the optimizations, and the re-evaluations of the optimal solutions in terms of various effectiveness indicators for the water quality monitoring system, prove that the first variant tends to produce better solutions in terms of detection likelihood and sensor redundancy. The second, instead, tends to produce better solutions in terms of contaminated population and event detection time. However, all the effectiveness indicators are well intercorrelated in the solutions of the optimizations. The ultimate choice of water utility managers is based on their

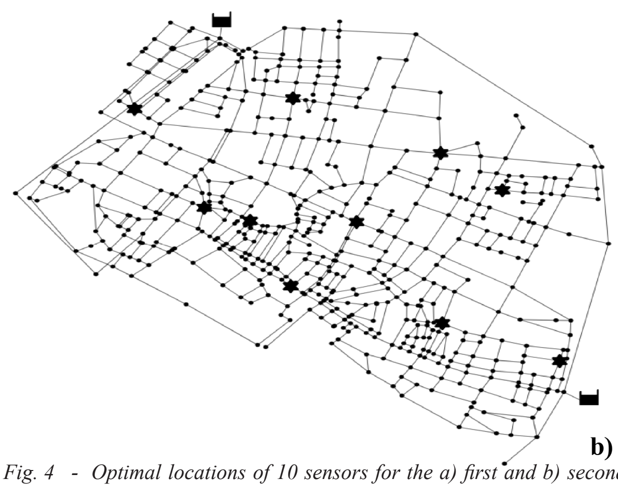

Fig. 4 - Optimal locations of 10 sensors for the a) first and b) second variant of optimization

preferences. In fact, minimizing the contaminated population yields benefits in terms of detection time and thus mainly contributes to the system's early warning capacity. On the other hand, maximizing the detection likelihood strongly impacts on the system redundancy and therefore contributes to the system safety.

A further difference between the two variants of optimization analyzed lies in the placement of sensors in the network layout. In fact, whereas the first variant tends to locate the sensors in the area where most water paths converge, the second produces a more scattered distribution over the layout.

REFERENCES

Berry J.W., Carr R.D., Hart W.E., Leung V.J., Phillips C.A. & Watson J.P. (2009) - *Designing contamination warning systems for municipal water networks using imperfect sensors.* J. Water Resour. Plann. Manage., 10.1061/(ASCE)0733-9496(2009)135:4(253), 253-263.

Berry J.W., Hart W.E., Phillips C.A., Uber J.G. & Watson J.P. (2006) - *Sensor placement in municipal water networks with temporal integer programming models.* J. Water Resour. Plann. Manage., 10.1061/(ASCE)0733-9496(2006)132:4(218), 218-224.

Cheifetz N., Sandraza A.C., Feliers C., Gilbert D., Piller O. & Lang A. (2015) - *An incremental sensor placement optimization in a large real-world water system.* Procedia Eng., **119**: 947-952.

Creaco E. & Franchini M. (2012) - *Fast network multi-objective design algorithm combined with an a-posteriori procedure for reliability evaluation under various operational scenarios.* Urban Water J., **9** (6): 385-399.

Creaco E., Franchini M. & Todini E. (2016) - *The combined use of resilience and loop diameter uniformity as a good indirect measure of network reliability*. Urban Water Journal, **13** (2): 167-181.

Deb K., Pratap A., Agrawal S. & Meyarivan T. (2002) - *A fast and elitist multiobjective genetic algorithm: NGSA-II.* IEEE Trans. Evol. Comput., **6** (2), 182-197.

EPANET 2.00.10 [Computer software]. DCUSEPA, Washington, D.C.

Guidorzi M., Franchini M., & Alvisi S. (2009) - *A multi-objective approach for detecting and responding to accidental and intentional contamination events in water distribution systems.* Urban Water J., **6** (2): 115-135.

Hart W.E. & Murray R. (2010) - *Review of Sensor Placement Strategies for Contamination Warning Systems in Drinking Water Distribution Systems.* J. Water Resour. Plann. Manage, 136(6):611-619.

Kessler A., Ostfeld A. & Sinai G. (1998) - *Detecting accidental contaminations in municipal water networks.* J. Water Resour. Plann. Manage., 10.1061/(ASCE)0733-9496(1998)124:4(192), 192-198.

Kumar A., Kansal M.L. & Arora G. (1997) - *Identification of monitoring stations in water distribution system.* J. Environ. Eng., 10.1061/(ASCE)0733-9372(1997)123:8(746), 746-752.

Lee B.H. & Deininger R.A. (1992) - *Optimal locations of monitoring stations in water distribution systems*. J. Environ. Eng., 10.1061/(ASCE)0733-9372(1992)118:1(4), 4-16.

McKenna S.A., Hart D.B., & Yarrington L. (2006) - *Impact of sensor detection limits on protecting water distribution systems from contamination events.* J. Water Resour. Plann. Manage., 10.1061/(ASCE)0733-9496(2006)132:4(305), 305–309.

Ostfeld A., & Salomons E. (2004) - *Optimal layout of early warning detection stations for water distribution systems security*. J. Water Resour. Plann. Manage., 10.1061/(ASCE)0733-9496(2004)130:5(377), 377-385.

Ostfeld A. & Salomons E. (2005) - *Securing water distribution systems using online contamination monitoring.* J. Water Resour. Plann. Manage., 10.1061/(ASCE)0733-9496(2005)131:5(402), 402-405.

Ostfeld A. *et alii* - *The battle of the water sensor networks (BWSN): a design challenge for engineers and algorithms*. J. Water Resour. Plann. Manage., 10.1061/(ASCE)0733-9496(2008)134:6(556), 556–568.

Preis A. & Ostfeld A. (2008) - *Multiobjective contaminant sensor network design for water distribution systems.* J. Water Resour. Plann. Manage., 10.1061/(ASCE)0733-9496(2008)134:4(366), 366–377.

Propato M. (2006) - *Contamination warning in water networks: general mixed-integer linear models for sensor location design.* J. Water Resour. Plann. Manage., 10.1061/(ASCE)0733-9496(2006)132:4(225): 225–233.

Rathi S. & Gupta R. (2014) - *Sensor placement methods for contamination detection in water distribution networks: a review.* Procedia Eng., **89**: 181-188.

Shastri Y. & Diwekar U. (2006) - *Sensor placement in water networks: A stochastic programming approach*. J. Water Resour. Plann. Manage., 10.1061/(ASCE)0733-9496(2006) 132:3(192), 192-203.

Tinelli S., Creaco E., & Ciaponi C. (2017) - *Sampling significant contamination events for optimal sensor placement in water distribution systems*. J. Water Resour. Plann. Manage., **143** (9): 04017058.

Walski M., Chase D., Savic D., Grayman W., Beckwith S. & Koelle E. (2003) - *Advanced water distribution modelling and management.* Haestad, Waterbury, CT.

Received April 2017 - Accepted November 2017