Validation of drinking water distribution monitoring schemes including water demand variability and minimum detectable concentrations

Marco Gabrielli[1], Alessandro Leoni[1], Laura Di Domenico[1], Manuela Antonelli[1*]

[1] Dipartimento di Ingegneria Civile e Ambientale – Sezione Ambientale, Politecnico di Milano, Milan, Italy

* Corresponding Author: **Manuela Antonelli**

Telephone: (+39) 02 2399 6407

E-mail: manuela.antonelli@polimi.it

Abstract

The selection of monitoring locations within drinking water distribution networks (DWDNs) is of primary importance to guarantee consumers safety. In order to estimate the monitoring performances of different potential layouts (i.e., configurations of monitoring locations) and select the best option, a simulation-based validation is required. While few methodologies to guide events' selection for this validation have been developed, most of them fail to properly capture the most impactful events and the heterogeneity in DWDNs operating conditions, not considering the inherent uncertainty in water demands within the DWDNs. In this study we extend a previously developed methodology to target the monitoring layout validation towards the most impactful events, and including water demand uncertainty. The developed methodology was tested on a literature benchmark DWDN, highlighting how optimal performances against generic events do not guarantee consumers' protection against the most impactful ones. Furthermore, using two real-world DWDNs, it was shown how neglecting water demand uncertainty leads to an overestimation of the detection likelihood of monitoring layouts. Finally, the effect of increasing minimum detection concentration of a generic harmful compound was assessed. Higher minimum detection concentrations resulted in lower detection probabilities, especially in case of larger DWDNs, while their effect on the average minimum detection time depended on the DWDN characteristics.

Keywords: Drinking water distribution networks; monitoring; water alteration events; water demand variability; scenario analysis

1. Introduction

The drinking water distribution networks (DWDNs) play a key role in ensuring consumers' safety, being located between the outlet of the treatment plants and the consumers' homes and potentially being subjected to contaminations (Robertson et al., 2003). The importance of DWDNs monitoring has been increasingly recognized with water utilities expanding monitoring programs into the networks, using both manual grab sampling programs (da Luz and Kumpel, 2020; World Health Organization, 2017) and with the installation of on-line sensors (Zulkifli et al., 2018) to catch water quality alterations due to backflows, cross-connections and other intrusions in DWDNs (USEPA, 2002).

However, monitoring such infrastructures is not trivial. Due to the DWDNs size, it is nor practical or feasible to monitor all possible locations. In addition, due to the high degree of interconnection, water and eventual contaminants might reach the consumers from multiple pathways, confounding the potential contamination source (Gong et al., 2023). Finally, DWDNs operating conditions, such as pressure, flow rate and direction, depend on the behavior of drinking water consumers and are both highly dynamic (Brentan et al., 2017) and characterized by an intrinsic degree of uncertainty (Shastri and Diwekar, 2006).

Therefore, simulation-based and graph-based strategies have been developed to identify monitoring layouts, i.e., configurations of monitoring locations, that maximize the coverage against water quality alteration events, while minimizing the number of monitoring locations. While the former set of techniques relies on hydraulic DWDN models and optimization algorithms, the latter group of methods is less computationally demanding and requires only details regarding the DWDN morphology (Hu et al., 2018). In any case, in order to select best layout among different potential alternatives, a set of water alteration events needs to be simulated to estimate the performance of each monitoring layout, for example evaluating the likelihood of detecting a water quality alterations and, especially in case of sensors are used, the time required to do so (Rathi et al., 2015; Rathi and Gupta, 2015).

A critical aspect in the validation process lies in the high computational requirements given by the number of operating conditions and water quality alterations that need to be simulated to cover all possible scenarios (Krause et al., 2008; Mu et al., 2021). To reduce such burden, previous studies have selected subsets of events either randomly (Rathi and Gupta, 2015) or following specified heuristics (Tinelli et al., 2017). However, such procedures are likely to miss rare and highly impactful events, leading to erroneous estimations of consumers' protection (Perelman and Ostfeld, 2010). For this reason, Weickgenannt et al. (2010) introduced a method, named Importance-Based Events Sampling (IBES), which guides the selection of validation events by assigning a simulation probability to each event based on a first estimation of its hydraulic importance. Notably, monitoring performances of a monitoring layout can change when focusing on highly impactful events, rather than generic ones. In light of this, it is not known whether monitoring layouts providing optimal performances for generic events will also result optimal to detect highly impactful events. This also extends to the comparison of the performance between simulation-based and graph-based layout identification strategies. In fact, even though comparable performances have been found between simulation-based layout identification strategies and graph-based ones (Giudicianni et al., 2020), it is not known whether this is also true for highly impactful events, representing an important knowledge gap for water utility managers.

In addition, even though calibrated models are used in the validation of monitoring layouts, such models might be not fully representative of the actual conditions within DWDNs. In fact, demand patterns can change in response to specific events and holidays, and also with seasons and weekdays (Alvisi et al., 2021; Brentan et al., 2017). Such variability can result in changes in the flow direction (Shastri and Diwekar, 2006), changing the water sources within the DWDN (Gabrielli et al., 2023). Hence, not taking into account this inherent stochasticity might provide inaccurate evaluation of the monitoring performances of a selected layout.

Finally, as the type of potential water quality alterations are not known in advance, validation procedures should be general enough to be adapted to the simulation of different contaminants without aggravating the computational burden. Methodologies dedicated to specific contaminants have been developed (Ohar et al., 2015), but most studies focus on monitoring layouts able to detect any non-zero concentrations (e.g., Giudicianni et al., 2020; Zhao et al., 2016) or a single fixed value (Zheng et al., 2018), providing either unrealistic estimates of events detection or neglecting the differences between different contaminants. This is not optimal for water utilities. In fact, the possibility to estimate and compare the performances of monitoring layouts with different minimum detection levels with limited computational effort would enable water utilities to not only select the best option for a given contamination scenario, but also to evaluate the differences that could be expected for the detection of specific contaminants or using different monitoring technologies.

For these reasons, this study extends the IBES method to account for the stochasticity of water demands and provides a conceptual model to estimate monitoring performances in case of different contaminants and/or sensitivity of the analytical methods. The eIBES (extended IBES) method was then used (i) on a literature benchmark DWDN to estimate the difference in performances due to the focus on important events, and on two real-world case studies to evaluate (ii) the effect of the inclusion of water demand uncertainty and (iii) the effect of increasing minimum detection concentration on monitoring layouts performances.

2. Materials and methods

2.1 Development of the extended Importance-Based Events Sampling Method

The IBES method proposed by Weickgenannt et al. (2010) focuses on the most impactful events by conducting a preliminary hydraulic simulation. From this simulation, the cumulated water volume flowing through node n_j from time t_i to time $t_i+\Delta t$, where Δt represents a specified time window, is estimated, representing the hydraulic importance value to each couple (t_i,n_j). These values are then converted to probability values by means of normalization with respect to the total sum of matrix entries and used to weight the probability of simulating the couples (t_i,n_j), representing pollution events' starting location and time. In such a way, while more emphasis is given to highly impacting events, also smaller ones are simulated, ensuring the protection of all consumers.

To take into account the inherent stochasticity of consumer water demands without the need of external data, the eIBES method exploits the periodic signal of the water demand patterns, commonly included in calibrated hydraulic models, to introduce the water demand uncertainty by a set of synthetic patterns. In order to generate the synthetic patterns, the original deterministic patterns were split in three additive components by the means of a decomposition in seasonal and trend patterns using Loess regression (STL decomposition; Cleveland et al., 1990): average daily patterns, trend and random variability. The random component was used to fit a Gaussian distribution to generalize the stochastic noise present in water demands. Random extraction of values from this distribution were summed to the other two components to generate Q novel synthetic patterns, considerable as alternative stochastic realizations of the original deterministic pattern. As a DWDN hydraulic model usually comes with a set of P demand patterns, each assigned to a district of nodes, such procedure was applied to all the demand patterns, resulting in a total amount of Q^P demand patterns due to the combination of the different patterns. A set of S demand scenarios was randomly selected out of all the Q^P generated, and each of them was used to perform a DWDN hydraulic simulation, computing the corresponding S importance sampling matrix. Then, the 95th percentile of the (t_i,n_j) couples were estimated and normalized by their sum, obtaining the percentages used to guide the random selection of N validation events, as in Weickgenannt et al. (2010). During validation, when a (t_i,n_j) couple was selected to be simulated, the corresponding simulation was conducted with the demand patterns which resulted in the importance closest to the 95th percentile. Doing so ensures the original aim to select a validation set composed predominantly of highly impacting events, without neglecting lesser impacting ones, and also considering the possible variability in water demands.

2.2 Adapting validation results for different water alteration scenarios

In order to estimate the monitoring layouts performances considering different water alteration scenarios, a proportion-based strategy was devised. $C_{X,DWDN}$ and $C_{X,intr}$ are the concentrations of pollutant X, respectively, in baseline DWDN conditions and in the intruding water, and $W_{\%}$ is a percentage ratio between the volume of intruding and distributed water at the intrusion node. Then, the pollutant concentration at the contaminated node can be calculated as a weighted average:

$$C_{X,node} = (C_{X,intr} W_{\%} + C^{repr}_{X,DWDN})/(1 + W_{\%}) \quad \text{Eq. 1}$$

where $C^{repr}_{X,DWDN}$ consists of a representative $C_{X,DWDN}$ value (i.e., the mean) (Teunis et al., 2010). Considering a variability around $C^{repr}_{X,DWDN}$, either due to the measurement uncertainty and/or the concentration variability within the DWDN, one can assume the minimum detectable concentration as a significant deviation from standard water quality conditions:

$$C_{X,min.det.} = C^{repr}_{X,DWDN} + z_{97.5} * \sigma(C_{X,DWDN}) \quad \text{Eq. 2}$$

where $z_{97.5}$ represents the 97.5th percentile of a standard Gaussian distribution percentile. Ultimately, a proportion factor between the expected concentration at the intrusion node and the minimum detectable concentration can be calculated as:

$$R = C_{X,node} / C_{X,min.det.} \quad \text{Eq. 3}$$

This value is, then, exploited to define an equivalent minimum detection concentration to be fixed as new detection limit in the estimation of the performances of the monitoring layouts without performing new simulations:

$$C_{X,min.det.equiv.} = C_{simulated} \; / \, R \qquad \text{Eq. 4}$$

2.3 Case studies

2.3.1 Literature benchmark case study

A literature benchmark DWDN (Ostfeld et al., 2008) was analyzed to compare the effect of importance-based sampling on the monitoring performance of monitoring layouts. The DWDN is composed of 126 nodes, 168 pipes, one water source and presents a daily dynamic water demand. Different monitoring layouts composed of five monitoring locations were included in the validation. Such layouts were obtained both from literature studies employing state-of-the-art simulation-based methods (Ostfeld et al., 2008; Zhao et al., 2016), and using several graph-based techniques obtained varying edges weights, graph clustering algorithms and centrality measures, similar to previous studies (Di Nardo et al., 2017; Giudicianni et al., 2020) and as explained in details in Section S1 in the Supplementary Information (SI). The combination of the selected edges weights, cluster algorithms and centrality measures resulted in a total of 160 layouts for a given number of locations. *NetworkX* v2.3 (Hagberg et al., 2008) and *python-igraph* v0.8.2 (Csardi and Nepusz, 2006) were used for the monitoring layouts estimation.

As only a daily water demand pattern is present within this DWDN model, precluding a successful STL decomposition, an uncertainty comparable to the ones of real-world DWDNs was applied (data not shown, see section 2.3.2 for details about real-world case studies) to generate the synthetic water demand patterns. After the generation of the synthetic water demand patterns, an eIBES validation was performed simulating 100 different events using *WNTR* v0.4.0 (Klise et al., 2017). Similarly to previous studies (Giudicianni et al., 2020; Ostfeld et al., 2008; Tinelli et al., 2017; Weickgenannt et al., 2010), the selected water quality alteration events were simulated considering the presence of a conservative (i.e., not degrading) pollutant introduced at a fixed concentration for a time window of 16 h at the outlet of the contaminated node. The duration of each quality simulation was set equal to 72 h with a discretization time step of 5 min. The simulated events were used to estimate monitoring layouts performances and comparing them with literature values.

2.3.2 Real-world case studies

Two real-world DWDNs supplying mid-to-large towns were used to assess the effect of the stochastic generation of synthetic water demand patterns on DWDN hydraulics and to estimate the difference in monitoring layout performances. The details of the hydraulic models provided by the managing water utility are presented in Table 1.

Table 1. Main features of the tested real-world case studies.

	DWDN C	DWDN R
Size	mid-scale	large-scale
Nodes	982	10,329
Drinking Water Treatment Plants (DWTPs)	2	10
Pipes	730	8,215
Pumps	4	0
Valves	273	2,393
Demand patterns	3	9

A set of 200 synthetic demand patterns was generated for each of the original patterns, which were then randomly combined to define a set of 100 demand scenarios used to estimate the 95th percentile hydraulic importances in each case study. Hydraulic importances were estimated considering a time window Δt of 16 hours and a temporal resolution of 5 min. These importances were then used to sample 250 and 2500 events for respectively DWDN C and R (i.e., approximately 25% of the total nodes), in order to achieve a set of simulations covering the most of the DWDN both with respect to the location of the water quality alteration and its starting time. Such events were simulated either with the use of synthetic patterns following the eIBES methodology (Section 2.1) or with the standard pattern present in the model, in order to check the difference in the results associated to the inclusion of demands uncertainty. Water quality simulation setting were kept equal to the ones used for the benchmark DWDN.

Such set of simulations was then used to validate the performances of several monitoring layouts identified through graph-based strategies (see section S1 in SI) and composed of a number of monitoring locations between three and seven. The performance of the monitoring layouts was tested both considering an ideal situation (i.e., detecting any concentration different from zero) and also with three scenarios with increasing equivalent minimum detection concentrations ($C_{X,min.det.equiv.}$). The features of each scenario are summarized in Table 2 assuming a 1% percentage ratio between the volume of intruding and distributed water at the intrusion node (Teunis et al., 2010) and using DWDN baseline and intruding concentrations either derived from data provided by the managing water utility or from literature (Besmer et al., 2017; Gabrielli et al., 2021; Henze et al., 2008).

Table 2. Main features of the contamination scenarios. $C_{min.\ eq.}$ values have been estimated considering $C_{simulated}$ = 1 g/L.

	Scenario I	Scenario II	Scenario III
Pollutant	Bacteria cells [cells/µl]	Total organic carbon [mg/l]	Ammonium [mg/l]
$C_{X,intr}$	147,000	100	20
$C^{repr}{}_{X,DWDN}$	73	0.2	0.05
$\sigma(C_{X,DWDN})$	51	0.1	0.05
$C_{X,node}$	1,542	1.20	0.25
$C_{X,min.\ det.}$	172.96	0.40	0.15
$C_{X,min.det.eqiv.}$	112.15	330.55	593.19

2.3.3 Monitoring layouts performance evaluation

For each simulated event, pollutants concentrations were computed at every couple (t_i,n_j) and then used to estimate the value of two different performance metrics for each monitoring layout (Rathi et al., 2015):

- Average detection likelihood (detLik) [-]: the percentage of detected events by a certain layout compared to all the simulated ones, considering an event as detected by the layout if at least one monitoring location observes a concentration above $C_{X,min.det.eqiv.}$.
- Average minimum detection time (detTime) [min]: the arithmetic mean of the minimum time needed by the monitoring layout to detect the water quality alteration in at least one node belonging to the monitoring layout.

In addition, to provide further opportunity for comparison with literature, the average consumed volume of contaminated water prior to detection (consVol) [m^3] was estimated for the benchmark DWDN.

3. Results and discussion

3.2 Effect of events selection using importance sampling

As detecting all the possible water quality alteration events with a limited number of monitoring locations is not feasible, priority should be given to the detection of the potentially most impactful ones (Perelman and Ostfeld, 2010; Weickgenannt et al., 2010). However, most studies have validated the proposed monitoring layouts using either all or a uniform subset of the events considered possible (i.e., traditional validation) (Rathi and Gupta, 2015; Tinelli et al., 2017), biasing the results of their evaluations toward the large number of smaller events simulated. The change in performances between a traditional or a eIBES validation was evaluated by re-estimating the performances of state-of-the-art simulation-based monitoring location strategies, obtained on the benchmark DWDN using layouts composed of five monitoring locations obtained from Ostfeld et al. (2008) and Zhao et al. (2016). Results deriving from events simulated through eIBES were, then, compared to the original publications and among the different layouts (Figure 1). As eIBES tends to select water quality alteration events with higher hydraulic importance, a clear difference can be seen in the performances estimated using the two different methods. Indeed, “more important” events affect larger areas of the DWDN, leading to higher detLik values. Simultaneously, thanks to their spread, such events are characterized by a greater effect on drinking water’s consumers (i.e. higher eIBES validation consVol), even though they are detected more rapidly (i.e. lower traditional validation detTime) than the generic event simulated using a traditional validation. Noteworthy, this fact has operational implications. As the former performance metric is more related to the direct effect on human health of drinking water consumers than the latter, this result stresses the importance to validate monitoring layouts with respect to the potentially most impactful events. However, within eIBES results a high correlation is present between detTime and consVol (Spearman correlation; $\rho = 0.9$, p-value < 0.001), remarking the importance of, not only, a timely detection of the water quality alterations, but also of their management (Ostfeld et al., 2008).

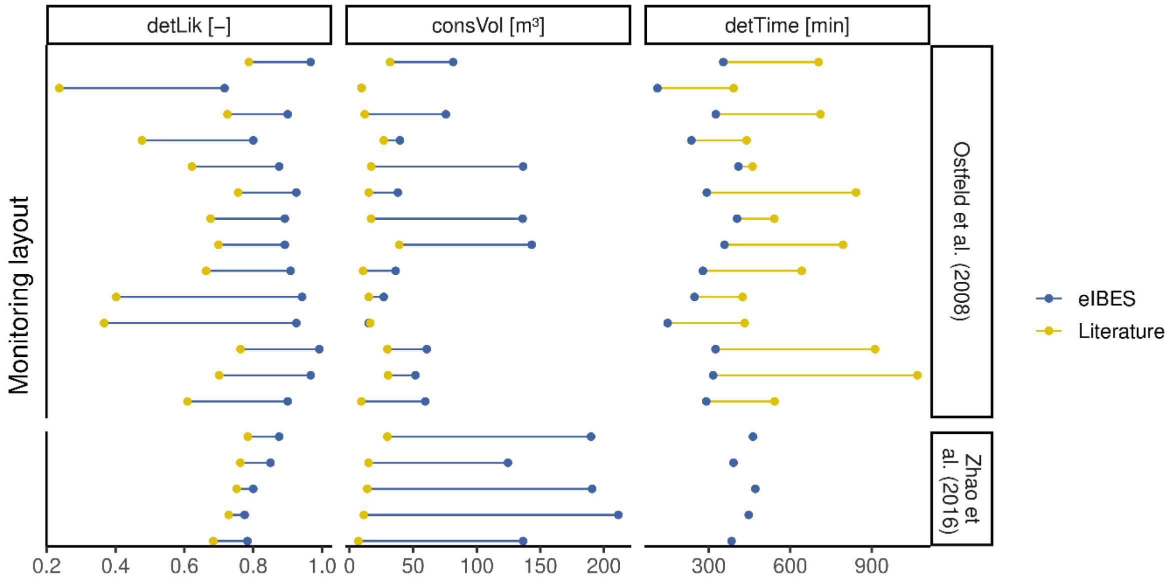

Figure 1. Performances of the monitoring layouts in Ostfeld et al. (2008) and Zhao et al. (2016) from the original literature study (marked as "Literature") and eIBES. No detTime values were provided by the layouts in Zhao et al. (2016). Markers indicate the performance obtained by the monitoring layouts through a literature or an eIBES validation, while horizontal lines highlight the difference between the two results, being colored as the validation obtaining higher values.

As it can be noted in Figure 1, the performance differences between the two validation sets varied greatly between the tested monitoring layouts, leading to different relative performances between them. In fact, the monitoring layouts presented by Zhao et al. (2016) provide strictly better performances than the ones provided by Ostfeld et al. (2008) with a traditional validation (Zhao et al., 2016), the opposite result was found when considering the impactful events selected by the eIBES validation, as shown in Figure 2: monitoring layouts from Ostfeld et al. (2008) present in general lower detTime and higher detLik than the ones provided by Zhao et al. (2016), indicating their Pareto-dominance. Furthermore, even within each single study, the relative performance of the different layouts differs between the traditional and the eIBES validation. In fact, limited (Spearman correlation; $\rho < 0.57$) or non-significant (p-values > 0.05) correlations were found between the different performance metrics (detLik, consVol, minTime) obtained using the traditional validation and the eIBES one. Such lack of correlation can lead the Pareto-optimal layouts previously identified within each study to become Pareto-dominated, as two of the solutions provided by Zhao et al. (2016) and highlighted in Figure 2. In addition, in contrast with the previous traditional validation by Perelman and Ostfeld (2013), the performances of several graph-based strategies presents lower detTime and similar detLik values than most simulation-based layouts. In details, while Perelman and Ostfeld (2013) reported the layouts of Krause et al. (2006) and Preis and Ostfeld (2006) as achieving more favorable performances than the graph-based methods tested, using the eIBES validation, these layouts resulted in higher detTime values than a large fraction of the graph-based ones (Figure 2). In addition, the detLik value achieved by Preis and Ostfeld (2006) resulted lower than most other graph-based layouts.

These results indicate how the optimization of the monitoring performances with respect to all the possible events does not necessarily result in enhanced protection against the highly impactful events. With respect to the "important" events, graph-based and current simulation-based strategies offer similar performances, supporting the application of graph-based methods to select monitoring layouts with only limited information, confirming the conclusion of Giudicianni et al. (2020). However, it needs to be noted that monitoring layouts with more favorable performances are likely to be expected using simulation-based methods developed specifically towards the detection of the "most important" events.

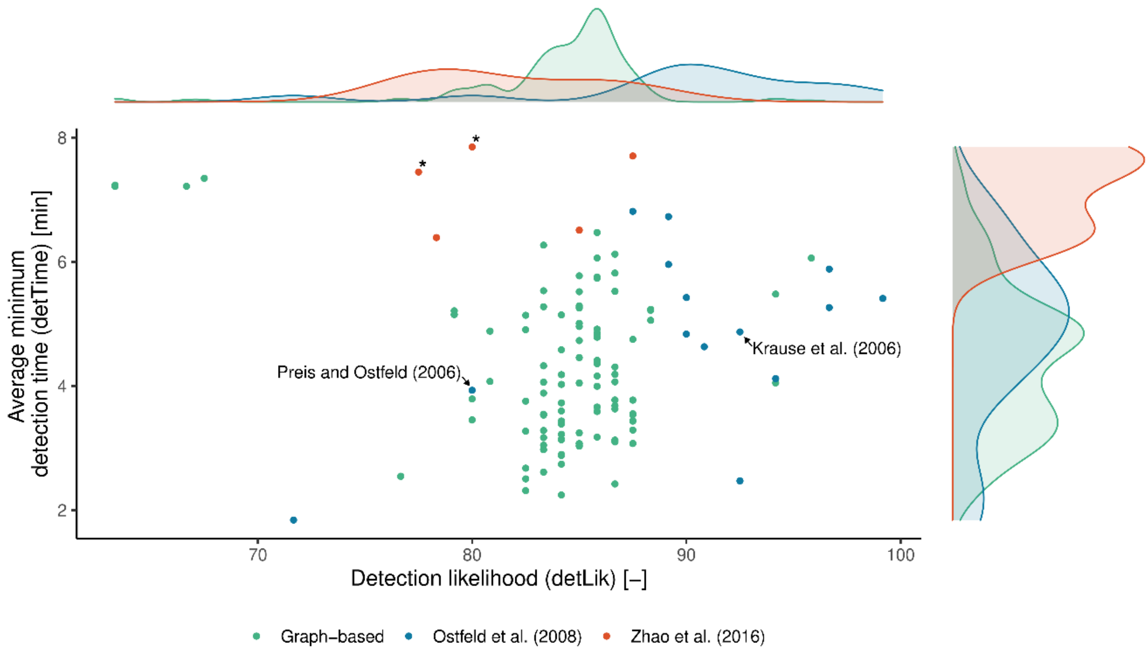

Figure 2. Detection likelihood (detLik) and average minimum detection time (detTime) of literature simulation-based and graph-based layouts. The density curves on the top and left of the main plot illustrate the probability density of the detTime and detLik for each group of layouts considered. Asterisks indicate the two monitoring layouts from Zhao et al. (2016) which result Pareto-dominated by the other layouts from the same study, while the references included in the plot highlight the performance of the two layouts discussed in the main text.

3.2 Effect of the inclusion of demand uncertainty on monitoring performances

The introduced eIBES validation fulfills two goals: (i) it selects primarily the most impactful events for the layout performance validation, (ii) it includes the uncertainty in water demand patterns, which can also potentially affect layouts performances (Shastri and Diwekar, 2006). The STL decomposition and synthetic pattern generation resulted in coefficients of variations of the water demand patterns ranging between 0.15 and 0.27. Such variability reflects on the flow conditions within the DWDNs, resulting also in the reversal of flow direction for more than 30 min in 13% and 28% of the pipes in case studies C and R, respectively (Figure 3). Noteworthy, such pipes are not distributed uniformly throughout the DWDNs, but they cluster in specific areas, where water flows regimes are highly variable among simulations likely due to the intersection of the water supply zones of different DWTPs. In fact, as shown by the higher fraction in case-study R, a higher number of DWTPs, pipes interconnections and also water demand patterns exacerbate the effect of water demand perturbations on water flows.

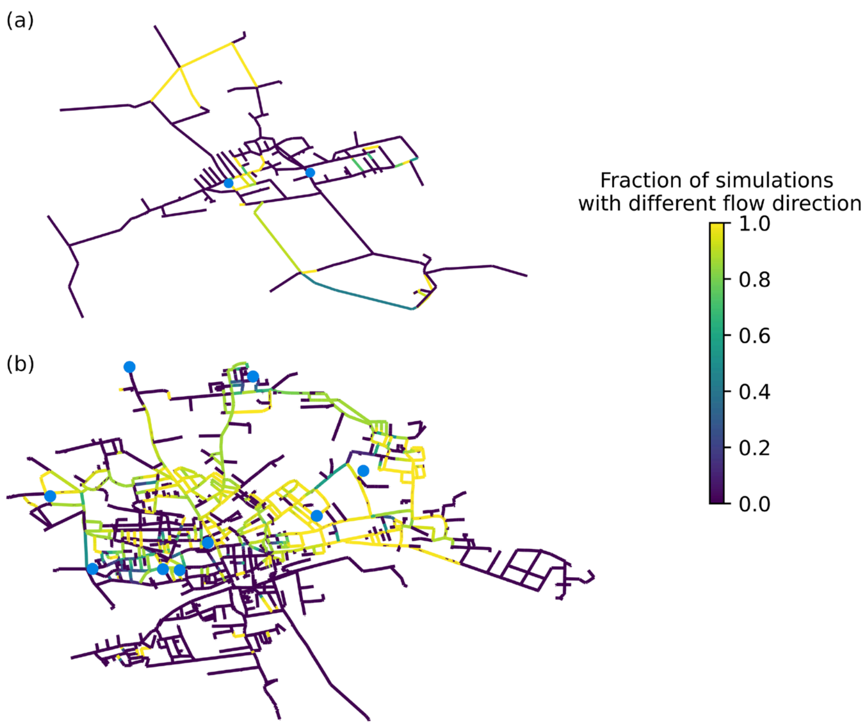

Figure 3. Fraction of simulation in which different flow directions were detected for at least 30 min for case studies C (a) and R (b). Blue markers indicate the presence of DWTPs.

Not accounting for the variability of flow regimes among the different simulations can provide non-representative performance estimates. In fact, the estimated detLik and detTime of the graph-based monitoring layouts differed between the use of stochastic patterns originated from the eIBES method and deterministic ones included in the models (i.e., neglecting the presence of stochasticity of water demands). This difference is shown in Figure 4 which displays the probability density of the difference between the performances obtained by the tested layouts including or neglecting the presence of stochasticity of water demands which were normalized by the 2.5^{th}-97.5^{th} percentiles range obtained with deterministic patterns for the performance metric and number of sensors considered. Values below zero indicate that larger values of the performance metrics were obtained using the deterministic patterns, while the opposite indicates that larger values were obtained in case of the inclusion of water demand stochasticity. The inclusion of water demand stochasticity resulted in detLik values showing systematically lower values in case of stochastic water demands (paired Wilcoxon test; p-values < 0.018; Figure 4a). This indicates that using a single set of water demand patterns leads to overestimate the confidence in the detection performance of the monitoring layouts. In addition, due to the concentration of flow differences in specific areas of the DWDNs, the detLik overestimation is not constant for all monitoring layouts, but it varies from case to case: while most of the monitoring layouts performances show normalized differences 25% in both case studies, as indicated by the higher probability densities in Figure 4a, in some cases the differences reach more than 50%, as occurring in case study C considering six or seven monitoring locations. As a result, only around 22% of the monitoring layouts with highest detLik are shared between the use of stochastic and deterministic water demand patterns. Moreover, comparing the probability densities of normalized difference in detLik values with different number of monitoring locations, it is possible to see how greater differences (in absolute terms) tend to be achieved in case of a larger number of monitoring locations. This fact indicates that while neglecting water demands stochasticity leads to an overestimation of detLik values regardless of the number of monitoring locations, this has stronger effects when deciding among monitoring layouts composed of a large number of monitoring locations. In practice, this neglection might lead to drastically different detLik estimates than what achievable in a real-world situation leading water utilities to select suboptimal monitoring layouts.

Differently from detLik, most of the variations in detTime showed values lower than the 10% range between the 2.5^{th} and 97.5^{th} percentiles of the range of the values obtained using the deterministic patterns (Figure 4b). Notably the two investigated DWDNs showed different results. Normalized detTime differences obtained in case study C indicates that higher detTime values were obtained for some monitoring layouts using deterministic water demands (except the layouts with three monitoring locations; paired Wilcoxon test; p-values = 0.255). Conversely, detTime variations in case study R are more centered around zero, with some cases (i.e., layouts with three and four monitoring locations) exhibiting higher detTime values in case of stochastic water demand patterns (paired Wilcoxon test; p-values < 0.002). Once again, this difference is likely to be caused by the distinct morphological and hydraulic characteristics of the two DWDNs. In any case, these results indicate a limited effect of water demands stochasticity on this performance metric.

Previous results by de Winter et al. (2019) also highlighted the effect of the inclusion of water demands stochasticity on water quality sensors placement, but outlined a relatively limited variation between the performance metrics using fixed and stochastic water demands. Even though the results are not directly comparable due to the different performance metric used, this disagreement is likely due to the differences between the methodologies used in the two studies. In fact, the results by de Winter et al. (2019) were obtained using static water demands and a skeletonized DWDN which simplify the water flow patterns within the DWDN and are not fully representative of real-world conditions. In addition, such study focuses on the detection of water quality alteration events originating from a limited number of locations within the DWDN, while in our case water quality alterations were allowed to occur potentially in all locations (e.g., due to back-siphonage or infiltrations; Robertson et al., 2003). These differences suggest how demand uncertainty seems to play a larger role when carrying out simulations with more realistic conditions (i.e., dynamic water demand patterns, models with more detailed DWDN structures).

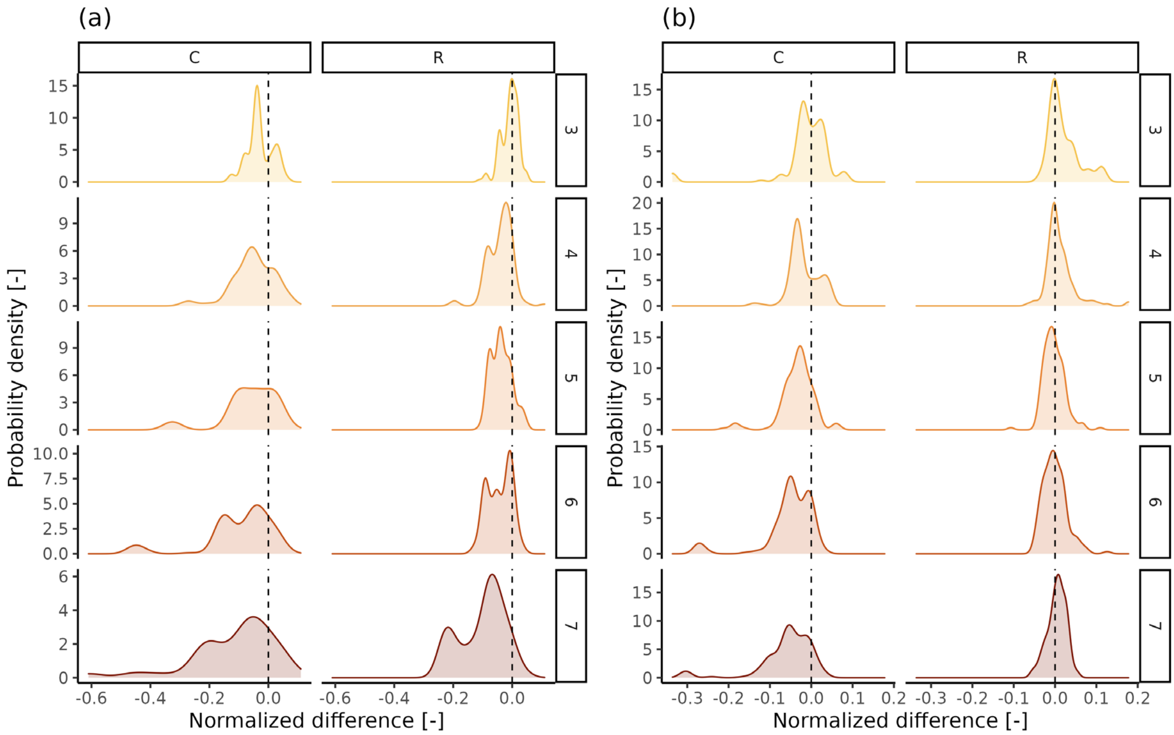

Figure 4. Probability density of the normalized difference of the values of detLik (a) and detTime (b) obtained using stochastic and standard water demand patterns in the two case studies (C, R) and at different number of monitoring locations (3 to 7) per monitoring layout.

3.3 Role of minimum detection concentration on monitoring performances

As highlighted by differences in detection performances in Figure 2 and in several other studies (e.g. Giudicianni et al., 2020; Ostfeld et al., 2008), the monitoring locations within a DWDN play a critical role in ensuring consumers' protection. However, other than the location in which DWDNs are monitored, the identification of water quality alterations depends also on the ability to distinguish them from baseline conditions. Such ability depends on both the precision of the instrument used for water monitoring, but also the actual baseline variability of the water quality within the DWDN (Ross et al., 2011): lower measuring precision and higher baseline variability will hamper the timely identification of water quality alterations. As a result, the same monitoring layout will likely achieve different performances with respect to different contaminants and/or using different monitoring technologies.

To evaluate this issue, we selected three graph-based strategies for the selection of monitoring locations in each case study, to evaluate the layouts performances under different scenarios. For case-study C, the selected strategies represent alternatives aimed at maximizing detLik (C1, based on D/L weighting, Spectral Random Walk clustering and Jordan Centrality), minimizing detTime (C3, based on D/L weighting, Edge Betweenness clustering and Closeness

Centrality), or offering a trade-off between the two alternatives (C2, based on L weighting, Edge Betweenness clustering and Betweeneess Centrality). Similarly in case-study R, the strategy R1 (based on D/L weighting, Edge Spectral clustering and Betweenness Centrality) was identified as maximizing detLik, while R3 (based on D weighting, Edge Betweenness clustering and Betweenness Centrality) as generically minimizing detTime and R2 (based on L weighting, Spectral Symmetric clustering and Closeness Centrality) as providing a trade-off between the two metrics. However, while in case of the simpler case-study C this characterization remains valid for all the numbers of the tested monitoring locations, for case-study R the characterization of R2 and R3 is not as clearly defined. Beside the ideal scenario where any non-zero contaminant concentrations are detected, the performances of the monitoring layouts obtained by these strategies were evaluated in three scenarios simulating an increasing $C_{X,min.det.eqiv.}$, as summarized in Table 2: they refer to the detection of a change in concentrations of, respectively, bacterial cells (Scenario I), total organic carbon (Scenario II) or ammonium (Scenario III), due to the intrusion of wastewater. As expected, higher $C_{X,min.det.eqiv.}$ resulted in lower detLik (Figure 5), especially in case-study R. In fact, the higher number of DWTPs and interconnections likely causes a greater dilution within the DWDN, hampering the detection of water quality alterations. Furthermore, detLik values in case-study C plateau after five monitoring locations and suggest how the addition of further monitoring locations would not further increase detLik values, contrarily to case study R where a constant increase is observed. Indeed, as shown in Table 1 and Figure 3, case-study R is approximately 10 times larger DWDN than case-study C, suggesting the need of a higher number of monitoring locations within the DWDN to effectively protect drinking water consumers in larger and more complex DWDNs.

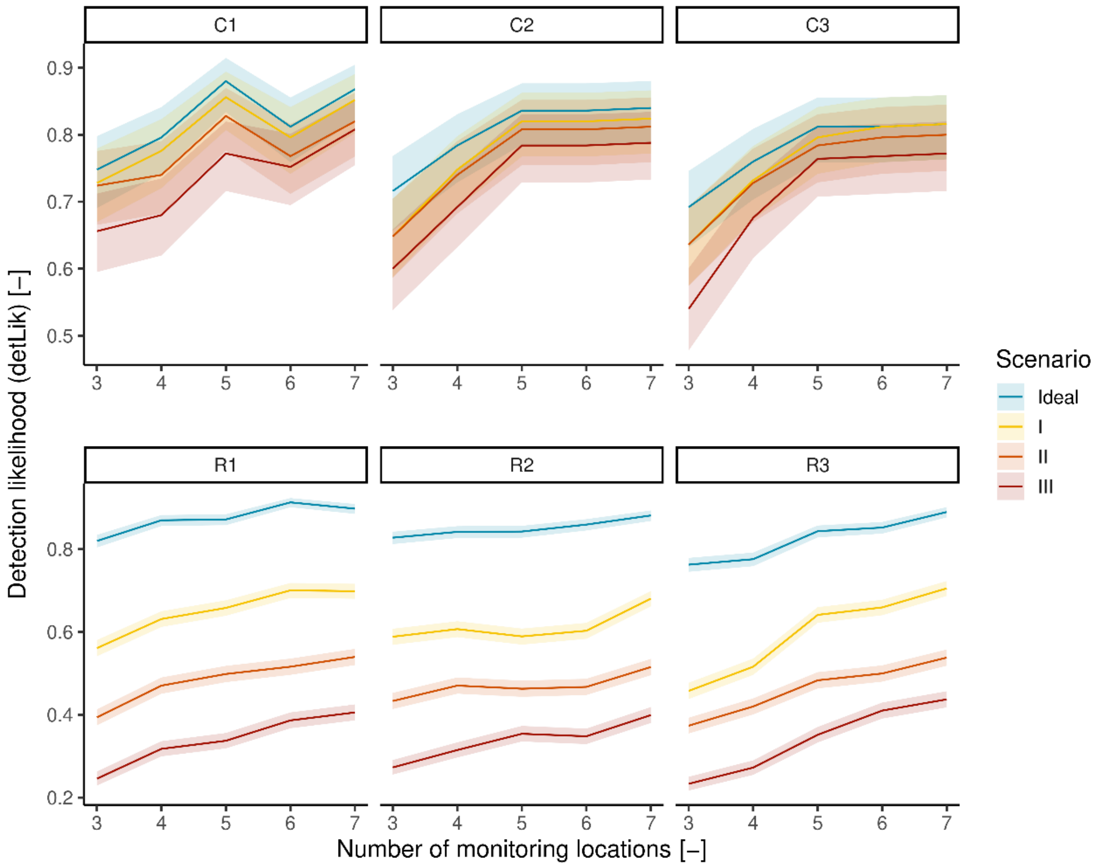

Figure 5. Detection likelihood (detLik) and its standard deviation in the different water quality alteration scenarios tested.

The effect of increasing $C_{X,min.det.eqiv.}$ has a less uniform effect on detTime, where the simulated scenarios present both lower and higher values than the ideal scenario (Figure 6). In case-study C the tested scenarios generically display detTime values greater than the ideal case, while most cases in case-study R show lower values and greater differences between the scenarios, likely due to the higher DWDN complexity. Actually, the values of detTime are calculated as the average detection time of the detected water quality alterations. In simpler DWDNs, as case-study C, limited dilution occurs, leading to the detection of a similar number of events across the scenarios, as shown in Figure 5. Compared to the ideal scenario, especially in Scenario III, certain events require longer times to expand with a sufficient concentration to be detected by the monitoring layouts, increasing the detTime values (Figure S1). On the contrary, in more complex DWDNs such as case-study R, events are either detected before excessive dilution or are not detected, especially with higher $C_{X,min.det.eqiv.}$, as shown by the negative correlation between the detTime deciles and the share of events detected both in the ideal and in the other scenarios (Figure S2).

detTime is a commonly used performance metric when evaluating DWDN monitoring layouts (Rathi et al., 2015). However, our results stress the need to include also other metrics, such as detLik, to properly contextualize the performances and account for the fact that different events are detected by distinct monitoring layouts and/or in different scenarios. Certain studies have addressed this issue by imposing a value equal to the simulation length in case of an undetected event (Watson et al., 2004). However, this solution introduces an arbitrary effect on detTime values, not allowing to compare the results of studies using different simulation lengths. One feasible alternative would be to estimate the harmonic mean of detTime values, considering undetected events with infinite detTime, as in fact not detected by the tested monitoring layout. Indeed, doing so results in higher detTime values (Figure S3) which, however, would provide a less explicit measure for a water utility operator.

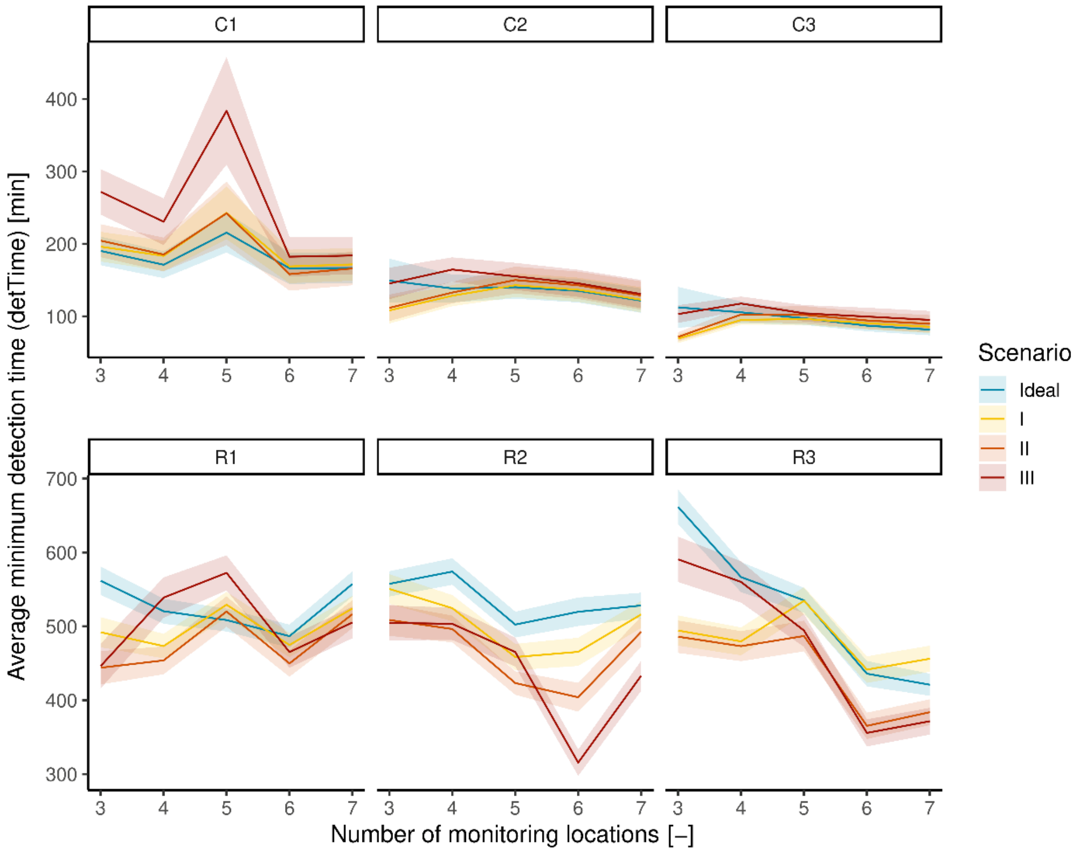

Figure 6. Average minimum detection time (detTime) and its standard deviation in the different water quality alteration scenarios tested.

Observing Figures 5 and 6, it is also possible to note how, on average, a higher number of monitoring locations improves all performance metrics, but this is not necessarily true for the single placement strategy, as for example moving from four to five monitoring locations using strategy C1. For this reason, similarly to DWDN partitioning (Khoa Bui et al., 2020), an optimal number of monitoring locations should be found, remarking the importance of an external validation using different performance metrics prior to the selection of sampling locations. In any case, the ability to compare different scenarios without excessive computational costs, like in the method proposed, allows water utilities to better characterize the performances expected from monitoring layouts in various conditions and can also aid in the selection of which monitoring technologies to deploy. In fact, following the proposed method, water utilities could compare easily the monitoring performances obtainable using different instruments, allowing to evaluate the different alternatives not only on the basis of the capital and operational costs, but also with regards to the protection of drinking water consumers' health.

While the previous results show the drop in performances which can be expected when monitoring different contaminants, different strategies can be used to preserve drinking water consumers health. In fact, even in case no measurements are specifically dedicated to the targeted contaminants, water quality alterations might interest multiple water quality parameters (Ikonen et al., 2017). As several commercial solutions for DWDN monitoring are able to measure multiple parameters at the same time, such signals could be combined to increase detection performances, even in case of unexpected water quality alterations (Perelman and Ostfeld, 2012). In addition, as water quality alterations might be detected in multiple monitoring locations, combining the result from these multiple sources and, possibly, comparing them with model-based predictions, can further aid events detection (Housh and Ohar, 2017; Oliker and Ostfeld, 2015).

3.4 Future developments

While the present work successfully expands the importance-based sampling methodology, evaluating its effects on the expected performance metrics, future work will be needed to further expand this methodology. In fact, since the current methodology to generate synthetic water demand patterns is entirely model-driven, it is possible that more systematic variations of such patterns (e.g., Alvisi et al., 2021; Brentan et al., 2017) are still missed. Future studies should include such variations within the eIBES methodology by including in the STL calculations water demand patterns estimated under different conditions (e.g., seasons), increasing the patterns variability. Furthermore, while the tested scenarios assume, in line with previous works (Giudicianni et al., 2020; Ostfeld et al., 2008; Tinelli et al., 2017; Weickgenannt et al., 2010), a constant water quality alteration for a given time, reduced

DWDN pressures, possibly leading to backflows, are often only temporary (i.e., under a minute), deriving from fluctuations in the operating conditions of the infrastructure (Teunis et al., 2010; USEPA, 2002). While flow directions are not expected to differ compared to the current approach, a higher degree of dilution might further reduce detLik values when considering realistic minimum detection concentrations, leading to consider the values estimated in this study as theoretical upper bounds. To achieve more realistic estimates of the performance metrics of the tested layouts, future studies should explicitly simulate water alterations temporal dynamics.

4. Conclusions

In this study a novel methodology, namely eIBES, for the validation of the performances of DWDN monitoring layouts was introduced in order to properly take into account water demand variability and, at the same time, focusing on the potentially most impactful water quality alteration events. The effect of both factors on the estimated validation performances was estimated using both real-world and benchmark DWDNs. Not taking into account water demand variability, while not affecting drastically minimum detection times, leads to an overestimation of the detection probabilities of the monitoring layouts. On the other hand, the results of this study highlight how Pareto-optimal performances with respect to all possible water quality alteration events does not necessarily protect drinking water consumers from the most impactful events. In fact, such events, even though are detected with higher frequency and with less delay, lead to a consumption of potentially contaminated water severely higher than a generic events, highlighting the importance of their detection. Finally, the effect of different minimum detection concentrations of several monitoring layouts was tested displaying its negative effect on detection likelihood, allowing to identify events closer to their occurrence and before their excessive dilution.

Author contributions

MG: Conceptualization, Methodology, Formal analysis, Writing - Original Draft, Visualization. **AL**: Formal analysis, Investigation. **LDD**: Formal analysis, Investigation. **MA**: Writing - Review & Editing, Supervision

Conflict of interest statement

There are no conflicts of interest to declare.

Acknowledgments

The Authors would like to thank CAP Holding S.p.A for funding the Ph.D. scholarship of M.G. and providing the models for the examined case-studies, and Gloria Salinetti, Cristina Floris and Ambra Lombardo for their helpful cooperation.

References

Alvisi, S., Franchini, M., Luciani, C., Marzola, I., Mazzoni, F., 2021. Effects of the COVID-19 Lockdown on Water Consumptions: Northern Italy Case Study. J. Water Resour. Plann. Manage. 147, 05021021. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001481

Besmer, M.D., Sigrist, J.A., Props, R., Buysschaert, B., Mao, G., Boon, N., Hammes, F., 2017. Laboratory-Scale Simulation and Real-Time Tracking of a Microbial Contamination Event and Subsequent Shock-Chlorination in Drinking Water. Front. Microbiol. 8, 1900. https://doi.org/10.3389/fmicb.2017.01900

Brentan, B.M., Luvizotto Jr., E., Herrera, M., Izquierdo, J., Pérez-García, R., 2017. Hybrid regression model for near real-time urban water demand forecasting. Journal of Computational and Applied Mathematics 309, 532–541. https://doi.org/10.1016/j.cam.2016.02.009

Cleveland, R.B., Cleveland, W.S., Terpenning, I., 1990. STL: A Seasonal-Trend Decomposition Procedure Based on Loess. Journal of Official Statistics 6, 3–33.

Csardi, G., Nepusz, T., 2006. The igraph software package for complex network research. InterJournal, complex systems 1695, 1–9.

da Luz, N., Kumpel, E., 2020. Evaluating the impact of sampling design on drinking water quality monitoring program outcomes. Water Research 185, 116217. https://doi.org/10.1016/j.watres.2020.116217

de Winter, C., Palleti, V.R., Worm, D., Kooij, R., 2019. Optimal placement of imperfect water quality sensors in water distribution networks. Computers & Chemical Engineering 121, 200–211. https://doi.org/10.1016/j.compchemeng.2018.10.021

Di Nardo, A., Di Natale, M., Giudicianni, C., Greco, R., Santonastaso, G.F., 2017. Weighted spectral clustering for water distribution network partitioning. Appl Netw Sci 2, 19. https://doi.org/10.1007/s41109-017-0033-4

Gabrielli, M., Pulcini, F., Barbesti, G., Antonelli, M., 2023. Source to taps investigation of natural organic matter in non-disinfected drinking water distribution systems. ChemRix. https://doi.org/10.26434/chemrxiv-2023-9w64x

Gabrielli, M., Turolla, A., Antonelli, M., 2021. Bacterial dynamics in drinking water distribution systems and flow cytometry monitoring scheme optimization. Journal of Environmental Management 286, 112151. https://doi.org/10.1016/j.jenvman.2021.112151

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. J. Water Resour. Plann. Manage. 146, 04020030. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001210

Gong, J., Guo, X., Yan, X., Hu, C., 2023. Review of Urban Drinking Water Contamination Source Identification Methods. Energies 16, 705. https://doi.org/10.3390/en16020705

Hagberg, A.A., Schult, D.A., Swart, P.J., 2008. Exploring Network Structure, Dynamics, and Function using NetworkX, in: Varoquaux, G., Vaught, T., Millman, J. (Eds.), Proceedings of the 7th Python in Science Conference. Pasadena, CA USA, pp. 11–15.

Henze, M., van Loosdrecht, M.C.M., Ekama, G.A., Brdjanovic, D., 2008. Biological Wastewater Treatment: Principles, Modelling and Design. IWA Publishing. https://doi.org/10.2166/9781780401867

Housh, M., Ohar, Z., 2017. Integrating physically based simulators with Event Detection Systems: Multi-site detection approach. Water Research 110, 180–191. https://doi.org/10.1016/j.watres.2016.12.003

Hu, C., Li, M., Zeng, D., Guo, S., 2018. A survey on sensor placement for contamination detection in water distribution systems. Wireless Netw 24, 647–661. https://doi.org/10.1007/s11276-016-1358-0

Ikonen, J., Pitkänen, T., Kosse, P., Ciszek, R., Kolehmainen, M., Miettinen, I.T., 2017. On-line detection of Escherichia coli intrusion in a pilot-scale drinking water distribution system. Journal of Environmental Management 198, 384–392. https://doi.org/10.1016/j.jenvman.2017.04.090

Khoa Bui, X., S. Marlim, M., Kang, D., 2020. Water Network Partitioning into District Metered Areas: A State-Of-The-Art Review. Water 12, 1002. https://doi.org/10.3390/w12041002

Klise, K.A., Bynum, M., Moriarty, D., Murray, R., 2017. A software framework for assessing the resilience of drinking water systems to disasters with an example earthquake case study. Environmental Modelling & Software 95, 420–431. https://doi.org/10.1016/j.envsoft.2017.06.022

Krause, A., Leskovec, J., Guestrin, C., VanBriesen, J., Faloutsos, C., 2008. Efficient Sensor Placement Optimization for Securing Large Water Distribution Networks. J. Water Resour. Plann. Manage. 134, 516–526. https://doi.org/10.1061/(ASCE)0733-9496(2008)134:6(516)

Krause, A., Leskovec, J., Isovitsch, S., Xu, J., Guestrin, C., VanBriesen, J., Small, M., Fischbeck, P., 2006. Optimizing Sensor Placements in Water Distribution Systems Using Submodular Function Maximization, in: Water Distribution Systems Analysis Symposium 2006. Presented at the Eighth Annual Water Distribution Systems Analysis Symposium (WDSA), American Society of Civil Engineers, Cincinnati, Ohio, United States, pp. 1–17. https://doi.org/10.1061/40941(247)109

Mu, T., Huang, M., Tan, H., Chen, G., Zhang, R., 2021. Pressure and Water Quality Integrated Sensor Placement Considering Leakage and Contamination Intrusion within Water Distribution Systems. ACS EST Water 1, 2348–2358. https://doi.org/10.1021/acsestwater.1c00209

Ohar, Z., Lahav, O., Ostfeld, A., 2015. Optimal sensor placement for detecting organophosphate intrusions into water distribution systems. Water Research 73, 193–203. https://doi.org/10.1016/j.watres.2015.01.024

Oliker, N., Ostfeld, A., 2015. Network hydraulics inclusion in water quality event detection using multiple sensor stations data. Water Research 80, 47–58. https://doi.org/10.1016/j.watres.2015.04.036

Ostfeld, A., Uber, J.G., Salomons, E., Berry, J.W., Hart, W.E., Phillips, C.A., Watson, J.-P., Dorini, G., Jonkergouw, P., Kapelan, Z., di Pierro, F., Khu, S.-T., Savic, D., Eliades, D., Polycarpou, M., Ghimire, S.R., Barkdoll, B.D., Gueli, R., Huang, J.J., McBean, E.A., James, W., Krause, A., Leskovec, J., Isovitsch, S., Xu, J., Guestrin, C., VanBriesen, J., Small, M., Fischbeck, P., Preis, A., Propato, M., Piller, O., Trachtman, G.B., Wu, Z.Y., Walski, T., 2008. The Battle of the Water Sensor Networks (BWSN): A Design Challenge for Engineers and Algorithms. J. Water Resour. Plann. Manage. 134, 556–568. https://doi.org/10.1061/(ASCE)0733-9496(2008)134:6(556)

Perelman, L., Ostfeld, A., 2013. Application of Graph Theory to Sensor Placement in Water Distribution Systems, in: World Environmental and Water Resources Congress 2013. Presented at the World Environmental and Water Resources Congress 2013, American Society of Civil Engineers, Cincinnati, Ohio, pp. 617–625. https://doi.org/10.1061/9780784412947.060

Perelman, L., Ostfeld, A., 2012. Extreme Impact Contamination Events Sampling for Real-Sized Water Distribution Systems. J. Water Resour. Plann. Manage. 138, 581–585. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000206

Perelman, L., Ostfeld, A., 2010. Extreme Impact Contamination Events Sampling for Water Distribution Systems Security. J. Water Resour. Plann. Manage. 136, 80–87. https://doi.org/10.1061/(ASCE)0733-9496(2010)136:1(80)

Preis, A., Ostfeld, A., 2006. Multiobjective Sensor Design for Water Distribution Systems Security, in: Water Distribution Systems Analysis Symposium 2006. Presented at the Eighth Annual Water Distribution Systems Analysis Symposium (WDSA), American Society of Civil Engineers, Cincinnati, Ohio, United States, pp. 1–17. https://doi.org/10.1061/40941(247)107

Rathi, S., Gupta, R., 2015. A critical review of sensor location methods for contamination detection in water distribution networks. Water Quality Research Journal 50, 95–108. https://doi.org/10.2166/wqrjc.2014.011

Rathi, S., Gupta, R., Ormsbee, L., 2015. A review of sensor placement objective metrics for contamination detection in water distribution networks. Water Supply 15, 898–917. https://doi.org/10.2166/ws.2015.077

Robertson, W., Stanfield, G., Howard, G., Bartram, J., 2003. Monitoring the quality of drinking water during storage and distribution, in: Assessing Microbial Safety of Drinking Water. Improving Approaches and Methods. IWA Publishing, pp. 179–204.

Ross, G.J., Tasoulis, D.K., Adams, N.M., 2011. Nonparametric Monitoring of Data Streams for Changes in Location and Scale. Technometrics 53, 379–389. https://doi.org/10.1198/TECH.2011.10069

Shastri, Y., Diwekar, U., 2006. Sensor Placement in Water Networks: A Stochastic Programming Approach. J. Water Resour. Plann. Manage. 132, 192–203. https://doi.org/10.1061/(ASCE)0733-9496(2006)132:3(192)

Teunis, P.F.M., Xu, M., Fleming, K.K., Yang, J., Moe, C.L., LeChevallier, M.W., 2010. Enteric Virus Infection Risk from Intrusion of Sewage into a Drinking Water Distribution Network. Environ. Sci. Technol. 44, 8561–8566. https://doi.org/10.1021/es101266k

Tinelli, S., Creaco, E., Ciaponi, C., 2017. Sampling Significant Contamination Events for Optimal Sensor Placement in Water Distribution Systems. J. Water Resour. Plann. Manage. 143, 04017058. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000814

USEPA, 2002. Potential Contamination Due to Cross-Connections and Backflow and the Associated Health Risks.

Watson, J.-P., Greenberg, H.J., Hart, W.E., 2004. A Multiple-Objective Analysis of Sensor Placement Optimization in Water Networks, in: Critical Transitions in Water and Environmental Resources Management. Presented at the World Water and Environmental Resources Congress 2004, American Society of Civil Engineers, Salt Lake City, Utah, United States, pp. 1–10. https://doi.org/10.1061/40737(2004)456

Weickgenannt, M., Kapelan, Z., Blokker, M., Savic, D.A., 2010. Risk-Based Sensor Placement for Contaminant Detection in Water Distribution Systems. J. Water Resour. Plann. Manage. 136, 629–636. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000073

World Health Organization, 2017. Guidelines for drinking-water quality: fourth edition incorporating first addendum, 4th ed + 1st add. ed. World Health Organization, Geneva.

Zhao, Y., Schwartz, R., Salomons, E., Ostfeld, A., Poor, H.V., 2016. New formulation and optimization methods for water sensor placement. Environmental Modelling & Software 76, 128–136. https://doi.org/10.1016/j.envsoft.2015.10.030

Zheng, F., Du, J., Diao, K., Zhang, T., Yu, T., Shao, Y., 2018. Investigating Effectiveness of Sensor Placement Strategies in Contamination Detection within Water Distribution Systems. J. Water Resour. Plann. Manage. 144, 06018003. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000919

Zulkifli, S.N., Rahim, H.A., Lau, W.-J., 2018. Detection of contaminants in water supply: A review on state-of-the-art monitoring technologies and their applications. Sensors and Actuators B: Chemical 255, 2657–2689. https://doi.org/10.1016/j.snb.2017.09.078