Title

Optimal Selective Withdrawal Rules Using a Coupled Data Mining Model and Genetic Algorithm

Journal
Journal of Water Resources Planning and Management, 142(12)

ISSN
0733-9496

Authors
Soleimani, Shima
Bozorg-Haddad, Omid
Saadatpour, Motahareh
et al.

Publication Date
2016-12-01

DOI
10.1061/(asce)wr.1943-5452.0000717

Peer reviewed

Optimal Selective Withdrawal Rules Using a Coupled Data Mining Model and Genetic Algorithm

Shima Soleimani[1]; Omid Bozorg-Haddad[2]; Motahareh Saadatpour[3]; and Hugo A. Loáiciga, F.ASCE[4]

Abstract: This work presents a methodology for extracting optimal operational rules for selective reservoir water withdrawal by considering fixed levels of reservoir water outlets for thermal control. The outlet water temperature of the Karkheh reservoir, Iran, is simulated with the CE-QUAL-W2 model. A data-mining model (the LIBSVM model) is applied as a surrogate model of the CE-QUAL-W2 model and coupled with a genetic algorithm (GA), resulting in the LIBSVM-GA algorithm. The selective withdrawal approach considered four fixed reservoir outlets, located at 120, 140, 163, and 181 m above sea level, to account for reservoir thermal stratification. This paper's methods are evaluated with nonselective and selective withdrawal operations through different scenarios in which single outlet, fixed withdrawal proportions, fixed monthly variable proportions, continually variable (10-day) proportions using total monthly LIBSVM input data, and continually variable (10-day) proportions using separated monthly LIBSVM input data are considered. The highest outlet (at 181 m) was found to be the best level for the nonselective withdrawal scenario. The best selective withdrawal operations scenario was the continually variable (10-day) proportions using separated monthly LIBSVM input data, which minimize the root-mean-square deviation (RMSD) between upstream and downstream temperatures during the operating period. **DOI: 10.1061/(ASCE)WR.1943-5452.0000717.**

Author keywords: CE-QUAL-W2 model; Environmental temperature regime; Genetic algorithm (GA); LIBSVM model; LIBSVM-GA algorithm; Selective withdrawal.

Introduction

Seasonal air temperature fluctuations lead to thermal stratification in reservoirs. This phenomenon can have major effects on reservoir water quality parameters, including water temperature. Thermal control in reservoirs can help manage the outlet water temperature and the water temperature regime in river reaches downstream of reservoirs (Hocking et al. 1988; Elçi 2008; Giuliani et al. 2013; Rheinheimer et al. 2015).

Key external factors affecting thermal stratification are the level at which reservoir water is withdrawn and the magnitude and timing of withdrawals. The latter forms a policy of selective withdrawal operations of a reservoir that can be optimized. Furthermore, withdrawing water from multilevel outlets causes mixing among reservoir thermal layers, which generally improves reservoir water quality (Çalışkan and Elçi 2009).

Fontane and Labadie (1981) applied the WESTEX model to simulate the thermal stratification cycle of a reservoir and linked the simulation model (WESTEX) to objective-space dynamic programming (OSDP) to achieve optimal control of reservoir discharge quality through selective withdrawal structure of multilevel water outlets and maintain the natural temperature regime downstream of the reservoir. Hanna et al. (1999) employed the CE-QUAL-W2 model to simulate thermal stratification coupled with a selective withdrawal system. Gelda and Effler (2007) defined different scenarios for reservoir release of the Schoharie reservoir in New York City in the United States. They linked the CE-QUAL-W2 simulation model to an evolutionary optimization algorithm and showed that using selective withdrawal can decrease the epilimnion and metalimnion in the reservoir. Saadatpour and Afshar (2011) applied simulation-optimization and selective withdrawal to calculate optimal rules of selective withdrawal operations in Ilam reservoir, Iran, to guide responses to pollutant spills.

Castelletti et al. (2014) developed a combinatorial simulation model which involved the DYRYSM model and a computational aquatic ecosystem dynamics model (CAEDYM) that simulated hydrodynamics and environmental processes. A batch-mode reinforcement learning algorithm called Fitted Q-Iteration was applied to calculate optimal rules of selective withdrawal operations in the multioutlet Tano reservoir, Japan. The research objective was to improve water quality parameters such as temperature and total suspended solids (TSS) in and below the reservoir. Soltani et al. (2010) applied a one-dimensional simulation model developed by Kerachian and Karamouz (2006) to simulate salinity in the 15-Khordad reservoir, Iran, and applied the hybrid genetic reservoir operation rules. The input and output of the simulation model were trained to adaptive neural fuzzy inference system (ANFIS), which is a data-mining algorithm, and four different ANFIS models were trained and tested to predict average salinity of the reservoir's outlet water, epilimnion, metalimnion, and hypolimnion at the end of each month. Mirfendereski and Mousavi (2011) optimally allocated water resources in the Atrak basin of Iran incorporating support vector machine (SVM) and response surface method as surrogates of the MODSIM model in the optimization module

[1]M.Sc. Student, Dept. of Irrigation and Reclamation, Faculty of Agricultural Engineering and Technology, College of Agriculture and Natural Resources, Univ. of Tehran, Karaj, 31587-77871 Tehran, Iran. E-mail: shimasoleimani@ut.ac.ir

[2]Professor, Dept. of Irrigation and Reclamation, Faculty of Agricultural Engineering and Technology, College of Agriculture and Natural Resources, Univ. of Tehran, Karaj, 31587-77871 Tehran, Iran (corresponding author). E-mail: obhaddad@ut.ac.ir

[3]Assistant Professor, School of Civil Engineering, Iran Univ. of Science and Technology, 16846-13114 Tehran, Iran. E-mail: msaadatpour@iust.ac.ir

[4]Professor, Dept. of Geography, Univ. of California, Santa Barbara, CA 93106-4060. E-mail: hugo.loaiciga@ucsb.edu

Note. This manuscript was submitted on February 9, 2016; approved on July 13, 2016; published online on August 24, 2016. Discussion period open until January 24, 2017; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Water Resources Planning and Management***, © ASCE, ISSN 0733-9496.

and comparing the performance of the data-mining models. Shokri et al. (2014) considered multiobjective quantity and quality operation of the Karaj reservoir located in Iran. The CE-QUAL-W2 model was applied to simulate methyl tertiary-butyl ether (MTBE) and benzene pollutants. The artificial neural network (ANN) was trained and tested with the simulation results and was coupled with the nondominated sorting genetic algorithm (NSGA-II) to yield the NSGA-II-ANN model. This scheme reduced the computational time required for model implementation.

Despite the numerous data-mining techniques that have been recently widely used in different fields of water resources investigations (Fallah-Mehdipour et al. 2012, 2013a, b, c; Akbari-Alashti et al. 2014; Ashofteh et al. 2015; Soltanjalili et al. 2013), the hybrid LIBSVM-GA has not been reported to address real-world problems, as done in the present study.

Storage of water and thermal stratification in reservoirs can render the water quality of reservoir outflow much different from that of reservoir inflow. The primary objective of reservoir releases in general are to meet water supply demands, generate hydropower, and manage flood risk, at least from an operator's perspective. Besides, releasing to maintain the natural temperature regime would typically be prioritized if legally mandated. Downstream water temperature is significant for pollution control, agriculture, and to support the downstream aquatic ecology (Fontane and Labadie 1981). The main objective of this study is to minimize the deviation between the river water temperature downstream from the reservoir and the natural temperature regime by applying selective withdrawal. The water quality CE-QUAL-W2 model requires time series of meteorological, hydrological, and water quality at short-term time steps (hourly) to simulate the reservoir's outlet water temperature. Therefore, solving the numerical equations of hydrodynamics and water quality with the CE-QUAL-W2 model is computationally burdensome and renders simulation-optimization approaches impractical. To avoid such a burden this study applies an algorithm as a surrogate of the CE-QUAL-W2 model. The algorithm is trained and tested with simulated data obtained from the CE-QUAL-W2 model and used to estimate outlet water temperature in the Karkheh reservoir, Iran. The best trained model is linked with a genetic algorithm (GA), which is used to calculate optimal selective withdrawal rules considering several water-release scenarios.

Methods

This section introduces the two-dimensional CE-QUAL-W2 model for simulating thermal stratification and reservoir outlet water temperature. The second part explains the SVM data-mining algorithm. The third part summarizes the GA's solution process as an evolutionary optimization algorithm. The fourth part defines the objective function of the optimization model and assesses the efficiency of the developed GA algorithm.

CE-QUAL-W2 Model

CE-QUAL-W2 is a two-dimensional hydrodynamics and water quality model developed and supported by the United States Army Engineering Waterways Experiment Station (WES). The spatial and temporal changes in reservoir surface elevation and temperature are simulated with the CE-QUAL-W2 Version 3.71 (Cole and Wells 2008). Multilevel outlets and selective withdrawal options are determined for CE-QUAL-W2 to release water from different elevations and proportions from outlets during reservoir operational period. Thermal stratification is simulated with

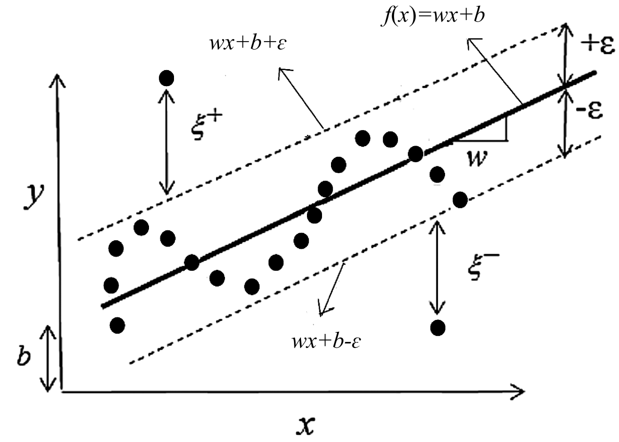

Fig. 1. Soft margin of ε-insensitive in the SVM_ε model

CE-QUAL-W2 to calculate the water temperature at both the outlets (each individual selective withdrawal outlet) and the main reservoir outlet. The temperature at the outlets is controlled by the temperature of the thermal layer at the outlet elevation (though there may be some local mixing), whereas the temperature at the reservoir outlet is the result of mixing of temperatures from the different layers.

Support Vector Machine

The SVM regression model approximates the functional dependence of the dependent variable y on independent variables $\boldsymbol{x}$. The function relating dependent and independent variables involves a structural term and an error term (or noise) as described by Eq. (1) (Vapnik 1995, 1998, 1999):

$$y = f(\boldsymbol{x}) + \text{noise} \tag{1}$$

The SVM estimates the function f to predict variables with new data. The function f is estimated by training the SVM on a set of data called the training (or calibration) data set using a process that optimizes the error function continuously. Vapnik (1999) introduced an error function called ε-*insensitive* in the SVM_ε model. Fig. 1 shows the soft margin in which the errors are not considered a linear SVM_ε model. The main objective in SVM_ε modeling is to minimize the error function given by Eq. (2) subject to the constraints expressed by Eq. (3) (Vapnik 1999):

$$\text{Minimize } \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} + C\sum_{i=1}^{I}(\xi_i + \xi_i^*) \tag{2}$$

$$\text{Subject to} \begin{cases} \boldsymbol{w}^T\phi(\boldsymbol{x}_i) + b - y_i \le \varepsilon + \xi^* \\ y_i - \boldsymbol{w}^T\phi(\boldsymbol{x}_i) - b \le \varepsilon + \xi \\ \xi_i, \xi_i^* \ge 0 \qquad i = 1, 2, 3, \ldots, I \end{cases} \tag{3}$$

where $\boldsymbol{w}$ = coefficient vector; $\boldsymbol{w}^T$ = transposed coefficient vector; ξ and ξ^* = slack variables; b = constant variable; I = total number of trained data; C = capacity coefficient; i = trained data counter; and ϕ = nonlinear mapping function. There is a lack of information about the choice of nonlinear mapping functions. Therefore, a kernel function equal to a radial basis function (RBF) is applied in this study (Chang and Lin 2011):

$$K(x_i, x) = \exp\left(-\frac{\|x_i - x\|^2}{2\sigma^2}\right) \qquad (4)$$

where K = kernel function; and σ = constant coefficient of the RBF kernel function.

In this study, the SVM_ε modeling was executed with the LIBSVM code, which was developed by Chang and Lin (2011). The LIBSVM code has been applied in many studies (Safavi and Esmikhani 2013; Dixon 2009)

Genetic Algorithm

The genetic algorithm is inspired by the theory of evolution and was introduced by Holland (1975). The steps of the GA are as follows:

At first, a set of random solutions called populations are generated by the GA. Each population contains many chromosomes, whereby each chromosome is a solution and contains many genes that are the decision variables. The generated population is replaced by a new population during the iterative procedure. In the each iteration, called generation, each of the solutions is evaluated by a fitness function and then some best solutions are selected by a selection operator and are promoted to the next generation. The selection operator uses a probability distribution function that increases the chance of promoting the best solutions to the next generation. Some chromosomes or solutions are sent to the next generation without any changes and the rest of the solutions are processed to produce children with the crossover and mutation operators. The crossover operator pairs parents to produce children (chromosomes). The crossover rate is named P_c and expresses the proportion of the number of produced children of each generation to the number of members in the present generation. The mutation operator is named P_m and expresses the proportion of the number of the mutated genes of each generation to the number of genes of the present generation. The mutation operator only uses one existing population member to produce a child. Each gene is randomly selected with a uniform distribution and the value of genes changes. Commonly the mutation operator is applied after the crossover operator.

The preceding process is repeated as long as the difference between the values of two consecutive objective functions pertaining to two consecutive iterations exceeds a convergence criterion. Otherwise, the iterations are terminated.

Objective Function

There are lots of aims to control the water temperature downstream of a reservoir. It is necessary to explain that the objective function can be selected to maximize or minimize each of the statistical criteria related to minimum, maximum, or average deviation of water temperature. According to the purpose of temperature control, some statistical criteria can be considered such as minimizing water temperature for fishery purpose in a downstream river. In this study, the purpose is to maintain upstream river temperature in downstream river temperature. Therefore, the objective function is defined to minimize the RMSD between the reservoir outflow temperature and the reservoir inflow temperature during the operational period. The GA's decision variables are the withdrawal proportion at each of the reservoir outlets, and it minimizes the nonlinear function defined by Eq. (5):

$$\text{Minimize RMSD} = \sqrt{\frac{1}{N}\sum_{t=1}^{t=N}(T^t_{\text{Outflow}} - T^t_{\text{Inflow}})^2} \qquad (5)$$

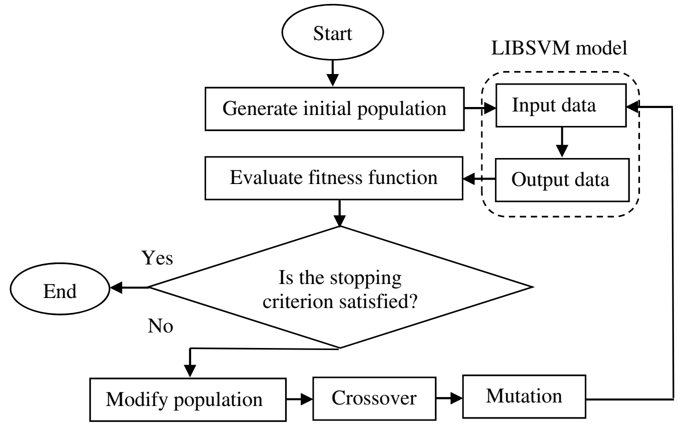

Fig. 2. LIBSVM-GA flowchart

where T^t_{inflow} = mean reservoir inflow temperature in the tth time step; T^t_{outflow} = mean reservoir outflow temperature in the tth time step; t = time interval counter; and N = total number of time steps during the period of selective withdrawal operations.

LIBSVM-GA Model

Several selective withdrawal operations options were defined independently of the CE-QUAL-W2 model to simulate water outflow temperature. The options were based on the withdrawal proportion at each of several outlets located at different, fixed elevations. The reservoir surface elevation, meteorological, and hydrological data used as input to the CE-QUAL-W2 model are selected so that they have the greatest effect on reservoir outflow temperature and are introduced as the input data to the LIBSVM model. In this manner the LIBSVM model is trained (that is, calibrated) and tested with inputs that, in turn, are selected outputs of the CE-QUAL-W2 model for different operational options with different withdrawal proportions of each reservoir outlet. The LIBSVM model with the best efficiency considering different statistic criteria is selected as a surrogate model of the CE-QUAL-W2 model. The GA is linked with the surrogate model (LIBSVM) to yield the developed LIBSVM-GA model, which is implemented to determine the optimal withdrawal outlets and the withdrawal proportions.

As shown in Fig. 2, at first an initial population is produced by the GA. The population consists of the withdrawal proportion at each reservoir outlet. Then, the withdrawal proportion at each reservoir outlet (GA population) with reservoir surface elevation, meteorological, and hydrological data that have the greatest influence on reservoir outflow temperature are defined as input data sets to the LIBSVM model. The LIBSVM model generates reservoir outflow temperature in every time step. The GA's fitness function is evaluated for every solution in the initial population based on the LIBSVM output. The population of solutions is modified based on the fitness function evaluation, and the modified population becomes input to the LIBSVM model for re-evaluation. This iterative procedure is repeated until a convergence criterion is satisfied.

Case Study

The Karkheh reservoir is the sixth largest reservoir in the world and the largest reservoir in Iran. The catchment area of the Karkheh reservoir is approximately 43,000 km^2 and is located between 46° 57′ to 49° 10′ eastern longitudes and 31° 48′ to 34° 58′ northern

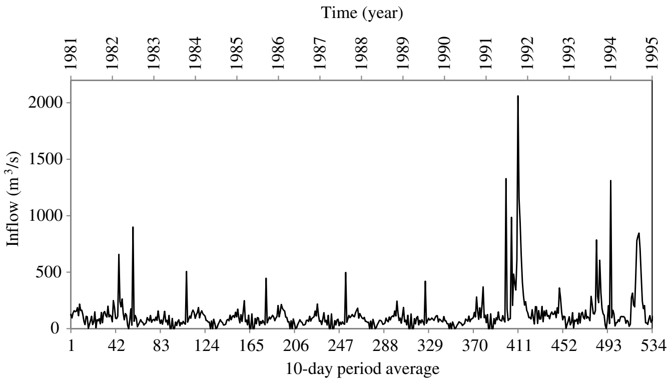

Fig. 3. 10-day average reservoir inflow during the operational period

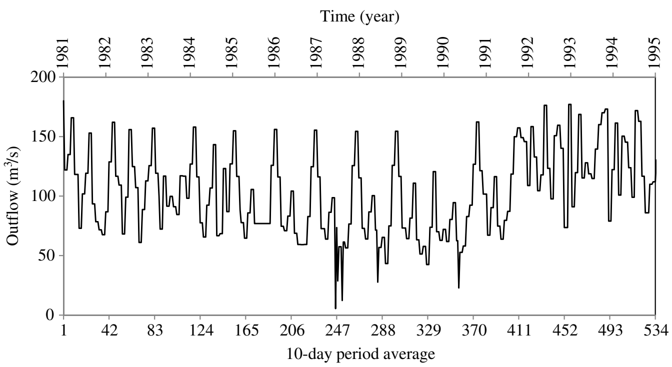

Fig. 4. 10-day average reservoir outflow during the operational period

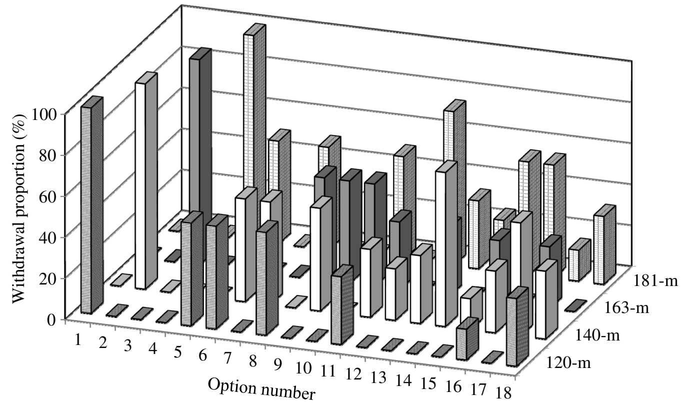

Fig. 5. Withdrawal proportion of each outlet for different operational options

latitudes. The Karkheh reservoir has a surface area of 162 km^2, a length of 64 km, and a volume of 5.9 km^3 at normal water elevation (220 m above sea level). The height of the Karkheh dam foundation is 127 m and the dam crest level is 234 m above sea level. The average and maximum depth of the Kharkhe reservoir are 61.8 and 117 m, respectively (Afshar et al. 2012).

It is essential to simulate outlet water temperature at different levels to apply selective withdrawal. The Karkheh reservoir has two outlets located at 163 and 181 m above sea level. Two other outlets located at 120 and 140 m above sea level were assumed. The shape type of all outlets was defined for the CE-QUAL-W2 model, which was calibrated and validated by Afshar and Saadatpour (2009). This study employs that calibrated model updated to version 3.71. The first and last simulation days in the CE-QUAL-W2 model were September 19, 1980, and September 12, 1995, respectively. These choices correspond to the available time series of meteorological and hydrologic data. The time span for reservoir operations was defined from January 1, 1981, to December 31, 1994. Ten-day average reservoir inflow and 10-day average reservoir outflow during the operational period are shown in Figs. 3 and 4, respectively.

The training and testing input data of the LIBSVM model are obtained from inputs and outputs of 18 defined operational options simulated independently using the CE-QUAL-W2 model in which the results of 12 and 6 operational options are used as training and testing data, respectively. The withdrawal proportions of each outlet for different operational options are depicted in Fig. 5. This study implemented as input data to the LIBSVM model the 10-day average air temperature, wind speed, input thermal flux to the reservoir (multiplication of inflow discharge times inflow water temperature), reservoir surface elevation, the withdrawal volume from the reservoir and the withdrawal proportion at each reservoir outlet. This input data was chosen from the CE-QUAL-W2 model's outputs for the Karkheh reservoir reported by Afshar and Saadatpour (2009). In addition, two different versions of the LIBSVM model are defined based on the selection of input data and time series length. Each of these two LIBSVM versions are independently trained and tested.

Version 1: In this version the time series input contains all 10-day time steps of total months during the operational period.

Version 2: In this version the LIBSVM models are independently trained and tested, one for each calendar month, using data from the operational period.

Scenario

Several scenarios of nonselective and selective withdrawal operations are defined as follows:

Scenario 1: Nonselective withdrawal

In this scenario release water temperature at each single outlet is simulated using CE-QUAL-W2 during the operational period. The sub-scenarios are defined as

- Scenario 1.1: 120-m outlet only
- Scenario 1.2: 140-m outlet only
- Scenario 1.3: 163-m outlet only
- Scenario 1.4: 181-m outlet only

Scenario 2: Selective withdrawal with fixed proportions

In this scenario the GA is linked to the LIBSVM version 1, and the release proportion at each outlet is fixed for the duration of operational period.

Scenario 3: Selective withdrawal with fixed monthly variable proportions

In this scenario the GA is linked to each calendar month of the LIBSVM version 2 and the release proportion at each outlet is allowed to vary for each calendar month, such that there is a monthly release proportion rule curve repeated every year.

Scenario 4: Selective withdrawal with continually variable (10-day) proportions using total monthly LIBSVM input data

In this scenario the GA is linked to the LIBSVM version 1, and the release proportion at each outlet is allowed to change each 10-day interval during the operational period.

Scenario 5: Selective withdrawal with continually variable (10-day) proportions using separated monthly LIBSVM input data

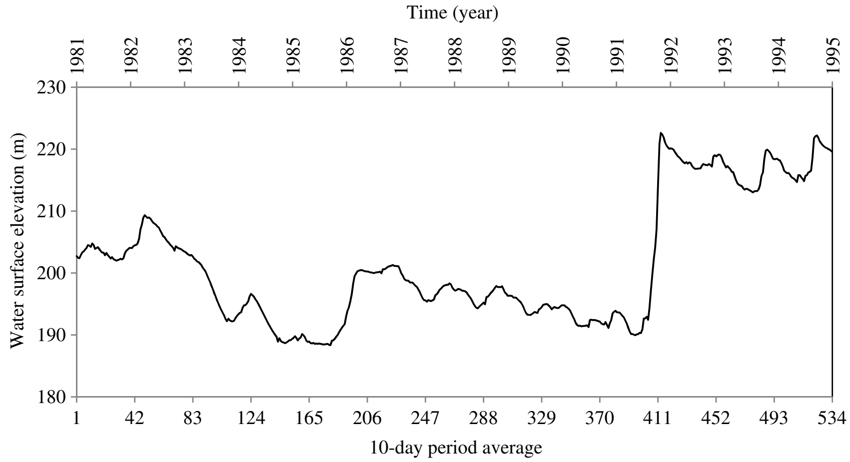

Fig. 6. Reservoir surface elevation of the reservoir during the operational period

In this scenario the GA is linked to the LIBSVM version 2, and the release proportion at each outlet is allowed to change each 10-day interval during the entire operational period.

Results

This section presents the LIBSVM verification results. In addition, the results of reservoir surface elevation simulated by CE-QUAL-W2 are considered. Finally, the results of each operational scenario are discussed.

Verification of the LIBSVM Model

The chosen statistical criteria were the RMSD, mean absolute deviation (MAD), and the Nash–Sutcliffe efficiency (NSE), which were used to evaluate the accuracy of testing for two versions of the LIBSVM (Moriasi et al. 2007). The RMSD, MAD, and NSE for version 1 were 0.49, 0.30, and 0.99°C, respectively, and the means of the RMSD, MAD, and NSE for all months of version 2 were 0.40, 0.30, and 0.97°C, respectively. The results of the statistical criteria for the two LIBSVM versions demonstrated the capacity of the LIBSVM model to approximate the outlet water temperature. For this purpose, the two versions of the LIBSVM model exhibited high reliability, and for this reason its results can be used as a surrogate of the CE-QUAL-W2 model to be linked with the GA.

Reservoir Surface Elevation and Withdrawal Options

The mean surface elevation of the Karkheh reservoir for each 10-day step during the operational period was simulated with the CE-QUAL-W2 model and is graphed in Fig. 6 in which the first and last numbers of time step are equal to January 1, 1981, and December 31, 1994, respectively. As Fig. 6 shows, the reservoir surface elevation was never lower than the highest outlet. Therefore, all outlets could be used during the entire operational period.

Scenario 1: Nonselective Withdrawal

In this scenario, it was assumed that there is no selective withdrawal for the entire operational period and, if there is only one outlet for the reservoir, which outlet is the optimal outlet to release water. The RMSD, MAD, and coefficient of determination (R^2) were used to select the best subscenario of Scenario 1 and the best entire scenario.

Table 1 shows that according to Scenario 1 the highest (worst) value of RMSD and MAD, which are 6.53 and 5.53°C, respectively, correspond to Scenario 1.2. And the lowest (best) values of RMSD and MAD, which are 5.71 and 4.70°C, respectively, correspond to Scenario 1.4. R^2 equals 99% for all outlets' elevations. Therefore, if only one outlet were used to release water, the 181-m outlet would be the best to minimize the deviation of the inflow and outflow water temperature of the reservoir.

Scenario 2: Selective Withdrawal with Fixed Proportions

According to first row of Table 1, the optimum fixed withdrawal proportions at each outlet calculated with the LIBSVM-GA were 0.35, 0.00, 0.00, and 0.65 at the 120, 140, 163, and 181-m outlets, respectively. The results show that when withdrawal proportions were fixed during the operational period, the optimal outlets were the 120 and 181-m outlets, which are the lowest and highest outlets among all outlets. In addition, the optimal proportion of the 181-m outlet (upper outlet) is approximately 1.85 times larger than the optimal proportion of the 120-m outlet.

Scenario 3: Selective Withdrawal with Fixed Monthly Variable Proportions

The optimal fixed monthly variable proportions at each outlet calculated with the LIBSVM-GA algorithm and the results of

Table 1. Calculated Statistical Criteria for Each Operational Scenario

Scenario	RMSD (°C)	MAD (°C)	R^2
1.1	6.36	5.38	0.99
1.2	6.53	5.53	0.99
1.3	6.30	5.31	0.99
1.4	5.71	4.70	0.99
2	5.53	4.51	0.99
3	5.32	4.58	0.99
4	5.20	4.27	0.99
5	5.15	4.14	0.99

J. Water Resour. Plann. Manage., 2016, 142(12): 04016064

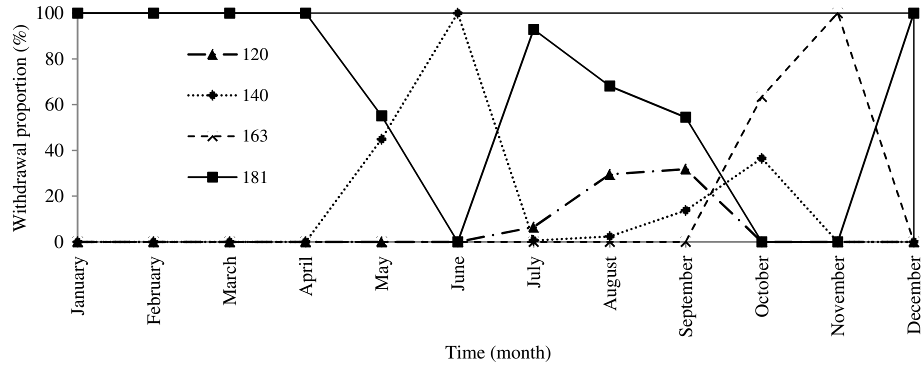

Fig. 7. Optimal withdrawal proportion at each outlet during the operational period based on Scenario 3

Scenario 3 are shown in Fig. 7. The results show that in some months all water is released just from one outlet, such as from December to April (cold months) all water is released at the 181-m outlet, and in some months water is released from two or three outlets such as May and September, respectively.

Scenario 4: Selective Withdrawal with Continually Variable (10-Day) Proportions Using Total Monthly LIBSVM Input Data

The optimal withdrawal proportions were calculated with the LIBSVM-GA. In this scenario proportions change for each 10-day time step during the operational period. To show this more clearly, Fig. 8 displays the 10-day withdrawal proportion average for all months of operational years. According to Fig. 8, the largest average withdrawal proportions are released from the 120 and 181-m outlets. In addition, the 140-m outlet has the minimum average withdrawal proportions during the operational period.

Scenario 5: Selective Withdrawal with Continually Variable (10-Day) Proportions Using Separated Monthly LIBSVM Input Data

The optimal withdrawal proportions that change for each 10-day time step were calculated with the LIBSVM-GA at each outlet during the operational period. Fig. 9 presents 10-day withdrawal proportion average from all months of the operational years. According to Fig. 9, the largest averaged withdrawal proportions are released from the 181-m outlets. The average withdrawal proportions at the 120, 140, and 163-m outlets are approximately the same during the operational period.

Comparison of Scenarios

Table 1 lists the statistical criteria used to compare the performance of scenarios, from which it follows that the highest (worst) and lowest (best) values of both RMSD and MAD correspond to Scenarios 1.2 and 5, respectively. The R^2 equals 99% for all scenarios, which shows a very high correlation between inflow and outflow temperatures for all scenarios.

The 181-m outlet was the best outlet if there were only one outlet to use. The RMSD decreased 0.18°C from Scenario 1.4 to Scenario 2 because the 120-m outlet was added for use in Scenario 2. In addition, when Scenario 3 was used, the RMSD reached 5.32°C, which is 0.21°C lower than Scenario 2. Using variable withdrawal proportions with Scenario 4, which changed in 10-day time steps, the RMSD reached 5.2°C. Finally, employing Scenario 5 with continual monthly variable (10-day) proportions, the RMSD reached the minimum value of 5.15°C.

The reservoir inflow water temperature and the calculated outflow water temperature are shown in Fig. 10 for each scenario. The outflow water temperature of Scenarios 4 and 5 better followed the trend and small changes in inflow water temperature because

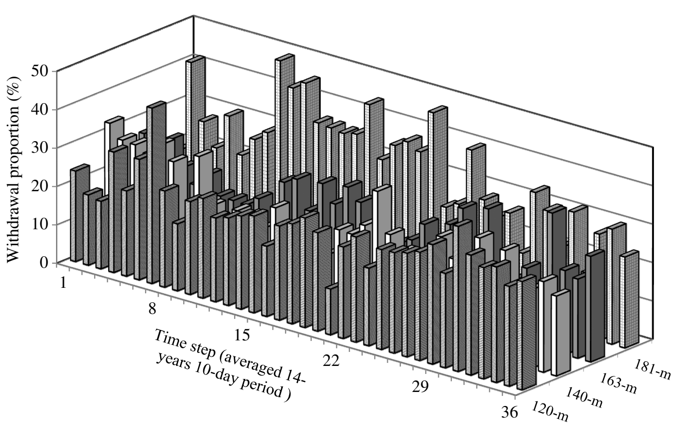

Fig. 8. Ten-day withdrawal proportion average for all months of operational years based on Scenario 4

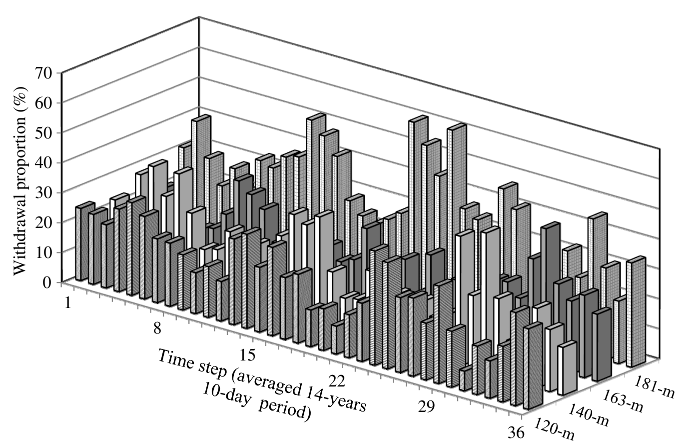

Fig. 9. Ten-day withdrawal proportion average for all months of operational years based on Scenario 5

J. Water Resour. Plann. Manage., 2016, 142(12): 04016064

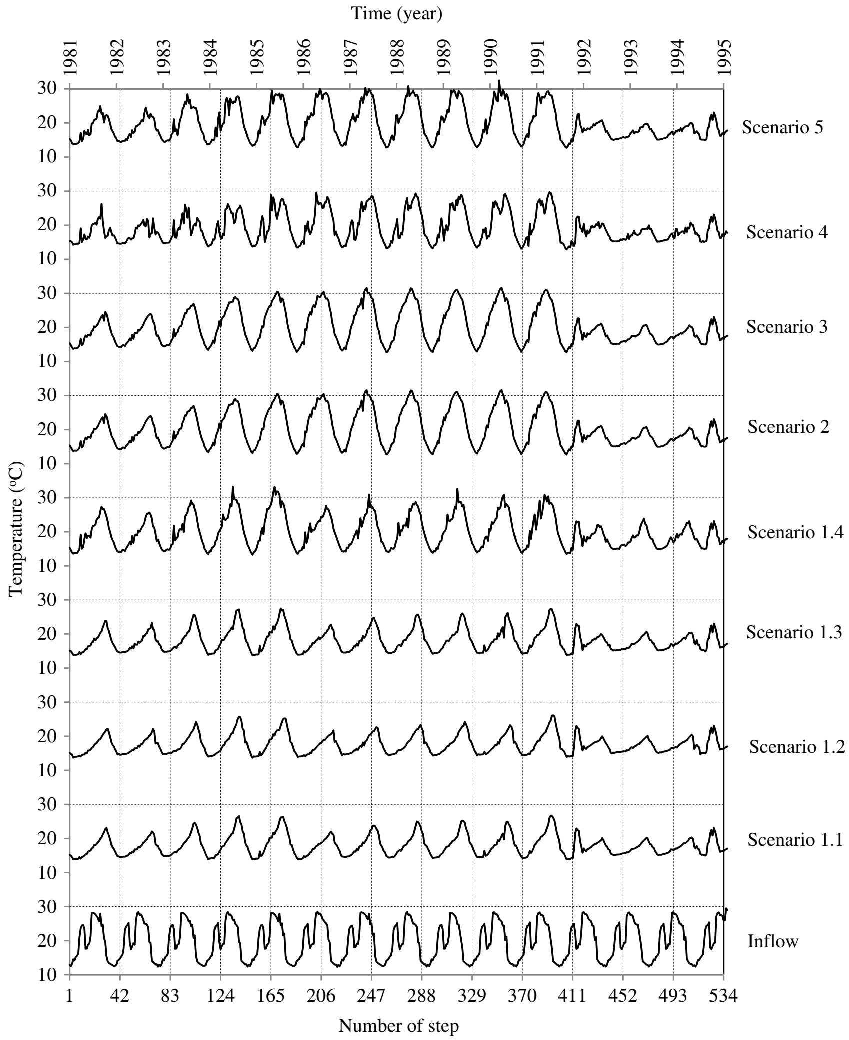

Fig. 10. Reservoir inflow and calculated reservoir outflow temperature for all scenarios

Scenarios 1.1, 1.2, and 1.3 only followed the trend of inflow water temperature.

Deviation of outflow temperature from inflow temperature for all scenarios is depicted in Fig. 11. It is evident that the differences between the actual inflow and calculated outflow temperature that occurred after 1991 were larger than those before 1991 because the reservoir surface elevation (Fig. 6) had a significant effect on reservoir thermal stratification after 1991, when the reservoir elevation exceeded 210 m. According to Fig. 11, the RMSD of the operational years 1981 to 1990 and 1991 to 1995 were equal to 4.5 and 6.41°C for Scenario 5 (best selected scenario), respectively. During the operational years 1981 to 1990 the reservoir water elevation was less than 210 m, and the RMSD was smallest. This implies that the reservoir surface elevation should be maintained below 210 m during the operational period.

Conclusions

This study calculated optimal selective withdrawal operations for thermal control of reservoir releases to downstream river. This study implemented a data mining model (LIBSVM) as a surrogate to the CE-QUAL-W2 simulation model to approximate reservoir outflow temperature. The data-mining model was combined with an evolutionary algorithm (LIBSVM-GA) to reduce optimization processing time. Four reservoir outlets at 120, 140, 163, and 181 m above mean sea level were considered in the Karkheh reservoir with 18 different options for selective withdrawal operations, in which each option comprised withdrawal proportions at each outlet. These options were input to the CE-QUAL-W2 model to produce the training and testing data for the LIBSVM model. The LIBSVM model with the best efficiency was linked to the GA to determine the optimal selective withdrawal rules.

Results show that the most important reservoir parameter is surface water elevation, which has a large effect on outlet water temperature. If only one outlet were used, the best location of the outlet would be on the upper thermal layer (epilimnion) to minimize the difference in temperature between reservoir inflow and downstream water temperature. The reservoir should have at least two outlets to control downstream temperature, and they should be located on the upper and lower reservoir thermal layers to control cold and warm water temperature. The possibility of selective withdrawal with continually variable (10-day) proportions using separated monthly LIBSVM input data produced minimal deviation between reservoir outflow and inflow temperatures. Optimal selective withdrawal rules efficiently control downstream water quality, which is vital to meet downstream ecological and environmental requirements.

J. Water Resour. Plann. Manage., 2016, 142(12): 04016064

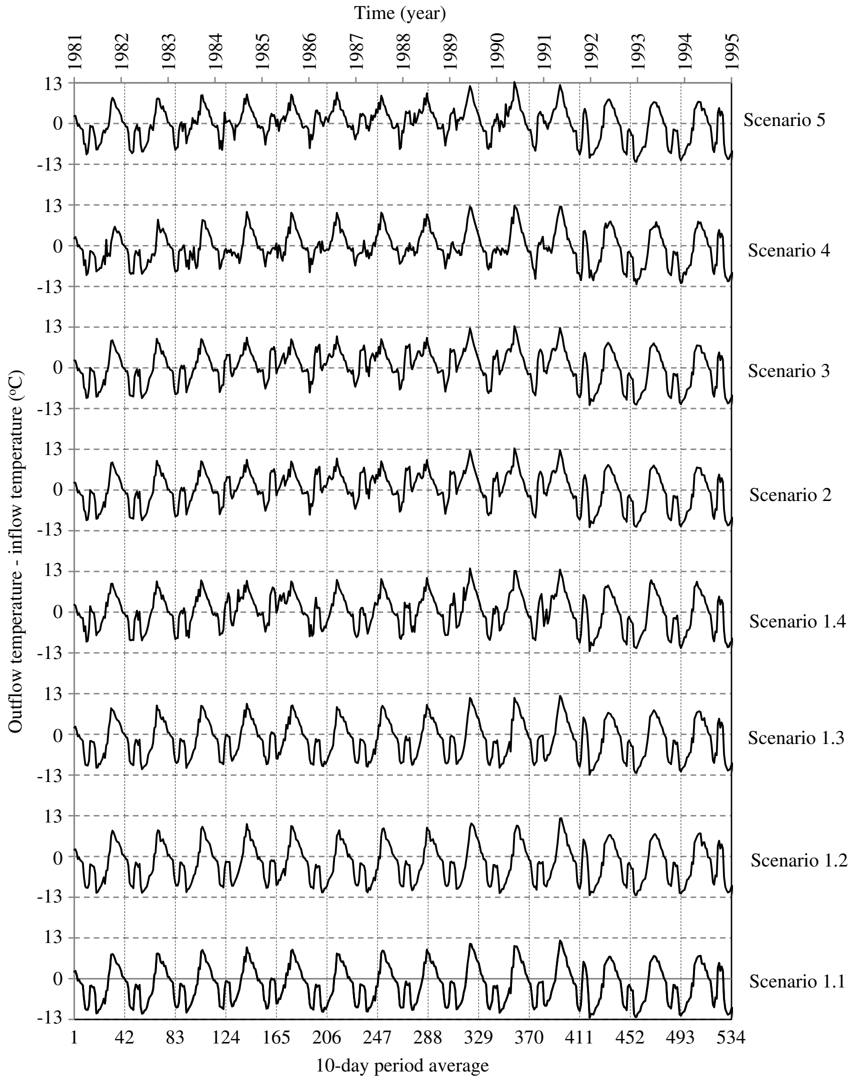

Fig. 11. Deviation of calculated reservoir outflow temperature from inflow temperature for all scenarios

Acknowledgments

The authors are thankful for the insightful and constructive comments submitted by the anonymous reviewers.

References

Afshar, A., and Saadatpour, M. (2009). "Reservoir eutrophication modeling, sensitivity analysis, and assessment: Application to Karkheh reservoir, Iran." *J. Environ. Eng. Sci.*, 26(7), 1227–1238.

Afshar, A., Saadatpour, M., and Mariño, M. A. (2012). "Development of a complex system dynamic eutrophication model: Application to Karkheh reservoir." *J. Environ. Eng. Sci.*, 29(6), 373–385.

Akbari-Alashti, H., Bozorg Haddad, O., Fallah-Mehdipour, E., and Mariño, M. A. (2014). "Multi-reservoir real-time operation rules: A new genetic programming approach." *Proc. Inst. Civ. Eng. -Water Manage.*, 167(10), 561–576.

Ashofteh, P. S., Haddad, O. B., and Loáiciga, H. A. (2015). "Evaluation of climatic-change impacts on multiobjective reservoir operation with multiobjective genetic programming." *J. Water Resour. Plann. Manage.*, 10.1061/(ASCE)WR.1943-5452.0000540, 04015030.

Çalışkan, A., and Elçi, Ş. (2009). "Effects of selective withdrawal on hydrodynamics of a stratified reservoir." *Water Resour. Manage.*, 23(7), 1257–1273.

Castelletti, A., Yajima, H., Giuliani, M., Soncini-Sessa, R., and Weber, E. (2014). "Planning the optimal operation of a multioutlet water reservoir with water quality and quantity targets." *J. Water Resour. Plann. Manage.*, 10.1061/(ASCE)WR.1943-5452.0000348, 496–510.

Chang, Ch.-Ch., and Lin, Ch.-J. (2011). "LIBSVM: A library for support vector machines." *ACM Trans. Intell. Syst. Technol.*, 2(3), 1–27.

Cole, T. M., and Wells, S. A. (2013). *CE-QUAL-W2: A two-dimensional, laterally averaged, hydrodynamic and water quality model, version 3.71. User manual*; U.S. Army Engineer Waterways Experiment Station, Louisville, KY.

Dixon, B. (2009). "A case study using support vector machines, neural networks and logistic regression in a GIS to identify wells contaminated with nitrate-N." *Hydrogeol. J.*, 17(6), 1507–1520.

Elçi, S. (2008). "Effects of thermal stratification and mixing on reservoir water quality." *Limnology*, 9(2), 135–142.

Fallah-Mehdipour, E., Bozorg-Haddad, O., and Mariño, M. A. (2012). "Real-time operation of reservoir system by genetic programming." *Water Resour. Manage.*, 26(14), 4091–4103.

Fallah-Mehdipour, E., Bozorg-Haddad, O., and Mariño, M. A. (2013a). "Developing reservoir operational decision rule by genetic programming." *J. Hydroinf.*, 15(1), 103–119.

Fallah-Mehdipour, E., Bozorg-Haddad, O., and Mariño, M. A. (2013b). "Extraction of optimal operation rules in aquifer-dam system: A genetic programming approach." *J. Irrig. Drain. Eng.*, 10.1061/(ASCE)IR.1943-4774.0000628, 872–879.

Fallah-Mehdipour, E., Bozorg-Haddad, O., and Mariño, M. A. (2013c). "Prediction and simulation of monthly groundwater levels by genetic programming." *J. Hydro-Environ. Res.*, 7(4), 253–260.

Fontane, D. G., and Labadie, J. D. (1981). "Optimal control of reservoir discharge quality through selective withdrawal." *Water Recour. Res.*, 17(6), 1594–1602.

Gelda, R. K., and Effler, S. W. (2007). "Simulation of operations and water quality performance of reservoir multilevel intake configurations." *J. Water Resour. Plann. Manage.*, 10.1061/(ASCE)0733-9496(2007)133:1(78), 78–86.

Giuliani, M., Castelletti, A., Galelli, S., Soncini-Sessa, R., Weber, E., and Alliance, S. D. W. (2013). "Many-objective operation of selective withdrawal reservoirs including water quality targets." *Proc., 2013 World Environmental and Water Resources Congress*, ASCE, Reston, VA.

Hanna, R. B., Saito, L., Bartholow, J. M., and Sandelin, J. (1999). "Results of simulated temperature control device operations on in-reservoir and discharge water temperatures using CE-QUAL-W2." *Lake Reservoir Manage.*, 15(2), 87–102.

Hocking, G. C., Sherman, B. S., and Patterson, J. C. (1988). "Algorithm for selective withdrawal from stratified reservoir." *J. Hydraul. Eng.*, 10.1061/(ASCE)0733-9429(1988)114:7(707), 707–719.

Holland, J. H. (1975). *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence*, University of Michigan Press, Ann Arbor, MI.

Kerachian, R., and Karamouz, M. (2006). "Optimal reservoir operation considering the water quality issues: A stochastic conflict resolution approach." *Water Resour. Res.*, 42(12), 1–17.

Mirfendereski, G. A., and Mousavi, S. J. (2011). "Comparison of support vector machine and response surface models in meta-modeling applied in basin-scale optimum water allocation." *MODSIM2011, Proc., 19th Congress on Modeling and Simulation*, Modelling and Simulation Society of Australia and New Zealand Canberra, ACT, Australia.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., and Veith, T. L. (2007). "Model evaluation guidelines for systematic quantification of accuracy in watershed simulations." *Trans. ASABE*, 50(3), 885–900.

Rheinheimer, D. E., Null, S. E., and Lund, J. R. (2015). "Optimizing selective withdrawal from reservoirs to manage downstream temperatures with climate warming." *J. Water Resour. Plann. Manage.*, 10.1061/(ASCE)WR.1943-5452.0000447, 04014063.

Saadatpour, M., and Afshar, A. (2011). "Multi objective simulation-optimization approach in pollution spill response management model in reservoirs." *Proc., Water Engineering and Management in a Changing Environment*, European Water Resources Association, Catania, Italy.

Safavi, H. R., and Esmikhani, M. (2013). "Conjunctive use of surface water and groundwater: Application of support vector machines (SVMs) and genetic algorithms." *Water Resour. Manage.*, 27(7), 2623–2644.

Shokri, A., Bozorg-Haddad, O., and Mariño, M. A. (2014). "Multi-objective quantity-quality reservoir operation in sudden pollution." *Water Resour. Manage.*, 28(2), 567–586.

Soltani, F., Kerachian, R., and Shirangi, E. (2010). "Developing operating rules for reservoirs considering the water quality issues: Application of ANFIS-based surrogate models." *Expert Syst. Appl.*, 37(9), 6639–6645.

Soltanjalili, M., Bozorg-Haddad, O., and Mariño, M. A. (2013). "Operating water distribution networks during water shortage conditions using hedging and intermittent water supply concepts." *J. Water Resour. Plann. Manage.*, 10.1061/(ASCE)WR.1943-5452.0000315, 644–659.

Vapnik, V. N. (1995). *The nature of statistical learning theory*, Springer, New York.

Vapnik, V. N. (1998). *Statistical learning theory*, Wiley, New York.

Vapnik, V. N. (1999). "An overview of statistical learning theory." *IEEE Trans. Neural Networks*, 10(5), 988–999.