Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 55-6 (2022) 482–489

Multi-objective, rule and preference-based placement of quality sensors in water supply networks

Bruno M. Brentan * **Silvia Carpitella** ** **Joaquín Izquierdo** *** **Idel Montalvo** ****

* *Engineering School, Federal University of Minas Gerais (UFMG), Belo Horizonte, Brazil (e-mail: brentan@ehr.ufmg.br).*
** *Department of Decision-Making Theory, Institute of Information Theory and Automation (UTIA) - Czech Academy of Sciences; 182 08 Prague, Czech Republic (e-mail: carpitella@utia.cas.cz)*
*** *Fluing-Institute for Multidisciplinary Mathematics (IMM), Universitat Politècnica de València; Cno. de Vera s/n, 46022 Valencia, Spain (e-mail: jizquier@upv.es)*
**** *Ingeniousware GmbH, Jollystraße 11, 76137 Karlsruhe, Germany (e-mail: imontalvo@ingeniousware.net)*

Abstract: To detect contaminant intrusion and, in general, to assess quality problems in their water distribution systems, water utilities need quality sensors that continuously measure, directly from the network, conductivity, PH, concentration of different substances, and other related parameters. Due to the nature of the objectives involved, the decision about where to place sensors in the network and the amount of them to be installed is a very challenging problem. In this investigation, we present a multi-objective approach to cast light on those decisions. Instead of a crisp solution, the multi-objective approach will provide a wide spectrum of solutions representing the best trade-off among all the decision criteria of the problem. This approach aims to integrate the practical experience of engineers into the decision-making process since, eventually, the solution will be selected among the Pareto front of solutions using the engineers' experience and the specific characteristics of their utility. To this end, the used algorithm adds agents based on both technical and user-preference rules on top of evolutionary search techniques to explore the decision space. The algorithm runs as a part of the Agent Swarm Optimization framework, a consolidated multi-objective software. Another novelty of this contribution is computational: the evaluation of the objective functions is executed directly in the MS SQL server and simulation data is never required to be loaded in their entirety. Without this important implementation detail, the solution for "large" water network models would not be affordable with the hardware typically used in desktop computers. To illustrate the solution process, a use case focused on a mid-size water supply network is addressed.

Keywords: water quality;contaminant intrusion;sensor placement;condition monitoring;active diagnosis;multi-objective optimization;rule-based agent; safety-critical networked system

1. INTRODUCTION AND LITERATURE REVIEW

It is difficult to ascertain the importance of drinking water supply infrastructures for the sustainable existence and development of modern cities (Gandy (2004); Hoekstra and Mekonnen (2012)). Water distribution systems (WDSs) are complex structures formed by pipes, valves, pumps, tanks, and other elements designed and erected to transport water of sufficient quality from water sources to consumers. The amount of the above elements, which can reach up to tens of thousands of links and junctions, their frequently wide spatial dispersion, and the WDS characteristic of being very dynamic structures make the management of real WDSs a complex problem (Perelman and Ostfeld (2012); Izquierdo et al. (2012); Ostfeld (2012); Diao et al. (2016)). However, although the main objective is to supply water in the quantity and quality required, other requirements are essential, namely maintaining conditions far from failure scenarios (Ostfeld et al. (2014); Herrera et al. (2016)), ability to quickly detect sources of contamination intrusion (Islam et al. (2015); Nafi et al. (2018)), minimization of leaks (Covas and Ramos (1999); Candelieri et al. (2014)), etc.

Most of these objectives may be achieved through the suitable location of sensors along the network and, currently, an increasing number of efforts are carried out in this direction (Zhao et al. (2016); Sarrate et al. (2014); Rathi and Gupta (2017); Antunes and Dolores (2016). The identification of potential contaminant intrusion in water networks is a crucial point to fully guarantee water quality in distribution systems. As a consequence, water utilities are bound to measure water quality parameters continuously, so that quality can be adequately monitored.

Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2022.07.175

To this end, an optimal lattice of sensors should be designed using strategic points of the water network (Oliker and Ostfeld (2015)). As this is a matter of safety and security arrangement in WDS management, sensors cannot be randomly placed along the network. Placing sensors may seem simple at the beginning; however, considering sensor costs and the extension of the pipes that should be covered, it turns out to be a challenging problem.

The plurality of potential contaminants, the identification of the potential contamination sources in the network, and the reaction time for the utilities to deal with a contamination event are also important elements to consider. Contamination spreads in WDSs derived, mainly, from the water flows along the system. Thus, a good understanding of water flows leads to a good understanding of contamination scenarios.

Several complexities inherent to WDSs should be considered additionally. For example, as suggested by Di Nardo et al. (2013, 2014) and Chang et al. (2011), among others, using a simplified system (for example a sectorization or division of the system in so-called DMAs, for district metered areas (Ilaya-Ayza et al. (2018); Ferrari et al. (2014); Campbell et al. (2016, 2015)), including dynamic sectorization (Wright et al. (2014)), may lead to quick responses to such extreme events as infection intrusion, and this would reduce the number of required elements for monitoring and, as a result, the operational expenses.

This work, however, is not intended to cover all the aspects related to network protection against potential contaminant intrusion. It will rather concentrate on proposing a solution just for the sensor placement problem, namely, optimally determining the number of sensors and their locations. And we address this optimization problem from a multi-objective perspective.

Several goals should be considered when placing water quality sensors. Optimal sensor placement aims to achieve early contaminant detection and seclusion of affected areas so that the public exposure to contamination is minimum. First, it is desired to identify quality problems as soon as possible; this means minimizing the detection time. Second, irrespective of the location of the contaminant source, at least one sensor should always be able to identify a quality problem; this amounts to minimizing the detection failure. Additionally, the bulk of poor or bad quality water consumed should be minimized; this, specifically, involves that high population density areas have to receive special attention compared to other areas with a much lower consumption rate. And, importantly, the cost, which is directly proportional to the number of installed sensors, should be kept to a minimum or at least should not go beyond some upper budget limit.

These objectives should ideally be reached with a suitable sensor placement solution. However, it is a fact that the mutually conflicting nature of the objectives makes them pull in opposite directions. Improving one of them will probably result in a detriment for another. The rationale is clear. For example, maximizing the protection coverage in the network will require either to increase the number of sensors (it means the cost) or to probably be bound to accept larger detection times. Consequently, the final solution will result from a compromise among the objectives rather than from a unique "best alternative". Suitably solving problems of this nature requires the use of a multi-objective approach. The key idea boils down to not finding one single optimal solution but to find (an approximation of) the Pareto front representing the best trade-off that can be achieved among all the objectives involved (Hurt and Murray 2010; Quiñones-Grueiro et al. (2019); Giudicianni et al. (2020); Wéber and Hős (2020)). The known alternative consisting in first somehow pondering the objectives and then adding their weighted influence into a single expression for solving a single objective problem will probably result in finding just one point of the Pareto front. This approach is equivalent to producing an a priori trade-off without having any clear idea of how the obtained solution relates to the rest of the potential solutions of the problem. Another approach, as the one used in (Brentan et al. (2021)), produces an entire Pareto front of non-dominated solutions which is subsequently clustered and organized using suitable multi-criteria decision-making methods. This can be of great help for the utility managers, who can easily be lost in a densely populated Pareto front, frequently formed by many potential solutions. However, those approaches are unable to answer marginal, what-if cost questions, such as if it is worth buying an additional sensor to get a reasonable improvement in another objective, because, for example, there is no way to know how much improvement in protection coverage and detection time would bring that additional sensor. Those are the kinds of questions a multi-objective, rule, and preferences-based approach helps answer. This implies the existence of some utility function (Fishburn (1970)) behind the decision-making process, which is derived from some compromise scheme allowing a constructive solution of the multi-objective problem (Voronin and Savchenko (2020)). We claim that those are the kind of questions and answers needed to eventually find a sensor placement solution that represents a good trade-off among all the objectives involved. The main novelties of the proposal are the following. First, the proposed algorithm adds agents based on both technical and user-preference rules on top of evolutionary search techniques to explore the decision space. The algorithm runs as a part of the Agent Swarm Optimization framework, a consolidated multi-objective software. Second, the evaluation of the objective functions is executed directly in the MS SQL server and simulation data is never required to be loaded in their entirety. Without this important implementation detail, the solution for "large" water network models would not be affordable with the hardware typically used in desktop computers.

The rest of the paper is organized as follows. The following two sections address the materials and methods issue. Section 2 develops on contamination scenarios and the evaluation of the considered objectives. Then, section 3 presents the multi-objective solution proposed based on the contamination matrix and the contaminated consumption matrix concepts and gives details about the used optimization algorithm, which includes specific rules for the agents that help reduce the search space. A use case corresponding to a medium-size water distribution network is presented in section 4 together with the obtained results and a thorough discussion. Finally, section 5 provides some interesting conclusions and the paper is closed by the references section.

2. CONTAMINATION SCENARIOS AND EVALUATION OF OBJECTIVES

WDSs are vulnerable against various sources of accidental and intentional contaminations. In its protocol, the US EPA (EPANET (2020)) considers three steps: (i) detection of contaminant presence, (ii) source identification and (iii) consequent management. To be able to develop Early Warning Systems (EWSs) that ensure early detection of contaminations it is necessary to monitor the water quality through as many as possible of the quality significant parameters such as pH, turbidity, etc. With no doubt, to be able to develop suitable EWSs for alerting the consumers and isolating contaminated areas, optimal location of measurement devices is paramount in order to accurately identify the source of contamination for the managers to develop suitable management actions. After reviewing several works on sensor location, Hart and Murray (2010) describe EWSs and conclude that sensor placement is one of the critical aspects of the design of EWSs.

2.1 The objectives

As stated in the Introduction, the objectives we consider to solve the sensor location problem are: detection time, detection failure, affected population and implementation costs. For an arbitrary sensor network layout, Λ, that is, a candidate solution to the problem, these objectives are described next.

Detection time. For a layout of sensors, Λ, to detect a contaminant in a distribution system having N nodes susceptible of contamination the (average) detection time is calculated through a set of hydraulic simulations, as follows: contamination is introduced in the first node of the network, then an extended period simulation (EPS) is started and the detection time for this contamination event will be the time elapsed between the beginning of the contamination and the instant when at least one sensor of is reached by the contaminant. Then a new contamination is introduced in the second node of the network, a new simulation is started, and the detection time will be calculated with the same idea. The same process is repeated for the rest of the network nodes in turn. After completing this iterative process, the average of all detection times is calculated. The calculation of this average detection time can be summarized in the following process:

(1) Initialize the sum of detection time as zero.
(2) Loop from 1 to the number of nodes considered as potential contamination points.
  a. Contaminate node j.
  b. Consider the time until one of the sensors of Λ is reached by the contaminant.
  c. Add the current simulation time to the sum of detection times.
  d. Eliminate the contamination at node j.
(3) Calculate the average detection time by dividing the sum of detection times obtained for the nodes considered as potential contamination points whose contamination event has been eventually detected.

This code may also be formulated in closed form as:

$$f_1(\Lambda) = \frac{\sum_{j=1}^{N}(\tau(j,t) > 0)\cdot(\tau(j,t))}{\sum_{j=1}^{N}(\tau(j,t) > 0)}; \quad (1)$$

where $\tau(j,t)$ is the time elapsed between the contamination time at node j and the time at which one of the sensors in Λ first detects the contaminant, or 0 if the contamination event is not detected. The logical condition $\tau(j,t) > 0$ equals 1 if the condition is true and 0 otherwise. Observe that non-detected contamination events do not contribute to this calculation.

A natural question to ask is what happens with those cases when none of the sensors in can detect the contaminant; for example, when sensors are located “very upstream” and a contaminant is injected “very downstream”. In these cases, this approach does not calculate any detection time. In (Berry et al. (2008)), for this case, it is assumed a detection time equal to the simulation time (the simulation time is introduced to define for how long the simulation should be run). Another possibility is to establish in these cases that the detection time is equal to zero, which is the approach followed in this research. Assuming detection time equals zero when there is no detection is not a problem if the calculation is based on a multi-objective approach where other objectives such as the detection failure and the contaminated water consumed are also considered. We consider these two objectives next.

Detection failure. The detection failure is calculated as the sum of the cases where none of the sensors in Λ is reached by the contaminant in the contamination events analyzed for estimating the detection time. In a closed form, this objective may be written as

$$f_2(\Lambda) = \sum_{j=1}^{N}(\tau(j,t) = 0); \quad (2)$$

which acts as a counter for non-detected contamination events.

Affected population. It is important to consider not only the detection time and the detection failure but also how many people would be affected before the contamination reaches at least one of the sensors. The affected population can be estimated either based on the number of people or on the consumption, that is, the water demand associated with each of the consumption nodes affected, since this is an information directly handled within the EPSs deployed to hydraulically analyze the problem.

In this research the calculations have been based on the water demand at the affected nodes. The following process illustrates how the calculation of the affected population is implemented in the objective function. Note that the number of people affected is not directly calculated. Instead, it is calculated the averaged contaminated water that has been supplied to the population.

(1) Initialize the sum of contaminated water consumption as equal to zero.
(2) Loop from 1 to the number of nodes considered as potential contamination points.
  a. Contaminate node j.
  b. Consider the time until one of the sensors of Λ is reached by the contaminant. Observe that the

calculation of this expression uses the detection time for sensor layout Λ has been previously calculated, as previously explained.

c. For all the nodes reached by the contaminant, add their real demand to the sum of contaminated water consumption.

d. Eliminate the contamination at node j.

(3) Return the sum of contaminated water consumption as a result.

A closed expression for this objective is the following:

$$f_3(\Lambda) = \sum_{j=1}^{N} \sum_{i=1}^{N} (d(i,j) > 0) \cdot (d(i,j)); \tag{3}$$

where $d(i,j)$ is the demand associated to node i during the contaminant even corresponding to node j, and $(d(i,j) > 0)$ is a logical condition which equals 1 if that demand is positive and 0 otherwise.

Of course, the estimation of the affected population could be formulated in a much more complex way. Insisting in calculating the number of persons affected will open the question of how much water should have consumed the consumers in order to be affected by the concentration of contaminants appearing in the water. Additionally, it should be considered that the actions taken directly by the utility after a contamination is detected do not happen instantaneously. It takes some time to inform the population, operate valves when necessary or do any additional action as a response to the contamination event. It means that the consequences of the contamination will be probably bigger than the calculation here. However, the way we are running the calculations makes this calculation to be proportional to what will happen in case of a real contamination event. This is why $f_3(\Lambda)$ can be used as an indicator of consumers affected to optimize the location of sensors, despite a different analysis is required to have an accurate calculation of how many people will be exactly affected.

Implementation costs. Costs are proportional to the number of sensors to be installed. Even if it is unknown how many sensors will be eventually installed, it can be obviously assumed that more sensors will imply more costs, and vice versa. The real total costs will be beyond the cost per sensor and will have to include also the installation and maintenance costs at least. A straightforward expression for this objective is:

$$f_4(\Lambda) = \alpha N_S; \tag{4}$$

A detailed calculation of costs will not be necessary for the optimization problem we are running because much of the cost's details have nothing to do directly with the decision variables in the problem. It can only be decided the number of sensors to be installed and it will be only influencing the costs of sensors proportionally to the number selected. In this research, as expressed by the previous formula, the costs were estimated as the number of sensors, N_S, in layout Λ, multiplied by an estimated average cost, α, per sensor.

3. ALGORITHM AND SOFTWARE FOR CALCULATIONS

The calculation of these objective functions is performed by using two matrices, namely the contamination matrix and the contamination consumption matrix, which are built after performing all the necessary contamination event simulations and suitably stored depending on the hydraulic network size. Then, the calculations of the first three objectives are performed through suitable look-ups to these matrices.

In this section we shortly describe these matrices and then the specificities about the used software are detailed.

The contamination matrix. This matrix stores, for every single contamination event, that is, the contamination of one susceptible node of the network, how long it takes in the corresponding EPS to reach each of the network nodes.

This matrix is used to calculate objective functions $f_1(\Lambda)$ and $f_2(\Lambda)$ for an arbitrary layout of sensors, Λ. It is also used, as a previous step to calculate $f_3(\Lambda)$, as seen next.

The contaminated consumption matrix. In a similar way as with the contamination matrix, the contaminated consumption matrix stores the amount of contaminated water consumed in each scenario by each consumption node of the network.

This matrix is used to calculate objective function $f_3(\Lambda)$ for the current layout of sensors Λ. Note that an element (i,j) of this matrix will represent how much contaminated water has been consumed in node j from the moment a contaminant is injected at node i till the moment a sensor in Λ detects the contaminant, and this last calculation uses the contamination matrix.

The used software. A number of approaches may be used to find the Pareto front in a multi-objective optimization problem (Ferreira et al. (2017); Kukkonen and Coello (2017); Jiang and Yang (2016)). In this research, the algorithm behind the solution search process is based on Agent Swarm Optimization (ASO) (Montalvo et al. (2014)). ASO is a distinctive combination of multi-objective evolutionary algorithms, rule-based agents and data analytics. ASO intelligently integrates problem-domain knowledge within the optimization process and learns engineer's preferences to achieve more real results. The algorithm has been integrated in Water-Ing (2020), a software package for analysis and decision support in water distribution systems. Water-Ing connects with the EPANET (2017) toolkit to perform the necessary hydraulic simulations that generate the contamination matrix and the contaminated consumption matrix. Using this information, ASO analyzes various alternatives for locating sensors and selects those alternatives representing non-dominated solutions.

Despite various evolutionary techniques can be used in ASO, for this research it was used a multi-objective version of Particle Swarm Optimization (PSO). A population of agents with a similar behavior to particles in PSO was introduced for solving the problem. Additionally, another population of rule-based agents were also included in the solution search process. Rule-based agents use rules designed to reduce the problem decision search space. In previous work, we have used various rules for reducing the problem decision space in the problem of water distribution system design (Montalvo et al. 2014; Izquierdo et al. (2016a,b)). In this research we introduce basic rules for reducing the search space for the water quality sensor placement problem. A "normal" agent based on the behavior of PSO can locate a sensor at virtually any node of the network. Based on the experience of the authors on solving several use cases it was found that:

- Locating sensors too much downstream of the network will probably guarantee good coverage of the network but unfortunately will result in unreasonable long detection times. Basically, the probability of detecting an event is relatively high. Nevertheless, detection will be definitively too late because most of the network will have already been contaminated when the detection happens, and there will be not much more to do in order to prevent people from using contaminated water.
- Locating sensors too much upstream of the network will detect events faster, but the coverage of the network will be seriously compromised. Many contamination events may happen downstream of the location of the sensors and no detection will happen in these cases.

Let us clarify several things regarding these two ideas.

First, even though the concepts of downstream and upstream are not clearly defined for a network, the averaged contamination times previously calculated and stored in the contamination matrix are a perfect guide to move upstream and downstream through the network and can be used as surrogates of the classical upstream and downstream concepts. Moreover, observe that, as the optimization process is performed after having run a suitable number of simulations leading to obtaining the contamination matrix, there is no need of extra hydraulic simulations and, as a result, those surrogate concepts of upstream and downstream are adequate.

Second, the above-mentioned two ideas suggest that boundaries should be drawn in order to define areas of higher interest where sensors should be placed. For obvious reasons, the nodes in this area of interest should be neither too close to the water sources nor at the very end of the piping network.

Another consideration should be taken into account: after a detection happens there is a reaction time from the water company to take suitable actions.

A first boundary should be designed to exclude too upstream nodes from being eligible. Points belonging to this boundary will be those located at a certain previously established distance from the water sources. Technicians must define that distance before running the algorithm, and instructed rule-based agents will be in charge of not selecting any node located closer to a water source than that distance.

A second boundary of nodes should consider the company reaction time to decide the operation actions to be executed, and the time to run those operation actions in the field. The reaction time is used by the rule-based agents to exclude candidate locations for sensor placement. Any node downstream of that boundary will not be eligible for hosting a water quality sensor.

The idea of establishing upstream and downstream frontiers for locating water quality sensors helps reduce the search space of the problem. Nevertheless, this has no influence on the size of the contamination matrix associated with the problem. Remember that this contamination matrix is the result of how long it takes to contaminate each node after contaminating each contamination-susceptible point in the water network. For "large" water network models it is not possible to keep the whole contamination matrix in RAM at a time. That is the reason why the software implementation associated with this research uses an MS SQL database for saving the contamination and the contamination consumption matrices. In other words, the evaluation of the objective functions is executed directly in the MS SQL server and these matrices are never required to be loaded in their entirety. Without this important implementation detail, the solution for "large" water network models would not be affordable with the hardware typically used in desktop computers.

We finally observe that Water-Ing is endowed with an advanced visualization environment to help insightful analyses of the obtained optimization solutions. Specifically, as many 2D Pareto charts as desired can be created in this environment showing the various relationships between pairs of objectives, namely cost, detection time, protection coverage and bad water quality impact, which can be simultaneously represented in a very friendly way. Selected solutions in one chart will be automatically selected in the rest of charts indicating how they behave with respect to all the aspects involved in the problem. Moreover, if more than one screen is used in the desktop, any chart can be detached from the application and moved to a different screen to expand the visualization capabilities.

4. USE CASE

In this research, for illustrating the solution process it has been used a modified version of the water network of San José de las Lajas. It is a small town in Cuba close to Havana with more than 24 km of pipes and one single entry point, as a consequence of the modifications in the original model.

Figure 1 represents the network with a solution integrated by 4 water quality sensors placed in the positions shown. For better interpretation of the results, the results for this 4-sensor layout solution is marked in red in the three following figures, corresponding to the execution of the sensor placement problem for the case of four sensors.

Figure 2 shows a 2D projection of the Pareto front representing average detection time versus contaminated water consumption. One can observe a number of solutions with zero detection time (note that the detection time is assumed equal to zero for non-detected events). Obviously, for most of these solutions the consumption of contaminated water is large. Observe the attractiveness of the 'red' solution.

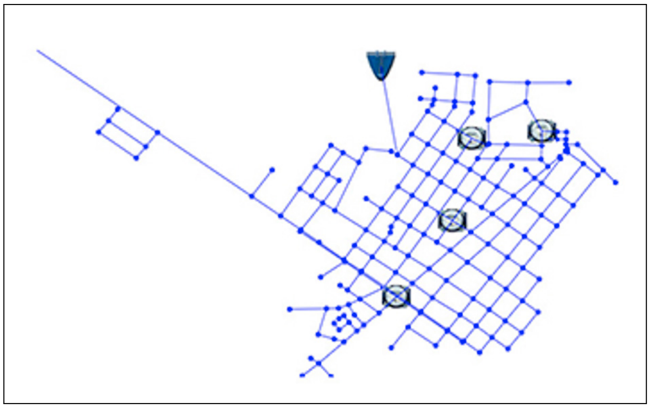

Fig. 1. Network model of San José with 4 water quality sensors

Figure 3 relates the amount of detection failures with the average detection time. Solutions with very high detection time represent layouts with sensors located at nodes very downstream in the network. In these cases, it takes longer to detect a contaminant (as an average, considering all possible contamination events) but the detection failure is much lower. In contrast, many solutions with too small detection times, exhibit large detection failure, and correspond to layouts with sensors too upstream or, as commented in the previous paragraph, to cases with detection time equal to zero, that is, undetected events. The 'red' solution presents a trade-off between both objectives.

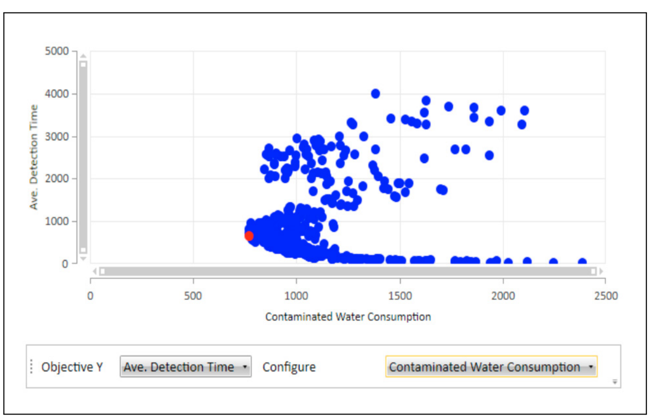

Fig. 2. Average detection time vs contaminated water consumption

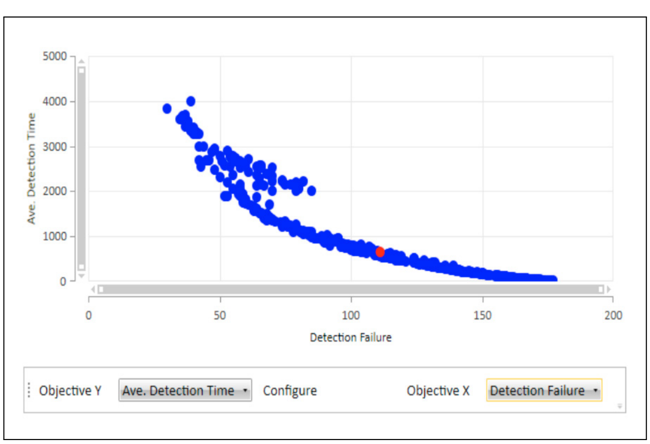

Fig. 3. Average detection time vs detection failure

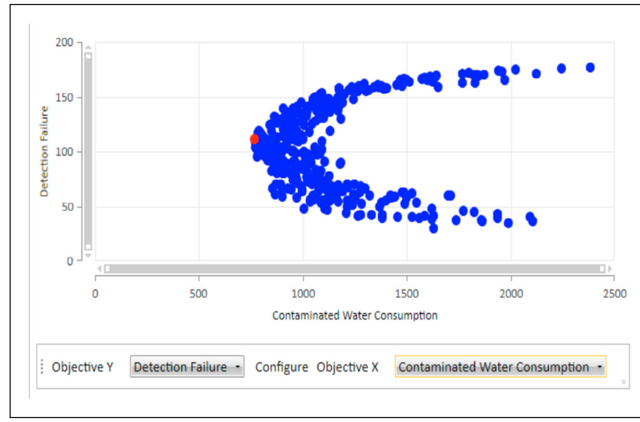

Fig. 4. Average detection failure vs average contaminated water consumed

Figure 4 shows that the average volume of contaminated water consumed can be increased because of two main reasons: either corresponds to solutions with high average detection failure (sensors located too close to the sources that cannot detect contamination downstream) or to solutions where sensors are located at nodes in very downstream positions, and it requires longer in average to receive the contamination effects. The relation between detection time and detection failure has been previously mentioned and can be seen in Figure 3. The 'red' solution once more exhibits a clear optimal position.

Figures 2-4 help locate a certain number of sensors in the network, four for the case of the mentioned figures. Nevertheless, additional actions are required to assess how many sensors should be installed in a network. Adding sensor stations implies an investment and this has a clear limitation. It will depend on the budget of the company and the rewards for installing them.

One approach would be to decide an objective budget and based on it, calculate the maximum number of sensors that can be installed and, of course, then run the approach to decide one optimal layout. Despite it is a realistic approach, it is worth asking, however, what about if, for example, for a 10% more of investment, the network coverage to detect contaminant intrusion is increased in 25%? Would the utility add this 10% to the budget? Also, what about if the network coverage was increased in a 30%, or if the detection time was reduced by a certain percentage? Following this line of thinking, the next question would be how much improvement/reward one would get by increasing the budget to acquire one more sensor? The same idea can bring the utility managers to ask how much improvement/reward would be lost if getting rid of one of the sensors to lower the budget? Answering these kinds of questions requires a representation of the improvement/rewards received as a function of the number of sensors installed.

Deciding a budget *a priori* without considering the improvement/reward received as a function of the number of sensors installed will not lead, in general, to a good decision. Despite an initial tentative budget can be decided, both aspects should be combined together before making a final decision. The improvements/rewards received will be expressed in terms of the objectives considered for placing sensors: reducing the detection time of contaminant intrusion or water quality problems; protecting the population from the consumption of water under minimum quality requirements; and maximizing the protection coverage in the network.

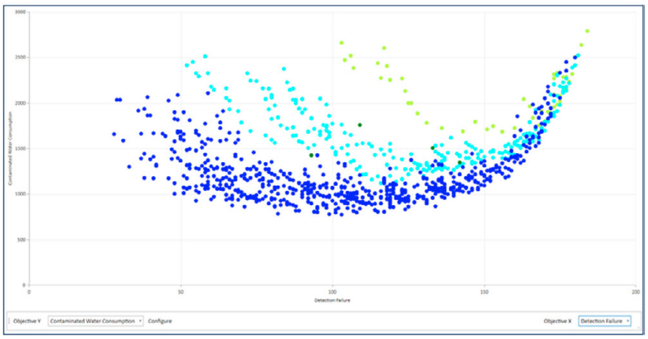

Fig. 5. Contaminated water consumption (Y axis) vs Detection Failure (X axis). Solutions for 3 (light green), 4 (light blue) and 5(dark blue) sensors

These three objectives have to be combined with the additional objective of maintaining the cost within the framework of a budget that can be afforded by the company. For visualization purposes, it would be better to add new 2D projections of the Pareto front where solutions for different numbers of sensors can be compared, as in figure 5.

With all the projections of the Pareto front it can be better decided not only where sensors should be located but also how many of them should be placed. In the case a budget can not be extended under any circumstances, the analysis will still be valid for estimating how much is needed for a realistic protection of the water network. It is true that some experience will definitively help in making these kinds of decisions. Nevertheless, even for experienced engineers, the Pareto charts shown in figures 2-5 constitute a great support for better evaluating alternatives.

5. CONCLUSIONS

To properly protect a water network against accidental or intentional contamination events and water quality problems, two important questions have to be answered: how many sensors are needed and where to place them. Answering these questions requires a decision about the criteria and the requirements to be considered for achieving a good solution, which inexorably has to obtained within a multi-objective approach for solving the problem. In this paper we have considered detection time, detection failure, consumed contaminated water and cost.

The final solution should be based on a trade-off among the objectives involved, with special emphasis on the tolerance to "fail" that the utility could afford in its water supply system. An improvement in all the objectives analyzed can be done by adding new sensors but this, of course, has the consequence of increasing the costs which can be a (hard) constraint for the implementation of the solution.

Additional information about the utility can also influence the final sensor layout adopted. This is the case, for example, of the average reaction time of the company when an event is detected. If the reaction time is relatively large then it may be convenient to use solutions with the sensors located a little more upstream if it is desired to avoid at least part of the population to receive contaminated water. Note that in this case there could be a higher number of detection failures too. Getting some improvement on both sides (reduced detection time and reduced detection failure) implies adding more sensors to the solution.

This paper shows a simplified overview about how to deal in practice with water quality sensor placement for protecting such important infrastructures as water distribution systems. It is fully recommended not to base decisions on just practical experience but to run hydraulic model calculations. We strongly believe that both computer models and experience should be jointly used for achieving better results.

ACKNOWLEDGEMENTS

The research was financially supported by the Czech Science Foundation under Grant No. 19-06569S.

REFERENCES

Antunes, C.H. and Dolores, M. (2016). Sensor location in water distribution networks to detect contamination events—a multiobjective approach based on nsga-ii. In *2016 IEEE Congress on Evolutionary Computation (CEC)*, 1093–1099. IEEE.

Berry, J., Boman, E., Riesen, L.A., Hart, W.E., Phillips, C.A., and Watson, J.P. (2008). User's manual: Teva-spot toolkit 2.0. *Sandia National Laboratories, Albuquerque, NM.*

Brentan, B., Carpitella, S., Barros, D., Meirelles, G., Certa, A., and Izquierdo, J. (2021). Water quality sensor placement: a multi-objective and multi-criteria approach. *Water Resources Management*, 35(1), 225–241.

Campbell, E., Izquierdo, J., Montalvo, I., Ilaya-Ayza, A., Pérez-García, R., and Tavera, M. (2015). A flexible methodology to sectorize water supply networks based on social network theory concepts and multi-objective optimization. *Journal of Hydroinformatics*, 18(1), 62–76.

Campbell, E., Izquierdo, J., Montalvo, I., and Pérez-García, R. (2016). A novel water supply network sectorization methodology based on a complete economic analysis, including uncertainties. *Water*, 8(5), 179.

Candelieri, A., Conti, D., and Archetti, F. (2014). A graph based analysis of leak localization in urban water networks. *Procedia Engineering*, 70, 228–237.

Chang, K., Gao, J.L., Wu, W.Y., and Yuan, Y.X. (2011). Water quality comprehensive evaluation method for large water distribution network based on clustering analysis. *Journal of Hydroinformatics*, 13(3), 390–400.

Covas, D. and Ramos, H. (1999). Practical methods for leakage control, detection, and location in pressurized systems. In *BHR group conference series publication*, volume 37, 135–152. Bury St. Edmunds; Professional Engineering Publishing; 1998.

Di Nardo, A., Di Natale, M., Musmarra, D., Santonastaso, G., Tzatchkov, V., and Alcocer-Yamanaka, V. (2014). A district sectorization for water network protection from intentional contamination. *Procedia Engineering*, 70, 515–524.

Di Nardo, A., Di Natale, M., Guida, M., and Musmarra, D. (2013). Water network protection from intentional contamination by sectorization. *Water Resources Management*, 27(6), 1837–1850.

Diao, K., Fu, G., Farmani, R., Guidolin, M., and Butler, D. (2016). Twin-hierarchy decomposition for optimal design of water distribution systems. *Journal of Water Resources Planning and Management*, 142(5), C4015008.

EPANET, U.E. (2020). Software that models the hydraulic and water quality behavior of water distribution piping systems.

Ferrari, G., Savic, D., and Becciu, G. (2014). Graph-theoretic approach and sound engineering principles for design of district metered areas. *Journal of Water Resources Planning and Management*, 140(12), 04014036.

Ferreira, J.C., Fonseca, C.M., Denysiuk, R., and Gaspar-Cunha, A. (2017). Methodology to select solutions for multiobjective optimization problems: Weighted stress function method. *Journal of Multi-Criteria Decision Analysis*, 24(3-4), 103–120.

Fishburn, P.C. (1970). Utility theory for decision making. Technical report, Research analysis corp McLean VA.

Gandy, M. (2004). Rethinking urban metabolism: water, space and the modern city. *City*, 8(3), 363–379.

Giudicianni, C., Herrera, M., Di Nardo, A., Greco, R., Creaco, E., and Scala, A. (2020). Topological placement of quality sensors in water-distribution networks without the recourse to hydraulic modeling. *Journal of Water Resources Planning and Management*, 146(6), 04020030.

Hart, W.E. and Murray, R. (2010). Review of sensor placement strategies for contamination warning systems in drinking water distribution systems. *Journal of Water Resources Planning and Management*, 136(6), 611–619.

Herrera, M., Abraham, E., and Stoianov, I. (2016). A graph-theoretic framework for assessing the resilience of sectorised water distribution networks. *Water Resources Management*, 30(5), 1685–1699.

Hoekstra, A.Y. and Mekonnen, M.M. (2012). The water footprint of humanity. *Proceedings of the national academy of sciences*, 109(9), 3232–3237.

Ilaya-Ayza, A.E., Martins, C., Campbell, E., and Izquierdo, J. (2018). Gradual transition from intermittent to continuous water supply based on multi-criteria optimization for network sector selection. *Journal of Computational and Applied Mathematics*, 330, 1016–1029.

Islam, N., Farahat, A., Al-Zahrani, M.A.M., Rodriguez, M.J., and Sadiq, R. (2015). Contaminant intrusion in water distribution networks: review and proposal of an integrated model for decision making. *Environmental Reviews*, 23(3), 337–352.

Izquierdo, J., Campbell, E., Montalvo, I., and Pérez-García, R. (2016a). Injecting problem-dependent knowledge to improve evolutionary optimization search ability. *Journal of Computational and Applied Mathematics*, 291, 281–292.

Izquierdo, J., Montalvo, I., Campbell, E., and Pérez-García, R. (2016b). A hybrid, auto-adaptive and rule-based multi-agent approach using evolutionary algorithms for improved searching. *Engineering Optimization*, 48(8), 1365–1377.

Izquierdo, J., Montalvo, I., Pérez-García, R., and Matías, A. (2012). On the complexities of the design of water distribution networks. *Mathematical problems in engineering*, 2012.

Jiang, S. and Yang, S. (2016). Evolutionary dynamic multiobjective optimization: Benchmarks and algorithm comparisons. *IEEE transactions on cybernetics*, 47(1), 198–211.

Kukkonen, S. and Coello, C.A.C. (2017). Generalized differential evolution for numerical and evolutionary optimization. In *NEO 2015*, 253–279. Springer.

Montalvo, I., Izquierdo, J., Pérez-García, R., and Herrera, M. (2014). Water distribution system computer-aided design by agent swarm optimization. *Computer-Aided Civil and Infrastructure Engineering*, 29(6), 433–448.

Nafi, A., Crastes, E., Sadiq, R., Gilbert, D., and Piller, O. (2018). Intentional contamination of water distribution networks: developing indicators for sensitivity and vulnerability assessments. *Stochastic environmental research and risk assessment*, 32(2), 527–544.

Oliker, N. and Ostfeld, A. (2015). Network hydraulics inclusion in water quality event detection using multiple sensor stations data. *Water research*, 80, 47–58.

Ostfeld, A. (2012). Optimal reliable design and operation of water distribution systems through decomposition. *Water Resources Research*, 48(10).

Ostfeld, A., Oliker, N., and Salomons, E. (2014). Multiobjective optimization for least cost design and resiliency of water distribution systems. *Journal of Water Resources Planning and Management*, 140(12), 04014037.

Perelman, L. and Ostfeld, A. (2012). Water-distribution systems simplifications through clustering. *Journal of Water Resources Planning and Management*, 138(3), 218–229.

Quiñones-Grueiro, M., Verde, C., and Llanes-Santiago, O. (2019). Multi-objective sensor placement for leakage detection and localization in water distribution networks. In *2019 4th Conference on Control and Fault Tolerant Systems (SysTol)*, 129–134. IEEE.

Rathi, S. and Gupta, R. (2017). Optimal sensor locations for contamination detection in pressure-deficient water distribution networks using genetic algorithm. *Urban Water Journal*, 14(2), 160–172.

Sarrate, R., Blesa, J., and Nejjari, F. (2014). Clustering techniques applied to sensor placement for leak detection and location in water distribution networks. In *22nd Mediterranean Conference on Control and Automation*, 109–114. IEEE.

Voronin, A. and Savchenko, A. (2020). A systematic approach to multiobjective optimization. *Cybernetics and Systems Analysis*, 56(6), 1000–1011.

Wéber, R. and Hős, C. (2020). Efficient technique for pipe roughness calibration and sensor placement for water distribution systems. *Journal of Water Resources Planning and Management*, 146(1), 04019070.

Wright, R., Stoianov, I., Parpas, P., Henderson, K., and King, J. (2014). Adaptive water distribution networks with dynamically reconfigurable topology. *Journal of Hydroinformatics*, 16(6), 1280–1301.

Zhao, Y., Schwartz, R., Salomons, E., Ostfeld, A., and Poor, H.V. (2016). New formulation and optimization methods for water sensor placement. *Environmental Modelling & Software*, 76, 128–136.