Spatio-Temporal Design for a Water Quality Monitoring Network Maximizing the Economic Value of Information to optimize the detection of accidental pollution

Destandau François, Zaiter Youssef

Spatio-Temporal Design for a Water Quality Monitoring Network

Maximizing the Economic Value of Information to optimize the detection of accidental pollution

DESTANDAU François

ZAITER Youssef

Abstract

The reduction of damage due to water pollution requires good knowledge of the quality of surface waters. The Water Quality Monitoring Networks (WQMNs) have evolved over time according to the objectives of each one of them: knowledge of long-term quality evolution, search for the origin of pollution, detection of accidental pollution, etc. Information provided by WQMNs could be improved by a spatial approach, optimizing the location or the number of monitoring stations, or by a temporal approach, optimizing the sampling frequency. However, there is a cost for monitoring water quality.

In this article, we show, for the first time, how the estimation of the Economic Value of Information (EVOI) can be used to determine the spatio-temporal design of the network. With the example of a network that aims to detect accidental pollution, we show how to calculate the EVOI according to the spatial and temporal network design (number and location of stations, temporal accuracy of measurement) and how to define this design by maximizing the EVOI. This will allow us to answer questions such as: Are the expenses invested in the networks justified? With an additional budget, is it better to add a station or to increase the temporal accuracy of the measurement of existing stations? What is the optimal spatial and temporal design of the network when working with a fixed budget?

Keywords Water Resource Management; Water Quality Monitoring Network; Economic Value of Information.

1. Introduction

A Water Quality Monitoring Networks (WQMN) is defined as the acquisition of information regarding chemical, biological, and physical characteristics of a water body over time and space [1]. WQMNs were developed later to answer different types of questions such as pollution trend, pollution source, and policies to be implemented to improve water quality. In 2000, Article 8 of the WFD [2] reaffirmed the need to "…ensure the establishment of programs for the monitoring of water status in order to establish a coherent and comprehensive overview of water." It imposes two types of networks: the Surveillance Control Network, a permanent monitoring network that aims to monitor the evolution of pollution over the long-term; and the Operational Control Network, a temporary monitoring network that aims to identify the pollution source and to make sure that the measures taken make it possible to achieve good status. However, other networks exist locally for other purposes, e.g., identification of seasonal variations in pollution or detection of accidental pollution. This article focuses on this last objective.

WQMNs have been the research topic for several studies. Part of these studies focused on the physical optimization of the WQMN, while the other part focused on the Economic Value of Information (EVOI) provided by the WQMN.

The first type focused on optimizing the network by taking hydrological considerations into account, with the aim of minimizing imprecisions regarding water quality monitoring. The optimization can be spatial and/or temporal. On the one hand, spatial issues comprise the optimization of the location and the number of monitoring stations. Alvarez-Vázquez et al. [3] determined the location of the monitoring stations that decreased the average deviation for the water pollution index. Telci et al. [4] focused on determining the location of the monitoring network that minimized the detection time for accidental pollution. For Destandau and Point [5], the WQMN aimed at the identification of blackspots. Finally, Park et al. [6] used a genetic algorithm to resolve a multi-objective network including: representativeness of the river system, compliance with water quality standards, supervision of water use, surveillance of pollution sources and examination of water quality changes. On the other hand, temporal issues deal with the optimization of the sampling frequency. These articles aim to minimize frequency by eliminating redundant information. In Liu et al. [7], the objective of the network was to determine the average annual pollution. In Naddeo et al. [8], the objective was to determine the pollution trend: upward, downward or no trend. In Kim et al. [9], the objective was to determine the seasonal variation of water quality. Finally, Zhou [10] implemented a multi-objective model with these three objectives: trend detection, determination of periodic fluctuation and estimation of mean values of the stationary component. More recently, Pourshahabi et al. [11] tried to find the optimal spatio-temporal design of a WQMN in the Karkheh reservoir. They presented a multi-criteria methodology based on maximizing a statistical value of information, minimizing the number of monitoring stations and the sampling frequency, and eliminating redundant information.

Other studies have focused on determining the EVOI provided by a WQMN. The EVOI analysis is based on the Bayesian method. The Bayesian method makes it possible to revise knowledge about the states of nature, more precisely, the probability of being in a state of nature, on the basis of initial beliefs and new information obtained from experimental data. Several authors have used this concept to determine the value of information provided by the WQMN. Bouma et al. [12] tried to find the best policy to manage eutrophication. They used satellite observation as additional information. Bouma et al. [13] used this concept to try to choose between a spatialized or uniform policy to protect the Great Barrier Reef. Destandau and Diop [14] identified the parameters that have an impact on the EVOI: prior probabilities on states of nature, costs linked to a bad decision and the accuracy of additional information. They constructed a theoretical model to study the impact of these parameters on the EVOI.

In this article, we combine the literature on physical network design optimization and EVOI in order to determine the spatio-temporal design of the WQMN that maximizes the EVOI. To our knowledge, this has never been done before. In the first case, the optimization of the network design was solely based on hydrological considerations, without regard for the economic benefit (or EVOI) generated by these networks. In the second case, the literature focused on the economic benefit generated by a network with a predefined design. Determining the design of a water quality monitoring network by maximizing the EVOI helps the system manager to answer questions such as: Is the money invested in the network justified by generating benefits in excess of costs? Is it better to increase the number of stations or the temporal accuracy of the measurement of existing stations? What is the optimal spatial and temporal design of the network when working with a fixed budget? In our article, we illustrate the method by which these questions can be answered by a network whose objective is to detect potential accidental pollution on a stretch of river. Two scenarios are studied: the case where the accidental pollution generates uniform damage along the river, and the case where the damage is decreasing from upstream to downstream.

In Section 2, we present the theoretical framework, the geographic context and our hypothesis. In Section 3 we define the optimal locations of the monitoring stations for our two damage scenarios. In Section 4, we analyse the tradeoff between spatial and temporal dimensions of the WQMN design. The Section 5 includes a discussion and concludes the paper.

2. Methods

2.1. EVOI and the Bayesian Method

The benefit generated by monitoring networks can be estimated by the EVOI ([12], [13], [14]). The EVOI is determined by implementing the Bayesian decision theory. Bayes' theorem is used in the field of decision-making under uncertainty where the Decision Maker (DM) updates his prior beliefs by acquiring new information. When a DM faces uncertainty, his decision will be based on the highest expected outcome, the expected value for his utility. This expected utility depends on the prior probabilities of the states of nature that make it possible to reach a certain outcome. If the DM is uncertain about the states of nature, he will try to obtain additional information on the likelihood of the potential states of nature. When the DM obtains new information on the states of nature, he uses this information to update his prior beliefs by implementing Bayes' theorem. Hence, the Bayesian method relies on the fact that additional information makes it possible to update the beliefs concerning the states of nature and, therefore, to make more appropriate decisions. The EVOI is calculated as the difference between utilities of the decisions after acquiring new information and decisions that could have been taken without additional information.

To each state of nature corresponds a preferable action.Thus, by considering K states of nature and K possible actions, the EVOI can be written as follows:

$$EVOI = \sum_{k=1}^{K} P(i_k).\left[\sum_{k'=1}^{K} P(s_{k'}/i_k).U_{a_k/s_{k'}}\right] \tag{1}$$

Where:

$P(i_k)$: Probability that the network indicates the state of nature s_k

$P(s_{k'}/i_k)$: Probability of being in the state of nature $s_{k'}$ when the network indicates i_k

$U_{a_k/s_{k'}}$: Utility of the action a_k (more appropriate for the state of nature *k)* when the state of nature is $s_{k'}$

In the context of the EVOI generated by a water quality monitoring network, the states of nature will depend on the objective of the network. The Water Framework Directive defines two types of networks: the Surveillance Control Network, which aims to monitor the evolution of water quality in the long term, and the Operational Control Network, which aims to understand why a water body does not reach good status and to act accordingly. In the first case, the states of nature may be the "improving", "stable" or "deteriorating" water quality; in the second case, it could be the origin of the pollution, i.e., "urban" or "agricultural". This information will help to better target policies. However, there are not only WFD networks but, in addition, other local networks with other objectives such as the identification of seasonal variations in quality [9], environmental vulnerability [14] or the detection of accidental pollution [4]. It is the latter case that interests us in this article. The states of nature inform us as to whether or not there is accidental pollution.

By applying Bayes' theorem, the EVOI (1) becomes:

$$EVOI = \sum_{k=1}^{K} P(i_k).\left[\sum_{k'=1}^{K'} \frac{P(s_{k'}).P(i_k/s_{k'})}{P(i_k)}.U_{a_{k/k'}}\right]$$

$$\Leftrightarrow$$

$$EVOI = \sum_{k=1}^{K}\sum_{k'=1}^{K} P(s_{k'}).\ P(i_k/s_{k'}).U_{a_{k/k'}} \tag{2}$$

To determine the EVOI, the network manager must define the *a priori* probability of being in each of the states of nature, the probability that his network will give him a good or wrong message, and the economic consequences of his decisions according to the states of nature.

2.2. Geographic context

In order to study a monitoring network aimed at detecting accidental pollution, we assume a stretch of river represented by a segment $[0,1]$, the potential target of this type of pollution and whose vulnerability[1] makes this detection a significant environmental issue. The location $l = 0$ is the most upstream point and location $l = 1$ is the most downstream point of the river (outlet).

We assume that accidental pollution could be emitted at any point in the stream with probability P. Accidents can be of various types: overflow from a waste water treatment plant in the case of an exceptional rain event, or an industrial accident as in the case of the chemical company Sandoz in Switzerland in 1986 where a fire caused the discharge of polluted water into the Rhine. There may, however, be other situations, e.g., where unidentified soil pollution may persist on industrial sites that have been closed for a long time and that may end up years later in a river after rainfall events and soil erosion. One example is JEC Industrie, a company specializing in the manufacture of metal furniture, which operated tetrachloroethylene metal degreasing installations in the Rhône department of France until 2004. However, it wasn't until 2012 that tetrachloroethylene was detected in the Saône.

This pollution generates an environmental damage D. We note $D_{l_x,out}$ the damage between the location of the emission l_x and the outlet of the river. We assume that downstream of our stretch of river, the water has a flow, purifying capacity or water use that makes the pollution non-damaging.

However, if a monitoring station is located in l_y, downstream from the location of the emission, the station could detect the accidental pollution. In that case, an action a at a cost C will be implemented to stop environmental damage. Environmental damage then becomes D_{l_x,l_y}. The effect of the action is therefore immediate but it cannot act retrospectively on the damage generated between the emission and the pollution detection. Thus, if pollution is detected at the outlet of the river, it will be too late to implement an action and stop environmental damage. We can mention, for example, the Huningue station in France. When pollution is detected, canals are cut off to prevent the pollution from flowing

[1] A stretch of river may be more or less vulnerable depending on hydrological (concentration of pollution according to flow), chemical (capacity for self-purification of the environment) or economic characteristics according to water use (an exceptional fish population, water used for swimming or to produce drinking water will require a higher quality).

into the groundwater or into canal systems that feed rivers in downstream Alsace. We can also cite the example of pollution by tetrachloroethylene in the Saône in 2013, mentioned above, where the immediate action was to close drinking water supply wells to prevent drinking water from being contaminated.

Nevertheless, the monitoring at each station is obviously not continuous. A probability α exists such that the monitoring stations may not detect pollution. This allows the introduction of the temporal issue in the design of the monitoring network. Obviously, this probability decreases when the sampling frequency increases[2]. Moreover, we assume that the probability α is the same for all stations of the network and that there is independence between the events. Specifically, the probability for detection/non-detection at a station does not depend on the number of stations that do not detect this pollution upstream.

2.3. EVOI in our geographical context

In our model, we assume that the WQMN comprises n monitoring stations. Hence, with n monitoring stations, we have $(n+2)$ states of nature depending on the location of the emission of the accidental pollution. If $l_{y/n}$ is the location of the y^{th} monitoring station for n monitoring stations, then $s_{[l_{(y-1)/n}, l_{y/n}]}$ is the state of nature where accidental pollution is emitted between $l_{(y-1)/n},\ l_{y/n}$, $s_{[l_{n/n}, out]}$ is the state of nature where accidental pollution is emitted between the last station and the outlet of the stream, and $s_{\emptyset}$ is the state of nature where no accidental pollution is emitted. The states of nature are then:

$$\left\{ s_{[0,l_{1/n}]}, s_{[l_{1/n},l_{2/n}]}, \dots, s_{[l_{(y-1)/n},l_{y/n}]}, \dots, s_{[l_{n/n},out]}, s_{\emptyset} \right\}$$

The network can deliver $(n+2)$ messages as well: pollution is detected by the y^{th} monitoring station, $i_{l_{y/n}}$, accidental pollution is detected at the outlet of the stream, i_{out}, and no pollution is detected, $i_{\emptyset}$:

$$\left\{ i_{l_{1/n}}, i_{l_{2/n}}, \dots, i_{l_{y/n}}, \dots, i_{out}, i_{\emptyset} \right\}$$

The probability of detecting accidental pollution depends on the state of nature. On the basis of our hypotheses, this can be written as follows:

$$\begin{gathered} P(i_{l_{y/n}} / s_{l_{(y-1)/n}, l_{y/n}}) = (1-\alpha) \\ P(i_{(l_{y+1})/n} / s_{l_{(y-1)/n}, l_{y/n}}) = \alpha.(1-\alpha) \\ P(i_{(l_{y+z})/n} / s_{l_{(y-1)/n}, l_{y/n}}) = \alpha^{z}.(1-\alpha) \end{gathered} \tag{3}$$

Obviously, if any pollution is emitted downstream of the station, or if no pollution is emitted, the probability of detection is null.

[2] To link the alpha probability to a number of measurements per year, it would be necessary to perform a non-parametric analysis based on data. This considerable task, which could itself be the subject of an article, is not useful for the purpose of our article.

We can deduce the probability that the network delivers the message $i_{l_{y/n}}$:

$$P(i_{l_{y/n}}) = (1-\alpha).P(s_{[l_{(y-1)/n},l_{y/n}]}) + \alpha.(1-\alpha).P(s_{[l_{(y-2)/n},l_{(y-1)/n}]}) + \cdots + \alpha^{(y-1)}.(1-\alpha).P(s_{[0,l_{1/n}]})$$

$$\Leftrightarrow P(i_{l_{y/n}}) = P.[(1-\alpha).(l_{(y-1)/n}, l_{y/n}) + \alpha.(1-\alpha).(\, l_{(y-2)/n}, l_{(y-1)/n}) + \cdots \alpha^{(y-1)}.l_{1/n}]$$

$$\Leftrightarrow P\left(i_{l_{y/n}}\right) = P.(1-\alpha)\sum_{j=0}^{y} \alpha^{j}\ .[l_{(y-j)/n} - l_{[(y-1)-j]/n}]$$

In Equation (2), utility "$U_{a_{k/k'}}$" is evaluated based on the status-quo ("no action is taken" in the example of detection of accidental pollution). We compute the utility as the difference between environmental damage saved thanks to the action taken following the detection of pollution: i_k, called D_{i_k}, and the cost C of implementing the action in the event of detection. Hence, the EVOI from Equation (2) becomes:

$$EVOI = \sum_{k=1}^{K} \sum_{k\prime=1}^{K} P(s_{k'}).\ P(i_k/s_{k'}).\left[D_{i_k} - C\right] \tag{4}$$

On the basis of our hypotheses of uniformity of accidental pollution distribution (the same probability of accidental pollution for areas as well as for lengths) and Equation (3), the EVOI (4) for n stations and a probability of non-detection α becomes:

$$EVOI_{n,\alpha} = P.(1-\alpha).\sum_{y=1}^{n} \left[D_{l_{y/n,out}} - C\right].\sum_{y'=1}^{y} \left[l_{y'/n} - l_{(y'-1)/n}\right].\alpha^{y-y'} \tag{5}$$

It appears from Equation (5) that the EVOI increases when the probability P and damage $D_{l_{y/n,out}}$ increase, and when the cost of action C decreases.

2.4. Network cost and net benefit

We assume a monitoring cost Θ_α for each station, given by the following function:

$$\Theta_\alpha = \frac{\lambda}{\alpha} \tag{6}$$

λ expresses the slope of the function. An increase in the value of this parameter means that it is more expensive to reduce α. The monitoring cost of a WQMN with n stations and a probability of non-detection α is:

$$\Theta_{n,\alpha} = n\left(\frac{\lambda}{\alpha}\right) \tag{7}$$

We compute the net benefit of monitoring $\pi_{n,\alpha}$ as the difference between the Economic Value of Information and the monitoring cost:

$$\pi_{n,\alpha} = EVOI_{n,\alpha} - \Theta_{n,\alpha}$$

With (5) and (7)

$$\pi_{n,\alpha} = P.(1-\alpha).\sum_{y=1}^{n}\left[D_{l_{y/n,out}} - C\right].\sum_{y'=1}^{y}\left[l_{y'/n} - l_{(y'-1)/n}\right].\alpha^{y-y'} - n\left(\frac{\lambda}{\alpha}\right) \tag{8}$$

2.5. Optimal spatio-temporal combination with a fixed budget

In this section, we suppose that the network manager has a fixed budget to establish the monitoring network. In that case, a combination of the optimal number of monitoring stations (with optimized locations) and the optimal probability of detection has to be chosen. The optimal spatio-temporal design must be defined, i.e., the one that maximizes the EVOI under the constraint of fixed budget β. The optimization problem is defined as follows:

$$max_{n,1-\alpha}\, EVOI_{n,\alpha}$$
$$subject\ to : \Theta_{n,\alpha} = \beta$$

The methodology consists of finding the optimum by combining the iso-cost curves and the indifference curves. The iso-cost curves give the combination of the number of stations and the probability of detection that can be obtained with the same budget. The indifference curves give the combination of the number of stations and the probability of detection that generated the same EVOI. The methodology then consists of choosing the combination in the iso-cost curve that reaches the highest indifference curve.

The iso-cost curve (9) is deduced from Equation (7):

$$\Theta_{n,\alpha} = n\left(\frac{\lambda}{\alpha}\right) = \beta$$
$$\Leftrightarrow\ 1-\alpha = 1-\left(\frac{\lambda}{\beta}\right)n \tag{9}$$

Based on our hypotheses, the iso-cost curve is a line with a decreasing slope in β and an increasing slope in λ.

An indifference curve is the combination of the number of monitoring stations and the probability of detecting accidental pollution that generates the same $\overline{EVOI}$.

$$EVOI_{n,\alpha} = \overline{EVOI} \tag{10}$$

Consequently, the most efficient combination in an indifference curve is the one that reaches the lowest iso-cost curve. The indifference curve has the form of the following function $g(.)$:

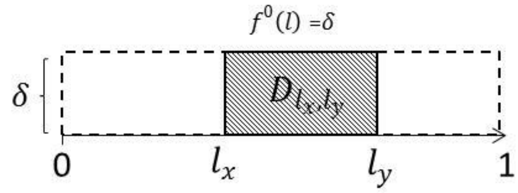
$$(1-\alpha) = g(n, P, D, C, \overline{EVOI}) \quad (11)$$

With a fixed budget, the optimal combination is the combination $(n, 1-\alpha)$ of the iso-cost curve (9) that reaches the highest indifference curve (11).

2.6. Damage scenarios

We will now study two scenarios. In the first one, "Uniform damage", the damage caused by pollution, D_{l_x,l_y}, is proportional to the distance between the location of the emission l_x and the location of the detection l_y (Figure 1).

Figure 1. Uniform damage

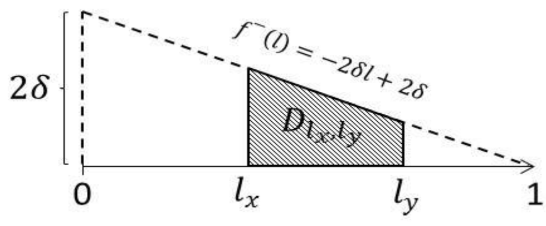


With this first hypothesis, the damage saved, $D_{l_{y/n,out}}$, of Equation (5) is:

$$D_{l_{y/n,out}} = \delta\left(1 - l_{y/n}\right) \quad (12)$$

In the second one, "Decreasing damage", the damage caused by pollution, D_{l_x,l_y}, is represented by Figure 2 below.

Figure 2. Decreasing damage



Decreasing damage can be explained by the presence of tributaries. Consequently, the flow of the river increases from upstream to downstream. Thus, the concentration of the same mass of pollution and, therefore, the damage, decreases from upstream to downstream.

With this second hypothesis, the damage saved, $D_{l_{y/n,out}}$, of Equation (5) is:

$$D_{l_{y/n,out}} = \int_{l_{y/n}}^{1} (-2\delta l + 2\delta) dl = \delta(1 - l_{y/n})^2 \quad (13)$$

3. Calculations

3.1. Location of the monitoring stations and EVOI

3.1.1. Scenario 1: Uniform damage

Close to the outlet of the river, the cost of the action could be higher than the damage avoided (12). Consequently, no station is useful in this part of the river. With the “Uniform damage” scenario, we can write this condition as follows:

$$\delta\left(1 - l_{y/n}\right) > C \qquad \forall n \in [1; +\infty[\;;\ \forall\, y \in [1, n] \tag{14}$$

$$\Leftrightarrow$$

$$l_{y/n} < 1 - \frac{C}{\delta} \qquad \forall n \in [1; +\infty[\;;\ \forall\, y \in [1, n] \tag{15}$$

It can be deduced that there is no station in the river if: $1 - \frac{C}{\delta} < 0 \quad \Leftrightarrow \quad C > \delta$

This means that there is no interest in implementing any monitoring station when the cost of action is higher than the marginal cost of the environmental damage.

It can be deduced from hypothesis of environmental damage uniformity that the locations of the monitoring stations have to be at equal distances (Appendix 1). Then:

$$l_{y/n} = \frac{y}{n+1}\left(1 - \frac{C}{\delta}\right) \qquad \forall n \in [1; +\infty[\;;\ \forall\, y \in [1, n] \tag{16}$$

Then, the environmental damage avoided when pollution is detected in $l_{y/n}$ is:

$$D_{l_{y/n,out}} = \delta.\left[1 - \frac{y}{n+1}.\left(1 - \frac{C}{\delta}\right)\right] \tag{17}$$

By implementing the conditions (16) and (17) in Equation (5), we rewrite EVOI as follows:

$$EVOI_{n,\alpha} = P.(1-\alpha).\sum_{y=1}^{n}\left[\delta.\left[1 - \frac{y}{n+1}.\left(1 - \frac{C}{\delta}\right)\right] - C\right].\sum_{y'=1}^{y}\left[l_{y'/n} - l_{(y'-1)/n}\right].\alpha^{y-y'}$$

$$\Leftrightarrow$$

$$EVOI_{n,\alpha} = P.(1-\alpha).\sum_{y=1}^{n}\left[\delta.\left[1 - \frac{y}{n+1}.\left(1 - \frac{C}{\delta}\right)\right] - C\right].\sum_{y'=1}^{y}\alpha^{y-y'}\left[\frac{y}{n+1}\left(1 - \frac{C}{\delta}\right) - \frac{y-1}{n+1}\left(1 - \frac{C}{\delta}\right)\right]$$

$$\Leftrightarrow EVOI_{n,\alpha} = \frac{P(1-\alpha)(\delta - C)}{\delta(n+1)} \sum_{y=1}^{n} \left[\delta.\left[1 - \frac{y}{n+1}.\left(1 - \frac{C}{\delta}\right)\right] - C\right]. \sum_{y'=1}^{y} \alpha^{y-y'}$$

$$\Leftrightarrow EVOI_{n,\alpha} = \frac{P(1-\alpha)(\delta - C)^2}{\delta(n+1)^2} \sum_{y=1}^{n} (n+1-y) \sum_{y'=1}^{y} \alpha^{y-y'}$$

$$\Leftrightarrow EVOI_{n,\alpha} = \frac{P(1-\alpha)(\delta - C)^2}{\delta(n+1)^2} \sum_{y=1}^{n} (n+1-y)\, \alpha^{y} \sum_{y'=1}^{y} \frac{1}{\alpha^{y'}}$$

$$\Leftrightarrow EVOI_{n,\alpha} = \frac{P(1-\alpha)(\delta - C)^2}{\delta(n+1)^2} \sum_{y=1}^{n} (n+1-y) \frac{(1-\alpha^{y})}{1-\alpha}$$

$$\Leftrightarrow EVOI_{n,\alpha} = \frac{P(\delta - C)^2}{\delta(n+1)^2} \sum_{y=1}^{n} (n+1-y)(1-\alpha^{y}) \tag{18}$$

3.1.2. Scenario 2: Decreasing damage

With Equation (13), the EVOI (5) becomes:

$$EVOI_{n,\alpha} = P.(1-\alpha) \sum_{y=1}^{n} \left[\delta(1 - l_{y/n})^2 - C\right]. \sum_{y'} \left[l_{y'/n} - l_{(y'-1)/n}\right].\alpha^{y-y'} \tag{19}$$

As in Scenario 1, close to the outlet of the river, the cost of the action could be higher than the damage avoided (13). Consequently, no station is useful in this part of the river. With the "Decreasing damage" scenario, this condition can be written as follows:

$$\delta(1 - l_{y/n})^2 > C \quad \forall n \in [1; +\infty[\ ;\ \forall\, y \in [1, n] \tag{20}$$

$$\Leftrightarrow$$

$$l_{y/n} < 1 - \sqrt{\frac{C}{\delta}} \quad \forall n \in [1; +\infty[\ ;\ \forall\, y \in [1, n] \tag{21}$$

With two stations located in $l_{(y-1)/n}$ and $l_{(y+1)/n}$, the optimal location $l_{y/n}$ of the y^{th} station can be found as follows :

$$Min\ D = \int_{l_{(y-1)/n}}^{l_{y/n}} \left[\int_{l_x}^{l_{y/n}} -2\delta l + 2\delta\, dl\right] dl_x + \int_{l_{y/n}}^{l_{(y+1)/n}} \left[\int_{l_x}^{l_{(y+1)/n}} -2\delta l + 2\delta\, dl\right] dl_x$$

$$\frac{\partial D}{\partial l_{y/n}} = -3l_{y/n}^2 + \left(4 + 2l_{(y-1)/n}\right) l_{y/n} - 2l_{(y-1)/n} + l_{(y+1)/n}^2 - 2l_{(y+1)/n} = 0$$

$$\frac{\partial^2 D}{\partial l_{y/n}^2} = -6.\,l_{y/n} + 2l_{(y-1)/n} + 4 > 0$$ [3]

$$\Leftrightarrow$$

$$l_{y/n} = \frac{1}{3}\left[2 + l_{(y-1)/n} - \sqrt{l_{(y-1)/n}^2 - 2l_{(y-1)/n} + 3l_{(y+1)/n}^2 - 6l_{(y+1)/n} + 4}\right] \quad (22)$$

$$\forall n \in [1; +\infty[\ ;\ \forall\, y \in [1, n]$$

The locations of stations in the EVOI (19) must comply with (21) and (22).

4. Results

In this section, we present our results. We also show how our work can help the network manager to answer questions such as: What benefits are generated by the network? Is the cost of the network justified? How to decide between increasing the spatial and temporal intensity of the measurement with a fixed budget? For this, we run a simulation, using Stata and Excel, by arbitrarily choosing parameters as follows:

$$P = 10\%;\ \delta = 100{,}000\ mu\ ; C = 1{,}000\ mu;\ \lambda = 100\ mu$$

Parameters C, δ and λ are expressed in any monetary unit (*mu*). Thereafter, we will not note the *mu* unit to simplify writing.

The above values of the parameters P, δ, C, and λ of the theoretical model in sections 2 and 3, used for our simulation, were chosen in order to have an order of magnitude that justifies the implementation of this monitoring network. Indeed, if the cost of action C is too high in relation to the damage, represented by the δ parameter, the action will be too costly in comparison with its benefit, making detection of the pollution unhelpful. Similarly, if the monitoring cost, represented by the λ parameter, is too high, the monitoring network is not justified in terms of the benefit generated.

4.1. What benefit does my network generate?

The benefit generated by the network, corresponding to the EVOI, is calculated for both scenarios using the results (18), (19), (21) and (22). We thus obtain the following Figures 3 and 4 that express the EVOI as a function of the number of stations (optimally located) and the temporal intensity of the measurement (probability of detection). The figures show both ways to increase the EVOI: by increasing the spatial accuracy or the temporal accuracy.

[3] This condition should be verified in the case of decreasing vulnerability to ensure that the location of the stations minimizes damage.

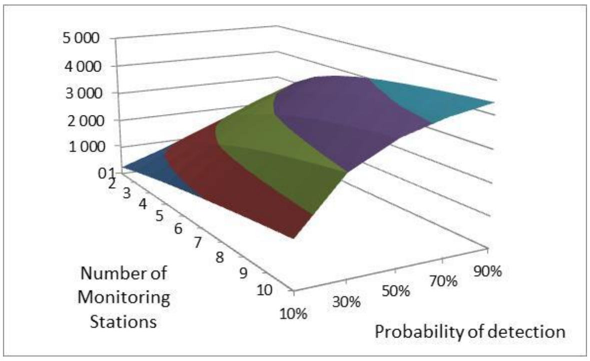


Figure 3. Variation of the EVOI in Scenario 1

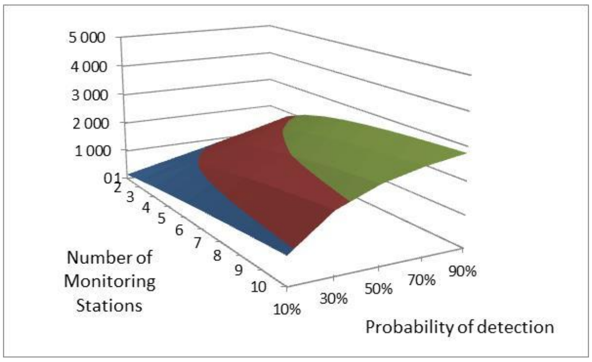


Figure 4. Variation of the EVOI in Scenario 2

In both cases, we observe that the higher the probability of detection $(1-\alpha)\%$ or the greater the number of monitoring stations is, the higher the EVOI will be, but at a decreasing rate. In relative terms, we earn less if the probability of detection or the number of stations increases.

Moreover, the EVOI is significantly lower with the decreasing damage scenario, about a third less. The benefit generated by the network is therefore directly linked to the socio-geographical context.

4.2. Is the cost of the network justified by generating benefits in excess of costs?

By subtracting the cost of the network (7) from the EVOI of Scenario 1 (18) and Scenario 2 (19) (taking the location constraints of stations (21) and (22) for Scenario 2 into account), we obtain the net benefit when the damage is uniform (Figure 5) and when the damage is decreasing (Figure 6). As we previously saw in Section 4.1, since the EVOI is lower in Scenario 2, we logically find a lower net benefit as well.

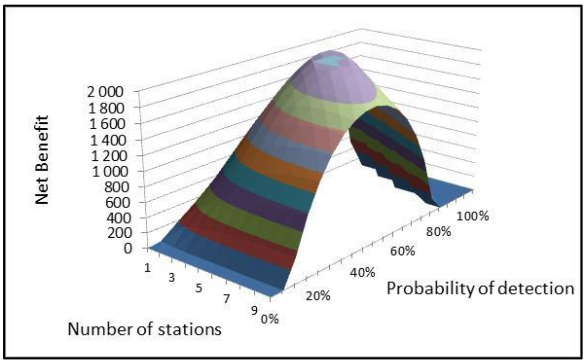


Figure 5. Net benefit in Scenario 1

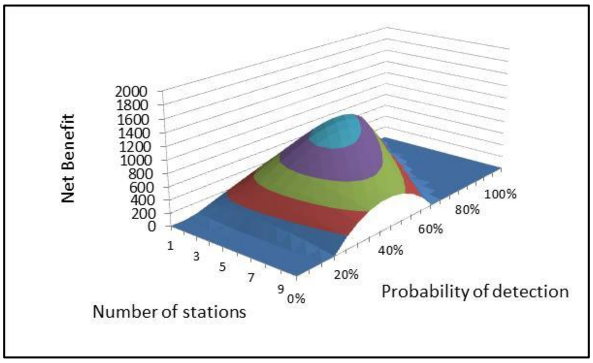
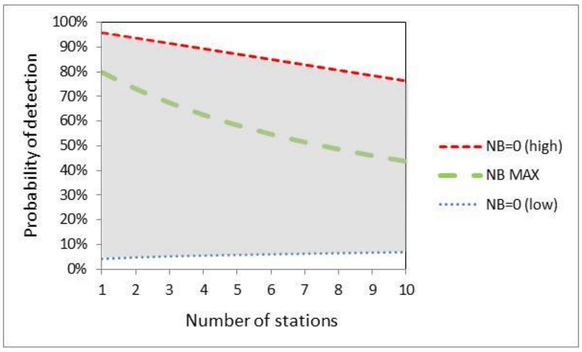
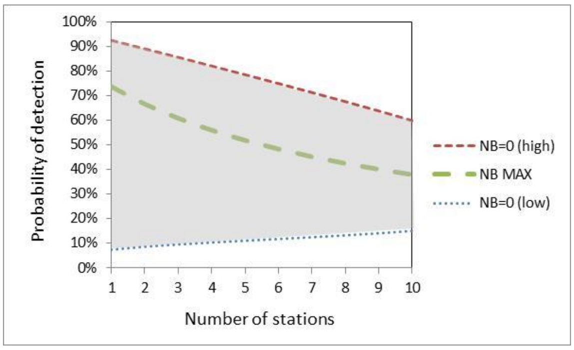


Figure 6. Net benefit in Scenario 2

.

In order to answer the question: "Is the cost of the network justified?", the network manager will be interested in the sign of the net benefit. A positive net benefit will justify the expenses to the funders. For our two scenarios, Figures 7 and 8 illustrate for which combinations (number of stations/probability of detection), the net benefit is positive, negative or zero. We note that there are two areas where the net benefit will be negative. In the upper part of Figures 7 and 8, the net benefit is negative due to the prohibitive cost of the network. In the lower part of these figures, the net benefit is negative due to the insufficient EVOI. Due to the lower EVOI for the decreasing damage scenario, the area where the net profit is positive is logically smaller.





Legend: The net benefit (NB) is positive in the gray part and negative in the white part

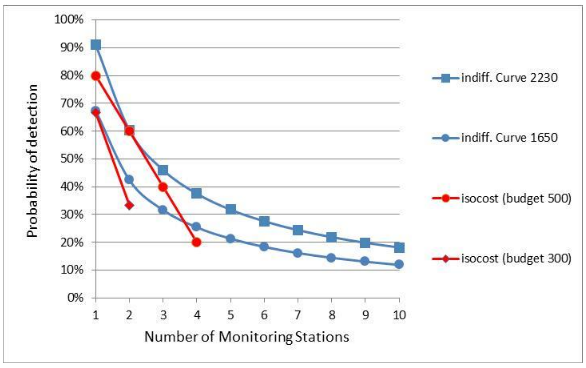
Figure 7. Sign of net benefit in Scenario 1

Figure 8. Sign of net benefit in Scenario 2

It can be observed in Figures 7 and 8 that the more (less) the number of monitoring stations increases, the more the maximum of the net benefit curve (NB MAX) is obtained with a lower (higher) probability of detection. For example, with two or eight monitoring stations, the maximum net benefit is obtained with detection probabilities of 79.8% and 48.6%, respectively, for the uniform damage scenario, and 66.7% and 42.5% for the decreasing damage scenario.

4.3. What is the optimal spatio-temporal design when working with a fixed budget?

After having seen which combination of spatial and temporal intensity the network manager must choose to maximize the net benefit of his network in the previous section, we will now see how this manager must define the network design if he is constrained by a fixed budget. To do this, we simulate the method developed in Section 2.5 using a solver and assuming two budgets: 300 and 500, for both scenarios.



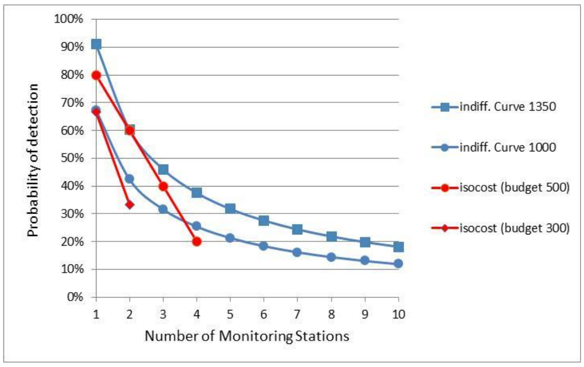


Figure 9. Optimal spatio-temporal design in Scenario 1

Figure 10. Optimal spatio-temporal design in Scenario 2

Figure 9 above shows that with a uniform damage scenario, a budget of 300 makes it possible to reach an indifference curve corresponding to an EVOI of 1650, whereas a budget of 500 makes it possible to reach an EVOI of 2230. For the decreasing damage scenario illustrated in Figure 10, the EVOI reached are lower, 1000 and 1350, respectively, with budgets of 300 and 500.

However, in both cases, the network manager seems to have to make the same choice of network design, namely for a budget of 300 to opt for a single station with a detection probability of 67%, and with a budget of 500 to opt for two stations with a detection probability of 60%. However, the choice will still differ as to the location of the monitoring stations. For the uniform damage scenario, a single station will be located at 0.495 on the river segment [0,1] according to Equation (16), and for two stations at locations 0.33 and 0.66. For the decreasing damage scenario, according to Equations (21) and (22), the sole station will be located at 0.328, and the two stations at 0.214 and 0.47 (see Appendix 2 for all the station locations in our simulation).

In both cases, going from one to two stations for the same budget of 300 would lead to a reduction in the probability of detection to 33%, and all this for a lower EVOI. For a budget of 500, an increase in the number of stations from two to three would make it necessary to decrease the probability of detection from 60% to 40%, and to increase the number of monitoring stations from three to four would make it necessary to decrease the probability of detection from 40% to 20%, once again for an ever lower EVOI.

It is therefore interesting to note that increasing the budget leads to a reduction in the probability of detection. It is therefore more interesting to increase the spatial intensity of the measurement in this case. This can be explained by the fact that the slope of the iso-cost line (9) decreases with the budget. However, as can be seen in Figures 9 and 10, a decreasing slope to the right of the iso-cost line leads to finding an optimal combination more to the "lower right" of the figures, which corresponds to more stations and a lower probability of detection.

5. Discussion and conclusions

To estimate the Economic Value of Information (EVOI) according to the location of the stations and the temporal intensity of the measurement, it is necessary to clearly define the objective of the network beforehand in order to identify the states of nature, namely the quality indicator useful for decision-making. In our example, of an objective of detecting accidental pollution, the states of nature refer to the presence and location of this pollution in order to act accordingly to stop the potential damage. The network design is thus chosen in such a way as to maximize the EVOI, i.e., to allow for economic considerations that are not taken into account in traditional network optimization where only statistical and hydrological considerations are important. We can see in Appendix 2 that the stations will not be located in the same place according to our two scenarios since the decreasing damage requires locations further upstream.

The knowledge of this "benefit" generated by the network allows the network manager to carry out a cost-benefit analysis by integrating the cost of the network according to the number of stations and the temporal intensity of the measurement. We can note that the network benefit calculated here only refers to one objective: the detection of accidental pollution to stop the damage in the short term. However, the information provided can provide other services in the medium or long term. In the case of accidental pollution, a follow-up in the medium or long term can indicate whether the number of accidental pollution events increases, stabilizes or decreases, or allows the sources and reasons for these accidents to be identified in order to make the appropriate decisions. Thus, the calculated EVOI is only a low estimate of the benefits generated by the network. However, this information remains valuable, as we have seen when defining the network design as well as when justifying the cost of the network. In Section 4.2, we see for which combinations of spatial and temporal intensity a network operator can justify the interest of these measures to its funders. Being in the low range of the benefit estimate does not mean that the network is not justified when the net benefit is negative.

This article shows how a network manager can optimize and, therefore, justify the choice of a spatio-temporal combination of the measurement by maximizing the net benefit (Section 4.2) or by maximizing the economic value of information with a fixed budget (Section 4.3). It is interesting to note results that may appear to be counter-intuitive, such as reducing the frequency of the measurement when the budget increases. Indeed, depending on the context, adding an optimally located monitoring station may be more advantageous, even if the frequency of the measurement is reduced over the entire network.

Acting on water quality means, above all, knowing about it. Water quality monitoring networks have been developed to provide this knowledge. However, perfect knowledge subsequent to continuous measurement in time and space is not possible. Thus, the network manager must define the network design, i.e., the number of stations, their location and the frequency of the measurement. A first category of literature has focused on optimizing this structure in order to minimize the statistical inaccuracy of the information. In a second category of literature, authors have estimated the economic value of information of a pre-defined network design using the Bayesian method.

In this article, we combine both categories of literature for the first time by optimizing the design of a monitoring network via the maximization of the economic value of the information generated. To develop our methodology, we take the example of a monitoring network whose sole objective is to detect any accidental pollution that may occur with a uniform probability over a stretch of river. We compare two damage scenarios, the first one where pollution generates the same damage regardless of the emission point, and the second one where this damage is decreasing downstream. This hypothesis can be explained by the increase in flow and, therefore, the dilution effect from upstream to downstream.

We first show how to estimate the Economic Value of Information (EVOI) according to the location of the stations and the temporal intensity of the measurement. The knowledge of this "benefit" generated by the network allows the network manager to carry out a cost-benefit analysis by integrating the cost of the network. The methodology outlined in this article provides network managers with tools to answer key questions such as: Are the costs of monitoring justified by generating benefits in excess of costs? What network design (spatial and temporal intensity of the measurement) should be adopted to maximize the net benefit generated? What is the optimal network design when working with a fixed budget?

In this paper, the choice of a single objective for the network: i.e; detecting an accidental pollution, allow us to assume simplified working hypotheses. The latter concern environmental damage that is inexistent downstream from the outlet of the river, and that can be stopped as soon as the pollution is detected. To justify this hypothesis, we relied on the Huningue alert station in France, where the canal's closure is based on the detection of accidental pollution to prevent the spread of pollution and downstream damages. However, as these hypotheses may limit the generalization of our results, it would be interesting, in a future study, to apply our model to other monitoring objectives or a multi-objective network. We could, for example, integrate the objectives of the Water Framework Directive, namely: does the water body have a good status or not? And if not, what are the types of pollution sources (urban or rural) that cause the downgrading of the water body? Based on a field study, the network design that offers the best trade-off between the different objectives will be determined. Moreover, it would also be interesting to associate the annual measurement frequency with the probability of identifying the real state of nature (probability of detecting accidental pollution in this paper). For this purpose, it would be necessary to perform a non-parametric analysis based on data. This would clarify the trade-off for a network manager to choose between increasing the measurement in space (number of stations) and in time (annual frequency).

Appendix 1: Location of the monitoring stations with uniform damage

We attempt to find the optimal locations of the monitoring stations, namely those that minimize environmental damage. According to our hypothesis, the damage is proportional to the distance between emission and detection. We then look for the location that minimizes this distance. We designate L as the whole distance between potential emissions and the n monitoring stations. The program is written as follows:

$$Min\, L = Min \sum_{y=1}^{n+1} \int_{l_{(y-1)/n}}^{l_{y/n}} \left(l_{y/n} - l_x\right) dl_x \quad \forall\, n \,\epsilon\, [1; +\infty[$$

where $l_{0/n}$ is the source of the river and $l_{(n+1)/n}$ is the outlet. Then: $l_{0/n} = 0$, and $l_{(n+1)/n} = 1$.

With two monitoring stations located at $l_{(y-1)/n}$ and $l_{(y+1)/n}$, the optimal location of the y^{th} station, $l_{y/n}$, is obtained as follows:

$$L = \int_{l_{(y-1)/n}}^{l_{y/n}} (l_{y/n} - l_x) dl_x + \int_{l_{y/n}}^{l_{(y+1)/n}} (l_{(y+1)/n} - l_x) dl_x$$

$$\Leftrightarrow\ L = l_{y/n}^2 - l_{(y-1)/n}.\, l_{y/n} - l_{y/n}.\, l_{(y+1)/n} + \frac{l_{(y-1)/n}^2}{2} + \frac{l_{(y+1)/n}^2}{2}$$

$$\Leftrightarrow\ \frac{\partial L}{\partial l_{y/n}} = 2l_{y/n} - l_{(y-1)/n} - l_{(y+1)/n} = 0$$

$$\Leftrightarrow\ \ l_{y/n} = \frac{l_{(y+1)/n} + l_{(y-1)/n}}{2}$$

Furthermore:

$$\frac{\partial^2 L}{\partial l_{y/n}^2} = 2 > 0$$

Hence, the optimal location is the one that divides the stream into two equal parts. Consequently, the stations must be located at equal distance.

Appendix 2: Location of the monitoring stations in the simulation

$$\delta = 100{,}000; C = 1{,}000$$

Uniform damage (according to Equation (16)):

n										
1	0.495									
2	0.330	0.660								
3	0.2475	0.495	0.743							
4	0.198	0.396	0.594	0.792						
5	0.165	0.330	0.495	0.660	0.825					
6	0.141	0.283	0.424	0.566	0.707	0.849				
7	0.124	0.248	0.371	0.495	0.619	0.7425	0.866			
8	0.110	0.220	0.330	0.440	0.550	0.660	0.770	0.880		
9	0.099	0.198	0.297	0.396	0.495	0.594	0.693	0.792	0.891	
10	0.090	0.180	0.270	0.360	0.450	0.540	0.630	0.720	0.810	0.900

Decreasing damage (according to Equations (21) and (22)):

n										
1	0.328									
2	0.214	0.470								
3	0.160	0.338	0.552							
4	0.128	0.266	0.421	0.606						
5	0.106	0.220	0.343	0.481	0.645					
6	0.091	0.188	0.290	0.402	0.527	0.674				
7	0.080	0.164	0.252	0.346	0.448	0.562	0.697			
8	0.071	0.145	0.223	0.304	0.392	0.486	0.592	0.716		
9	0.064	0.130	0.199	0.272	0.348	0.429	0.518	0.616	0.731	
10	0.058	0.118	0.181	0.245	0.313	0.385	0.461	0.544	0.636	0.745

References

[1] Sanders, T.G., Ward, R.C., Loftis, J.C., Steele, T.D., Adrian, D.D., Yejevich, V., 1983. Design of Networks for Monitoring Water Quality. Water Resources Publication.

[2] European Commission, 2000. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000: establishing a framework for Community action in the field of water policy.

[3] Alvarez-Vázquez, L.J., Martínez, A., Vázquez-Méndez, M.E., Vilar, M.A., 2006. Optimal location of sampling points for river pollution control. Mathematics and Computers in Simulation. 71, 149–160. https://doi: 10.1016/j.matcom.2006.01.005.

[4] Telci, I.T., Nam, K., Guan, J., Aral, M.M., 2009. Optimal water quality monitoring network design for river systems. Journal of Environmental Management. 90, 2987–2998. https://doi: 10.1016/j.jenvman.2009.04.011.

[5] Destandau, F., Point, P., 2000. Cheminement d'impact et tarification efficace sur un basin versant. Revue économique. 51(3), 609-620.

[6] Park, S.Y., Choi, J.H., Wang, S., Park, S.S., 2006. Design of a water quality monitoring network in a large river system using the genetic algorithm. Ecological Modelling. 199, 289-297.

[7] Liu, Y., Zheng, B., Wang, M., Xu, Y., Qin, Y., 2014. Optimization of sampling frequency for routine river water quality monitoring. Science China Chemistry. 57, 772–778. https://doi: 10.1007/s11426-013-4968-8.

[8] Naddeo, V., Scannapieco, D., Zarra, T., Belgiorno, V., 2013. River water quality assessment: Implementation of non-parametric tests for sampling frequency optimization. Land Use Policy. 30, 197–205. https://doi: 10.1016/j.landusepol.2012.03.013.

[9] Kim, G.B., Lee, K.K., Lee, J.Y., Yi, M.J., 2007. Case study for determination of a water level monitoring frequency for nationwide groundwater monitoring network in Korea. Journal of hydrology. 342, 223-237.

[10] Zhou, Y., 1996. Sampling frequency for monitoring the actual state of groundwater systems. Journal of Hydrology. 180, 301–318. https://doi: 10.1016/0022-1694(95)02892-7.

[11] Pourshahabi, S., Talebbeydokhti, N., Rakhshandehroo, G., Nikoo, M.R., 2018. Spatio-Temporal Multi-Criteria Optimization of Reservoir Water Quality Monitoring Network Using Value of Information and Transinformation Entropy. Water Resources Management. 32, 3489–3504. https://doi: 10.1007/s11269-018-2003-8.

[12] Bouma, J.A., Van der Woerd, H.J., Kuik, O.J., 2009. Assessing the value of information for water quality management in the North Sea. Journal of Environmental Management. 90, 1280–1288. https://doi: 10.1016/j.jenvman.2008.07.016.

[13] Bouma, J.A., Kuik, O., Dekker, A.G., 2011. Assessing the value of Earth Observation for managing coral reefs: An example from the Great Barrier Reef. Science of the Total Environment. 409, 4497–4503. https://doi: 10.1016/j.scitotenv.2011.07.023.

[14] Destandau, F., Diop, A.P., 2016. An Analysis of the Value of Additional Information Provided by Water Quality Measurement Network. Journal of Water Resource and Protection. 08, 767–776. https://doi: 10.4236/jwarp.2016.88062.