

Stakeholder engagement in multi-objective optimization of water quality monitoring network, case study: Karkheh Dam reservoir

Mohammad Reza Nikoo, Shokoufeh Pourshahabi, Najmeh Rezazadeh and M. Ehsan Shafiee

ABSTRACT

Reservoir water quality is important for water quality management downstream. A hierarchical approach is developed to present the monitoring locations within a format that satisfies the objectives of social stakeholders for making final decisions. First, a CE-QUAL-W2 model is applied to simulate water quality variables in the reservoir for a long time using a set of historic data. Second, transinformation entropy theory is used to quantify mutual information among a set of monitoring stations for each water quality variable. Then, a non-dominating sorting genetic algorithm-based model is developed for multi-objective optimization of the water quality monitoring network. Finally, a social choice method is applied to the identified non-dominated solutions to achieve a strategy that is compromised among stakeholders. The variations of the water quality variables at different depths and different seasons are investigated. The proposed approach is illustrated for Karkheh Reservoir in Iran. The number of optimized monitoring stations is the same for all seasons (three out of 22 potential stations) using different social choice methods. The results show the appropriate performance of the proposed methodology for optimization of reservoir water quality monitoring stations.

Key words | CE-QUAL-W2 model, NSGA-II multi-objective optimization method, reservoir monitoring network, social choice, transinformation entropy

INTRODUCTION

Sampling locations are designed to monitor reservoir water quality. The design of sampling locations is difficult due to a wide range of water quality variables that can be used to present water quality, the temporal and spatial characteristics of sampling, and the duration and objectives of sampling (Harmancioglu *et al.* 1999). For optimization of a water quality monitoring program, various regulations and the physical and geographical properties of reservoirs must be considered in formulating this non-linear, complex problem (Behmel *et al.* 2016).

The study of the placement of monitoring stations in a reservoir is limited in the literature, although similar studies

can be found for identifying locations to monitor water quality in river and groundwater resources. Lee & Kwon (2009) and Lee *et al.* (2011) proposed methodologies using statistical analysis to reduce redundant sampling locations in a reservoir. Based on information theory, Lee *et al.* (2014) proposed an approach for optimization of water quality monitoring stations in a reservoir. They used water quality data at similar depths of Lake Yongdam. Instead of using an optimization approach, they tested all possible combinations to find the optimal solution among potential combinations of stations. These studies are not suitable for the large reservoirs, at which the potential sampling

Mohammad Reza Nikoo (corresponding author)

Shokoufeh Pourshahabi
Department of Civil and Environmental
Engineering,
Shiraz University,
Shiraz,
Iran
E-mail: nikoo@shirazu.ac.ir

Najmeh Rezazadeh
Civil Engineering Faculty,
Ferdowsi University,
Mashhad,
Iran

M. Ehsan Shafiee
Sensus,
Morrisville, NC 27650,
USA

locations increase significantly. *Yenilmez et al. (2015)* proposed a methodology to cope with this problem by deploying an approach based on kernel density estimation and ordinary kriging to optimize the number of monitoring stations in the Porsuk dam reservoir for a single water quality variable. They used dissolved oxygen (DO) at the surface layer of the reservoir. Their approach neglects the effect of reservoir depth on the water quality.

In general, a series of data can be translated into information using the Shannon theory (*Harmancioglu 1981*). The entropy quantities include marginal entropy, joint entropy, conditional entropy and transinformation. The use of information theory was explored to assess the groundwater-monitoring network. *Mogheir & Singh (2002)*, *Mogheir et al. (2009)*, *Masoumi & Kerachian (2010)*, and *Mondal & Singh (2012)* evaluated the groundwater-monitoring network using information theory. This theory has also been used to evaluate the effectiveness of monitoring stations in rivers. *Ozkul et al. (2000)*, *Karamouz et al. (2009)*, *Mahjouri & Kerachian (2011)*, and *Memarzadeh et al. (2013)* employed entropy theory to assess water quality monitoring networks in rivers. To the best of our knowledge, *Lee et al. (2014)* used entropy theory to identify water quality monitoring locations in a reservoir, however, as mentioned earlier, the effect of reservoir depths on the water quality was not considered. They modeled water quality by observing water quality data that was collated at nine locations at the surface of reservoir. These studies considered the technical aspects of reservoir water quality and failed to include stakeholders and decision makers to evaluate the plausibility and practicality of a design to monitor reservoir water quality.

Stakeholder engagement can lead to more sustainability for a water resources system as they can input their expert knowledge during the design process (*Loucks et al. 2005*). To address the choices of decision makers, which are most likely conflicting, a set of equations is formulated to mathematically capture each choice as an objective. Each equation determines a way to evaluate the goodness of a solution based on the stakeholder's choice. Using a heuristics approach, a set of non-dominated solutions is identified to represent the relationship among objectives. Each solution from the non-dominated front provides the best for an objective given the constraints of the other objectives. To find the

best solution among a set of non-dominated solutions, a social method can be used to rank solutions based on the preferences of decision makers. Social choice theory is a useful method to encompass the preferences of engagement stakeholders over the available options (*Sheikhmohammady & Madani 2008*).

In this study, a new methodology is developed to identify a set of locations to monitor water quality in a reservoir. The methodology contains two main steps. First a non-dominated sorting genetic algorithm (NSGA-II)-based optimization model is applied to find a set of non-dominated solutions. Second, the social choice theory-based approach is used to identify a solution with the highest rank based on the preferences of participating stakeholders. The linkage between the technical aspects of this problem with the social aspects is the contribution of this study. This study introduces and evaluates a new approach to include the decision-makers' cooperation in the design process. In addition, this study considers the effect of reservoir depth on water quality. The presented optimization model is employed to identify monitoring locations for water quality sampling from the reservoir of the Karkheh Dam in Iran.

MODEL FRAMEWORK

The flowchart for this study is illustrated in *Figure 1*. The proposed approach contains two steps: (a) the optimization part to identify non-dominated solutions and (b) the social choice method. First, water quality variables are selected. The water quality of the reservoir water body is simulated by a calibrated and verified CE-QUAL-W2 model. The governing equations in the CE-QUAL-W2 model are the continuity equation, momentum equation and advection-diffusion equation. Based on some experiments (*Debele et al. 2008*), this model is able to simulate the water quality of reservoirs with an accuracy that is comparable to the WASP model (Supplementary Material Table S1, available with the online version of this paper). The geometric data, meteorological parameters, hydraulic coefficients, inflow/outflow data, and flow sources and sinks are inputs into the CE-QUAL-W2 model. Once the CE-QUAL-W2 model is calibrated and validated using the historic dataset, it is

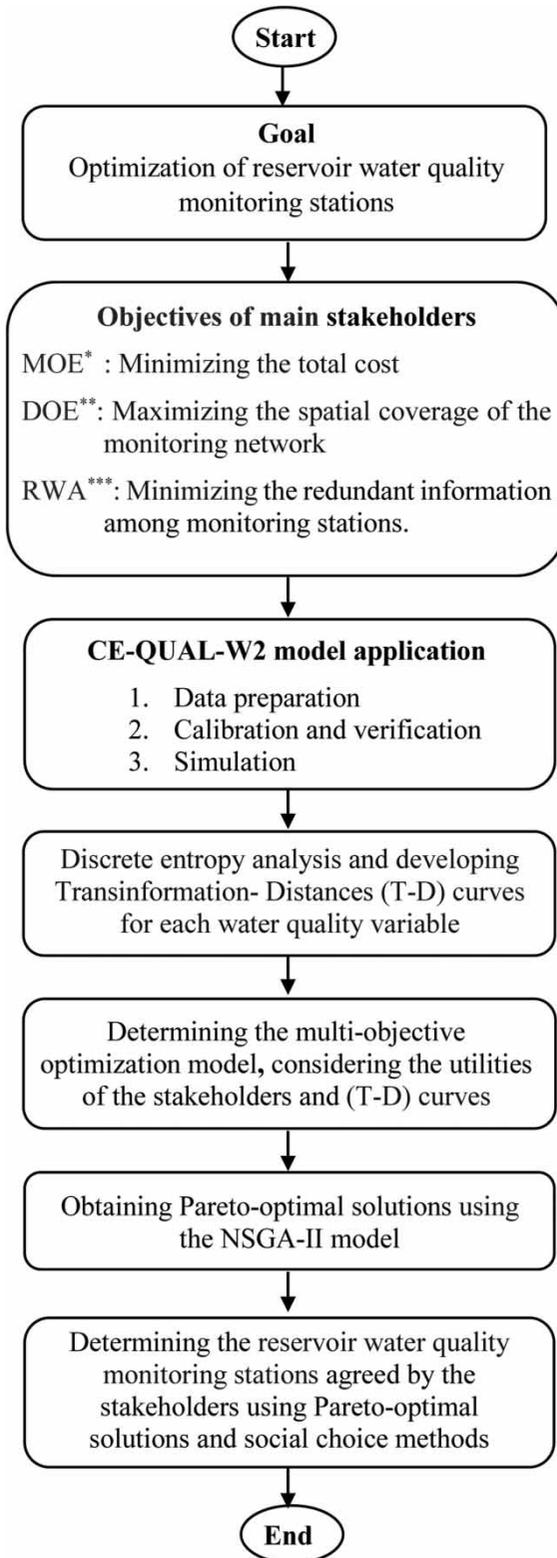


Figure 1 | Flowchart to identify locations to optimally monitor reservoir water quality.
*Ministry of Energy; **Department of Environment; ***Regional Water Authority.

used to simulate reservoir water quality for each water quality variable. The transinformation entropy quantity ($Te(x,y)$) is calculated for each pair of potential monitoring stations, represented as station x and station y as follows:

$$Te(x, y) = - \sum_{i=1}^n \sum_{j=1}^m p(x_i, y_j) \ln \left[\frac{p(x_i, y_j)}{p(x_i)p(y_j)} \right] \quad (1)$$

where $p(x_i)$ and $p(y_j)$ are the occurrence probability of x_i and y_j , respectively, and $p(x_i, y_j)$ is the joint probability between x_i and y_j (Mogheir et al. 2004a, 2004b). A T - D curve is obtained by plotting the transinformation quantities (T) with respect to the spatial distances (D) among each pair of monitoring stations.

Stakeholders

Three stakeholders, i.e. the Ministry of Energy (MOE), the Department of Environment (DOE), and the Regional Water Authority (RWA), are involved in the operation of a reservoir in Iran. The MOE provides financial support and balances the budget while the DOE is responsible for water quality assessment and the RWA aims to increase the efficiency of the monitoring network by minimizing the redundant information among stations. The objectives of these stakeholders are summarized as follows:

1. The MOE objective is to minimize the cost.
2. The DOE objective is to maximize the spatial coverage of the water quality monitoring network.
3. The RWA objective is to minimize the redundant information among water quality monitoring stations.

Considering the utilities of these stakeholders and the T - D curves, a multi-objective optimization model is formulated and solved by a conventional heuristic algorithm to find a set of non-dominated solutions. Due to the complexity of this problem and the large size of the decision space, the multi-objective algorithm likely converges to non-dominated solutions instead of a true Pareto-front, therefore, the non-dominated term is used instead of the Pareto-front. In this study, the formulated problem is solved using NSGA-II (Deb et al. 2002).

A preference matrix is developed by ranking m non-dominated solutions based on the choice of three

Table 1 | The preference matrix

	1 2 3..... m
MOE	Each of m solutions in order of priority for MOE
DOE	Each of m solutions in order of priority for DOE
RWA	Each of m solutions in order of priority for RWA

stakeholders. The preference matrix is a $3 \times m$ matrix as demonstrated in Table 1, in which m is the number of optimal solutions. Finally, the best solution is determined by the social choice method which selects a solution based on its popularity among the stakeholders.

Multi-objective optimization model: NSGA-II

The objectives of stakeholders are formulated into a multi-objective optimization model as follows:

$$\text{Minimize } y_1 = \frac{n_{opt} - n_{min}}{n_p - n_{min}} \tag{2}$$

$$\text{Maximize } y_2 = \sum_{q=1}^Q \gamma_q \sum_{i=1}^{n_p} c_i / n_p \times \begin{cases} \frac{D_{max} - D_i}{D_{max} - D_{opt,q}} & \text{if } D_i > D_{opt,q} \\ 1 & \text{if } D_i \leq D_{opt,q} \end{cases} \tag{3}$$

$$\text{Minimize } y_3 = \sum_{q=1}^Q \gamma_q \sum_{i=1}^{n_p} c_i / n_p \times \begin{cases} \frac{Te_{i,q} - Te_{min,q}}{Te_{max,q} - Te_{min,q}} & \text{if } D_i < D_{opt,q} \\ 1 & \text{if } D_i \geq D_{opt,q} \end{cases} \tag{4}$$

$$n_{min} = \frac{D_{max}}{\min(D_{opt,q})} + 1 \tag{5}$$

$$\sum_{i=1}^{n_p} c_i = n_{opt} \tag{6}$$

$$\sum_{q=1}^Q \gamma_q = 1, \tag{7}$$

where the variables in Equations (2)–(7) are:

- $y_1, y_2, \& y_3$ Normalized objectives between 0 and 1,
- n_{opt} Number of optimized stations,
- n_p Number of potential stations,
- n_{min} Minimum number of stations,
- Q Number of water quality variables,
- γ_q Importance coefficient for a water quality variable q in a monitoring program, $q = 1, \dots, Q$,
- c_i Binary variable (if there is a monitoring station in potential station i , c_i is 1, otherwise, it is 0),
- $Te_{i,q}$ Transinformation entropy (Te in Equation (1)) associated with the distance d_i in station i and water quality variable q ,
- $Te_{min,q}$ Minimum transinformation entropy quantity for water quality variable q ,
- $Te_{max,q}$ Maximum transinformation entropy quantity for water quality variable q ,
- D_{max} Maximum value of the distances, taken pairwise, among all pairs of the potential stations (in kilometers),
- D_i Distance between station i and the closest station (in kilometers),
- $D_{opt,q}$ Optimized distance associated with $Te_{min,q}$ for water quality variable q (in kilometers).

Social choice theory

Social choice theory considers the involvement of stakeholders with different preferences over the available options (solutions obtained from NSGA-II) in the process of decision-making. The common social choice methods are as follows (Sheikhmohammady & Madani 2008).

Condorcet choice

This method elects an option that is preferred more times to the other options by the stakeholders with pairwise comparisons (McLean 1990; Shalikaran et al. 2011).

Borda count method

A score is assigned to each option based on the stakeholder preference. The solutions are ranked based on

their preference. The highest score is n (the count of non-dominated solutions) for the most preferable option and this score decreases incrementally all the way to 1 for the least preferable option. A total score is calculated by summing scores across the stakeholders. The option that has the highest total score is selected as the compliant option for the stakeholders (McLean 1990; Shalikaran *et al.* 2011).

The plurality rule

The plurality rule selects the option that receives the most votes by the stakeholders in the first level of preference (Sheikhmohammady & Madani 2008; Shalikaran *et al.* 2011).

Majority voting rule

The majority voting rule (MVR) is a decision rule that selects options that have the majority votes (more than half) at the highest possible level of preference. In this way, only a majority is important, not the number of votes (Bassett & Persky 1999; Shalikaran *et al.* 2011).

More details about the social choice methods and their applications are shown in the supplementary material (Supplementary Material Example S1, Tables S2 to S4, available with the online version of this paper).

CASE STUDY

The reservoir of Karkheh Dam is selected due to numerous water quality problems such as eutrophication. It is a large multi-purpose earth dam in Iran on the Karkheh River. The capacity of the reservoir (at normal level) is 5,900 million cubic metres. Figure 2 shows the Karkheh River and its reservoir. Five quality variables – including total dissolved solids (TDS), total phosphorus (TP), total nitrogen (TN), DO and temperature – are selected for this study. Twenty-two potential monitoring stations are selected along the length of the reservoir to sample from different depths (Supplementary Material Figure S1 and Table S5, available with the online version of this paper). The importance coefficients of the water quality variables (γ_q in Equations (3) and (4)) are shown in Table 2. Each coefficient prioritizes a water quality variable in the quality monitoring program. Once the CEQUAL-W2 model is calibrated and validated using the historic dataset, it is used to simulate reservoir water quality for 40 years in the period of 1968–2008.

RESULTS AND DISCUSSION

The transinformation quantity is calculated (Equation (1)) for each potential monitoring station when it is paired with another potential monitoring station for each water quality variable in

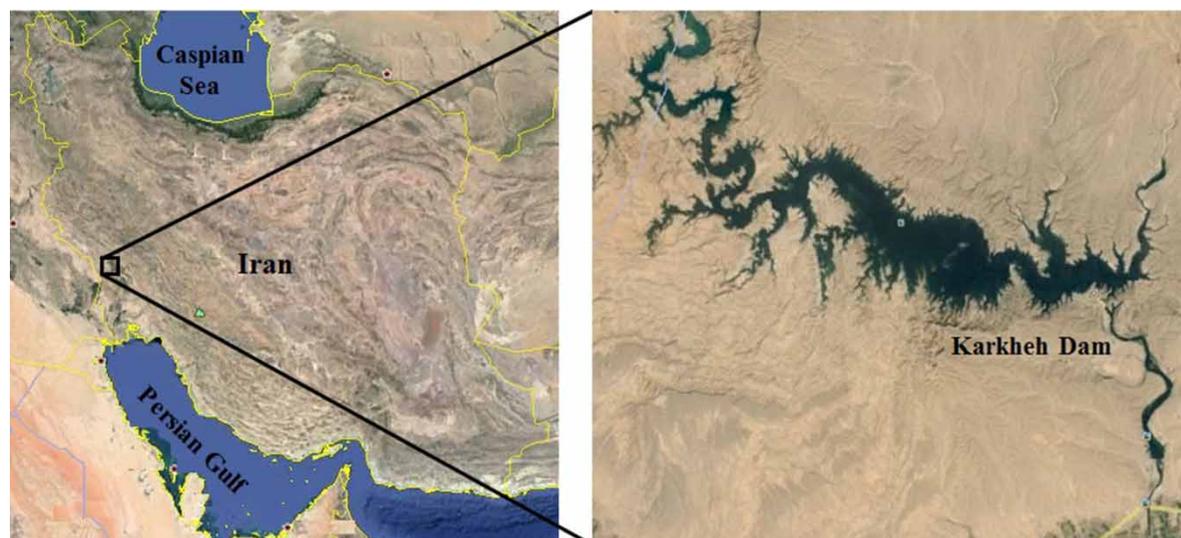


Figure 2 | Google Earth view of the Karkheh dam, 2016.

Table 2 | The importance coefficients of the water quality variables (adapted from Guideline for Water Quality Monitoring of Dam Reservoirs 2011)

Variables	Aquatic life and fisheries	Drinking water	Irrigation	Power generation	Importance coefficient
TDS	*	*	***	***	0.30=8/27
TP	***	***			0.22=6/27
TN	*	***			0.15=4/27
DO	***	*	*	*	0.22=6/27
Temperature	***				0.11=3/27

Low to high importance is represented by * to ***.

four seasons. The transinformation quantities (T) are plotted versus the spatial distances (D) among monitoring stations to obtain the T - D curve and determine the mutual information for each variable across seasons. The T - D curves that are obtained for TDS for each season are demonstrated in Figure 3. As shown in this figure, the redundant or mutual information (transinformation) among monitoring stations decreases as the distance between two monitoring stations increases. The

D_{opt} (the optimal distance) is selected when the transinformation entropy reaches an equilibrium (insignificant changes when distance increases). If the spatial distances among monitoring stations are less than D_{opt} , there is mutual or redundant information among monitoring stations. For spatial distances more than D_{opt} , the spatial coverage of the water quality monitoring network is not enough. The distance D_{opt} corresponds to the minimum value of transinformation on the T - D curve (the

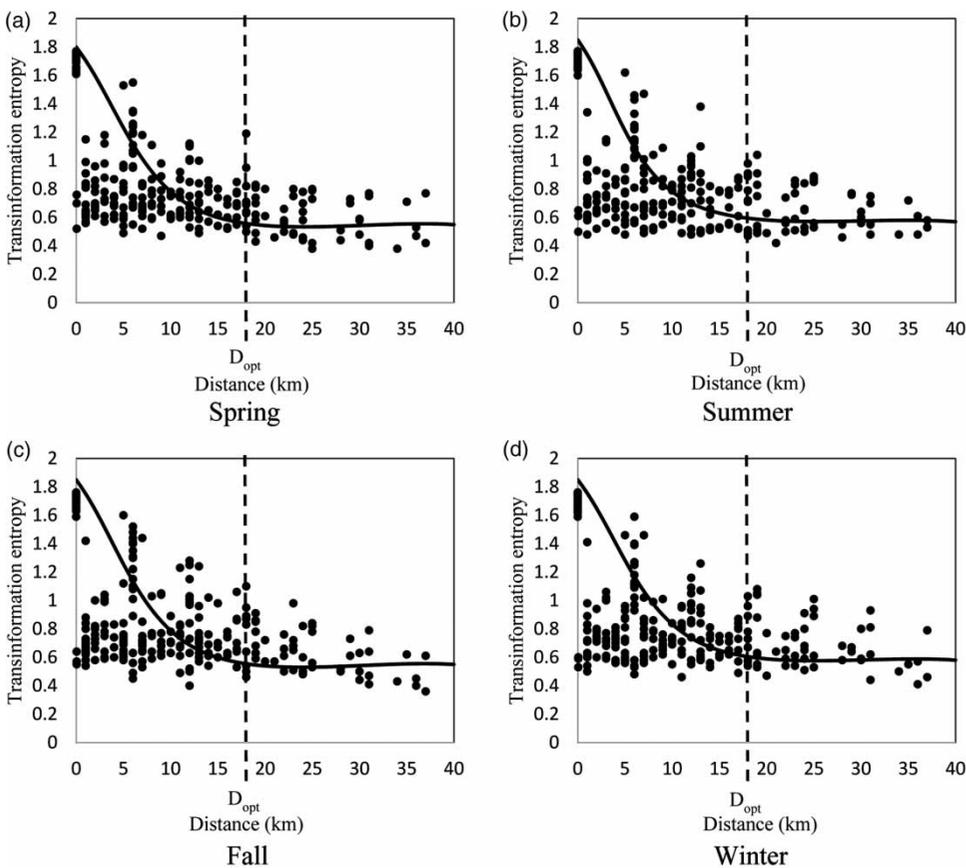


Figure 3 | TDS T - D curves represent transinformation entropy variation with respect to the distance among monitoring stations. These curves are obtained using the simulated data for 40 years (1968–2008).

dashed line in Figure 3). The optimal distance is approximately 18 km for TDS for all seasons. The optimal distance is between 17 and 18 km for each water quality variable across all seasons. Due to the space available in this paper, the minimum distance (17 km) is defined as the optimal distance for all water quality variables and used in Equations (2)–(7) (Supplementary Material Figures S2 to S5, available with the online version of this paper).

Considering the T - D curves and the objectives of the stakeholders, the multi-objective optimization model (described above) is characterized and solved by the NSGA-II. The number of binary decision variables is 22 (equal to 22 potential stations). The NSGA-II multi-objective optimization algorithm is tested for various population and generation numbers. After

an experiment for the size of population and generation, a population size of 330 ($15 \times 22 = 330$) individuals is selected when the NSGA-II algorithm is stopped after 200 generations.

By obtaining the non-dominated solutions, the social choice methods are applied and the preference of the solutions from each of the three stakeholders. The results of the compromise solution based on different social choice methods are presented in Table 3. As shown in this table, the number of optimized monitoring stations is the same (three out of 22 potential stations) across all seasons using the four social choice methods (described above).

This result shows that the monitoring locations vary across seasons while the number of locations remains the same. The result shows that a more complex case study is required to analyze the sensitivity of the social choice method on the final solution, because all social choice methods provide three locations. Obtaining the same results using different social choice methods in this study is in agreement with the research conducted by Sheikhmohammady & Madani (2008). The locations of the optimal water quality monitoring locations are shown in Figure 4. As shown in this figure, most optimal monitoring locations are at lower depths.

Table 3 | Results of the compromise solution based on different social choice methods

Method	Season	Number of optimized stations	Optimal locations
Condorcet,	Spring	3	(64,27) (65,45) (59,45)
Borda,	Summer	3	(57,33) (62,39) (50,39)
Plurality rule,	Fall	3	(47,21) (52,27) (50,39)
and MVR	Winter	3	(64,27) (50,39) (64,51)

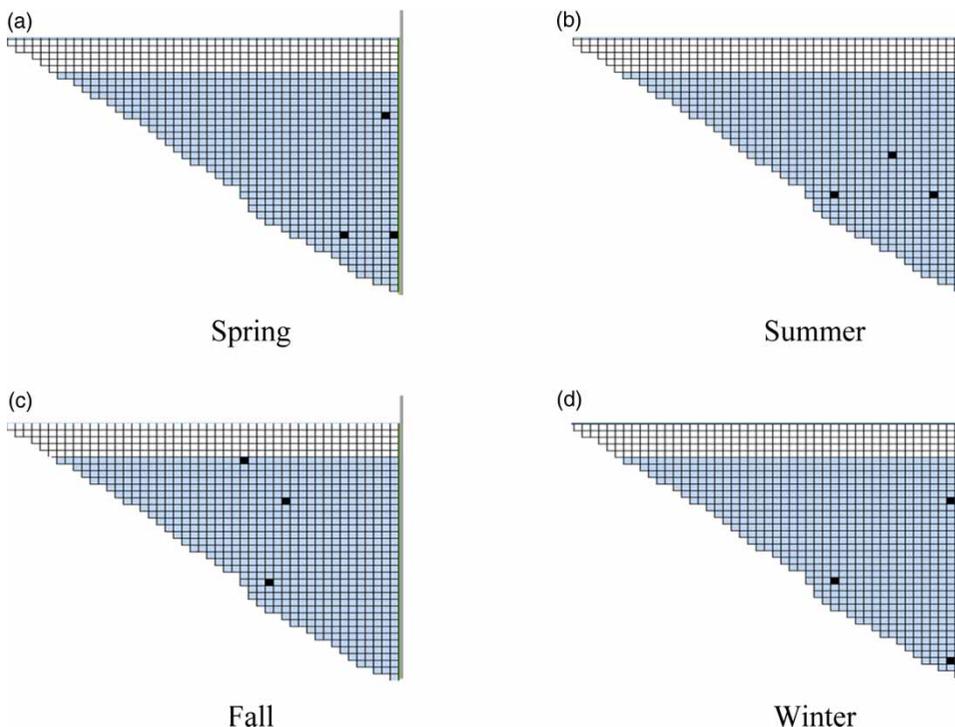


Figure 4 | Optimal locations to sample reservoir water quality.

SUMMARY AND CONCLUSION

In this study, a new methodology based on the CE-QUAL-W2 model, NSGA-II optimization method, transinformation entropy theory, and social choice theory is proposed for optimization of water quality monitoring stations of dam reservoirs. The transinformation entropy quantities and the $T-D$ curves are obtained to determine the mutual information among monitoring stations. Considering the utilities of the existing stakeholders, a multi-objective NSGA-II optimization model is applied to find Pareto-optimal solutions. Then, the best solution is determined based on social choice methods to achieve a common option that is agreed among all stakeholders. The variations of water quality variables in different seasons are investigated. The potential monitoring stations are located at different depths along the length of the reservoir. The proposed approach is applied for optimization of water quality monitoring stations for the Karkheh reservoir in Iran. The optimal distance among stations is considered to be 17 km. The number of optimized monitoring locations is the same (three out of 22 potential stations) across all seasons using different social choice methods, however, these locations vary across seasons. The results show that NSGA-II is able to converge to a set of non-dominated solutions and social choice methods are effective ways for finding a compromise solution.

In future work, the temporal frequency of monitoring locations can be studied, and also, a spatial variogram can be used to determine the nugget (minimum value of mutual information and its corresponding distance) instead of transinformation entropy. The effect of a more accurate water quality model such as WASP should also be examined.

REFERENCES

- Bassett, G. W. & Persky, J. 1999 Robust voting. *Public Choice* **99**, 299–310.
- Behmel, S., Damour, M., Ludwig, R. & Rodriguez, M. J. 2016 Water quality monitoring strategies – a review and future perspectives. *Science of the Total Environment* **571**, 1312–1329.
- Deb, K., Agrawal, S., Pratap, A. & Meyarivan, T. 2002 A fast and elitist multi-objective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* **6** (2), 182–197.
- Debele, B., Srinivasan, R. & Parlange, J. Y. 2008 Coupling upland watershed and downstream water body hydrodynamic and water quality models (SWAT and CE-QUAL-W2) for better water resources management in complex river basins. *Environmental Monitoring and Assessment* **13**, 135–153.
- Guideline for Water Quality Monitoring of Dam Reservoirs. 2011 Issue No. 551, Iran (in Persian).
- Harmancioglu, N. B. 1981 Measuring the information content of hydrological processes by the entropy concept. *Journal of Civil Engineering* (Ege University, Izmir, Turkey, Special Issue for the Centennial of Atatürk's Birth), 13–40.
- Harmancioglu, N. B., Fistikoglu, O., Ozkul, S. D., Singh, V. P. & Alpaslan, N. 1999 *Water Quality Monitoring Network Design*. Kluwer, Boston, MA, USA.
- Karamouz, M., Nokhandan, A. K., Kerachian, R. & Maksimovic, C. 2009 Design of on-line river water quality monitoring systems using the entropy theory: a case study. *Environmental Monitoring and Assessment* **155** (1–4), 63–81.
- Lee, Y. & Kwon, S. 2009 A study on measuring the similarity among sampling sites in Lake Yongdam with water quality data using multivariate techniques. *Environmental Impact Assessment* **18** (6), 401–409.
- Lee, Y., Kwon, S., Lee, S. & Ban, Y. 2011 Construction and application of network design system for optimal water quality monitoring in reservoir. *Journal of Korea Water Resources Association* **44** (4), 295–304.
- Lee, C., Paik, K. & Lee, Y. 2014 Optimal sampling network for monitoring the representative water quality of an entire reservoir on the basis of information theory. *Journal of Water and Climate Change* **5** (2), 151–162.
- Loucks, D. P., Beek, E. V., Stedinger, J. R., Dijkman, J. P. M. & Villars, M. T. 2005 *Water Resources Systems Planning and Management*. UNESCO Publishing, Paris, France.
- Mahjouri, N. & Kerachian, R. 2011 Revising river water quality monitoring networks using discrete entropy theory: the Jajrood River experience. *Environmental Monitoring and Assessment* **175**, 291–302.
- Masoumi, F. & Kerachian, R. 2010 Optimal redesign of groundwater quality monitoring networks: a case study. *Environmental Monitoring and Assessment* **161** (1–4), 247–257.
- McLean, I. 1990 The Borda and Condorcet principles: three medieval applications. *Social Choice and Welfare* **7**, 99–108.
- Memarzadeh, M., Mahjouri, N. & Kerachian, R. 2013 Evaluating sampling locations in river water quality monitoring networks: application of dynamic factor analysis and discrete entropy theory. *Environmental Earth Sciences* **70** (6), 2577–2585.
- Mogheir, Y. & Singh, V. P. 2002 Application of information theory to groundwater quality monitoring networks. *Journal of Water Resources Management* **16**, 37–49.
- Mogheir, Y., de Lima, J. L. M. P. & Singh, V. P. 2004a Characterizing the spatial variability of groundwater quality using the entropy theory: I. Synthetic data. *Hydrological Processes* **18**, 2165–2179.

- Mogheir, Y., de Lima, J. L. M. P. & Singh, V. P. 2004b [Characterizing the spatial variability of groundwater quality using the entropy theory: II. Case study from Gaza Strip](#). *Hydrological Processes* **18**, 2579–2590.
- Mogheir, Y., de Lima, J. L. M. P. & Singh, V. P. 2009 [Entropy and multi-objective based approach for groundwater quality monitoring network assessment and redesign](#). *Water Resources Management* **23** (8), 1603–1620.
- Mondal, N. C. & Singh, V. P. 2012 [Evaluation of groundwater monitoring network of Kodaganar River basin from Southern India using entropy](#). *Environmental Earth Sciences* **66**, 1183–1193.
- Ozkul, S., Harmancioglu, N. B. & Singh, V. P. 2000 [Entropy-based assessment of water quality monitoring networks](#). *Journal of Hydrologic Engineering* **5** (1), 90–100.
- Shalikaran, L., Madani, K. & Naeeni, S. 2011 [Finding the socially optimal solution for California's Sacramento–San Joaquin Delta problem](#). In: *World Environmental and Water Resources Congress 2011: Bearing Knowledge for Sustainability*, R. E. Beighley II & M. W. Killgore (eds), ASCE, Reston, VA, USA, pp. 3190–3197.
- Sheikhmohammady, M. & Madani, K. 2008 [Bargaining over the Caspian Sea – the largest lake on the Earth](#). In: *World Environmental and Water Resources Congress 2008: Ahupua'a*, R. W. Babcock & R. Walton (eds), ASCE, Reston, VA, USA.
- Yenilmez, F., Düzgün, S. & Aksoy, A. 2015 [An evaluation of potential sampling locations in a reservoir with emphasis on conserved spatial correlation structure](#). *Environmental Monitoring and Assessment* **187**, 4216.

First received 17 August 2016; accepted in revised form 21 November 2016. Available online 8 December 2016