Environ Model Assess (2015) 20:29–39
DOI 10.1007/s10666-014-9413-7

Hybrid ACO–ANN-Based Multi-objective Simulation–Optimization Model for Pollutant Load Control at Basin Scale

Mohammad J. Emami Skardi · Abbas Afshar · Motahareh Saadatpour · Samuel Sandoval Solis

Received: 11 September 2012 / Accepted: 14 May 2014 / Published online: 29 May 2014

Abstract The systematic and integrative approach to optimum watershed management couples a watershed simulation model and an efficient optimization algorithm for evaluating great number of "what if" scenarios in the search domain. This study integrates a multi-objective Non-dominated Archiving Ant Colony Optimization (NA-ACO) algorithm as an optimization tool with Soil and Water Assessment Tool (SWAT) as the simulation module for optimum management of total suspended solids (TSS) loading to downstream water bodies. The resulting NA-ACO–SWAT model is computationally and experimentally expensive because of the large number of required function evaluations which demands repetitive execution of SWAT simulation model. To increase the computational efficiency of the watershed simulation model, the SWAT model is replaced by a trained artificial neural network (ANN) model to form a hybrid NA-ACO–SWAT–ANN model to efficiently develop the set of optimum non-dominated solutions for configuration and design of detention ponds in basin scale. The applicability of the proposed method is evaluated at Gharesou watershed in the northwest of Iran. The outcomes of the proposed approach is further analyzed and compared in terms of their quality of solutions and computational efficiencies. Results show that the proposed hybrid approach may reduce the computational time by 90 % while keeping the accuracy of the results in the same order.

Keywords Non-dominated Archiving Ant Colony Optimization algorithm · Artificial neural network · SWAT · Total suspended solids (TSS) · Wet detention pond · Gharesou watershed

M. J. Emami Skardi (✉)
School of Civil Engineering, Iran University of Science and Technology, Tehran, Iran
e-mail: m.j.e.skardi@gmail.com

A. Afshar · S. Sandoval Solis
Department of Land, Air and Water Resources, University of California, Davis, CA, USA

A. Afshar
e-mail: aafshar@ucdavis.edu

S. Sandoval Solis
e-mail: samsandoval@ucdavis.edu

M. Saadatpour
School of Environmental Engineering, University of Environment, Karaj, Iran
e-mail: msaadatpour@iust.ac.ir

1 Introduction

Structural Best Management Practices (BMPs) are often used for sediment and nutrient control in both urban and agricultural watersheds. Wet detention ponds, as one of the most efficient elements in BMPs, have successfully been used for many years. They are known as the storm water control structures that can be used for both water quality and quantity management.

In general, a higher level of nutrient removal and better storm water quantity control can be gained in wet detention ponds than other BMPs. Wet ponds are often used to improve runoff water quality through controlling the total suspended solid (TSS) loading to downstream receiving water bodies. The ponds are usually designed as single elements, and their interactions in an integrative environment are often disregarded. Generating cost-effective pond size and configurations that satisfy system-wide targets for total target sediment removal believed to be much more effective [20, 22].

In general, the systematic and integrative management approach may be developed which applies watershed intensive mathematical simulation model and evolutionary (or meta-heuristic) search-based algorithms to evaluate great number of "what if" scenarios in the search domain.

Simulation–optimization (S–O) approach which links a detailed simulation model with a meta-heuristic or population-based evolutionary algorithm (EA) for solution of verity of optimization problems has received considerable attention. As highlighted by Wagner [50], the meta-heuristic and/or EAs have considerable priority over implicit enumeration techniques in locating good near-optimal solutions to combinatorial optimization problems with greater efficiency. They handle the discontinuities and nonlinearities existing in most of the real-world problems more easily than gradient-based techniques. What makes use of gradient-free meta-heuristic and/or evolutionary optimization techniques more attractive in S-O problems lies in the possibility of using any kind of built-up simulation package without having access to its embedded source codes [41]. The S–O problems employing any of the available search-based algorithms such as Genetic Algorithms (GA) [3], Ant Colony Optimization (ACO) [23], Particle Swarm Optimization (PSO) [31], Honey Bees Mating Optimization (HBMO) [12], etc. The S–O approach may lose their merits when a computationally expensive simulation model is employed for objective function evaluations. Although, one may partially overcome this problem by increasing the efficiency of the search algorithm, the main ongoing challenge is to reduce the computational cost of the simulation process within the optimization framework. The ongoing efforts are grouped into methods which either reduce the execution time for each simulation run through parallel computing and enhanced computer architecture or using a meta-model as an approximation to the real system to quickly supply predictions during the course of the search [24]. The main advantage of a surrogate model is to retain the size and scope of the original problem while capitalizing on the simpler empirical relationships between decision variables and selected model outcomes [50]. This paper employs the latter approach to partially overcome the problem of computationally expensive simulation in the proposed S–O problem.

Although the idea of replacing an extensive simulation model with an approximate one is not new, the concept receives increasing popularity in the water resources simulation–optimization literature as advance computing technologies and meta-modeling approaches emerge [30]. Artificial neural networks (ANNs), as function approximators and meta-models, have proved to be a logical choice for application when the complexity of the mapping is difficult to anticipate. Different versions of meta-models have successfully been used to replace computationally expensive simulation models in support of engineering design optimization [41, 51, 55].

This paper presents a hybrid modeling framework which couples Soil and Water Assessment Tool (SWAT) and ANN as watershed simulation models with a multi-objective version of ACO (NA-ACO) to minimize both the total suspended sediment loading to downstream receiving water body and the construction cost of the wet detention ponds in an integrative scheme. The NA-ACO–SWAT is computationally and/or experimentally expensive because of computational burden of SWAT simulation model and large number of required function evaluations to locate good near-optimal solutions. To increase the computational efficiency of the proposed simulation–optimization scheme, the SWAT model is replaced by a trained ANN model to form a hybrid NA-ACO–SWAT–ANN model to find the set of optimum non-dominated solutions for detention ponds design in basin scale. The applicability of the proposed method is evaluated using Gharesou watershed in the northwest of Iran. The outcomes of both models are further analyzed and compared in terms of their quality of solutions and computational loads.

2 Literature Review

Although, the integration of any multi-objective version of ACO with the SWAT and ANN for watershed management is of very recent origin, extensive applications of any of the three models in different areas of water resources management have been reported.

As documented by Gassman et al. [21], over 250 peer-reviewed published papers have reported SWAT applications, components' reviews, or cited SWAT. It is a comprehensive spatial and temporal simulation model, developed by Arnold et al. [8] for USDA-ARS (USDA-Agricultural Research Service). The SWAT model has been extensively tested for hydrologic modeling at different spatial scales for flow simulation [15, 38, 46]. Tolson and Shoemaker [48] applied SWAT model to a New York City water supply reservoir (Cannonsville Reservoir). They found that SWAT is a valuable tool that can be used to help evaluating the long-term effects of various phosphorus management options for mitigating pollutant loading to the reservoir. In an interesting work, Zhang et al. [54] applied SWAT model for snowmelt-driven flow in the 114,345-km^2 headwaters of the Yellow River. It has been successfully used to simulate water and pollutant load in various watersheds around the world [2, 36, 40].

Different versions of ANNs have successfully been used in modeling complex systems such as conjunctive use modeling and groundwater systems operation [34, 35], water distribution design with water quality consideration [14], rainfall–runoff modeling [45], prediction of daily stream flow [47], groundwater simulation [52], inferring reservoir operating rules [30], and water quality model calibration [6]. Yan and Minsker [52] replaced a computationally expensive groundwater simulation model with a dynamic modeling approach in which ANNs are adaptively trained within a genetic algorithm. In an interesting work, Johnson and Rogers [24] employed ANNs to examine their impacts on the quality and quantity of solutions obtained from simulated annealingdriven searches on two different groundwater remediation problems. The quality of results obtained when ANNs served as substitutes for the full model was consistently comparable to those obtained when the full model itself was called in the course of the search.

During the last decade, application of different versions on ANNs to approximate SWAT model for prediction of watershed responses to various scenarios and/or human activities has received considerable attention. Zhang et al. [55] used ANNs and support vector machine (SVM) to evaluate their performances as a surrogate model for approximating the SWAT model. They applied ANN as the surrogate models to speed up the determination of proper hydrological calibration parameters in SWAT model. It was illustrated that surrogate models may efficiently be used to approximate the computationally intensive models. Demirel et al. [17] compared the performance of ANN and SWAT model in daily flow forecasting of the Pracana Basin in Portugal. Their results show, in general, ANNs can be powerful tools in daily flow forecasts.

The systematic watershed management approaches such as coupling watershed simulation model with heuristic optimization techniques in deriving optimal BMPs have been tackled in various researches. Successful simulation of various scenarios on different basins around the world has identified SWAT as an effective method for evaluating alternative watershed management practices [13, 7, 26, 25, 27, 32, 33, 38, 39, 49]. To find the optimal BMP in watershed scale at different hydrological settings, SWAT and GAs have successfully been coupled to locate near-optimal solutions [25]. Recently, Kaini et al. [26] developed a multi-objective optimization approach to find a set of near-optimal solutions to trade off between cost and sediment control for constructing BMPs in watershed scale. In an interesting research, Qi and Altinakar [32] coupled a process-based watershed simulation model with a modern heuristic optimization technique to manage watershed land use allocation for achieving sustainable development. In an extension to their previous research, they included the placement of vegetation of buffer strips (VBS) in the watershed to guarantee the efficiency of the VBS applications in varying geological and economic conditions [33]. In a most recent work, Skardi et al. [44] used SWAT with game theory concept in a simulation–optimization approach for nonpoint source pollution management.

ACO algorithm was initially proposed by Dorigo et al. [18] as a meta-heuristic approach to solve different problems including the traveling salesman and the quadratic assignment optimization problems. Although, ACO was basically developed for the discrete optimization problems, different versions of ACO have successfully been applied in both discrete and continuous domains in various water resource problems such as the following: parameter estimation in unsaturated soils [1], water distribution system optimization [4, 29, 53], single and multi-objective optimization of reservoir operations [5, 23], and continuous reservoir operation [28].

3 Model Description and Development

In this research, an effective and integrative methodology is developed to identify the optimal size and configuration of detention ponds in watersheds to control the TSS pollutant loading to downstream receiving water bodies. The SWAT is a careful spatial and temporal analysis tool and high fidelity simulation model for predicting the important characteristics of watershed system behavior. However, it remains computationally far too expensive, especially in coupling with search-based algorithms to derive optimum watershed management strategy. To overcome this bottleneck, a data-driven ANN model is used to considerably reduce the computational time while maintaining the prediction accuracy in acceptable range. The structured ANN model suggests a functional relationship between selected decision variables and watershed system responses. The surrogate ANN model is coupled with multi-objective ACO algorithm to derive optimal location and sizes of detention ponds in watershed scale. The developed integrative framework in this research minimizes the practice costs and annual TSS pollutant loads. The detail description of the integrative watershed management model components are presented in the following subsections.

3.1 SWAT-ANN Simulation Model

The SWAT is computationally and/or experimentally expensive which may reduce the merits of the proposed S–O scheme in the proposed watershed management optimization. The computational efficiency of the scheme is increased by replacing SWAT model with trained ANN model to form a hybrid SWAT–ANN model. This model is applied to predict the watershed responses to management strategies. ANN emulates the behavior of SWAT as the highly detailed hydrological and water quality simulation model. This fast analysis tool is implemented in optimization and exploration of the design space by using approximations in lieu of the computationally expensive analysis codes (SWAT).

Data-driven ANN model is able to learn from examples and respond to functional relationships within the data, even when the underlying relationships are unknown or difficult to describe. In addition, ANN is able to mimic the behavior of nonlinear multivariate functions with high accuracy. This feature is very important since very large problems such as real-world case studies encompass complicated unknown nonlinear patterns and/or behaviors.

Input variable selection is an important part of the identification of ANN models since the form of the model is derived purely from the available data. In real-world applications, such as SWAT analysis, there are potentially many variables that could be used as inputs to the ANN model. However, for the development of ANN models, the minimum number of variables should be used as inputs to the ANN in order to (1) increase computational efficiency, (2) minimize redundancy, (3) reduce noise, and (4) increase the interpretability of the model [9, 42].

The existence of height–width relationship in a design of wet detention ponds provides unique information about pond's geometry. The area and volume as the geometry data are estimated using the mathematical equations. Generally, geometric characteristics have important role in sediment entrap and settlement. Average annual TSS and economic cost of pond constructions and maintenance are state variables associated with any given management option.

Due to dependency of geometric parameters (i.e., height, volume, and surface area), height vector of detention ponds, as a collection of numbers between zero and given ranges, forms the ANN input data. Each member of this vector shows the geometric property of a given detention pond in a known position. The zero values mean that no pond is nominated for that specific location. The volume and surface area of the pond is calculated using the known height–width–volume relationships. These geometric characteristics are required by the SWAT model to estimate the annual average TSS concentration. The structure of the ANN input data is presented in Fig. 1.

Multilayer feed forward network is used for function approximation in the proposed ANN. Feed forward layered networks have the flexibility to approximate smooth functions arbitrarily well through providing sufficient nodes and layers [11]. The performance of ANN network in this research is evaluated with various training functions, learning functions, and neuron numbers in the hidden layer. The optimal structure of the ANN is selected based on the approximation accuracy of the validation data among various defined ANN structure alternatives.

3.2 Non-dominated Ant Colony Optimization Algorithm

3.2.1 General Aspects of ACO Algorithms

It is shown that a colony of ants is able to find the shortest route from their nest to a food source via an indirect form of communication. This communication involves deposition of a chemical substance, called a pheromone, on the paths as they travel. Over time, shorter and more desirable paths are reinforced with greater amounts of pheromone, thus becoming the dominant path for the colony. Ant Colony Optimization (ACO) is inspired by the foraging behavior of ant colonies, and this collective trail-laying and trail-following in which an ant is influenced by the pheromone trail left by other ants. Ant algorithms benefit from artificial ants which deposit pheromone based on the fitness and goodness of the identified trial solutions.

A graph is very helpful for successful application of the ACO algorithms to combination of optimization problems. Consider $G=(D,L,C)$, in which $D=\{d_i\}$ is the set of decision points at which some decisions are to be made, $L=\{l_{ij}\}$ is defined as the set of the options $j=1,2,\ldots,NC$, at each decision point $i=1,2,\ldots,NT$, and $C=\{c_{ij}\}$ is the set of costs associated with $L=\{l_{ij}\}$. An acceptable solution based on the graph is called an answer $(\varphi)^k$, and the path associated with minimum cost is called the optimum solution $(\varphi^*)^k$.

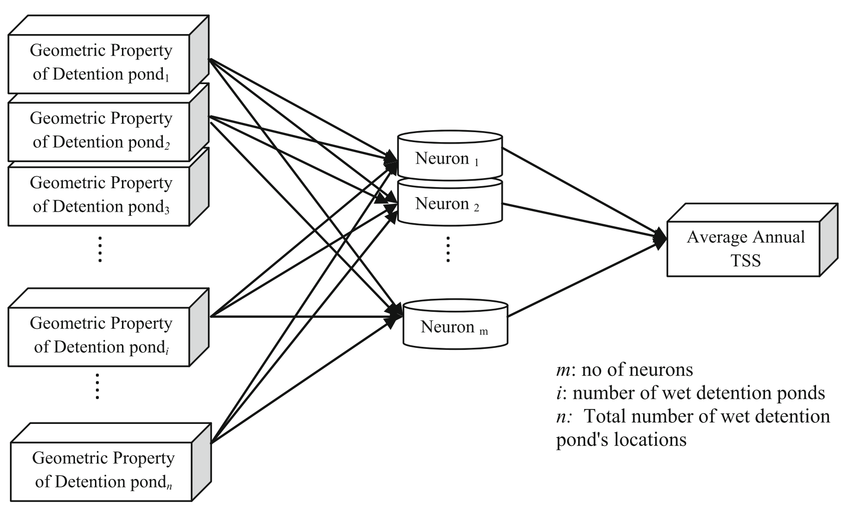

Fig. 1 Schematic structure of the proposed ANN model

The decision policy $p_{ij}(k,t)$ in ant system (AS) defines the probability that edge (i,j) will be selected at decision point i as [18] the following:

$$P_{ij}(k,t) = \begin{cases} \dfrac{\left[\Gamma_{ij}(t)\right]^{\alpha}\left[\eta_{ij}(t)\right]^{\beta}}{\sum\limits_{j=1}^{J}\left[\Gamma_{ij}(t)\right]^{\alpha}\left[\eta_{ij}(t)\right]^{\beta}} & \text{if } j \in \text{Allowed}\,k \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

Here, $P_{ij}(k,t)$ is the probability that the ant number k selects option j at decision point i and iteration t. $\Gamma_{ij}(t)$ is the concentration of pheromone on arc (i,j) at iteration t; $\eta_{ij}(t)=(c_{ij})^{-1}$ is the heuristic value representing the cost of choosing option j at decision point i, and α and β are two parameters that control the relative importance of the pheromone trail and heuristic value. The heuristic value η_{ij} is analogous to providing the ants with sight and is something called visibility. The total number of ants, k, as well as α and β are tunable model parameters. For $\alpha>>\beta$, the decisions will be based mainly on the learned information, as represented by the pheromone. For very large values of β (i.e., $\beta>>\alpha$), the algorithm will set a greedy heuristic and disregards the impact of decisions on final solution quality. In any iteration, the pheromone is updated through evaporation. Pheromone evaporation has the advantage of avoiding the rapid premature convergence to a local optimal solution. Without pheromone evaporation, the paths chosen by the very first ants would be excessively attractive to the following ants, and consequently, the exploration of the solution space would be constrained. The pheromone may be updated using the following relation:

$$\Gamma_{ij}(t+1) = (1-\rho)\Gamma_{ij}(t) + \rho\Delta\Gamma_{ij}(t) \tag{2}$$

Here, $\Gamma_{ij}(t+1)$ is the amount of pheromone trail on option j of the i^{th} decision point at iteration $t+1$; $0\leq\rho\leq1$ is the pheromone evaporation coefficient, and finally, $\Delta\Gamma_{ij}(t)$ is the change in pheromone concentration associated with *arc* (i,j) at iteration t. Although, several methods are available to determine the value of $\Delta\Gamma_{ij}(t)$, this study benefits from one used by Afshar et al. [5]:

$$\Delta\Gamma_{ij}(t) = \begin{cases} \dfrac{\gamma}{1+G^{k_{\text{gb}}^{*}}} & \text{if} \quad (i,j)\in \text{ tour done by Ant } k_{\text{gb}}^{*} \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

Here, $G^{k_{\text{gb}}^{*}}$ is value of the objective function corresponding to the ant k_{gb}^{*} as the ant with the best performance within the past total iterations, and γ is a constant coefficient.

3.2.2 Multi-objective Ant Colony Optimization Method

In the single objective problem, the optimal solution is unique. In the multi-objective ones, however, there is a set of non-dominated solutions that is known as the Pareto front. Deb [16] discussed the concept of domination and non-dominant solution in multi-objective optimization problems as the following:

"A solution $\times_1$ dominates a solution $\times_2$, if solution $\times_1$ is strictly better than solution $\times_2$ in all M objectives," and "Among a set of solutions P, the non-dominated set of solution P' are those that are not dominated by any member of the set P."

Multi-objective Ant Colony Optimization (MACO) algorithms have widely been used to investigate optimization problems [4, 10, 19]. A detailed background about the application of multi-objective ACO is presented by Dorigo et al. [18]. Afshar et al. [5] proposed a new version of ACO-based algorithm for multi-objective problems known as "Non-dominated Archiving ACO" (NA-ACO). In this algorithm, a colony of agents with the same population is assigned to each objective and the pheromone updating is performed by the best ant in each colony. All ants in each colony simultaneously explore a solution which is evaluated according to the objective defined for that colony. These trial solutions are then transferred to the next colony to be evaluated and being updated according to the objective of the next one. Then, new found solutions on the basis of the updated pheromone trail in the second colony are transferred to the first colony, in the case of two objective function problems. This cycle is repeated for a predefined number of iterations. After ending the predefined iteration, the non-dominated solutions are transferred to an offline archive for more pheromone updating. Then, the pheromones of two colonies are updated according to the non-dominated solutions in the offline archive. Finally, the algorithm returns to the starting point for another iteration. This process continues to a predefined number of total iterations or a favorable number of archived Pareto answers. In this research, the NA-ACO is implemented to address the set of non-dominated solutions for a real-life watershed management problem.

3.3 Simulation–Optimization (S-O) Model

As depicted in Fig. 2, the proposed methodology consists of three main elements: Non-dominated ACO algorithm, SWAT, and ANN model. The optimization algorithm is programmed in MATLAB environment and coupled with ANN model that is included as a toolbox in the same programming language.

In the first step of the proposed S–O framework, ANN is trained based on the results provided by SWAT simulation model according to various conservation practices in the watershed. Various conservation practices consisting of variant

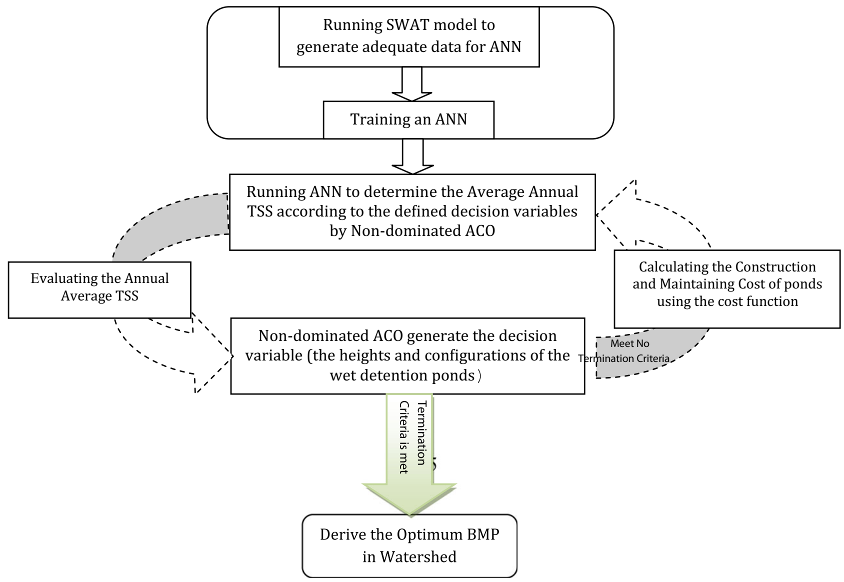

Fig. 2 A scheme of the proposed methodology

size and configurations of wet detention ponds are randomly defined to train the ANN model. The surrogate ANN model is trained, tested, and validated based on intensive combination of various training functions, learning functions, and neuron numbers. Finally, the optimal structure of ANN, as the surrogate model, is chosen among various ones based on approximation accuracy in validation step. In the second step, ANN as the simulation model is coupled with the NA-ACO as the optimization algorithm to derive optimal and/or near-optimal watershed management practice. Average annual TSS is approximated by ANN model for any trial solution identified by the optimizer (NA-ACO).

The proposed methodology is expanded to solve the typical multi-objective watershed management problem involving both TSS loading and economic goals. The methodology is designed to yield directly the size and configuration of wet detention ponds that simultaneously minimize TSS loading and construction and maintenance cost subject to specified constraints.

The optimization formulation of the proposed methodology may mathematically be expressed as follows:

$$\begin{aligned} &\text{Minimize} \quad \text{PONDC} = \sum_i \sum_j \text{PONDC} \\ &\text{Minimize} \quad \sum_i \text{TSS}^i \,\text{load} \end{aligned} \qquad (4)$$

Subject to

$$A_{ij} \leq A_{\max}$$

$$S_j \leq S_{j\max}$$

Here, PONDC_{ij} is the construction and maintenance cost of pond number j in subbasin number i; A_{ij} is the surface area of the pond number j in subbasin number i which should be less than a maximum value of $A_{\max}$. S_j is the storage of the ponds which should be less than a predetermined maximum volume of $S_{j\max}$. The units of cost and TSS loading in the objective function (Eq. 4) are dollar and ton per hectare per year, respectively. $\text{TSS}^i_{\text{load}}$ is the loading TSS from the subbasin number i. The cost value (PONDC_{ij}) is evaluated based on the mathematical function of pond's volume. The other objective function namely, average annual TSS ($\text{TSS}^i_{\text{load}}$) is approximated by ANN model instead of the SWAT model to save the computational time.

The construction and maintenance cost of the pond in subbasin i is site dependent and a functional relation (and/or a look up table) should be made available as input data to the model. A typical form of this functional relation is introduced and used in the case study and application of the model. The TSS loading for any trial solution (or strategy) will be determined by running the simulation model (i.e., SWAT and/or proposed data-driven meta-model). In other words, for any trial solution which is addressed by a set of decision variables, the simulation model will be executed to estimate the value of the resulting TSS loading. This means that if the optimization is performed with 500 ants and 2,000 iterations (i.e., total of 1,000,000 function evaluations), then the simulation model has to be executed 1,000,000 times to address the trial values of TSS loadings. Although, we explicitly consider the height of the ponds as decision variables, we implicitly account for their length, width, volume, and surface area in the functional relationship to their heights. In other words, for a given site with known geometry and longitudinal slope of the reach, as well as the side slopes of the embankment, any combination of pond's volume, surface area, length, and width will be uniquely defined by the height.

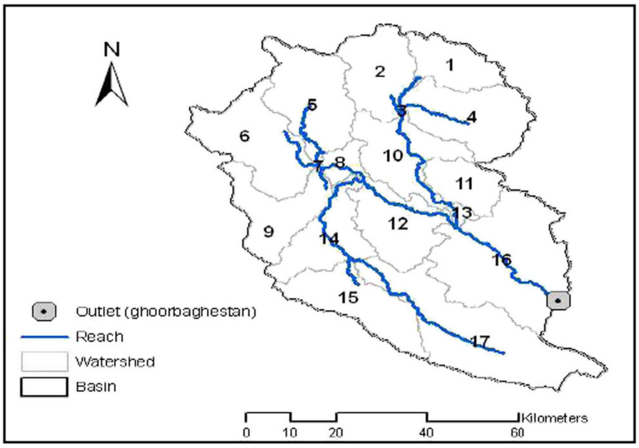

Fig. 3 The Gharesou watershed in Iran

4 Case Study

Gharesou watershed with a 5,793-km^2 area is located in the northwest of Iran (Fig. 3). Approximately 48 % of the watershed is mountainous and the remaining 52 % is plain. The elevation of the watershed ranges between 1,300 and 3,351 m from the sea level. The climatic condition of the watershed varies from cold and dry in the central part of the plain to cold and wet near the border line.

The available data to run the SWAT model consist of digital elevation model (DEM), soil and land use maps, climate data records from four precipitation gauges, and two air temperature gauges during the years 1990 to 2000, which were obtained from the Iran Metrological Organization. The highland soil consists of sand and gravel besides the silt with substantial erosion. The runoff and TSS loading from the watershed, to downstream water bodies, has been calculated based on calibrated model by Sadeghi [37] with CUP_SWAT. The data from 1990 to 1996 has been used for the calibration, and the data from 1997 to 2000 has been used for the model validation. To facilitate the application and maintain the desired accuracy, the watershed was subdivided into 17 subbasins and 113 Hydrological Response Units (HRUs). More information about Gharesou watershed simulation model could be found in research by Sadeghi [37]. In this study, the potential locations of the ponds for each subbasin are assumed to be known. The pond location in any subbasin determines the fraction of the runoff water which will potentially drain into it. These fractions for potential pond locations for each subbasin are presented in Table 1.

The relative height–width relation for each pond is presented in Table 2. The height of the proposed ponds in the trial solutions are determined by the optimization algorithm. Having the height and its associated width (Table 2), the surface area and storage volume for the ponds may be estimated using the mathematical equation (Eqs. 5 and 6).

$$\begin{aligned}\text{Volume}_{\text{Pond}} = {} & (a.b.h) + \frac{h^3}{3 \times \text{slop}_{\text{pond}_{\text{border}}} \times \text{slop}_{\text{pond}_{\text{dam}}}} \\ & + \frac{h^3}{3 \times \left(\text{slop}_{\text{pond}_{\text{border}}}\right)^2} + \frac{a.h^2}{3 \times \text{slop}_{\text{pond}_{\text{border}}}} \\ & + \frac{b.h^2}{2 \times \text{slop}_{\text{pond}_{\text{border}}}} + \frac{b.h^2}{2 \times \text{slop}_{\text{pond}_{\text{dam}}}}\end{aligned} \tag{5}$$

$$\begin{aligned}\text{Surface}_{\text{Pond}} = {} & (a.b) + \frac{b.h}{\text{slop}_{\text{pond}_{\text{border}}}} + \frac{2h^2}{\left(\text{slop}_{\text{pond}_{\text{border}}}\right)^2} \\ & + \frac{a.h^2}{\text{slop}_{\text{pond}_{\text{border}}}} + \frac{2h^2}{\text{slop}_{\text{pond}_{\text{border}}} \times \text{slop}_{\text{pond}_{\text{dam}}}} + \frac{b.h}{\text{slop}_{\text{pond}_{\text{dam}}}}\end{aligned} \tag{6}$$

Here, $\text{Volume}_{\text{Pond}}$ is the principle volume of the ponds; $\text{Surface}_{\text{Pond}}$ is the principle surface for the ponds; $\text{slop}_{\text{pond}_{\text{border}}}$ is the slope of the ponds' border; $\text{slop}_{\text{pond}_{\text{dam}}}$ is the slope of the embankment; and a, h, and b are the length, height, and width of the pond, respectively.

Brief descriptions of the cost function and simulation time horizon are described in Table 3. The cost function of BMP in Gharesou watershed is related to construction and maintenance cost and the maintenance duration. The construction and maintenance costs are functions of the volume of wet detention ponds.

5 Results and Discussion

In order to train the proposed surrogate ANN model for prediction of the Gharesou watershed response to management strategies, data from application of the SWAT model to the same watershed under various strategies are used. For the purpose of reliable data generation, the SWAT model which was calibrated and verified for the same watershed was employed [37]. Calibration and verification of the SWAT

Table 1 The potential locations for construction of ponds in different subbasin and the fraction of the draining runoff water

The subbasin's number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Fraction of the basin that drains into the pond	0.4	0.2	0.1	0.4	0.3	0.5	0.1	0.2	0.2	0.3	0.1	0.3	0.2	0.1	0.5	0.2	0.5

Table 2 The height–width relation for ponds

Height (m)	0.0	1.2	1.8	2.4	3.0
Related width (m)	0.0	70.0	80.0	90.0	100.0

model was performed automatically using field data for the periods of 1990–1996 and, 1997–2000, respectively. It was shown that the calibrated model may be used to estimate the responses of the watershed to varying range of management options [37].

In this research, ANN as an approximation tool is replaced by watershed comprehensive hydrological and water quality simulation model, namely SWAT. ANN is trained based on finite conservation practices simulated with the SWAT model as the input data. The structure of ANN is assessed based on various training functions, learning functions, and several numbers of neurons to construct the more accurate approximation model. The final functions are assigned to ANN structure based on intensive sensitivity analysis on numerous training and learning functions in MATLAB environment. The number of neurons in hidden layer was allowed to vary between 4 and 20. Finally, based on intensive sensitivity analysis, "*learncon*" and "*trainbr*" functions have been applied as the learning and training functions of ANN model in MATLAB environment, respectively. The total of 10 neurons has finally been included in the hidden layer as the optimum number. The best ANN structure is selected to provide insight on the behavior of the watershed system to repeat "what if" analyses quickly. The final structure is selected among various structure combinations of ANN model based on approximation accuracy in validation step. The structure of ANN model is proportional to the final optimal structure derived based on the intensive sensitivity analysis. The annual average TSS load is modeled with the process-based model, namely SWAT, as the target values. Also, the annual average TSS loads as the output values are derived with ANN model.

To exhibit the ANN accuracy in approximation of average annual TSS based on various conservation practices in the watershed, the statistical criteria based on training, validation, and testing data are represented in Table 4. As presented in Table 4, the accuracy of ANN model to predict watershed responses according to various conservation practices is satisfactory. The high correlation coefficients in training, test, and validation data show the satisfied approximation accuracy.

In the next step, the integrative watershed management framework is applied. The proposed methodology relies on the interface between ANN and non-dominant ACO algorithm as illustrated in Fig. 2. The optimization model developed here, however, requires an iterative search for which a number of function evaluations, or ANN calls are necessary. The surrogate ANN as the approximation model is implemented to considerably reduce the computational time through rapid reporting of the watershed response. Rapid response reporting will highly enhance the efficiency of the model in accounting for repeated what if analysis of the changes in management options.

To exhibit the computational advantages of proposed methodology in this research, the computational performance of the following alternatives are compared. The performance of the proposed methodology in this research is compared with the coupling SWAT and non-dominant ACO technique [43]. The computational times of S–O approach in this study are about 6 h, whereas the combination of SWAT and non-dominant ACO requires more than 64 h for the same problem. Both S–O approaches have been terminated with the same number of function evaluations (i.e., 100,000). Very large percentage of the execution time in ANN–NA-ACO model is related to the SWAT excitation to prepare train data for ANN model. In other words, the execution of S–O model with ANN as the simulation model requires a very small computer run time. The final Pareto front derived by ANN–NA-ACO model is simulated by SWAT model to determine the annual average TSS based on accurate process-based model. The comparison of two alternatives reveals that the ANN model may be considered as an accurate and effective surrogate model substituting the process-based SWAT model. The ANN–NA-ACO model is recognized as a highly effective and efficient systematic watershed management tool for deriving optimal and/or near optimal conservative practice

Table 3 Brief discussion of the Gharesou watershed

Decision variable	Cost per volume of the ponds ($)	The maintenance duration (years)	The maintenance cost (% of construction cost)	Operation period (years)
Height	1.5	10	3	11

Table 4 The relation between the output variables and target for the training, test, and validation data

	Train	Test	Validation
Regression equation	y=0.999×	y=1.00×−0.001	y=0.998×+0.002
R^2	0.999	0.997	0.999

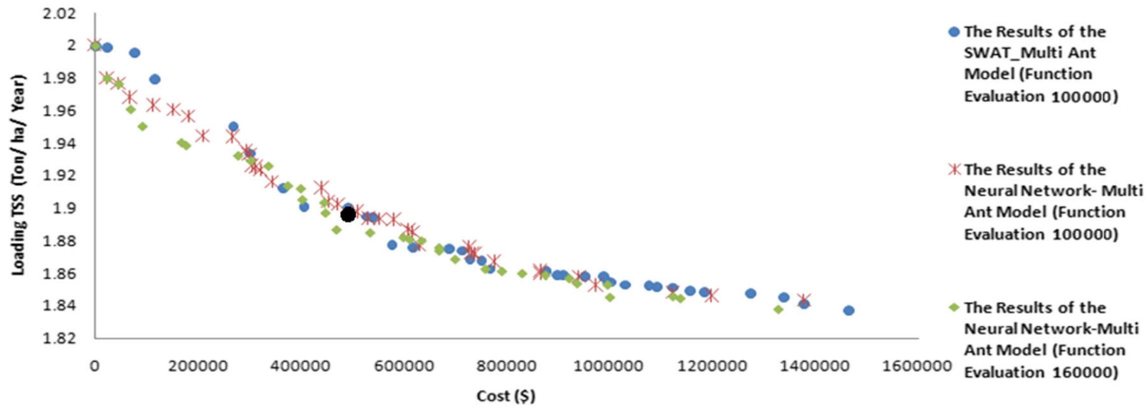

Fig. 4 The Pareto front of the annual average loading TSS to the construction cost for the ponds (the *black point* is presented in details)

in large watersheds where extremely large number of function evaluations might be required.

The derived Pareto front based on effective alternative (ANN–NA-ACO) is achieved in 160,000 function evaluations. The comparison results are shown in Fig. 4. Like other evolutionary algorithms, a number of parameters of ACO are needed to be set before its application. In this study, the tuned values for the most important parameters, namely β, α, and ρ are obtained after performance evaluation of extensive combinations of the parameters and the reference values [43]. Finally, values of β=0, α=1, ρ=0.9, and $\gamma_{f1(\mathrm{TSSS\_loading})}$=0.1 and $\gamma_{f2(\mathrm{cost})}$=10.0 ($\gamma$ has been defined in the pheromone concentration changing equation) were selected for model implementation. To further analyze the solutions on the Pareto front (Fig. 4), one of the solutions is selected and detailed in Table 5. As presented in Table 5, the decision variables, corresponding to one of the Pareto front's members, consist of the size and configuration of wet detention ponds in each subbasin. The zero values of geometric characteristic in subbasins 1, 3, 7 to 9, 12 to 14, and 16 are representative of lack of detention ponds in those subbasins. The nonzero values represent the geometric characteristics of the detention ponds in the subbasins.

As expected, increasing the cost of BMP implementations decreases the annual TSS load from the watershed. The BMP cost in this research depends on the size and number of wet detention ponds in the watershed. Properly located and sized wet detention ponds would improve the sediment deposition process rate, hence reducing the annual TSS load to downstream.

6 Conclusion

This study presented a multi-objective simulation–optimization model for optimum watershed management to develop a trade-off between the adverse environmental impacts and cost of implementing various management options. Realizing the extensive computational time required to solve the coupled SWAT–NA-ACO due to large number of function evaluations, the SWAT model was efficiently replaced by a surrogate ANN model. It was illustrated that the proposed data-driven ANN model may effectively reproduce the watershed responses under various management options. Including the surrogate ANN model in the proposed S–O scheme considerably improved the computational efficiency and practically removed the prohibitive computational bottleneck of the original S–O with SWAT as the simulator. Application of the proposed hybrid S–O model to Gharesou watershed revealed that the scheme may efficiently be used to develop the set of optimum non-dominated solutions for large-scale watershed management problems in a relatively small computer execution time. Although, this study was limited to optimal location and sizing of detention ponds, the approach may easily be extended to include other potential elements of the BMP. The accuracy of the surrogate model in those cases with multiple input and/or output vectors remains to be tested.

Table 5 The detailed results of one of the Pareto front corresponding to loading TSS 1.90 and related cost $431,323.5 (trade off between cost and TSS loading)

Subbasin no.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Height of the ponds (m)	0.0	1.8	0.0	1.2	3.0	2.4	0.0	0.0	0.0	1.2	1.2	0.0	0.0	0.0	1.8	0.0	1.8
Width of the ponds (m)	0.0	80	0.0	70	100	90	0.0	0.0	0.0	70	70	0.0	0.0	0.0	80	0.0	80
Length of the ponds (m)	0.0	160	0.0	140	200	180	0.0	0.0	0.0	140	140	0.0	0.0	0.0	160	0.0	160

References

1. Abbaspour, K. C., Schulin, R., & Genuchten, M. T. V. (2001). Estimating unsaturated soil hydraulic parameters using ant colony optimization. *Advances in Water Resources, 24*(8), 827–841.
2. Abbaspour, K. C., Yang, J., Maximov, I., Siber, R., Bogner, K., Mieleitner, J., Zobrist, J., & Srinivasan, R. (2007). Modelling hydrology and water quality in the pre-alpine/alpine thur watershed using SWAT. *Journal of Hydrology, 333*(2–4), 413–430.
3. Afshar, M. H., Afshar, A., Marino, M. A., & Darbandi, A. A. S. (2006). Hydrograph-based storm sewer design optimization by genetic algorithm. *Canadian Journal of Civil Engineering, 33*(3), 319–325.
4. Afshar, M. H., & Marino, M. A. (2006). Application of an ant algorithm for layout optimization of tree networks. *Engineering Optimization, 38*(3), 353–369.
5. Afshar, A., Sharifi, F., & Jalali, M. R. (2009). Non-dominated archiving multi-colony ant algorithm for multi-objective optimization: application to multi-purpose reservoir operation. *Engineering Optimization, 41*(4), 313–325.
6. Afshar, A., & Kazemi, H. (2012). Multi objective calibration of large scaled water quality model using a hybrid particle swarm optimization and neural network algorithm. *KSCE Journal of Civil Engineering, Environmental Engineering, 16*(6), 913–918.
7. Arabi, M., Govindaraju, R. S., Hantush, M. M., & Engel, B. A. (2006). Role of watershed subdivision on modeling the effectiveness of best management practices with SWAT. *Journal of American Water Resources Association, 42*(2), 513–528.
8. Arnold, J. G., Srinisvan, R., Muttiah, R. S., & Williams, J. R. (1998). Large area hydrologic modeling and assessment, part I: model development. *Journal of American Water Resource Association, 34*(1), 73–89.
9. Back, A. D., & Trappenberg, T. P. (2001). Selecting inputs for modeling using normalized higher order statistics and independent component analysis. *IEEE Transactions in Neural Networks, 12*(3), 612–617.
10. Baran, B., Schaerer M., (2003). A multi-objective ant colony system for vehicle routing problem with time windows. *Twenty first IASTED International Conference on Applied Informatics, Insbruck, Austria*, February 10–13, pp. 97–102.
11. Barton, R. R., (1998). SIMULATION METAMODELS. *Proceedings of the 1998 Winter Simulation Conference*, Medeiros, D.J., Watson, E.F., Carson, J.S., and Manivannan, M.S. as editors. pp. 167–174.
12. Bozorg Haddad, O., Afshar, A., & Marino, M. A. (2006). Honey-bees mating optimization (HBMO) algorithm: a new heuristic approach for water resources optimization. *Water Resources Management, 20*(5), 661–680.
13. Bracmort, K. S., Arabi, M., Frankenberger, J. R., Engel, B. A., & Arnold, J. G. (2006). Modeling long-term water quality impact of structural BMPs. *Transactions American Society of Agricultural and Biological Engineers, 49*(2), 367–374.
14. Broad, D., Dandy, G. C., & Maier, H. R. (2005). Water distribution systems optimization using metamodels. *Water Resources Planning and Management, ASCE, 131*(3), 172–180.
15. Chu, T. W., & Shirmohammadi, A. (2004). Evaluation of the SWAT model's hydrology component in the piedmont physiographic region of Maryland. *Transaction of ASAE, 47*(4), 1057–1073.
16. Deb, K. (2001). *Multi objective optimization using evolutionary algorithms*. New York: Wiley.
17. Demirel, M. C., Venancio, A., & Kahya, E. (2009). Flow forecast by SWAT model and ANN in pracana basin, Portugal. *Advances in Engineering Software, 40*(7), 467–473.
18. Dorigo, M., Maniezzo, V., & Colorni, A. (1996). The ant system: optimization by a colony of cooperating ants. *IEEE Transaction on System Management Cybernetics, 26*(1), 29–41.
19. Doerner, K., Gutjahr, W. J., Hartl, R. F., Strauss, C., & Stummer, C. (2004). Pareto ant colony optimization: a meta-heuristic approach to multi-objective portfolio selection. *Annals of Operations Research, 131*(1–4), 79–99.
20. Elliot, A. H. (1998). Model for preliminary catchment scale planning of urban storm water quality controls. *Journal of Environmental Management, 52*(3), 273–288.
21. Gassman, P. W., Reyes, M. R., Green, C. H., & Arnold, J. G. (2007). The soil and water assessment tool: historical development, applications and future research directions. *Invited Review Series, American Society of Agricultural and Biological Engineers, 50*(4), 364–381.
22. Harrell, L. J., & Ranjithan, S. R. (2003). Detention pond design and land Use planning for watershed management. *Journal of Water Resources Planning and Management, 129*(2), 98–106.
23. Jalali, M. R., Afshar, A., & Marino, M. A. (2007). Multi-colony ant algorithm for continuous multi-reservoir operation optimization problem. *Water Resources Management, 21*(9), 1429–1447.
24. Johnson, V. M., & Rogers, L. L. (2000). Accuracy of neural network approximators in simulation-optimization. *Water Resources Planning and Management, ASCE, 126*(2), 48–56.
25. Karamouz, M., Taheriyoun, M., Baghvand, A., Tavakolifar, H., & Emami, F. (2010). Optimization of watershed control strategies for reservoir eutrophication management. *Journal of Irrigation and Drainage Engineering, 136*(12), 847–861.
26. Kaini, P., Artita, K., & Nicklow, J. W., (2009). Generating different scenarios of BMP designs in a watershed scale by combining NSGA-II with SWAT. *World Environmental and Water Resources Congress 2009: Great Rivers*, ASCE.,1-9.
27. Kaur, R., Singh, O., Srinivasan, R., Das, S. N., & Mishra, K. (2004). Comparison of a subjective and a physical approach for identification of priority areas for soil and water management in a watershed: a case study of Nagwan watershed in Hazaribagh District of Jharkhand, India. *Environmental Modelling and Assessment, 9*(2), 115–127.
28. Madadgar, S., & Afshar, A. (2009). An improved continuous Ant algorithm for optimization of water resources problems. *Water Resources Management, 23*(10), 2119–2139.
29. Maier, H. R., Simpson, A. R., Zecchin, A. C., Foong, W. K., Phang, K. Y., Seah, H. Y., & Tan, C. L. (2003). Ant colony optimization for the design of water distribution systems. *Journal of Water Resources and Planning Management, 129*(3), 200–209.
30. Mousavi, S. J., Ponnambalam, K., & Karray, F. (2007). Inferring operating rules for reservoir operation using fuzzy regression and ANFIS. *Fuzzy Sets Systems, 158*(10), 1064–1082.
31. Ostadrahimi, L., Marino, M. A., & Afshar, A. (2012). Multi-reservoir operation rules: multi-swarm PSO-based optimization approach. *Water Resources Management, 26*(2), 407–427.
32. Qi, H., & Altinakar, M. S. (2011). A conceptual framework of agricultural land use planning with BMP for integrated watershed management. *Journal of Environmental Management, 92*(1), 149–55.
33. Qi, H., & Altinakar, M. S. (2011). Vegetation buffer strips design using an optimization approach for Non-point source pollutant control, of an agricultural watershed. *Water Resources Management, 25*(2), 565–578.
34. Ranjithan, S., Eheart, J. W., & Garrett, J. H. (1993). Neural network based screening for groundwater reclamation uncertainty. *Water Resources Research, 29*(3), 563–574.
35. Rogers, L. L., & Dowla, F. U. (1994). Optimization of groundwater remediation using artificial neural networks with parallel solute transport modeling. *Water Resources Research, 30*(2), 457–481.
36. Rostamian, R., Jaleh, A., Afyuni, M., Mousavi, S. F., Heidarpour, M., Jalalian, A., & Abbaspour, K. C. (2008). Application of a SWAT model for estimating runoff and sediment in two mountainous basins in central Iran. *Hydrological Sciences-Journal, 53*(5), 977–988.

37. Sadeghi, S., (2010). Identifying and prioritizing causing erosion areas in the Gharesou Watershed. M.S. Dissertation, *Department of civil engineering of Islamic Azad University, Science and Research Branch*, Iran, Tehran.
38. Santhi, C., Arnold, J. G., Williams, J. R., Dugas, W. A., Srinivasn, R., & Hauck, L. M. (2001). Validation of the SWAT model on a large river basin with point and nonpoint sources. *Journal of American Water Resources Association, 37*(5), 1169–1188.
39. Santhi, C., Srinivasan, R., Arnold, J. G., & Williams, J. R. (2006). A modeling approach to evaluate the impacts of water quality management plans implemented in a watershed in Texas. *Environmental Modelling and Software, 21*(8), 1141–1157.
40. Shepherd, B., Harper, D., & Millington, A. (1999). Modeling catchment-scale nutrient transport to watercourses in the UK. *Hydrobiologia, 395–396*, 227–238.
41. Shourian, M., Mousavi, S. J., Menhaj, M., & Jabbari, E. (2008). Neural network-based simulation optimization model for optimal water allocation planning at basin scale. *Journal of Hydroinformatics, 10*(4), 331–343.
42. Sindelar, R., & Babuska, R. (2004). Input selection for nonlinear regression models. *IEEE Transactions on Fuzzy Systems, 12*(5), 688–696.
43. Emami Skardi, M. J., (2011). Design of detention ponds for watersheds, using ant colony algorithm (AC) and SWAT model; a cooperative game theory approach, MS Dissertation, Afshar, A., as Advisor, *Iran University of Science and Technology.*
44. Skardi, M. J., Afshar, A., & Sandoval Solis, S. (2013). Simulation-optimization model for Non-point source pollution management in watersheds: application of cooperative game theory. *KSCE Journal of Civil Engineering, 17*(6), 1232–1240.
45. Smith, J., & Eli, R. N. (1995). Neural network models for rainfall runoff Process. *Water Resources Planning and Management, ASCE, 121*(6), 499–508.
46. Spruill, C. A., Workman, S. R., & Taraba, J. L. (2000). Simulation of daily and monthly stream discharge from small watersheds using the SWAT model. *Transaction of ASAE, 43*(6), 1431–1439.
47. Sureerattanan, S., & Phein, H. N., (1997). Back propagation networks for daily stream flow forecasting. *Water Resources Research*, C/195, 1–7.
48. Tolson, B. A., & Shoemaker, C. A., (2004). Watershed modeling of the Cannonsville Basin using SWAT2000: model development, calibration and validation for the prediction of flow, sediment and phosphorus transport to the Cannonsville Reservoir. Technical Report, *School of Civil and Environmental Engineering, Cornell University.*, Ithaca, N. Y.
49. Vaché, K. B., Eilers, J. M., & Santelman, M. V. (2002). Water quality modeling of alternative agricultural scenarios in the U.S. Corn Belt. *Journal of American Water Resources Association, 38*(2), 773–787.
50. Wagner, B. J. (1995). Sampling design methods for groundwater modeling under uncertainty. *Water Resources Research, 31*(10), 2581–2591.
51. Wang, G. G., & Shan, S. (2006). Review of meta-modeling techniques in support of engineering design optimization. *Journal of Mechanical Design, 76*(5), 34–56.
52. Yan, S., & Minsker, B., (2006). Optimal groundwater remediation design using an adaptive neural network genetic algorithm. *Water Resources Research*, 42(5), 1–14. W05407.
53. Zecchin, A. C., Maier, H. R., Simpson, A. R., Roberts, A. J., Berrisford, M. J., & Leonard, M., 2003. Max-min ant system applied to water distribution system optimization. *International Congress on Modeling Simulation (MODSIM)*, Vol. 2, Townsville, Australia, pp. 795–800.
54. Zhang, X., Srinivasan, R., Debele, B., & Hao, F. (2008). Runoff simulation of the headwaters of the yellow river using the swat model with three snowmelt algorithms. *Journal of American Water Resources Association, 44*(1), 48–61.
55. Zhang, X., Srinivasan, R., & Van Liew, M. (2009). Approximating SWAT model using artificial neural network and support vector machine. *Journal of American Water Resources Association, 45*(2), 460–474.