Optimum Water Quality Monitoring Network Design for Bidirectional River Systems Using Modified MOPSO

Xiaohui Zhu[1,4], Yong Yue[1,3], Prudence W.H. Wong[5], and Yixing Zhang[2,3]

[1]Department of Computer Science and Software Engineering, Xi'an-Jiaotong Liverpool University, Suzhou, Jiangsu Province, 215123 P.R.China
[2]Department of Environmental Science, Xi'an-Jiaotong Liverpool University, Suzhou, Jiangsu Province, 215123 P.R.China
[3]Research Institute of New-type Urbanization, Huai'an, Jiangsu Province, 223005 P.R.China
[4]Department of Computer Science, Nantong University, Nantong, Jiangsu Province, 226019 P.R. China
[5]Department of Computer Science, University of Liverpool, Liverpool, L69 3BX United Kingdom

Correspondence to: Yong Yue (yong.yue@xjtlu.edu.cn)

Abstract. Designing an optimum water quality monitoring network will not only minimize the pollution detection time and maximize the detection probability in river systems, but also reduce the redundant monitoring nodes and save the investment and costs for building and running the network. We propose a novel method for the optimum water quality monitoring network design and identification of the influence of bidirectional water flows which has not be studied in the literature. In order to

5   handle discrete issues of designing an optimum water quality monitoring network for bidirectional rivers, we have modified the Multi-Objective Particle Swarm Optimization (MOPSO) algorithm and developed new fitness functions. The Storm Water Management Model (SWMM) is used to simulate pollution events of a hypothetical river network which was studied in the literature for comparative analysis of our work. Simulation results show that the modified MOPSO can obtain a better Pareto frontier whilst bidirectional water flows have a significant effect on the optimization monitoring network design. We achieve

10   a different optimum deployment from unidirectional water flow for the same river system. We also find that the probability of bidirectional water flows has no effect on the optimum monitoring network design but the pollution detection threshold of the monitoring devices can affect the network design when the threshold is high.

1   Introduction

River systems play a crucial role in the sustainable development of a community. Water quality is influenced simultaneously

15   by both anthropogenic and natural activities. Overexploitation and increasing pollution of this vital resource is threatening our ecosystems and even the life of future generations. In recent 50 years, with rapid development of world economy, on one hand we need more and more clean water, and on the other hand, industry and living activities create more and more pollutants to freshwater sources. It is estimated that 280 billion yuan is lost each year in China for freshwater pollution events. Water quality monitoring has become one of the routine efforts for environmental protection all over the world. However, monitoring water

20   quality remains a very complex process due to the large number of factors to consider such as monitoring locations, selection of water quality parameters, monitoring frequency, identification of monitoring objectives (Behmel et al., 2016). The problem

of planning and optimizing water quality monitoring programs (WQMPS) has been addressed since 1940s and a great many papers have been published on this subject (Park et al., 2006; Chilundo and Kelderman, 2008; Chang et al., 2014).

With the rapid development of computer science and communication technology, more water quality parameters can be remotely detected and transmitted by automatic monitoring stations resulting in a much higher monitoring frequency, more
5   monitoring data and better monitoring efficiency. However, the cost of building and operating an automatic monitoring station is very high (about 500,000-600,000 dollars per station for construction and 14,000 dollars per year for operating and maintain). The successful water quality monitoring relies on the availability of low-cost and high efficient monitoring network to collect appropriate and reliable data. Optimization design of the water quality monitoring network cannot only help us to build a cost-effective and logistically adaptable monitoring network but also improve the monitoring performance and reduce the
10   construction and operating cost, which is very essential for the sustainable development of water quality monitoring network. Many researchers have studied the optimum design of water quality monitoring network for river systems. Ouyang et al. (2008) used a single objective generic algorithm (GA) to design an optimum monitoring network based on a geometric analysis and a simple application in a hypothetical river system. However, only the spatial distribution of the monitoring stations was considered as an optimization objective in this algorithm. Practical river systems are complex and other factors such as
15   flow rate, river depth and width should also be considered while designing the monitoring network. Telci et al. (2008) argued that the design of an optimal water quality monitoring network should mainly focus on two objectives: minimal pollution detection time and maximal detection reliability and calculated the optimal placement of monitoring devices using the GA under relatively simple discrete uniform distributions on spill events. They also applied this methodology to the Altamaha river basin to identify the locations of the best monitoring stations in the river system (Telci et al., 2009). However, the Pareto frontier
20   for the optimization results was not mentioned in this paper resulting in difficulties for evaluating all the optimization results. Park (2013) used stochastic discrete optimization via a simulation (QvS) algorithm and a penalty function with memory (PFM) to the optimal location of a finite number of monitoring positions that minimize the expected detection time of a contaminant spill event while guarantee a higher detection probability. However, the penalty value significantly increased the detection time of a deployment solution when the detection probability is less than 100%. Chang and Lin (2014) selected seven criteria to
25   evaluate the suitability of the water quality monitoring design and used fuzzy theory to improve the objectivity in the data classification and ranking. However, it is very difficult for researchers to collect detailed information and data (e.g. percentage of farmland and built-up area and green cover ratio) to satisfy all the criteria for the algorithm.

Most of the research simplifies the model of river systems and only unidirectional water flow is considered to design the optimum water quality monitoring network. However, affected by regular tides, some river systems have bidirectional water
30   flows. To the best of our knowledge, the effect of bidirectional water flows has not been studied on the optimum network design for a water quality monitoring system and we do not know how far the bidirectional water flows will affect the optimal deployment.

2   Methodology

The purpose of designing an optimum water quality monitoring network is: given a river system being monitored and a definite number of monitoring devices, try to find a deployment solution to maximize the pollution detection probability and minimize the pollution detection time within all the potential monitoring locations. In this study we emphasize the dynamic behavior

5   of water flow directions affected by tides, pollution transport along the river system and the multi-objective particle swarm optimization. The Pareto frontier is also used to evaluate optimization results as well as compare the results to the literature.

2.1   Hydrodynamic Simulations

The Storm Water Management Model (SWMM) is a dynamic rainfall-runoff simulation model used for single event or long-term (continuous) simulation of runoff quantity and quality from primarily urban areas. It can track the quantity and quality of

10   runoff generated within each sub-catchment, and the flow rate, flow depth, and quality of water in each pipe or channel during a simulation period (Rossman et al., 2010). Here we use SWMM to simulate the pollution events and pollutants transport along the river system. To compare our study results with the achievements given by Ouyang et al. (2008) and Telci et al. (2008, 2009), we use the same hypothetical river network they used in Figure 1. We build the river network using SWMM shown in Figure 1. The hypothetical river network A has six inlet nodes of 1, 3, 5, 8, 10 and 11, five intermediate nodes of 2, 4, 6, 7 and

15   9, as well as one outlet node of 12. We assume that a pollution event can occur at any node randomly with the same amount of pollutant spilling and there is only one pollution event at each time. In order to get steady water flows when a pollution event occurs, we simulate the water flows for 24 hours from 00:00 to 23:59 with a steady water flow of $10ft^3/s$ for each inlet node and the pollution event occurs at 10:00 and lasts for 1 hour. We also assume that the pollutant concentration is $10mg/L$ when a pollution event occurs at inlet nodes, resulting in $10.19kg$ of the total amount of pollutant spilling for each pollution

20   event. In order to get the same pollution level during simulations, we also set the same amount of pollutant spilling of $10.19kg$ for intermediate nodes of 2, 4, 6, 7 ,9 and the outlet node of 12 when a pollution event occurs at these nodes. The remaining characteristics of the river network A is shown in Table 1, which is the same as Telci used.

2.2   Optimum objectives

The optimum design for water quality monitoring network has several optimization objectives. Here we consider two optimal

25   objectives of minimum pollution detection time and maximum pollution detection probability, which is the same as Telci's paper.

2.2.1   minimum pollution detection time

Assume that we will deploy n monitoring devices in a river system within m potential monitoring locations ($n \leq m$), which means n special monitoring locations will be selected to deploy monitoring devices from m potential monitoring locations. It
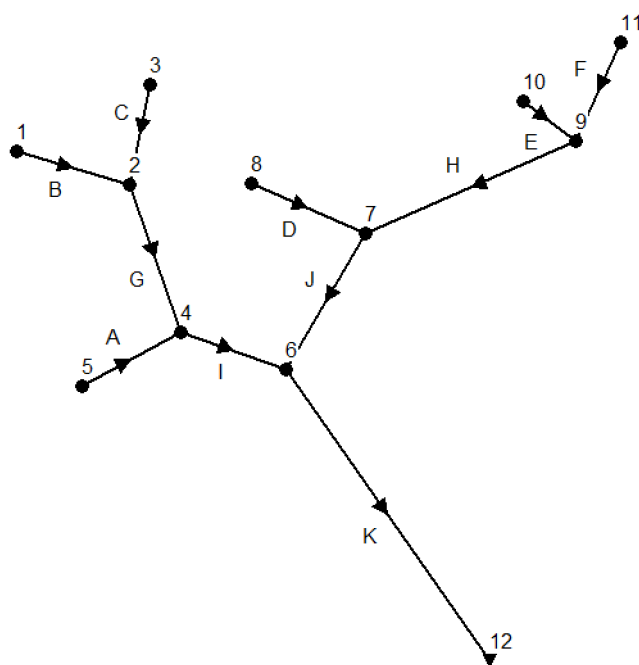
Figure 1. Hypothetical river network A

Table 1. Hydraulic characteristics of the river network A

Catchment	Width (ft)	Channel slope	Manning's coefficient	Length (ft)	Flow rate (ft^3/s)
A	10	0.0001	0.02	2000	10
B	10	0.0001	0.02	2000	10
C	10	0.0001	0.02	2000	10
D	10	0.0001	0.02	2000	10
E	10	0.0001	0.02	1000	10
F	10	0.0001	0.02	2000	10
G	10	0.0001	0.02	3000	20
H	10	0.0001	0.02	4000	20
I	10	0.0001	0.02	2000	30
J	10	0.0001	0.02	3000	30
K	10	0.0001	0.02	5000	60

is easy to know that the total number of potential deployment solutions T is:

$$T = \prod_{i=1}^{n}(m - i + 1) \tag{1}$$

where m is the number of potential monitoring locations, n is the number of monitoring devices deployed in a river system. For a given optimum deployment solution $S_k = [s_{k1}, s_{k2}, s_{ki}, ..., s_{kn}]$, where s_{ki} is the index of selected monitoring locations, $k \leq T$ and $s_{ki} \leq m$. Let $d_i^j(S_k)$ be the pollution detection time of monitoring location i when a pollution event occurs at location j. The minimum pollution detection time for location j is:

$$d^j(S_k) = \min\{d_1^j(S_k), d_2^j(S_k), ..., d_n^j(S_k)\} \tag{2}$$

where $j \leq m$. For a definite optimum deployment solution S, the set of minimal pollution detection time for all potential locations is $d(S_k) = [d^1(S), d^2(S_k), ..., d^j(S_k)..., d^m(S_k)]$. Assume $\overline{d(S_k)}$ is the mean value of all minimal pollution detection time at all m potential monitoring locations for the given solution S_k, $\overline{d(S_k)}$ is:

$$\overline{d(S_k)} = \frac{1}{m}\sum_{j=1}^{m}d^j(S_k) \tag{3}$$

Let $\overline{d(S)}$ be the minimal mean pollution detection time for all potential deployment solutions, we can get the following equation:

$$\overline{d(S)} = \min\{\overline{d(S_1)}, \overline{d(S_2)}, ..., \overline{d(S_T)}\} \tag{4}$$

where T is the total number of deployment solutions shown in Formula (1). One of our two objectives is to find a deployment solution which has the minimum mean pollution detection time as Formula (4) shows.

2.2.2 maximum pollution detection probability

Let $R(S_k)$ be the ratio of successful pollution detection scenarios to all potential detection scenarios for a given deployment solution S_k. We get $R(S_k)$ as:

$$R(S_k) = \frac{1}{m}\sum_{i=1}^{m}r_i \tag{5}$$

where $k \leq T$, m is the amount potential monitoring locations, $r_i = 1$ if the pollution event at location i can be detected by the deployment solution S_k or $r_i = 0$ if the pollution event cannot be detected. Let $R(S)$ be the maximum pollution detection probability within all the potential deployment solutions:

$$R(S) = \max\{R(S_1), R(S_2), ..., R(S_T)\} \tag{6}$$

where T is the total number of potential deployment solutions. Our second objective is to find a proper deployment solution which has a maximum pollution detection probability as Formula (6) shows.

2.3 MOPSO Algorithm

On one hand, we can find from Formula (1) that when we increase the value of m and/or n, the number of deployment solutions will be increased exponentially. Assume we will deploy 20 monitoring devices within 100 potential locations, the number of the deployment combinations is about 10^{40}, which is too large to get the optimum deployment results using enumeration search methods within a reasonable time. On the other hand, these two optimum objectives above are normally conflict with each other, which means we aim to find some good trade-off solutions among these objectives (Reyes-Sierra et al., 2006). So some optimization methodologies should be used here to save the computing time and converge to optimum results in a reasonable period of time.

The Multi-objective particle swarm optimization algorithm (MOPSO) is one of the popular evolution algorithms used in recent years (Coello and Lechuga, 2002). The Pareto dominance was used in the MOPSO to handle multi-objective functions to improve the PSO algorithm to be able to deal with multi-objective optimization problems (Coello et al., 2004). The algorithm uses a secondary repository of particles that is later used by other particles to guide their own flight and the special mutation operator to enrich the exploratory capabilities. Compared to some special multi-objective evolutionary algorithms known to date, MOPSO has a highly competitive performance and can be considered a viable alternative to solve multi-objective optimization problems. Literature research shows that it can cover the full Pareto frontier of all the potential solutions with low computational time. The velocity and position of particles during the computing iteration are updated by the following equations:

$$
\begin{aligned}
V_i(t+1) = & wV_i(t) + c_1 r_1(pbest(i,t) - p_i(t)) \\
& + c_2 r_2(gbest(t) - p_i(t))
\end{aligned}
\tag{7}
$$

$$
p_i(t+1) = p_i(t) + V_i(t+1)
\tag{8}
$$

where V denotes the particle's velocity, w is an inertia weight constant, r_1 and r_2 are uniformly distributed random variables within range $[0, 1]$, $pbest(i,t)$ is the best position that the particle i has had, $gbest(t)$ is the best position in all current particles, and c_1 and c_2 are positive constant coefficients for acceleration. The pseudocode of MOPSO is shown in Algorithm 1.

The classical MOPSO is a powerful algorithm to find global optimum results for continuous definition domains, however it cannot be applied to discrete problems directly. Here we define a new fitness function to calculate the cost of each particle by the means of using a round function to map the continuous value of a particle to a discrete space, which represents the number of potential monitoring locations. Assume that we will deploy n monitoring devices in the hypothetical river system shown in Figure 1, each particle is composed of n different values and each value represents a location chosen for monitoring. The main idea of the fitness function is: First, we decompose the particle into n separate real values and then get n integers using a round function. The n integers represent the number of n potential monitoring locations respectively. Second, we will search each row in pollution detection time table gotten from the pollution simulation by SWMM (e.g. Table 2) and get the minimal detection time for each potential pollution event. Finally, we can calculate the mean detection time and the detection probability for this particle.

As we mentioned above, we try to find an optimum monitoring deployment solution with minimum mean pollution detection time and maximum detection probability. However, MOPSO always requires minimal parameter values to calculate the Pareto frontier. So we calculate the mean pollution detection time and the reciprocal of pollution detection probability in our fitness function to meet the requirement of MOPSO. The pseudocode for the fitness function is as Algorithm 2 shows. In our fitness
5 function, if a pollution event cannot be detected in a deployment scenario ($detectTime = Inf$) we will not count it in the mean pollution detection time but will calculate it in the pollution detection probability. It is different from Telci's paper. They used a penalty value for non-detection scenario which will significantly increase the final pollution detection time when the pollution detection probability is less than 100%.

Algorithm 1 MOPSO

Step 1. Algorithm initialization

    (1) Initialize all parameters (e.g. size of population and repository, maximum number of iterations and search space)

    (2) For each *particle* do

        (a) Initialize the particle's position randomly

        (b) Initialize *pbest* with its initial position

        (c) Initialize particle's velocity V_i=0

    (3) Calculate non-domination particles using fitness function

    (4) Initialize *gbest* with a particle selected from non-domination particles using a roulette wheel selection.

Step 2. Repeat until the termination criteria is met or to the maximum number of iterations

    (5) For each *particle* do

        (a) Calculate particle's new velocity using formula (1)

        (b) Calculate particle's new position using formula (2)

        (c) Update particle's *pbest*

        (d) Calculate non-domination particles using fitness function

        (e) *gbest* = a particle selected from non-domination particles using a roulette wheel selection.

Step 3. Output non-domination particles.

---

3 Simulations and analysis

10 In order to make a deep understanding about how the dynamic characteristics of a river system affect the optimum design of water quality monitoring network, we carry out 4 groups of simulations in the following sections. We also assume that only 3 monitoring devices will be deployed within the 12 potential monitoring locations showed in Figures 1 and 3.

3.1 Simulation for river network A with a pollution detection threshold of 0.01mg/L

At first, we set the pollution detection threshold to $0.01mg/L$ and run the hydraulic simulation in SWMM. Table 2 shows the
15 simulation results of pollution detection time for each potential monitoring locations. The value of Inf in Table 2 represents an

---

Algorithm 2 Fitnesscost ($Particle\ p$)

Array $pos \leftarrow$ [n position values in particle p]

Array $loc \leftarrow [\ \ ]$

for each $element$ in pos **do**

   $node \leftarrow round(element)$

   $loc \leftarrow node$

end for

$meanTime \leftarrow 0$

$count \leftarrow 0$

$probability \leftarrow 0$

 for each row in $TableII$ **do**

   $detectTime \leftarrow Inf$

    for each l in loc **do**

      $detectTime = \min(detectTime, \text{row}[l\ \ ])$

    end for

   if $detectTime \neq Inf$ **then**

     $meanTime \leftarrow meanTime + detectTime$

     $count \leftarrow count + 1$

   end if

end for

$meanTime \leftarrow meanTime / count$

$probability \leftarrow row.length / count$

return$(meanTime, probability)$

---

infinite value, which means the pollution event cannot be successfully detected at a monitoring location. For example, the first row in Table 2 demonstrates a scenario that a pollution event occurs at location 1 and can be detected at location 1, 2, 4, 6 and 12. The pollution detection time for these locations are 0 (detected immediately), 27, 81, 118 and 198 minutes respectively. However, this pollution event cannot be detected at location 3, 5, 7, 8, 9, 10 or 11 because the polluted water flow cannot reach

5 these locations.

    We run the MOPSO algorithm based on the data in Table 2. For the validation of MOPSO to confirm whether the simulation results are steady or not, we run the simulation several times. The simulation results show that though the main particles are quite different from each other, their Pareto frontiers are almost the same. Figure 2 shows 4 Pareto frontiers in 4 different sub-diagrams with 8 non-dominated particles. The mean pollution detection time, pollution detection probability and optimum

10 monitoring locations for each non-dominated particle are shown in Table 3.

    Table 3 indicates that if we deploy 3 monitoring devices at locations 6, 9 and 12 respectively, all the pollution events can be detected, which is the same as the result in Telci's paper. If monitoring devices are deployed at locations 2, 6 and 9, the

Table 2. Pollution detection time for river network A with a detection threshold of $0.01mg/L$

Pollution	Pollution detection time for potential monitoring locations											
locations	1	2	3	4	5	6	7	8	9	10	11	12
1	0	27	Inf	81	Inf	118	Inf	Inf	Inf	Inf	Inf	198
2	Inf	0	Inf	40	Inf	75	Inf	Inf	Inf	Inf	Inf	152
3	Inf	27	0	81	Inf	118	Inf	Inf	Inf	Inf	Inf	198
4	Inf	Inf	Inf	0	Inf	23	Inf	Inf	Inf	Inf	Inf	96
5	Inf	Inf	Inf	28	0	62	Inf	Inf	Inf	Inf	Inf	139
6	Inf	Inf	Inf	Inf	Inf	0	Inf	Inf	Inf	Inf	Inf	62
7	Inf	Inf	Inf	Inf	Inf	38	0	Inf	Inf	Inf	Inf	113
8	Inf	Inf	Inf	Inf	Inf	79	27	0	Inf	Inf	Inf	157
9	Inf	Inf	Inf	Inf	Inf	111	57	Inf	0	Inf	Inf	190
10	Inf	Inf	Inf	Inf	Inf	133	78	Inf	10	0	Inf	213
11	Inf	Inf	Inf	Inf	Inf	156	99	Inf	27	Inf	0	236
12	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	0

pollution detection probability will be slightly decreased to 91.7% while the mean pollution detection time is also reduced from 45.8 minutes to 26.6 minutes. It is also the second best choice in the Pareto frontier. However, the second highest detection probability in Telci's paper is 83% and the monitoring locations are 4, 7 and 9, which also can be found in our main particles in Figure 2 but is not a non-dominated particle. Based on this observation, we confirm that using MOPSO can get a better Pareto

5 frontier and more detailed optimal deployment solutions.

Telci el at. (2008) used a penalty for non-detected pollution scenarios resulting in a much higher pollution detection time for non-100% detected pollution scenarios. We argue that it is not reasonable, because the mean detection time represents how long the pollution event will be detected if it can be detected by current monitoring network. On the contrary, if a pollution event cannot be detected, the detection probability will be decreased to reflect this non-detected scenario. So we ignore these

10 non-detected pollution events when calculate the mean pollution detection time, which results in a shorter mean pollution detection time than in Telci's paper.

Comparing Table 3 to Figure 2, we find that there are 13 different monitoring deployment solutions mapping to 8 non-dominated particles. This is because some deployment solutions with different monitoring locations have the same mean detection time and detection probability, and they map to a same non-dominated particle.

15 ## 3.2 Simulation for a reversed river network B with pollution detection threshold of $0.01mg/L$

Most of literature only considers the unidirectional water flow. However, influenced by tides, some river systems have bidirectional water flows. In order to observe how far the bidirectional water flows can affect the monitoring network optimization,
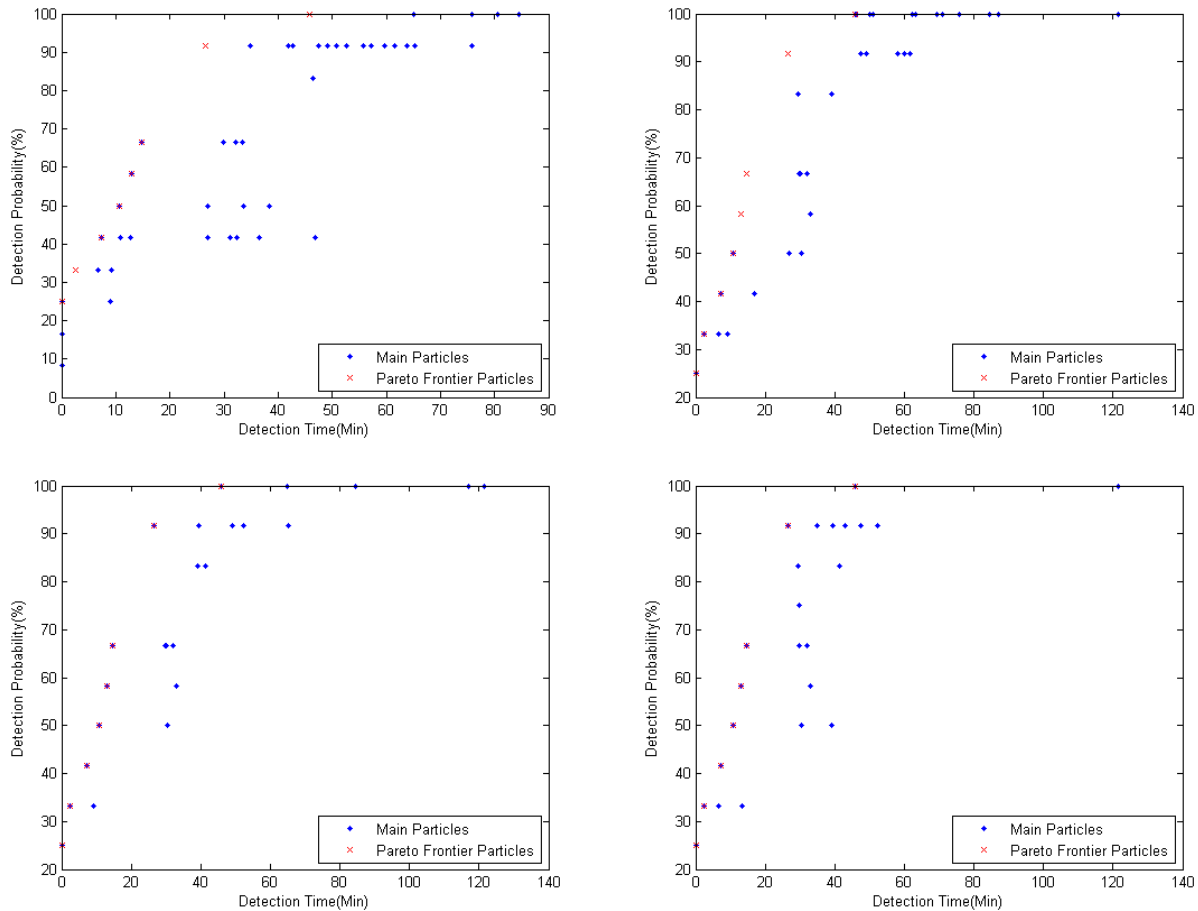
Figure 2. Pareto frontier for river network A with 3 monitoring nodes and a detection threshold of $0.01 mg/L$

we create river network B shown in Figure 3 with the same parameters and settings as river network A in Figure 1 but have a reversed water flow direction resulting in a new river network with 6 outlet nodes, 5 intermediate nodes and only 1 inlet nodes. We set the water flow rate of inlet node 12 to $60 ft^3$/s, which is as same as the water flow rate at outlet node of 12 in Figure 1. We run the hydraulic simulation using SWMM again and get pollution detection times shown in Table 4. we can find from Tables 2 and 4 that when we reverse the water flow, we get a transposed pollution detection time matrix.

Due to the page limit, only one MOPSO Pareto frontier is shown here in Figure 4. The optimum deployment solutions are shown in Table 5.

We find that when we reverse the water flow direction, there are 7 non-dominated particles in Pareto frontier and there is no 100% detection probability solution for river network B. The maximum pollution detection probability is decreased to 75% with a mean pollution detection time of 38.2 minutes and the optimization monitoring locations are 3, 5 and 10. This is because

Table 3. Pollution detection time and probability in Pareto frontier for river network A with a detection threshold of $0.01 mg/L$

Monitoring locations	Detection time (min)	Detection probability
6, 9, 12	45.8	100%
2, 6, 9	26.6	91.7%
2, 7, 9	14.8	66.7%
2, 8, 9	13	58.3%
3, 7, 9	10.7	50.0%
5, 7, 9	10.7	50.0%
5, 8, 9	7.4	41.7%
3, 8, 9	7.4	41.7%
5, 9, 11	2.5	33.3%
5, 8, 11	0.0	25%
1, 5, 10	0.0	25%
5, 8, 10	0.0	25%
1, 5, 8	0.0	25%

there are 6 outlet locations in river network B and only 3 monitoring devices cannot detect all the pollution scenarios occurred randomly in 12 potential locations.

Comparing Table 5 to Table 3, we find that the optimization results for both water flow directions are quite different. Based on this observation, we argue that the water flow direction has a significant effect on optimization results of monitoring network design even for the same river system and we should consider the bidirectional water flows when we design an optimization monitoring network for a river system affected by tides regularly.

3.3 Simulation with bidirectional water flows

In order to have a deep insight of the influence of bidirectional water flows for an optimum monitoring network design, we calculate the mean pollution detection time for both the original river network A (Figure 1) and the reversed river network B (Figure 3) at the same time based on the data of pollution detection time in Tables 2 and 4. As water flow directions can be changed regularly due to tides and the duration for each flow direction may not be equal in a river system. So, we consider two scenarios here when a pollution event occurs:

  – both water flows have the same probability in a river system.

  – the probability of two water flows are different.

We slightly modify the previous fitness function in Figure 2 and add two extra parameters of $probA$ and $probB$ in the procedure to represent the probability of the two water flows in the river system. We calculate the pollution detection time and
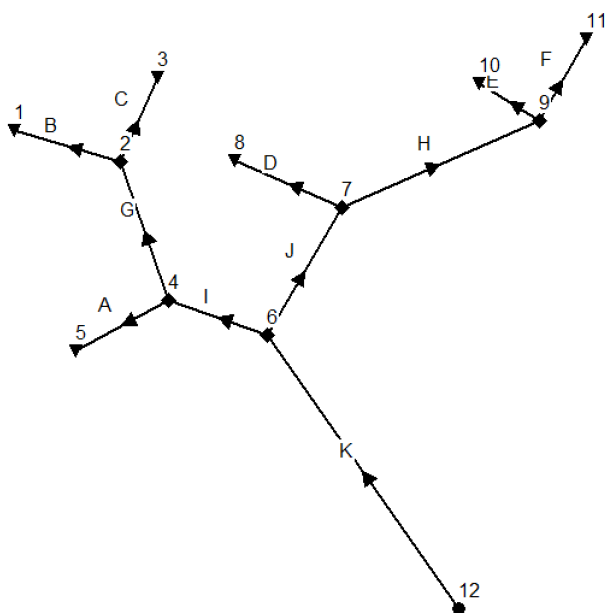
Figure 3. Hypothetical river network B

Table 4. Pollution detection time for river network B with a detection threshold of $0.01 mg/L$

Pollution	Pollution detection time for potential monitoring locations											
locations	1	2	3	4	5	6	7	8	9	10	11	12
1	0	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
2	27	0	27	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
3	Inf	Inf	0	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
4	81	40	81	0	28	Inf	Inf	Inf	Inf	Inf	Inf	Inf
5	Inf	Inf	Inf	Inf	0	Inf	Inf	Inf	Inf	Inf	Inf	Inf
6	118	75	118	23	62	0	38	79	111	133	156	Inf
7	Inf	Inf	Inf	Inf	Inf	Inf	0	27	57	78	99	Inf
8	Inf	Inf	Inf	Inf	Inf	Inf	Inf	0	Inf	Inf	Inf	Inf
9	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	0	10	27	Inf
10	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	0	Inf	Inf
11	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	0	Inf
12	198	152	198	96	139	62	113	157	190	213	236	0
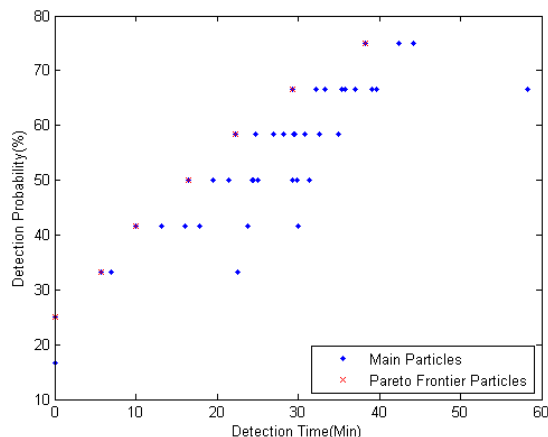
Figure 4. Pareto frontier for river network B with 3 monitoring nodes and a detection threshold of $0.01mg/L$

Table 5. Pollution detection time and probability in Pareto frontier for river network B with a detection threshold of $0.01mg/L$

Monitoring locations	Detection time (min)	Detection probability
3, 5, 10	38.2	75%
3, 4, 10	29.3	66.7%
4, 8, 10	22.3	58.3%
6, 8, 10	16.5	50.0%
4, 8, 12	10.0	41.7%
4, 5, 12	5.8	33.3%
4, 7, 12	5.8	33.3%
6, 7, 12	0.0	25%
4, 6, 12	0.0	25%

pollution detection probability for bidirectional water flows respectively and get the final mean pollution detection time and probability for two water flows at last. The new fitness function is shown in Algorithm 3.

3.3.1 Bidirectional water flows with the same probability

We let $probA$ and $probB$ in Algorithm 3 be 0.5 separately to assume that each water flow with a different direction has the same probability. The simulation results of Pareto frontier and optimization monitoring locations are shown in Figure 5 and Table 6.

Comparing Table 6 to Table 3 and Table 5, we observe that when we consider the bidirectional water flows, the maximum detection probability is decreased from 100% (in Table 3) and 75% (in Table 5) to 66.7% respectively while the mean pollution
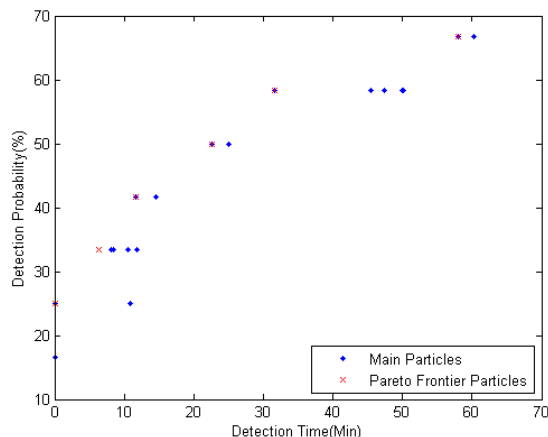
Figure 5. Pareto frontier for bidirectional water flows with 3 monitoring nodes and a detection threshold of $0.01mg/L$

Table 6. Pollution detection time and probability in Pareto frontier for bidirectional water flows with a detection threshold of $0.01mg/L$

Monitoring locations	Detection time (min)	Detection probability
3, 10, 12	57.9	66.7%
3, 6, 10	31.6	58.3%
3, 6, 8	22.6	50.0%
5, 6, 8	11.6	41.7%
5, 6, 7	6.4	33.3%
4, 7, 8	0	25%
4, 8, 10	0	25%
3, 7, 9	0	25%
3, 9, 11	0	25%
6, 7, 9	0	25%

detection time is increased from 45.8 and 38.2 minutes to 57.9 minutes. This is because we take into account two water flows with different directions at the same time and calculate their mean detection time, which will significantly increase the pollution detection time and decrease the detection probability. We also find that the optimum deployment solutions are quite different from the previous results in Table 3 and Table 5 and monitoring locations 3, 10, 12 have the best optimum result with 57.9 minutes of mean pollution detection time and 66.7% pollution detection probability.

Algorithm 3 BidirectionalFitnessCost ($Particle\ p$, $probA$, $probB$)

Array pos = n position values in particle p

Array loc =

For **each** element in pos

node = round(element)

loc = node

EndFor

 meanTime = 0

count = 0

probability = 0

For **each** row1 in TableII and row2 in TableIV

detectTimeA = Inf

detectTimeB = Inf

For **each** l in loc

detectTimeA = min (detectTimeA, row1[l)]

detectTimeB = min (detectTimeB, row1[l)]

$EndFor$

$a \neq b$

avgTime = detectTimeA*porbA+detectTimeB*probB

meanTime = meanTime+avgTime

count = count+1

EndIf

EndFor

meanTime = meanTime / count

probability = row.length / count

Return (meanTime, probability)

3.3.2 Bidirectional water flows with different probabilities

Here we assume two water flows in a river system have different probabilities. We consider two scenarios: (1) the probability of the water flow as river network A is 70% and the opposite water flow as river network B is 30%. (2) the probability of the water flow as river network A is 30% and the opposite water flow as river network B is 70%. We set the parameter of $probA$ to

5 0.7 and $probB$ to 0.3 for the first scenario and exchange the value with each other for the second scenario. We get two Pareto frontiers in Figure 6 and two pollution detection time and probabilities in Table 7.

We find from Table 7(a) and Table 6, that though we set 70% probability for river network A and 30% probability for river network B, we get the same optimization monitoring locations and detection probabilities while the pollution detection time is slightly increased. This is because the pollution detection time for river network A (Table 3) is slightly higher than for river
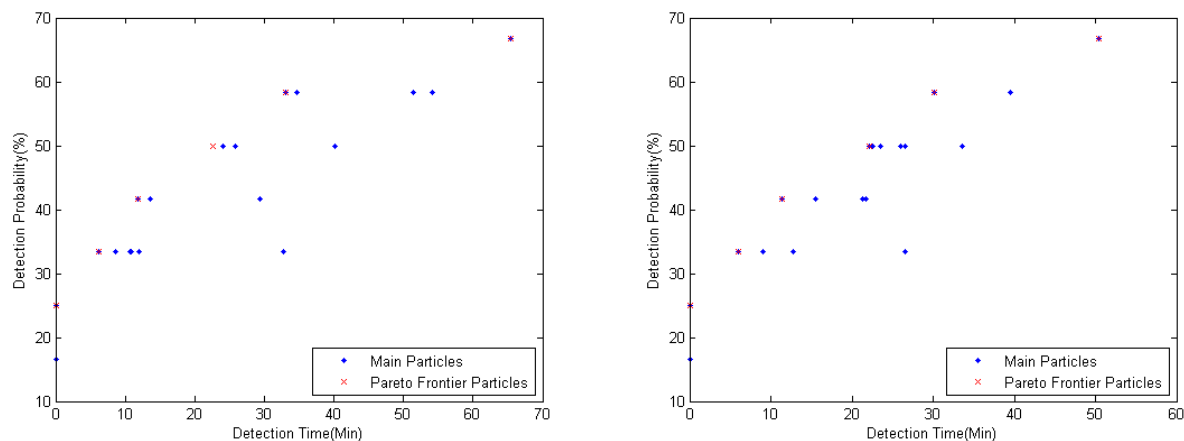
Figure 6. Pareto frontiers for bidirectional water flows with 3 monitoring nodes and a detection threshold of $0.01mg/L$

network B (Table 5) resulting in a higher mean pollution detection time. When we reverse the probability of the two water flows, we get similar results but with a lower mean pollution detection time in Table 7(b).

Comparing Table 7(a) to Table 7(b), we observe that though we exchange their probabilities for the two water flows, we get the same optimization monitoring locations and the same detection probability while the pollution detection time is slightly increased.

Based on the observations above, we draw a conclusion that the bidirectional water flows have a significant effect on an optimal monitoring network design. However, the different ratio of probability for bidirectional water flows has no effect on the optimization results of monitoring location selection or the pollution detection probability but slightly influences the pollution detection time.

3.4 Higher pollution detection threshold for bidirectional water flows

In order to observe how far the pollution detection threshold can affect the optimum deployment solution for a bidirectional water flow river system, we assume two bidirectional water flows have the same probability and set the pollution detection threshold to $1mg/L$ and $2mg/L$ respectively. We run the hydraulic simulation using SWMM again based on river networks A and B. Tables 8 and 9 show four pollution detection time tables for both detection thresholds.

We find that all the pollution detection time in Table 8(a) are much higher than in Table 2 except for non-detected scenarios. This is because when we increase the pollution detection threshold from $0.01mg/L$ to $1mg/L$ for river network A, it will spend more time to accumulate to a certain pollutant concentration at each potential monitoring location before it can be detected, which will significantly increase the pollution detection time.

Comparing Table 8(a) to Table 8(b), we find that all the pollution events can be successfully detected at location 12 when the pollution detection threshold is $1mg/L$. However, no pollution event can be detected at location 12 when we set pollution

Table 7. Pollution detection time and probability in Pareto frontier for bidirectional water flows with a detection threshold of $0.01mg/L$

(a) Probability ratio for river network A and B is 70%:30%			(b) robability ratio for river network A and B is 70%:30%		
Monitoring locations	Detection time (min)	Detection probability	Monitoring locations	Detection time (min)	Detection probability
3, 10, 12	65.4	66.7%	3, 10, 12	50.5	66.7%
3, 6, 10	33.1	58.3%	3, 6, 10	30.2	58.3%
3, 6, 8	22.6	50.0%	5, 6, 10	22.1	50.0%
5, 6, 8	11.8	41.7%	5, 6, 8	11.4	41.7%
5, 6, 7	6.1	33.33%	5, 6, 7	6.0	33.33%
4, 7, 8	0	25%	4, 7, 8	0	25%
4, 8, 10	0	25%	4, 8, 10	0	25%
3, 7, 9	0	25%	3, 7, 9	0	25%
3, 9, 11	0	25%	3, 9, 11	0	25%
6, 7, 9	0	25%	6, 7, 9	0	25%

detection threshold to $2mg/L$, even the pollution event occurs at location 12 itself. This is because the pollution detection threshold is so high that it is even higher than the maximum pollutant concentration at location 12 when any pollution event occurs.

Figure7(a) shows the pollutant degradation along downstream locations when a pollution event occurs at the upstream location of 1 in the hypothetical river network A (Figure 1). We can find that the pollutant concentration is decreased from maximal value of $10mg/L$ at location 1 to minimal value of $1.67mg/L$ at outlet location 12 along the downstream. Figure 7(b) demonstrates the pollutant accumulation progress in river network A at location 12 when a pollution event occurs at monitoring locations 1, 6, 11 and 12 respectively. We can see from Figure 7(b) that when a pollution event occurs at location 1, the pollutant will arrive at location 12 in 198 minutes and will be completely discharged in 368 minutes with a maximum pollutant concentration of $1.44mg/L$ in the pollution event duration. When the pollution occurs at location 12 itself, the pollutant will be diluted by upstream water flows and the maximum pollutant concentration is only $1.67mg/L$. That's why none of the pollution events can be detected when we set pollution detection threshold to $2mg/L$. We get similar results in Table 9 when we increase the pollution detection threshold to $1mg/L$ and $2mg/L$ respectively for river network B.

Figure 8 shows the Pareto frontier for bidirectional water flows based on the pollution time data in Tables 8 and 9. We can find that Figure 8(a) is quite different from Figure 8(b) and there are 5 Pareto frontier particles in Figure 8(a) but only 2 Pareto frontier particles in Figure 8(b). Table 10 shows the detailed pollution detection time and probability.

Comparing the monitoring location distribution in Table 6 and Table 10(a), we observe that though we increase the pollution detection threshold from $0.01mg/L$ to $1mg/L$, the two optimum deployment solutions are the same while the detection time is slightly increased. However, from Table 10(b) we know, when we continue to increase the pollution detection threshold to $2mg/L$, which is higher than the maximum pollutant concentration in pollution events, the pollution detection probability is

Table 8. Pollution detection time for river network A with higher pollution detection thresholds

(a) Pollution detection threshold=$1mg/L$

Pollution locations	Pollution detection time for potential monitoring locations											
	1	2	3	4	5	6	7	8	9	10	11	12
1	0	44	Inf	112	Inf	165	Inf	Inf	Inf	Inf	Inf	253
2	Inf	0	Inf	61	Inf	110	Inf	Inf	Inf	Inf	Inf	199
3	Inf	44	0	112	Inf	165	Inf	Inf	Inf	Inf	Inf	253
4	Inf	Inf	Inf	0	Inf	42	Inf	Inf	Inf	Inf	Inf	131
5	Inf	Inf	Inf	47	0	97	Inf	Inf	Inf	Inf	Inf	186
6	Inf	Inf	Inf	Inf	Inf	0	Inf	Inf	Inf	Inf	Inf	90
7	Inf	Inf	Inf	Inf	Inf	62	0	Inf	Inf	Inf	Inf	152
8	Inf	Inf	Inf	Inf	Inf	116	47	0	Inf	Inf	Inf	205
9	Inf	Inf	Inf	Inf	Inf	153	82	Inf	0	Inf	Inf	242
10	Inf	Inf	Inf	Inf	Inf	181	108	Inf	20	0	Inf	269
11	Inf	Inf	Inf	Inf	Inf	208	134	Inf	44	Inf	0	297
12	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	0

(b) Pollution detection threshold=$2mg/L$

Pollution locations	Pollution detection time for potential monitoring locations											
	1	2	3	4	5	6	7	8	9	10	11	12
1	0	50	Inf	124	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
2	Inf	0	Inf	69	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
3	Inf	50	0	124	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
4	Inf	Inf	Inf	0	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
5	Inf	Inf	Inf	55	0	Inf	Inf	Inf	Inf	Inf	Inf	Inf
6	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
7	Inf	Inf	Inf	Inf	Inf	Inf	0	Inf	Inf	Inf	Inf	Inf
8	Inf	Inf	Inf	Inf	Inf	Inf	55	0	Inf	Inf	Inf	Inf
9	Inf	Inf	Inf	Inf	Inf	Inf	92	Inf	0	Inf	Inf	Inf
10	Inf	Inf	Inf	Inf	Inf	Inf	119	Inf	24	0	Inf	Inf
11	Inf	Inf	Inf	Inf	Inf	Inf	146	Inf	50	Inf	0	Inf
12	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
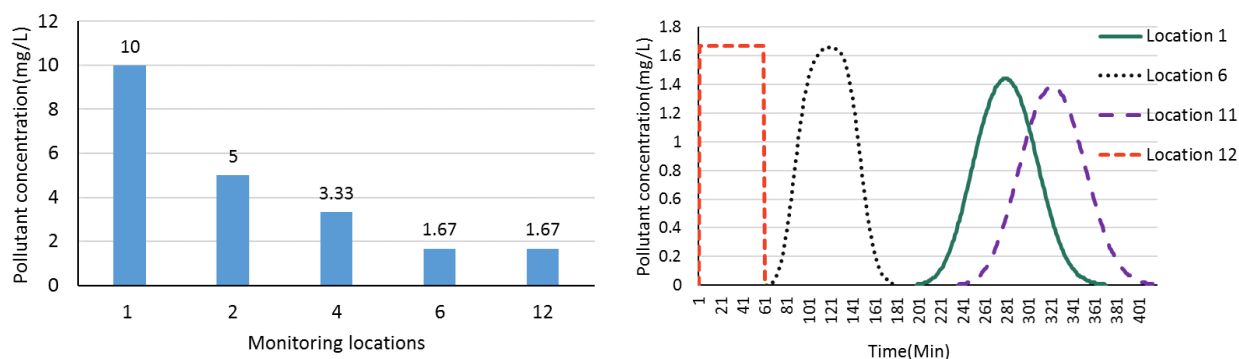
Figure 7. Pareto frontier for bidirectional water flows with 3 monitoring nodes and $0.01mg/L$ detection threshold

significantly decreased and we get quite different optimum solutions. Based on the observation above, we make a conclusion that a slight change of monitoring device's pollution detection threshold may not affect the design of optimum monitoring network if the threshold is smaller than the maximal pollutant concentration in the pollution events.

Table 9. Pollution detection time for river network B with higher pollution detection thresholds

(a) Pollution detection threshold=$1mg/L$

Pollution locations	Pollution detection time for potential monitoring locations											
	1	2	3	4	5	6	7	8	9	10	11	12
1	0	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
2	44	0	44	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
3	Inf	Inf	0	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
4	112	61	112	0	47	Inf	Inf	Inf	Inf	Inf	Inf	Inf
5	Inf	Inf	Inf	Inf	0	Inf	Inf	Inf	Inf	Inf	Inf	Inf
6	165	110	165	42	97	0	62	116	153	181	208	Inf
7	Inf	Inf	Inf	Inf	Inf	Inf	0	47	82	108	134	Inf
8	Inf	Inf	Inf	Inf	Inf	Inf	Inf	0	Inf	Inf	Inf	Inf
9	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	0	20	44	Inf
10	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	0	Inf	Inf
11	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	0	Inf
12	253	199	253	131	186	90	152	205	242	269	297	0

(b) Pollution detection threshold=$2mg/L$

Pollution locations	Pollution detection time for potential monitoring locations											
	1	2	3	4	5	6	7	8	9	10	11	12
1	0	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
2	50	0	50	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
3	Inf	Inf	0	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
4	124	69	124	0	55	Inf	Inf	Inf	Inf	Inf	Inf	Inf
5	Inf	Inf	Inf	Inf	0	Inf	Inf	Inf	Inf	Inf	Inf	Inf
6	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
7	Inf	Inf	Inf	Inf	Inf	Inf	0	55	92	119	146	Inf
8	Inf	Inf	Inf	Inf	Inf	Inf	Inf	0	Inf	Inf	Inf	Inf
9	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	0	24	50	Inf
10	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	0	Inf	Inf
11	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	0	Inf
12	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf

Figure 8. Pareto frontier for bidirectional water flows with 3 monitoring nodes and higher detection threshold

4 Conclusion and future work

We have presented a novel method based on a modified MOPSO algorithm for the optimum water quality monitoring network design and identification of the influence of bidirectional water flows. We develop new fitness functions for MOPSO to handle the discrete issues and speed up the convergence. Simulation results show that the modified MOPSO can get a better Pareto

Table 10. Pollution detection time and probability in Pareto frontier for bidirectional water flows

(a) Pollution detection threshold=$1mg/L$			(b) Pollution detection threshold=$2mg/L$		
Monitoring locations	Detection time (min)	Detection probability	Monitoring locations	Detection time (min)	Detection probability
3, 10, 12	78.9	66.7%	5, 7, 10	14.5	33.3%
3, 6, 10	46.5	58.3%	4, 7, 10	14.5	33.3%
3, 6, 8	34.8	50.0%	3, 7, 10	14.5	33.3%
5, 6, 8	19.8	41.67%	3, 4, 9	14.5	33.3%
5, 6, 7	11.1	33.33%	3, 4, 7	14.5	33.3%
4, 7, 8	0	25%	4, 7, 9	0	25%
4, 8, 10	0	25%	4, 8, 10	0	25%
3, 7, 9	0	25%	3, 7, 9	0	25%
3, 9, 11	0	25%	3, 9, 11	0	25%
6, 7, 9	0	25%	5, 7, 9	0	25%

frontier than GA. A bidirectional fitness function is also developed to handle the bidirectional water flows with different probabilities. We find that bidirectional water flows in a river system have a significant effect on the optimum design of water quality monitoring network and the deployment result is quite different from the same river system with a unidirectional water flow. However, the probability of bidirectional water flows in a river system has no effect on the optimum monitoring network design but will slightly affect the mean pollution detection time. We also find that the monitoring sensor's pollution detection threshold also has little effect on the design of the optimum water quality monitoring network if it has a smaller value than the maximal pollutant concentration of a pollution event. However, the sensor's pollution detection threshold will evidently affect the monitoring network design when it is larger than the maximal pollution concentration.

This novel approach will be applied to a water quality monitoring network to optimize the network design and improve the monitoring performance. Further research is planed to explore the feasibility of integrating graph theory and priority coefficients into MOPSO to guide the convergence processing. Finally, it is desirable to redesign the velocity and position functions with a fully discrete method to improve the computing performance.

Acknowledgements. This work was partly supported by Jiangsu Natural Science Foundation (BK20151245), Huai'an Natural Science Foundation (HAG2015007), Nantong Natural Science Foundation (MS12016048) and Key Laboratory of High Performance Computing of Jiangxi Province (PKLHPC1303).

References

Behmel S, Damour M, Ludwig R, et al. Water quality monitoring strategies - A review and future perspectives.[J]. Science of The Total Environment, 2016: 1312-1329.

Strobl, Robert O., and Paul D. Robillard. "Network design for water quality monitoring of surface freshwaters: A review." Journal of environmental management 87.4 (2008): 639-648.

Ouyang, H. T., Yu, H., Lu, C. H., & Luo, Y. H. (2008). Design optimization of river sampling network using genetic algorithms. Journal of Water Resources Planning and Management, 134(1), 83-87.

Telci, I. T., Nam, K., Guan, J., & Aral, M. M. (2008). Real time optimal monitoring network design in river networks. In World Environmental and Water Resources Congress 2008: Ahupua'A (pp. 1-10).

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(10), 2987-2998.

Park, C. (2013). Discrete optimization via simulation with stochastic constraints (Doctoral dissertation, Georgia Institute of Technology).

Chang, C. L., & Lin, Y. T. (2014). A water quality monitoring network design using fuzzy theory and multiple criteria analysis. Environmental monitoring and assessment, 186(10), 6459-6469.

Rossman, Lewis A. Storm water management model user's manual, version 5.0. Cincinnati, OH: National Risk Management Research Laboratory, Office of Research and Development, US Environmental Protection Agency, 2010.

Reyes-Sierra, M., & Coello, C. C. (2006). Multi-objective particle swarm optimizers: A survey of the state-of-the-art. International journal of computational intelligence research, 2(3), 287-308.

Coello C A, Lechuga M. MOPSO: a proposal for multiple objective particle swarm optimization[C]. congress on evolutionary computation, 2002: 1051-1056.

Coello, C. A. C., Pulido, G. T., & Lechuga, M. S. (2004). Handling multiple objectives with particle swarm optimization. IEEE Transactions on evolutionary computation, 8(3), 256-279.

Park S Y, Choi J H, Wang S, et al. Design of a water quality monitoring network in a large river system using the genetic algorithm[J]. Ecological modelling, 2006, 199(3): 289-297.

Chilundo M, Kelderman P. Design of a water quality monitoring network for the Limpopo River Basin in Mozambique[J]. Physics and Chemistry of the Earth, Parts A/B/C, 2008, 33(8): 655-665.

Guestrin C, Krause A, Singh A P. Near-optimal sensor placements in gaussian processes[C]. Proceedings of the 22nd international conference on Machine learning. ACM, 2005: 265-272.

Chen Q, Wu W, Blanckaert K, et al. Optimization of water quality monitoring network in a large river by combining measurements, a numerical model and matter-element analyses[J]. Journal of Environmental Management, 2012: 116-124.

Chang, Chia-Ling, and You-Tze Lin. A water quality monitoring network design using fuzzy theory and multiple criteria analysis[J]. Environmental monitoring and assessment 186.10 (2014): 6459-6469.