

## Water quality sensor placement: a multi-objective and multi-criteria approach

**Bruno Brentan · Silvia Carpitella · Daniel Barros · Gustavo Meirelles · Antonella Certa · Joaquín Izquierdo**

**Abstract** To satisfy their main goal, namely providing quality water to consumers, water distribution networks (WDNs) need to be suitably monitored. Only well designed and reliable monitoring data enables WDN managers to make sound decisions on their systems. In this belief, water utilities worldwide have invested in monitoring and data acquisition systems. However, good monitoring needs optimal sensor placement and presents a multi-objective problem where cost and quality are conflicting objectives (among others). In this paper, we address the solution to this multi-objective problem by integrating quality simulations using EPANET-MSX, with two optimization techniques. First, multi-objective optimization is used to build a Pareto front of non-dominated solutions relating contamination detection time and detection probability with cost. To assist decision makers with the selection of an optimal solution that provides the best trade-off for their utility, a multi-criteria decision-making technique is then used with a twofold objective: 1) to cluster Pareto solutions according to network sensitivity and entropy as evaluation parameters; and 2) to rank the solutions within each cluster to provide deeper insight into the problem when considering the utility perspectives. The clustering process, which considers features related to water utility needs and available information, helps decision makers select reliable and useful solutions from the Pareto front. Thus, while several works on sensor placement stop at multi-objective optimization, this work goes a step further and provides a reduced and simplified Pareto front where optimal solutions are highlighted. The proposed methodology uses the NSGA-II algorithm to solve the optimization problem, and clustering is performed through ELECTRE TRI. The developed methodology is applied to a very well-known benchmarking WDN, for which the usefulness of the approach is shown. The final results, which correspond to four optimal solution clusters, are useful for decision makers during the planning and development of projects on networks of quality sensors. The obtained clusters exhibit distinctive features, opening ways for a final project to prioritize the most convenient solution, with the assurance of implementing a Pareto-optimal solution.

---

Bruno Brentan  
Dept. de Engenharia Hidráulica e Recursos Hídricos  
Federal University of Minas Gerais  
CEP 31270-901 Belo Horizonte, Brazil  
E-mail: brentan@ehr.ufmg.br

Silvia Carpitella  
Department of Decision-Making Theory, Institute of Information Theory and Automation  
Czech Academy of Sciences  
18208 Prague, Czech Republic

Daniel Barros  
Dept. de Engenharia Hidráulica e Recursos Hídricos  
Federal University of Minas Gerais  
CEP 31270-901 Belo Horizonte, Brazil

Gustavo Meirelles  
Dept. de Engenharia Hidráulica e Recursos Hídricos  
Federal University of Minas Gerais  
CEP 31270-901 Belo Horizonte, Brazil

Antonella Certa  
Dipartimento di Ingegneria  
Università degli Studi di Palermo  
90128 Palermo, Italy

Joaquín Izquierdo  
Fluing - Institute for Multidisciplinary Mathematics  
Universitat Politècnica de València  
46022 Valencia, Spain.

**Keywords** Water distribution systems · Water quality sensor placement · Optimization · ELECTRE TRI

## 1 Introduction

Water distribution networks (WDNs) are core infrastructures for transporting water of acceptable quality from sources to final customers, and are designed to assure sustainability and the development of modern cities (Gandy 2004). These systems are very complex and dynamic due to their wide spatial dispersion, uncertainties in consumption and leakage, vulnerabilities, and so on. As a consequence, many planning and operational tasks, such as maintenance for failure minimization (Herrera et al. 2016) and leak detection (Candelieri et al. 2014), as well as detection and identification of contamination sources (Nafi et al. 2018), are objects of intensive research. Additionally, as underlined by de Winter et al. (2019), WDNs are susceptible to malicious contamination events, potentially leading to poisoned water, deaths, and huge economic losses.

Water utilities devote considerable efforts to keeping the most significant water quality parameters under continuous monitoring, so that the safety and security of WDNs can be guaranteed. It goes without saying that this aspect becomes increasingly challenging when covering widely distributed WDNs. In this context, making effective decisions about how to design and implement optimal networks of quality sensors placed at strategical nodes (Oliker and Ostfeld 2015) plays a crucial role in protecting human communities against intrusion or attacks using water contamination. However, optimal sensor placement (OSP) is a multi-objective optimization problem.

Many practical optimization problems related to WDNs were solved in the past by traditional optimization tools. Recently, hydraulic optimization, ranging from the optimal design of water systems (Montalvo et al. 2014) to the process of leakage detection (Kapelán et al. 2003), has been widely solved using heuristic algorithms. Moreover, in water distribution engineering, as in many other fields, optimization problems are typically characterized by the presence of several conflicting objectives that need to be jointly analyzed. Regarding the problem of sensor placement, a clear trade-off exists, for example, between costs and coverage rate. Cost, directly linked to the number of sensors to be installed, must obviously be minimized (Berry et al. 2005). System coverage, in turn, must be maximized.

Network monitoring is performed using various methods applied to sets of quality parameters, such as free chlorine, pH, etc. Moreover, to maximize network coverage, and guarantee maximal protection for the system, the dynamic nature of that system (e.g. consumption oscillation during the day, and maneuvers of control devices) must be considered since, as a consequence, these dynamics modify the topology of the water system, and the mass transport capacity.

In this paper, to deal with the quality OSP problem, contamination by pathogen intrusion is considered, as presented in (Rathi et al. 2016), where the authors compute a probabilistic distribution function based on the hydraulic behavior of the network, and considering that some nodes are more subject to intrusions. Other works also use the variation of a specific quality parameter with respect to the variation of flow in a certain pipe, or related to a specific nodal demand, to place quality sensors on strategic nodes. This concept is known as sensitivity. Maximization of the sensitivity of a network of sensors has been applied in water quality problems in WDNs (Cheung et al. 2005).

Maximizing network sensitivity has, however, an intrinsic drawback. Usually, in WDN models, a highly-sensitive node has, as neighbors, other highly-sensitive nodes. With this evidence, a sensitivity-solely-based algorithm will concentrate sensors in a reduced region, and this will clearly impair network coverage. To lower this agglomeration effect and distribute sensors more homogeneously throughout the water network, redundancy information analyses that make use of entropy information are herein considered. Traditionally, entropy provides a measure to quantify the degree of disorder in a system, but it is also used in the field of information theory to quantify the efficacy of the information provided by a given data set to a system (Shannon 1948). According to Lee (2013), information can be quantitatively measured on the indirect basis of the entropy characterizing the data set of interest. The larger the entropy, the lower the information redundancy and thus the higher the degree of uncertainty. The WDN security level can be improved by using entropy information from sensors strategically placed, thus helping identify intrusions (Weickgenannt et al. 2010) or leaks (Christodoulou et al. 2013) in the water system.

Expanding the concepts of sensitivity, Ostfeld and Salomons (2005) consider, for the OSP problem, the response delay for contamination detection. Risk-based approaches have also been proposed in the literature, considering the uncertainties associated with water systems (Broad et al. 2008). The authors propose a single objective approach together with Monte Carlo simulations for calculating the uncertainties associated with demand, pipe roughness, and chlorine decay.

The Battle of Water Sensors Network (BWSN) (Ostfeld et al. 2008) proposes a challenge for sensor placement. Four main objective functions are considered: minimizing the detection time; minimizing the portion of affected population before detection; minimizing the consumption of contaminated water before detection; and maximizing the detection probability.

Following some objectives from the BWSN, Ohar et al. (2015) developed a single objective optimization, minimizing the consumption of contaminated water, and linking the optimization process to the Epanet-MSX. This is an important work since it uses the link between multi-species quality models and optimization, which is a complex problem from the computational point of view.

The multi-objective approach presented in the BWSN, which uses single species modelling, has been widely explored in the literature. Various multi-objective problems have been designed to handle the global goal of a network of sensors. Huang et al. (2008) use the detection time, the detection probability, and the consumed contaminated water to set sensors in the network provided by (Ostfeld et al. 2008). The demand coverage is used as an external criterion to rank the solutions of a Pareto front. In (Hart and Murray 2010) a literature review presents many of the methodologies applied to optimal water quality sensor placement. The authors analyze more than 90 published articles and, obviously, as the multi-objective problem is solved by means of multi-objective optimization tools, results consist of sets of non-dominated solutions (Wéber and Hős 2020; Giudicianni et al. 2020; Quiñones-Grueiro et al. 2019) and not just a single solution. A Pareto front is indeed very convenient, since decision makers do not generally seek fast and efficient optimization algorithms, but effective strategies that minimize uncertainty and enable them to evaluate which solutions present better trade-offs.

In this context, multi-criteria analyses offer a wide range of tools capable of ranking and/or clustering the Pareto front of solutions to assist decision makers in implementing a solution. The existing literature attributes to multi-criteria decision-making (MCDM) methods a highly positive impact for resolving diverse decision-making problems in a structured way (Barak and Mokfi 2019; Štirbanović et al. 2019; Seiti et al. 2019; Wang et al. 2019; Mohammed et al. 2019). Many MCDM methods have been proposed and applied so far. Broadly speaking, they are aimed at:

- selecting the alternative (option) representing the best trade-off with respect to the various aspects considered;
- ranking various solutions to understand their priority or degree of importance;
- grouping alternatives, according to their similarities, into suitable clusters.

Regarding the last objective, ELECTRE TRI, a technique which belongs to the family of methods ELimination Et Choix Traduisant la REalité (ELECTRE) (Roy 1968; Figueira et al. 2013), has proven to be a powerful tool to assign alternatives to specific categories (Liu and Ming 2019) defined according to suitable reference profiles. Specifically, the ELECTRE TRI technique enables directly visualizing the assignment of clustered solutions on the basis of the mutual importance of certain established evaluation criteria. Given its main features, ELECTRE TRI has been studied in depth and applied in the literature (Dias et al. 2018; Reginaldo 2015; Neto et al. 2017; Sánchez-Lozano et al. 2016; Bouyssou and Marchant 2015).

In (Costa et al. 2018) it is underlined as MCDM analyses enable an easy evaluation of the degree of performance of alternatives according to various criteria of interest, based on preference judgments elicited by one or more experts. The authors propose an ELECTRE TRI-based approach to classify suppliers into predefined ordered clusters in the context of supply chain resilience for emerging economies. The work (Fernández et al. 2017) affirms that ELECTRE TRI is the most widely used method among the various existing outranking approaches, and underlines (Fernandez and Navarro 2011) the difficulty in assuming a single limiting reference profile to acceptably determine its related category. For this reason, the authors propose an extension of the classical version of ELECTRE TRI to resolve problems of ordinal classification with multiple criteria by considering richer relations of preference among alternatives and reference profiles. In (Corrente et al. 2016) an extended version of the ELECTRE TRI to consider criteria according to a defined hierarchical structure is presented. The authors highlight that a hierarchy of criteria may better reflect the complexity of practical decision problems. In (Brito et al. 2010), ELECTRE TRI is used in the risk management of natural gas pipeline networks to provide support in planning and implementing correct control actions, and maintenance to prevent and mitigate risks.

Given its utmost importance, the OSP problem in WDNs was discussed by the authors in a previous research (Francés-Chust et al. 2020) to identify the most suitable nodes for placing pressure sensors aimed at effectively detecting leakage. In the present research, we are interested, instead, in dealing with quality sensors (rather than pressure sensors), and approach the OSP problem for water quality monitoring by means of a new hybrid perspective that integrates multi-objective optimization and multi-criteria analysis.

The optimization problem is addressed by means of the non-dominated sorting genetic algorithm II (NSGA-II). For abnormal scenarios, where a contaminant is flowing in the water network, detection time, probability detection, and the number of sensors are used as objective functions. NSGA-II will build the Pareto front of non-dominated solutions. ELECTRE TRI then clusters those solutions into ordered classes, thus highlighting their distinctive features. Accordingly, the objective of the paper consists in: i) providing a structured tool capable of dealing with huge sets of (Pareto) solutions of a quality OSP; and ii) identifying those alternatives that simultaneously better satisfy the given evaluation criteria. Thus, Pareto optimal solutions will be classified into the following performance categories according to how they perform with respect to several aspects: low; medium; medium-high; and high. The hybrid approach proposed can be seen as an important contribution to the quality OSP literature, since

h

**Table 1.** Degradation of the parathion

| Reaction Stoichiometry   | Rate/Equilibrium Coefficient at 25°C  |
|--|---|
| $5.HOCL + PA \xrightarrow{K_{HOCL,OP}} PAO + 5.H^+ + 5.CL^- + SO_4^{2-}$ | $(K_{HOCL,PA} = 2, 2(\pm 0, 53) \cdot 10^6 M^{-1} h^{-1})$  |
| $PA \xrightarrow{K_{h,PA}} PAH$  | $K_{h,PA} = K_{N,PA} + K_{B,PA} \cdot [OH^-]$<br>$K_{N,PA} = 2, 66 \cdot 10^{-4} h^{-1}$<br>$K_{B,PA} = 4, 3 \cdot M^{-1} h^{-1}$   |
| $PAO \xrightarrow{K_{h,PAO}} PAH$  | $K_{h,PAO} = K_{N,PAO} + K_{B,PAO} \cdot [OH^-]$<br>$K_{N,PAO} = 2 \cdot 10^{-4} h^{-1}$<br>$K_{B,PAO} = 46, 1 \cdot M^{-1} h^{-1}$ |
| $PA + OCL^- \xrightarrow{K_{OCL,PA}} PAH$                                | $K_{OCL,PA} = 37(\pm 10) M^{-1} h^{-1}$   |
| $PAO + OCL^- \xrightarrow{K_{OCL,PAO}} PAH$                              | $K_{OCL,PAO} = 48(\pm 10) M^{-1} h^{-1}$  |
| $HOCL \rightarrow H^+ + OCL^-$   | $pka = 7, 5$  |

clustering solutions with ELECTRE TRI facilitates the selection of a solution for implementation in the WDN. This procedure faces the challenge posed by the analysis of a Pareto front, and helps planners and managers of water systems make better decisions.

A real-world case study is solved and discussed to check the validity of the proposed approach. In addition to clustering a Pareto front of 173 solutions, we conclude our case study by ranking the solutions in each class. This last effort is aimed at additionally showing the diverse response to the evaluation criteria of solutions within the same class.

## 2 Methodology

This section provides a description of the methodologies integrated to deal with the problem object of analysis. redFirst, we present the optimization problem, and then the two methodologies, namely the NSGA-II to solve the optimization problem, and the ELECTRE TRI to cluster the optimal solutions obtained in the multi-objective stage.

### 2.1 The optimization problem

In this work, the main objective of the network monitoring problem is the fast and accurate detection of possible chemical intrusions. To this end, parathion intrusion is simulated to obtain the quality state of the network. Parathion is a species of organophosphate that could be used as a contaminant in water systems (Ohar et al. 2015). To model the intrusion of parathion and reactions with the free chlorine present in the water bulk, this work uses Epanet-MSX (Shang et al. 2008). redEpanet-MSX simulates chemical and mass transportation of multiple species in the water network and is linked in use to Epanet 2.0, which is responsible for calculating the hydraulic network state (e.g. nodal pressure, and flowrate in pipes). Both simulators are integrated in the MATLAB programming environment, using the wrapper developed by Eliades et al. (2016).

Chemical reactions of multi-species are based on (Ohar et al. 2015) for generating contamination scenarios. Table 1 presents the degradation system used for parathion modelling.

The efficiency of a monitoring system can be evaluated under anomaly conditions by means of response time and accuracy of detection. The BWNS (Ostfeld et al. 2008) proposes various objective functions to pursue OSP. One of these is the minimal time elapsed between the beginning of a contamination event, and the related detection averaged over the number of detected contamination scenarios. The objective function related to detection time,  $F_1$ , can be expressed as:

$$F_1 = \frac{\sum_{i=1}^N \sum_{t=1}^T \alpha(\gamma) \cdot \gamma(i, t)}{\sum_{i=1}^N \sum_{t=1}^T \alpha(\gamma)}, \quad (1)$$

with the definition

$$\alpha(\gamma) = \begin{cases} 1, & \gamma(i, t) > 0 \\ 0, & \gamma(i, t) \text{ is not available} \end{cases}, \quad (2)$$

where  $\gamma(i, t)$  is the time elapsed from the beginning of a contamination event at node  $i$  in a network with  $N$  nodes till the first identification of the contaminant concentration, during a total simulation duration of  $T$ .  $\alpha(\gamma)$  is a Boolean variable that identifies if contamination has been detected by at least one sensor.

The second objective function ( $F_2$ ) is related to the number of simulated contamination scenarios. During the contamination simulations, each node of the network is selected for parathion injection. The solutions that maximize  $F_2$  are directly linked to the capacity of contamination identification. This objective function is mathematically expressed by

$$F_2 = \sum_{i=1}^N \sum_{t=1}^T \alpha(\gamma) \frac{1}{T}. \quad (3)$$

Finally, considering the cost generated from the process of network monitoring, the objective function  $F_3$  minimizes the number of sensors  $N_s$ :

$$F_3 = \min(N_s). \quad (4)$$

The three mentioned objective functions are used to find optimal positions for the sensors in anomaly quality scenarios. A multi-objective algorithm is applied to find the Pareto front, that is, the set of non-dominated solutions for the problem. Nevertheless, the monitoring system can be also used for normal conditions. The approach proposed by De Schaetzen et al. (2000) has been adapted to evaluate the Pareto front with external parameters. In this work, the authors use two objective functions for optimal redwater quality sensors placement. The first function maximizes the sensitivity of the monitoring network, calculated by the relation between concentration and nodal demand variations:

$$s(i, j) = \frac{\delta C_i}{\delta q_j} = \frac{C_i - C_i^*}{q_j - q_j^*}, \quad (5)$$

where  $C_i^*$  is the free chlorine concentration at node  $i$  for the scenario with the new nodal demand  $q_j^*$  at node  $j$ .  $C_i$  is the free chlorine concentration at node  $i$  for the basic scenario. For  $N_s$  sensors, the accumulated sensitivity corresponding to the entire monitoring network may be evaluated by function  $E_1$ , which can be written as:

$$E_1 = \sum_{k=1}^{N_s} a_k, \quad (6)$$

where  $a_k$  is the maximal value of the sensitivity matrix on line  $k$  corresponding to the position of the sensor.

The second objective function used by De Schaetzen et al. (2000) involves the entropy parameter. Entropy improves sensor distribution in the network by eliminating redundant information. The entropy,  $E_2$ , can be calculated as:

$$E_2 = - \sum_{k=1}^{N_s} p_k \times \ln(p_k), \quad (7)$$

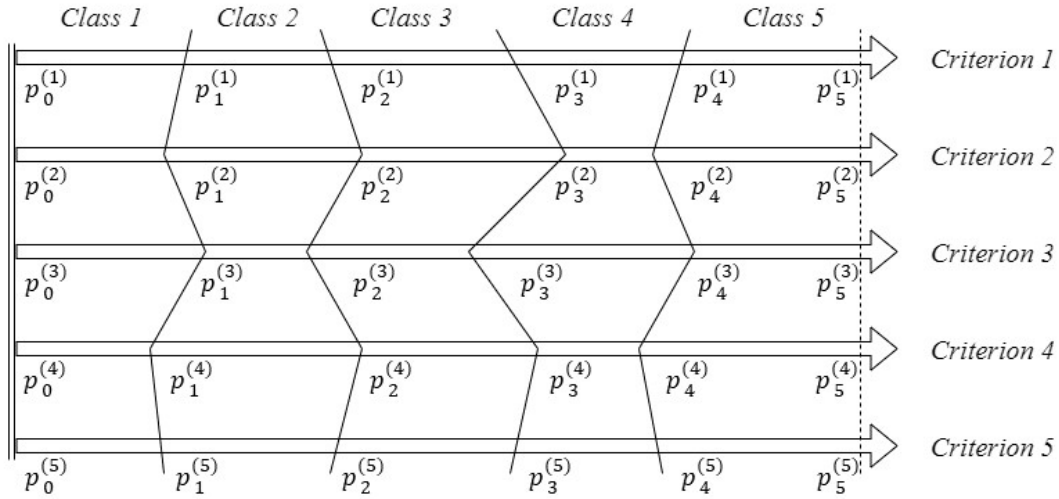
where

$$p_k = \frac{a_k}{\sum_{k=1}^{N_s} a_k}. \quad (8)$$

## 2.2 The multi-objective optimization algorithm - NSGA- II

The algorithm NSGA-II, a fast and elitist multi-objective genetic algorithm developed by Deb et al. (2002), is an adaptation of a genetic algorithm for multi-objective problems. The algorithm generates a set of solutions as the best trade-off among the objective functions throughout the optimization process. Such a set of solutions, the so-called Pareto front, provides relevant information for generating new solutions with suitable crossing and mutation operators.

When the user selects the number of initial solutions, the algorithm is randomly initialized. The objective function values are calculated for each solution provided. Based on their fitness values, the non-dominated solutions are then identified and used to create a new generation, based on crossover, mutation, and elitism mechanisms [17]. The objective functions are evaluated again, and a new Pareto front is progressively defined. The process is repeated until convergence.



**Fig. 1.** ELECTRE TRI classes representation

### 2.3 ELECTRE TRI to cluster optimal solutions

ELECTRE TRI is a non-compensatory MCDM method specifically used for supporting categorization problems (Ramezani 2019). The objective of the method consists in redclustering alternatives into ordered classes defined by threshold values, also called reference profiles. Specifically, alternatives are not pairwise compared with each other. Instead, the clustering procedure, aimed at assigning the alternatives to classes, is carried out by pairwise comparisons of each alternative with the thresholds characterizing the classes. Non-overlapping classes (Certa et al. 2017) are considered in ELECTRE TRI. Figure 1 exemplifies the case of five classes defined by four reference profiles with relation to five evaluation criteria.

redELECTRE TRI requires the development of a specific outranking relation (Figueira et al. 2010); in other words, a particular relation has to be established when comparing alternatives with reference profiles. Such a relation can be expressed in three main ways, namely: indifference, when an alternative outranks a reference profile and vice versa; preference, when an alternative outranks a reference profile and not vice versa; and incomparability, when an alternative and reference profile cannot be compared because they diverge too much with each other.

The following input data must first be collected:

- set of  $K$  evaluation criteria,  $B_k$  ( $k = 1, \dots, K$ ), under which alternatives have to be evaluated;
- criteria weights  $w_k$ , expressing certain relative importance among criteria;
- set of  $J$  reference profiles  $P_j$  ( $j = 1, \dots, J$ ), each one characterized by specific evaluations under each criterion, and defined by limits  $p_0^{(k)} < \dots < p_{J+1}^{(k)}$ ;
- $J + 1$  number of classes determined by the  $J$  profiles;
- set of  $I$  alternatives  $A_i$  ( $i = 1, \dots, I$ ), with the related evaluations  $B_k(A_i)$  assumed under each criterion;
- a threshold value  $\lambda$ , known as *cutting level*, needed to complete the first stage;
- values of indifference, strong preference and veto thresholds, namely  $I_k$ ,  $S_k$ , and  $V_k$ , related to the outranking relations.

$I_k$  is the minimal significant difference to express a preference between two elements,  $S_k$  is the minimal difference to express a strong preference between two elements, and  $V_k$  is the minimal difference expressing incomparability between two elements (Carpitella et al. 2018).

redOnce the listed input data is collected, the ELECTRE TRI clustering procedure follows two main stages: 1. development of outranking relations, and 2. exploitation of the defined outranking relations to classify alternatives. redIn the first stage, the outranking relation characterizing comparisons between alternatives and reference profiles is based on the calculation of suitable concordance and discordance indexes, as shown in (Roy 1990). In the second stage, the clustering is performed by making use of the previously established outranking relation, and alternatives are assigned to classes following two possible procedures, namely pessimistic and optimistic procedures.

redA complete description of the ELECTRE TRI application may be found in (Certa et al. 2017).

### 3 Case study: solution and discussion

The methodology developed in this work is applied to the D-town network (Marchi et al. 2013). The network is composed of 388 nodes, 429 pipes, 13 pumps, 4 pressure reducing valves, 1 reservoir, and 7 tanks; and is divided into 5 district metered areas.

Chlorination is modelled by constant chlorine injection in reservoir and tanks with a concentration of  $1.5 \text{ mg/L}$ . Contamination is modelled with injections of parathion after every 24 hours of normal simulation. Injections have a duration of 12 hours at a concentration of  $12.4 \text{ mg/L}$ , corresponding to the saturation concentration of parathion. To build a database for the optimization process, each node is contaminated in turn, and the corresponding contamination event is simulated. redHydraulic and water quality simulations are run in an extended period simulation for a lapse of 48h, with time steps of 1 minute.

The application of NSGA-II results in a Pareto front with 173 non-dominated solutions. It is clear that such huge numbers of solutions, all of them optimal, represent a problem for decision makers. It is in this situation that this work applies the ELECTRE TRI method to cluster the solutions, thus providing the decision makers with a reduced set of representative layouts that eases the decision-making problem.

When the ELECTRE TRI application handled the large set of 173 mentioned optimal solutions ( $A_i, i = 1, \dots, 173$ ), the solutions were clustered into four ordered classes according to five evaluation criteria, namely: detection time ( $B_1$ ); detection probability ( $B_2$ ); number of sensors ( $B_3$ ); sensitivity ( $B_4$ ); and entropy ( $B_5$ ). The first three criteria are used in the optimization process, while the last two are external criteria, used to assess good solutions both for anomaly detection and for normal scenarios. Regarding the preference directions, criteria  $B_1$  and  $B_3$  have to be minimized, whereas criteria  $B_2$ ,  $B_4$  and  $B_5$  have to be maximized. The five criteria have been assumed as having the same importance for the clustering process, and thus assigned equal weights,  $w_k = 0.2$ . Figure 2a presents the Pareto front, and Figure 2b shows the external criteria according to the number of sensors.

For better understanding of the hydraulic problem, Figure 3 presents the relation, with respect to the number of sensors, of the other four criteria. Observe that by increasing the number of sensors, the detection time is reduced approximately from 3300s to 2700s when the sensors are better placed. However, the Pareto front still has solutions with few sensors and better detection times. These solutions, however have poor values for the other criteria. Also observe that probability detection and sensitivity increase with the number of sensors. Since sensitivity is a cumulative feature, increasing the number of sensors increases this parameter. For the probability detection, more sensors in the network mean a higher probability of at least one sensor detecting contaminants. Finally, it can be seen how entropy slightly increases with the number of sensors, reaching a maximal value of around 6.

Table 2a presents the three reference profiles  $P_1, P_2, P_3$ , characterized by their criterion-dependent limits  $p_0^{(k)} < \dots < p_4^{(k)}$ , derived from the available range of alternative evaluation values. These values identify the four original classes  $C_j$  ( $j = 1 \dots 4$ ). These classes have been ordered from worst to best, in the following way: low ( $C_1$ ); medium ( $C_2$ ); medium-high ( $C_3$ ); and high ( $C_4$ ). The alternatives are then classified on the basis of the outranking relation established with the reference profiles. As a consequence, a solution may be eventually assigned to a class even though, for some criteria, its evaluations do not exactly match that specific class. In some sense, the assignment procedure is carried out by globally evaluating solutions according to the entire set of criteria, and with respect to the reference profiles delimiting the original classes.

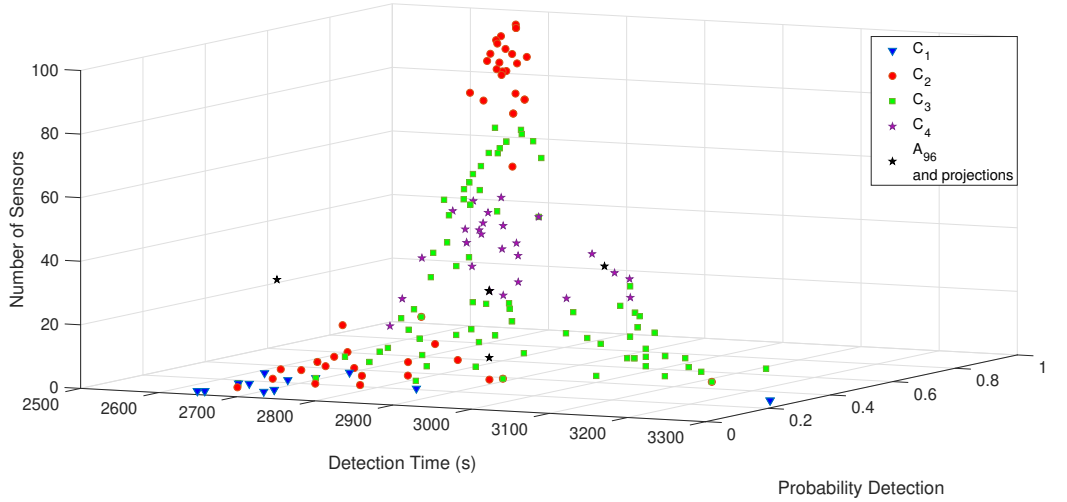
The discrimination thresholds must be established by the decision maker (Mousseau et al. 2000) to opportunely calibrate the method to the specific problem. No veto conditions have been established, whereas indifference and preference thresholds have been determined by first setting larger values and progressively reducing them until considered appropriate for each criterion. The validity of the achieved threshold values was finally checked and confirmed by an external expert of ELECTRE TRI applications. These values are given in Table 2a for all the criteria.

Supplementary files provide the tables summarizing all the achieved results, obtained by fixing the value of the cutting level  $\lambda$  to 0.80 (see (Certa et al. 2017)), and show the final allocation of the optimal solutions to the classes. We note here that most solutions belong to the medium-high class.

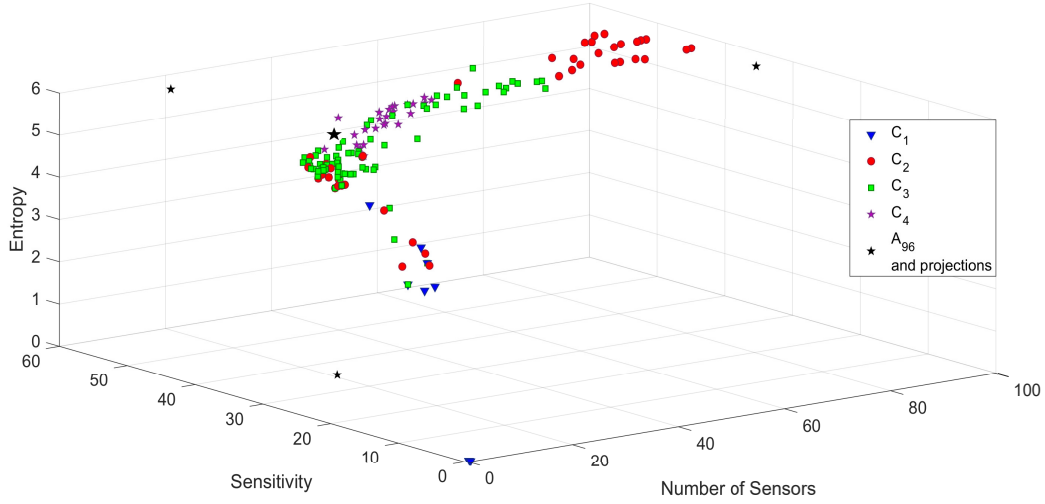
Once the clustering process has been carried out by means of the ELECTRE TRI procedure, we conclude the case study by ranking alternatives from best to worst within the classes. Table 2b presents, for each final class, the five first and the five last solutions of the rankings obtained by applying the technique for order of preference by similarity to ideal solution (TOPSIS), a well-known MCDM method used by the authors to rank large sets of alternatives in previous works (Brentan et al. 2019; Carpitella et al. 2018a; Carpitella et al. 2018b). Data corresponding to these solutions may also be found in supplementary materials.

For the sake of completeness, Table 2c presents the criteria values for the best solution of each final class.

$A_{70}$  is the best solution of class  $C_1$ . This solution has a small detection time and the lowest number of installed sensors. However, its other indicators, specifically probability detection and sensitivity, place this solution in the worst class. The best ranked solution in  $C_2$ ,  $A_{87}$ , presents the lowest detection time and the best probability detection. Furthermore, this solution has the highest sensitivity and entropy values. However, the solution also entails the



(a) Pareto front resulted from NSGA-II algorithm



(b) Evaluation of optimal solutions under normal scenario criteria

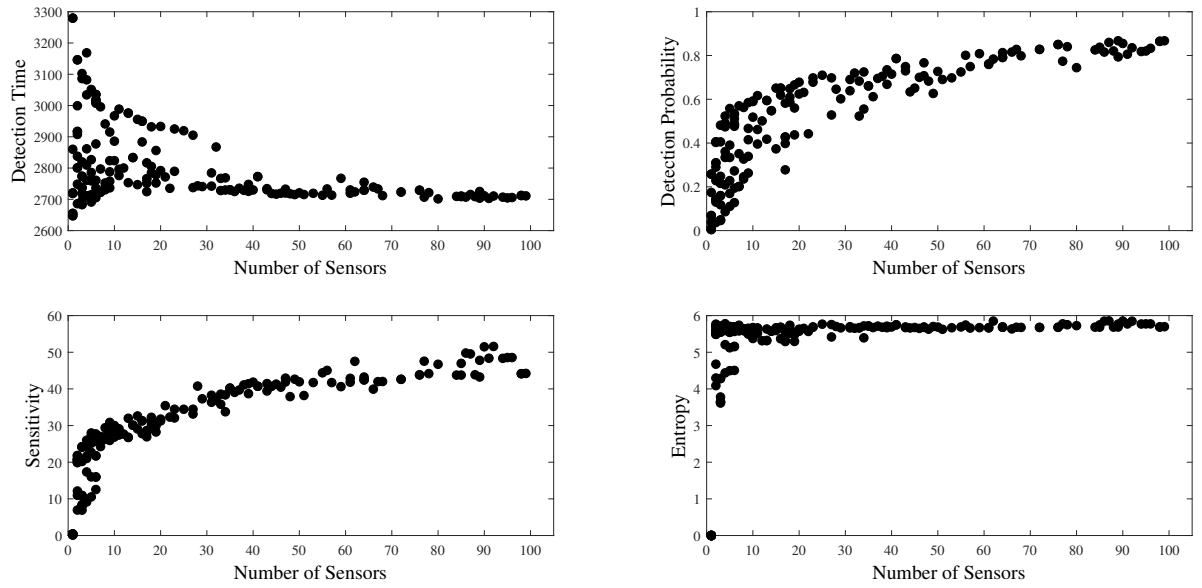
**Fig. 2.** Pareto front and evaluation criteria visualisation after ELECTRE TRI clustering.

highest number of sensors, usually far from the economic criteria for water companies. The lead solution from  $C_3$ ,  $A_{111}$ , has the worst detection time; however, it is a detection time just five minutes longer than the best detection time. Nevertheless, with only one sensor more than solution  $A_{70}$  from  $C_1$ , probability detection and sensitivity increase significantly. This solution looks desirable for small budgets, and provides a good trade-off among the five criteria. Finally, solution  $A_{96}$ , which is represented, together with its projections to the coordinate planes, by black stars, is the best solution of  $C_4$ .  $A_{96}$  has 37.5% fewer sensors than the first ranked solution in  $C_2$ , while achieving greater sensitivity and detection probability. For a moderate to high budget, this solution is extremely appealing. Figure 4 shows the layout of the monitoring network for each these best ranked solutions. One can observe the uniform distribution of the solution from  $C_3$ , and the high redundancy provided by the solution from  $C_2$ .

#### 4 Conclusions

Water distribution system managers worldwide have increased budgets for monitoring and data acquisition. However, two main issues still remain regarding the problem of sensor placement: the optimal number of sensors to be installed, and the positions to place those sensors. Three research fields have been linked and jointly explored in this work in an attempt to give answers to both problems. Water quality simulation using EPANET-MSX is undertaken to generate a database of chlorine and parathion concentration after simulating a parathion intrusion. The database is used during the optimization process, where the optimal quality sensors are mathematically described by three





**Fig. 3.** Relation between objective functions and evaluation criteria and number of sensors

**Table 2.** ELECTRE TRI application

(a) Input data

| Criterion | $I_k$  | $S_k$  | $C_1$             | $C_2$             | $C_3$             | $C_4$             |
|-----------|--------|--------|-------------------|-------------------|-------------------|-------------------|
| $B_1$     | 409.95 | 819.89 | 3279.57 - 3121.55 | 3121.55 - 2963.54 | 2963.54 - 2805.52 | 2805.52 - 2647.50 |
| $B_2$     | 6E-4   | 12E-4  | 0.00 - 0.22       | 0.22 - 0.44       | 0.44 - 0.65       | 0.65 - 0.87       |
| $B_3$     | 12     | 24     | 100 - 75          | 75 - 50           | 50 - 25           | 25 - 0            |
| $B_4$     | 6.45   | 12.90  | 0.15 - 13.01      | 13.01 - 25.87     | 25.87 - 38.73     | 38.73 - 51.58     |
| $B_5$     | 0.73   | 1.46   | 0.00 - 1.46       | 1.46 - 2.93       | 2.93 - 4.39       | 4.39 - 5.86       |

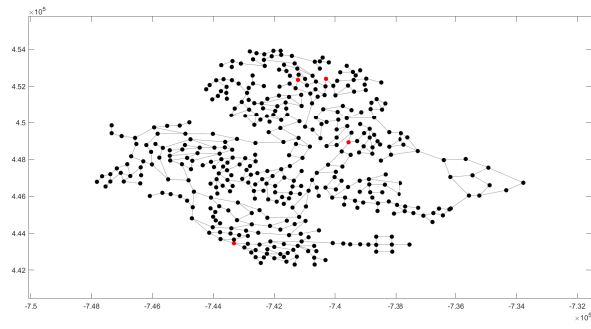
(b) The five best and worst solutions within each class

|                 | $C_1$     | $C_2$     | $C_3$     | $C_4$     |
|-----------------|-----------|-----------|-----------|-----------|
| best trade-off  | $A_{70}$  | $A_{87}$  | $A_{111}$ | $A_{96}$  |
|                 | $A_{143}$ | $A_{137}$ | $A_{131}$ | $A_{84}$  |
|                 | $A_{75}$  | $A_{65}$  | $A_{30}$  | $A_{132}$ |
|                 | $A_{95}$  | $A_{46}$  | $A_{164}$ | $A_{25}$  |
|                 | $A_{169}$ | $A_{125}$ | $A_{32}$  | $A_{127}$ |
| worst trade-off | $A_{109}$ | $A_{89}$  | $A_{54}$  | $A_{16}$  |
|                 | $A_5$     | $A_{63}$  | $A_{145}$ | $A_{72}$  |
|                 | $A_{155}$ | $A_{113}$ | $A_{73}$  | $A_{83}$  |
|                 | $A_4$     | $A_{86}$  | $A_{44}$  | $A_{101}$ |
|                 | $A_{148}$ | $A_{43}$  | $A_{135}$ | $A_{106}$ |

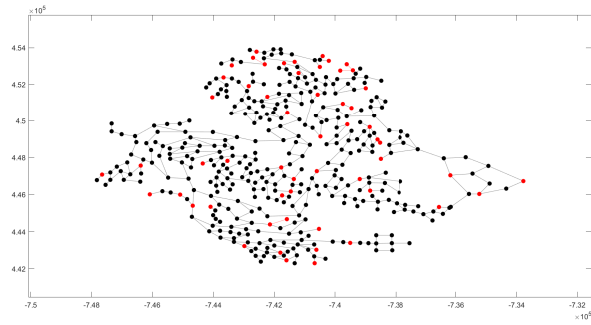
(c) Best solutions for each class

| Class | Best Solution | $B_1$   | $B_2$ | $B_3$ | $B_4$ | $B_5$ |
|-------|---------------|---------|-------|-------|-------|-------|
| $C_1$ | $A_{70}$      | 2760.88 | 0.21  | 4     | 17.34 | 5.21  |
| $C_2$ | $A_{87}$      | 2733.22 | 0.80  | 56    | 45.04 | 5.74  |
| $C_3$ | $A_{111}$     | 3051.41 | 0.56  | 5     | 27.98 | 5.70  |
| $C_4$ | $A_{96}$      | 2771.67 | 0.63  | 21    | 35.43 | 5.65  |

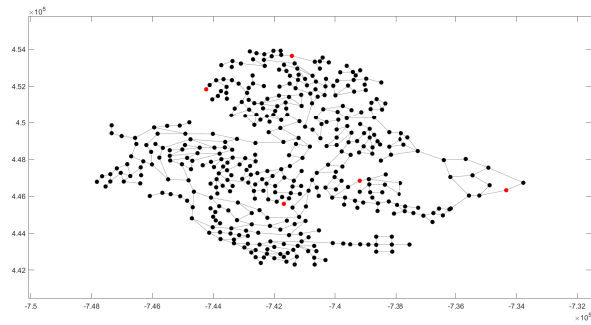
objective functions. Detection time and detection probability are related to the capability of the monitoring network to be useful in anomalous scenarios, while the number of sensors to be placed represents the budget limitations for water companies. The 173 Pareto-optimal solutions have been evaluated under two external parameters: sensitivity and entropy. Both parameters are often applied for sensor placement under normal conditions, and are useful for decision makers to select a trade-off solution.



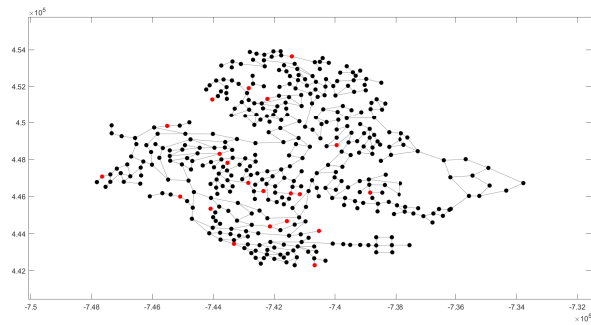
(a) Layout of best ranked solution for  $C_1$



(b) Layout of best ranked solution for  $C_2$



(c) Layout of best ranked solution for  $C_3$



(d) Layout of best ranked solution for  $C_4$

**Fig. 4.** Layout of best ranked solution for each class determined by ELECTRE TRI

With the aim of reducing the number of solutions among which decisions have to be made, a clustering process has been carried out by applying the MCDM ELECTRE TRI method. While classical multi-objective methods result in Pareto fronts, offering decision makers large sets of good (non-dominated) solutions, this work clusters the solutions to reduce that set of options for decision makers. ELECTRE TRI is far more than a mere clustering tool. It constitutes a dialogue framework for decision makers to perform more informed, and thus better, solutions.

By means of this application, four clusters have been identified for the D-town network, which represents a considerable support for analysts when deciding about which solution to implement on the basis of the available budget. Solutions have been ranked within each class to enable managers to easily identify the most suitable layout for the monitoring network. This reduction, keeping sensitivity and entropy as external evaluation parameters, makes the optimization process smoother, while guaranteeing better and faster convergence.

Regarding possible future developments, uncertainty affecting demand could be added to the model, and the fuzzy set theory could be coupled with the clustering method to obtain even better informed results. Furthermore, a reverse analysis using the normal condition as a basis for optimization may be implemented, and anomalous parameters may be used for evaluation, thus allowing robust comparisons among different layouts.

### Conflict of interest

The authors declare that they have no conflict of interest.

### Acknowledgments

The prose of this paper has been revised by John Rawlins (a qualified member of the UK Institute of Translation and Interpreting).

### References

1. Barak S, Mokfi T (2019) Evaluation and selection of clustering methods using a hybrid group mcdm. *Expert Systems with Applications* 138:112817
2. Berry JW, Fleischer L, Hart WE, Phillips CA, Watson JP (2005) Sensor placement in municipal water networks. *Journal of Water Resources Planning and Management* 131(3):237–243
3. Bouyssou D, Marchant T (2015) On the relations between electre tri-b and electre tri-c and on a new variant of electre tri-b. *European Journal of Operational Research* 242(1):201–211
4. Brentan B, Carpitella S, Izquierdo J, Luvizotto Jr E, Meirelles G (2019) A multi-objective and multi-criteria approach for district metered area design: water operation and quality analysis. In: *International Conference on Mathematical Modeling in Engineering & Human Behaviour* (2019), pp 110–117
5. Brito AJ, de Almeida AT, Mota CM (2010) A multicriteria model for risk sorting of natural gas pipelines based on electre tri integrating utility theory. *European Journal of Operational Research* 200(3):812–821
6. Broad DR, Maier HR, Dandy GC, Nixon JB (2008) Optimal design of water distribution systems including water quality and system uncertainty. In: *Water Distribution Systems Analysis Symposium 2006*, pp 1–17
7. Candelieri A, Conti D, Archetti F (2014) A graph based analysis of leak localization in urban water networks. *Procedia Engineering* 70:228–237
8. Carpitella S, Brentan B, Montalvo I, Izquierdo J, Certa A (2018) Multi-objective and multi-criteria analysis for optimal pump scheduling in water systems. *EPiC Series in Engineering* 3:364–371
9. Carpitella S, Certa A, Izquierdo J, La Fata CM (2018) k-out-of-n systems: an exact formula for the stationary availability and multi-objective configuration design based on mathematical programming and topsis. *Journal of Computational and Applied Mathematics* 330:1007–1015
10. Carpitella S, Ocaña-Levario SJ, Benítez J, Certa A, Izquierdo J (2018) A hybrid multi-criteria approach to gpr image mining applied to water supply system maintenance. *Journal of Applied Geophysics* 159:754–764
11. Certa A, Enea M, Galante GM, La Fata CM (2017) Electre tri-based approach to the failure modes classification on the basis of risk parameters: An alternative to the risk priority number. *Computers & Industrial Engineering* 108:100–110
12. Cheung P, Piller O, Propato M (2005) Optimal location of water quality sensors in supply systems by multiobjective genetic algorithms. In: *Eight International Conference on Computing and Control in the Water Industry CCWI05*, vol 1, p 2
13. Christodoulou SE, Gagatsis A, Xanthos S, Kranioti S, Agathokleous A, Fragiadakis M (2013) Entropy-based sensor placement optimization for waterloss detection in water distribution networks. *Water resources management* 27(13):4443–4468

14. Corrente S, Greco S, Słowiński R (2016) Multiple criteria hierarchy process for electre tri methods. *European Journal of Operational Research* 252(1):191–203
15. Costa AS, Govindan K, Figueira JR (2018) Supplier classification in emerging economies using the electre tri-nc method: A case study considering sustainability aspects. *Journal of Cleaner Production* 201:925–947
16. De Schaetzen W, Walters G, Savic D (2000) Optimal sampling design for model calibration using shortest path, genetic and entropy algorithms. *Urban Water* 2(2):141–152
17. Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE transactions on evolutionary computation* 6(2):182–197
18. Dias LC, Antunes CH, Dantas G, de Castro N, Zamboni L (2018) A multi-criteria approach to sort and rank policies based on delphi qualitative assessments and electre tri: The case of smart grids in brazil. *Omega* 76:100–111
19. Eliades DG, Kyriakou M, Vrachimis S, Polycarpou MM (2016) Epanet-matlab toolkit: An open-source software for interfacing epanet with matlab. In: *Proceedings of the 14th International Conference on Computing and Control for the Water Industry, CCWI*
20. Fernandez E, Navarro J (2011) A new approach to multi-criteria sorting based on fuzzy outranking relations: The theseus method. *European Journal of Operational Research* 213(2):405–413
21. Fernández E, Figueira JR, Navarro J, Roy B (2017) Electre tri-nb: A new multiple criteria ordinal classification method. *European Journal of Operational Research* 263(1):214–224
22. Figueira JR, Greco S, Roy B, Słowiński R (2010) Electre methods: main features and recent developments. In: *Handbook of multicriteria analysis*, Springer, pp 51–89
23. Figueira JR, Greco S, Roy B, Słowiński R (2013) An overview of electre methods and their recent extensions. *Journal of Multi-Criteria Decision Analysis* 20(1-2):61–85
24. Francés-Chust J, Brentan BM, Carpitella S, Izquierdo J, Montalvo I (2020) Optimal placement of pressure sensors using fuzzy dematel-based sensor influence. *Water* 12(2):493
25. Gandy M (2004) Rethinking urban metabolism: water, space and the modern city. *City* 8(3):363–379
26. Giudicianni C, Herrera M, Di Nardo A, Greco R, Creaco E, 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
27. Hart WE, 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
28. Herrera M, Abraham E, Stoianov I (2016) A graph-theoretic framework for assessing the resilience of sectorised water distribution networks. *Water Resources Management* 30(5):1685–1699
29. Huang JJ, McBean EA, James W (2008) Multi-objective optimization for monitoring sensor placement in water distribution systems. In: *Water Distribution Systems Analysis Symposium 2006*, pp 1–14
30. Kapelan ZS, Savic DA, Walters GA (2003) A hybrid inverse transient model for leakage detection and roughness calibration in pipe networks. *Journal of Hydraulic Research* 41(5):481–492
31. Lee JH (2013) Determination of optimal water quality monitoring points in sewer systems using entropy theory. *Entropy* 15(9):3419–3434
32. Liu Z, Ming X (2019) A methodological framework with rough-entropy-electre tri to classify failure modes for co-implementation of smart pss. *Advanced Engineering Informatics* 42:100968
33. Marchi A, Salomons E, Ostfeld A, Kapelan Z, Simpson AR, Zecchin AC, Maier HR, Wu ZY, Elsayed SM, Song Y, et al. (2013) Battle of the water networks ii. *Journal of water resources planning and management* 140(7):04014009
34. Mohammed A, Harris I, Soroka A, Nujoom R (2019) A hybrid mcdm-fuzzy multi-objective programming approach for a g-resilient supply chain network design. *Computers & Industrial Engineering* 127:297–312
35. Montalvo I, Izquierdo J, Pérez-García R, Herrera M (2014) Water distribution system computer-aided design by agent swarm optimization. *Computer-Aided Civil and Infrastructure Engineering* 29(6):433–448
36. Mousseau V, Slowinski R, Zielniewicz P (2000) A user-oriented implementation of the electre-tri method integrating preference elicitation support. *Computers & operations research* 27(7-8):757–777
37. Nafi A, Crastes E, Sadiq R, Gilbert D, 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
38. Neto JGD, Machado MAS, Gomes LFAM, Caldeira AM, Sallum FSV (2017) Investments in a new technological infrastructure: Decision making using the electre-tri methodology. *Procedia computer science* 122:194–199
39. Ohar Z, Lahav O, Ostfeld A (2015) Optimal sensor placement for detecting organophosphate intrusions into water distribution systems. *Water research* 73:193–203
40. Olikar N, Ostfeld A (2015) Network hydraulics inclusion in water quality event detection using multiple sensor stations data. *Water research* 80:47–58

41. Ostfeld A, Salomons E (2005) Optimal early warning monitoring system layout for water networks security: Inclusion of sensors sensitivities and response delays. *Civil Engineering and Environmental Systems* 22(3):151–169
42. Ostfeld A, Uber JG, Salomons E, Berry JW, Hart WE, Phillips CA, Watson JP, Dorini G, Jonkergouw P, Kapelan Z, et al. (2008) The battle of the water sensor networks (bwsn): A design challenge for engineers and algorithms. *Journal of Water Resources Planning and Management* 134(6):556–568
43. Quiñones-Grueiro M, Verde C, 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), IEEE, pp 129–134
44. Ramezani R (2019) Estimation of the profiles in posteriori electre tri: A mathematical programming model. *Computers & Industrial Engineering* 128:47–59
45. Rathi S, Gupta R, Kamble S, Sargaonkar A (2016) Risk based analysis for contamination event selection and optimal sensor placement for intermittent water distribution network security. *Water resources management* 30(8):2671–2685
46. Reginaldo F (2015) Portfolio management in brazil and a proposal for evaluation and balancing of portfolio projects with electre tri and iris. *Procedia Computer Science* 55:1265–1274
47. Roy B (1968) Classement et choix en présence de points de vue multiples. *Revue française d'informatique et de recherche opérationnelle* 2(8):57–75
48. Roy B (1990) The outranking approach and the foundations of electre methods. In: *Readings in multiple criteria decision aid*, Springer, pp 155–183
49. Sánchez-Lozano J, García-Cascales M, Lamata M (2016) Comparative topsis-electre tri methods for optimal sites for photovoltaic solar farms. case study in spain. *Journal of Cleaner Production* 127:387–398
50. Seiti H, Hafezalkotob A, Najafi SE, Khalaj M (2019) Developing a novel risk-based mcdm approach based on d numbers and fuzzy information axiom and its applications in preventive maintenance planning. *Applied Soft Computing* p 105559
51. Shang F, Uber JG, Rossman LA, et al. (2008) Epanet multi-species extension user's manual. Risk Reduction Engineering Laboratory, US Environmental Protection Agency, Cincinnati, Ohio
52. Shannon CE (1948) A mathematical theory of communication. *The Bell system technical journal* 27(3):379–423
53. Štirbanović Z, Stanujkić D, Miljanović I, Milanović D (2019) Application of mcdm methods for flotation machine selection. *Minerals Engineering* 137:140–146
54. Wang H, Jiang Z, Zhang H, Wang Y, Yang Y, Li Y (2019) An integrated mcdm approach considering demands-matching for reverse logistics. *Journal of cleaner production* 208:199–210
55. Wéber R, 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
56. Weickgenannt M, Kapelan Z, Blokker M, Savic DA (2010) Risk-based sensor placement for contaminant detection in water distribution systems. *Journal of Water Resources Planning and Management* 136(6):629–636
57. de Winter C, Palleti VR, Worm D, Kooij R (2019) Optimal placement of imperfect water quality sensors in water distribution networks. *Computers & Chemical Engineering* 121:200–211