Dear Dr Christian Stamm,

We greatly appreciate your great effort and helpful comments. Please see our response below.

Response to Editor's Comments

C1: p. 1, L: 15: Replace activities by processes.

A1: We accept this suggestion. Word "activities" was replaced by "processes".

C2: p. 1, L: 19: Sometimes, you use yuan, sometimes dollars as currencies. You might use both for the first instance and then consistently use one currency in the remainder of the text.

A2: Thank you. We will consistently use the currency of dollar.

C3: p. 1 - 2; Introduction: As in the abstract, the context and motivation for the specific research question is only poorly provided. Depending on the objectives for the monitoring program, the development of the design can be rather straightforward not requiring any complicated numerical optimization procedure. If a country like Switzerland for example, has a program that aims at quantifying the loads of major water constituents such as nutrients that are discharged from the country, it is sufficient to locate monitoring sites at the four main rivers leaving the country. Accordingly, you should describe what kind of general monitoring objectives (may) require a complex optimisation procedure. This gives the motivation for actually using them. Subsequently, you may report on the current state of the art in that field (including the pros and cons of the existing optimisation algorithms). This will lead to the open questions that you would like to address with your paper.

A3: Thank you for your suggestion. We will further revise the abstract to clearly provide the context and motivation for this research question. We will also extend the introduction section to give state of the art achievements and different opinions in water quality monitoring network optimization.

C3: p. 2, L. 7-8: What is the argument for this statement?

A3: This statement is from a reference "Strobl, Robert O., and Paul D. Robillard. "Network design for water quality monitoring of surface freshwaters: A review." Journal of environmental management 5 87.4 (2008): 639-648.".

We will explicitly mark the reference in our paper.

C4: p. 2, L. 11: The optimum design depends on the actual objectives for the monitoring network. Accordingly, the means to find such an optimal solution may also change with these objectives.

A4: Objectives for monitoring network design can affect the optimum deployment solution. However, to a definite multi-object optimization algorithm, we only need to update the fitness functions when we change the optimization objectives.

We will replace the original sentence by "Many researchers have studied the optimum design of water quality monitoring network for river systems based on varies of optimization objectives and approaches".

C5: p. 2, L. 15: Why are these factors relevant?

A5: The flow rate, river length and width can affect the flow speed and pollutant diffusion speed when a pollution event occurs, which results in an influence on the objective of minimal pollution detection time.

C6: p. 2, L. 27: What is the relevance of farmland in this context? Explain.

A6: The farmland area is one of seven criteria used by Chang and Lin (2014) to evaluate the suitability of the water quality monitoring network design. The authors argued that many factors such as farmland area, Green cover ratio, landslide areas and so on represent the pollution potential and vulnerability of areas with different land-use conditions. The areas with higher pollution potential will impact water quality and should be considered in the design of water quality monitoring network.

We will add a brief explanation of the reliance between farmland and the water quality monitoring network design.

References:

[1] Chang C L, Lin Y T. USING THE VIKOR METHOD TO EVALUATE THE DESIGN OF A WATER QUALITY MONITORING NETWORK IN A WATERSHED[J]. International Journal of Environmental Science and Technology, 2013, 11(2): 303-310.)

[2] 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.

C7: p. 2, L. 30: Can you provide some quantitative data on the (global) length of river sections influenced by tides? This might be interesting for readers to put your work in context.

A7: Tai Lake basin in China is one of the important tide-affected area with many rivers distributed in this 36,900 km^2 basin. Local governments have built hundreds of water quality monitoring stations to collect and evaluate varies of water quality data for several years. We will try to find if the local governments have the quantitative data on length of river sections influenced by tides. We have also searched by google. Unfortunately, we did not find the accurate data of global length of tide-influenced rivers yet.

C8: p. 3, L. 2 - 6: This should be in the Introduction, not in the Method section.

A8: Thank you. We agree and move it to the Introduction section.

Q9: p. 3, L. 16: How did you simulate the bidirectional flow where you have different flow directions at the same time (p. 14, L. 2)? This cannot be steady-state, can it? What is the governing equation of solute transport and how did you parameterise this?

A9: The phrase “at the same time” may not be appropriate here. As a matter of fact, we calculate the pollution detection time and detection probability for each water flow respectively. Then, we combine them together to get the mean pollution detection time and probability based on the time ratio of two reversed water flows. We simulate the water flow route in SWMM using kinematic wave routing model and Horton infiltration model.

C10: p. 3, L. 17ff: Use only SI units throughout the paper.

A10: It is corrected. We have changed all the units to SI units throughout the paper.

C11: p. 3, L. 17ff: Why do you use (arbitrary) absolute masses and concentrations? Your entire analysis can be dimensionless by just using for example concentrations relative to LOQ or LOD.

A11: This is because we should demonstrate whether different pollution concentrations will affect the optimal deployment solutions or not. When we simulate pollution events in SWMM, we should also set a concentration value.

C12: p. 3, L. 28: Why are there only m potential monitoring locations? There is an infinite number of potential locations along such a network. What are the actual locations you have in mind? This is not clear.

A11: In theory, pollution events can occur at anywhere along the river and any location in a river network can be a potential monitoring location. So, there is an infinite number of potential monitoring locations. However, infinite potential monitoring locations cannot be dealt with by computer. We should convert this continuous problem to a discrete domain and assume there is only *m* potential monitoring locations. As we know, the more potential monitoring locations we set, the more accurate the optimal deployment solutions are. However, for the simplicity of demonstration and the comparison to the literature, we select inlet, outlet and intersection nodes as potential monitoring locations and assume all the pollution events only occur at these locations.

C12: p. 4, Table 1: Is it reasonable to assume a constant width although the flow rate varies by a factor of six?

A12: For the comparison of our results to the literature, parameters in Table I are the same as Telci's paper. The reason for higher flow rates in catchments G, H, I, J and K is that several upstream (inlet) flows from catchments A, B, C, D, E and F converge to one river channel.

Q13: p. 6, L. 9 - 16, 24 - 25: These paragraphs report the state of the art. As such, it should be presented already in the Introduction and needs references for the statements about the performance. In the Intro you should also explain what PSO is compared to MOPSO.

A13: Thank you for your reasonable suggestion. We will introduce MOPSO and explain the difference between PSO and MOPSO in the Introduction and add references for the statements of the performance. This comment is similar with the first comments of referee 2 (C1 in Referee Comments 2).

C14: p. 6, L. 25 - 30: The new fitness function and how it differs from others is not very well described. Make it more prominent.

A14: we will explain the fitness function in more detail.

C15: p. 7, Algorithm 1: Label this as a table. The same holds for the other two algorithms you present.

A15: We have labelled all algorithms as tables.

C16: p. 7, L. 10: Replace make a deep by gain deeper.

A16: Thank you for your suggestion. We have replaced "make a deep" by "gain deeper".

C17: p. 8, L. 7: What are the main particles? Explain for a non-specialist.

A17: The main particles are the particles in a repository. MOPSO uses the repository to calculate the new particles and get the non-dominated particles (Pareto frontier) from these main particles. when the MOPSO comes to an end, we list all the main particles in the repository and the particles on Pareto frontier in the same figure. We will explain it in more detail for a non-specialist in our paper.

C18: p. 9, Table 2: This table contains little information (per area of page). Please consider to put it into Supporting Information. This applies to Tables 4 and 9 as well.

A18: We accept your suggestion. We will put tables 2, 4 and 9 into Supporting Information.

C19: p. 9, L. 2: How can you have a second best choice on a Pareto front?

A19: The sentence is not appropriate here. We should replace it by "it is also the second maximal pollution detection probability on a Pareto frontier".

C20: p. 10, Fig. 2: The caption does not explain what the four figures are. How do the figures relate to Table 3? The symbols are too small and hard to read and to distinguish from each other.

A20: Sub-captions for these 4 figures were lost. There are several similar mistakes in the paper. We will add these sub-captions and enlarge the symbols to make it easier to read and distinguish.

Figure 2 shows pollution detection time and probability of the particles on Pareto frontier (red cross) and other main particles (blue square). Assume we select 3 optimal monitoring locations out of 12 locations. Each particle will be composed by 3 positions with random value between 1 and 12 before computing. When the algorithm comes to an end, the 3 position values of each particle on the Pareto frontier are the optimal monitoring locations, which is shown in Table 3.

C21: p. 10, L. 10: Why only one? What is special about this example?

A21: We are not very sure if this comment is for P10. L.6 (only one MOPSO Pareto frontier is shown here in Figure 4.)? in Figure 2 we run the algorithm for 4 times and get 4 different simulation results in 4 subfigures, however, we can find that though the main particles in 4 subfigures are quite different, the Pareto frontiers are the same. It means that our algorithm can get a steady Pareto frontier. Based on this observation, we only show one result of Pareto frontier in the subsequent experiments.

C22: p. 10, Table 3: The caption does not explain what monitoring locations represent. What is the meaning of a zero detection time? How does it come that the combination 3,5,8 is not listed? It is equivalent to 1,5,8 from a geometric perspective.

A22: The monitoring locations in Table 3 represent the locations we should deploy monitoring devices. The zero detection time means it can detect the pollution event immediately when a pollution event occurs. For example, last row in Table 3 means that if we deploy monitoring devices at locations 1, 5 and 8, the detection probability is only 25%. This is because we deploy monitoring devices at the inlet nodes and only pollution events at node 1, 5 and 8 can be detected by this deployment solution (3/12=25%). However, we can detect the pollution events immediately (detection time =0.0) if the pollution event occurs at locations 1, 5 or 8. As we think the combinations with zero detection time and lower detection probability are almost useless in practice, we did not list all this kind of combinations in Table 3.

However, we can list all the optimal deployment solutions on Pareto frontier. We will also add a brief explanation of the zero pollution detection time as well as replace all the captions in Tables 3, 5, 6, 7 and 10 by “Optimal deployment solutions on Pareto frontier”.

C23: p. 17, L. 4: Degradation is the wrong word here. It is the decrease of the maximum concentration, I assume.

A23: Thank you. We have replaced “Degradation” by “dilution”.

C24: p. 17, L. 7: accumulation is the wrong word here.

A24: Thank you. We have replaced “accumulation” by “reach”.

C25: p. 18, Fig. 7: The figure caption is not properly describing the content of the figures.

A25: Thank you. We have updated several similar errors in our paper.

C26: p. 19, L. 3: Where did you report on the speed up of convergence?

A26: Because we use a discrete method to calculate the fitness cost for MOPSO. It need less iterations to search the domain than continuous method resulting in a higher convergence. We will add a brief interpretation in our paper.

C27: p. 20, L. 1: Where did you show that MOPSO outperforms GA in general terms (for what kind of problems)? Can you explain why this should be so? The discussion and conclusion lacks the entire aspect that you have worked on one single, rather artificial model system. What would happen if one considers other network topologies or spatially continuous instead of distinct possibilities for locating monitoring sites?

A27: Coello et al. (2004) compared MOPSO against three state-of-art multi-objective evolutionary algorithms of Nondominated Sorting Genetic Algorithm II (NSGA-II), Pareto Archived Evolution Strategy (PAES) and Microgenetic Algorithm for Multi-objective Optimization (MicroGA) using 5 different test functions. Experiment results show that MOPSO has a highly competitive performance and can be considered a viable alternative to solve multi-objective optimization problems.

We also compared our results to Telci's paper in P.20, L1-5. And verified that GA algorithm used by Telci did not get the full Pareto frontier (on line 2 page 9 in our paper). In addition, we developed an enumeration search method to confirm that our algorithm can get a full Pareto frontier.

We will add a comparison of the results between our algorithm and the enumeration research method as well as add some explanations on why we use MOPSO to design an optimal network for water quality monitoring in Introduction section.

C28: p. 20, L. 11: Why would you like to use graph theory? Which problems do you imagine to use with such an approach? Again, you have to link this to aspects you have already discussed previously. Otherwise, it is a rather arbitrary addition to the text.

A28: Thank you for your suggestion. We found that in practical water quality monitoring network design, some locations have special management features (e.g. deploying a monitoring device at an intersection of two cities can count the amount of pollutant discharge from one city to another) and we should deploy monitoring devices at these locations no matter they are in the optimal deployment solution or not. We are further developing our algorithm based on MOPSO to support reserved monitoring locations in optimal monitoring network design. Graph theory and priority coefficient are used to guide the convergence processing and ensure the final optimal deployment solutions include these reserved locations.

We will replace "Further research is planned to explore the feasibility of integrating graph theory and priority coefficients into MOPSO to guide the convergence processing" by "Further research is planned to explore the feasibility of integrating reserved monitoring locations beforehand into MOPSO using graph theory and priority coefficients to guide the convergence processing"