

Comparison of Support Vector Machines and Response Surface Models in Meta-Modeling Applied in Basin-Scale Optimum Water Allocation

G.A. Mirfendereski^a and S. Jamshid Mousavi^b

^a MSc. Student, Department of Civil Engineering, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran, Email: ginaz_2005@aut.ac.ir

^b Associate Professor, Department of Civil Engineering, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran, Email: jmosavi@aut.ac.ir

Abstract: Basin-scale water planning and management problems may be approached by linking a river basin simulation model to a global heuristic optimization algorithm. Global optimization algorithms, however, need at least thousands of function evaluations each requires the river basin simulation model to run. This makes the resulting simulation-optimization model computationally expensive. Meta-modelling can be used to deal with the burden of computations in which a surrogate model, running of which is much faster than the exact simulation model, is used instead of the simulation model.

The surrogate or meta-model is built by using a function-approximation technique by which the expensive simulation model is approximated. Support Vector Machines (SVMs), a novel artificial intelligence based method, and Response surface modeling techniques are adopted and tested in this study as the meta-models approximating MODSIM river basin simulation model. PSO (particle Swarm Optimization) algorithm is linked first to the MODSIM DSS resulting in PSO-MODSIM model that can be used in solving a variety of water resource problems at basin scale. Then an adaptive sequentially space filling meta-modeling approach is developed in which SVM and polynomial-based surrogate models replace MODSIM. In this approach the accuracy of the approximate model is sequentially improved in course of optimization in an adaptive way. Finally, the performance of the PSO-MODSIM, PSO-MODSIM~SVM and PSO-MODSIM~Polynomial models are evaluated by their application to integrated water resources planning problem of Atrak river basin located in north-east of Iran and the results are analysed and compared.

Keywords: River Basin, Optimization, Support Vector Machines, Response surface, Meta-modelling

1. INTRODUCTION

Modeling water resources at basin scale calls for the need of using different kind of mathematical models representing complex hydrologic, socio-economic and management processes taking place in a river basin. This need may be dealt with through linking a river basin simulation model to a global heuristic optimization algorithm. Global optimization algorithms, however, need at least thousands of function evaluations each requires the river basin simulation model to run. This makes the resulting simulation-optimization model computationally expensive. Meta-modeling can be used to lower the burden of computations in which a surrogate model, running of which is much faster than the exact simulation model, is used instead of the simulation model. The surrogate or meta-model is built by using a function-approximation technique by which the expensive simulation model is approximated. Artificial Neural Network (ANN), Support Vector Machine (SVM), Kriging and Polynomial models are some of the most common meta-models used.

Robinson and Keane (1999) presented a case for employing variable-fidelity analysis models and approximation techniques to improve the efficiency of evolutionary optimization for complex design tasks. Computational frameworks for integrating a class of single-point approximation models with Evolutionary Algorithms (EAs) were proposed by Nair and Keane (2001). However, such frameworks are restricted to a special class of approximation models that are domain specific. For more general approximation models, Ratle (2001) examined a strategy for integrating evolutionary search with Kriging models. Numerical studies were presented for certain pathological cases to show that the idea of constructing an accurate global surrogate model may be fundamentally flawed due to the curse of dimensionality. Liang et al. (2000) proposed a strategy for coupling EAs with local search and quadratic response surface methods. However, when working with multimodal problems, the accuracy of quadratic models may become questionable. Farhang-Mehr and Azarm (2005) used a sequential approach with adaptation to irregularities in the response behavior for Bayesian meta-modeling in engineering design simulations. Shourian et al. (2008) linked MODSIM river basin DSS to the PSO algorithm to simulate a river basin system operation and to evaluate the fitness of each set of selected design and operational variables with respect to the model's objective function, which is the minimization of the system's design and operational cost. Since the direct incorporation of MODSIM into a PSO algorithm is computationally prohibitive, they trained an ANN model as a meta-model to approximate the MODSIM modeling tool. Mousavi et al. (2010) presented a methodology in which the problems of design of experiments, function approximation and function optimization in a surrogate optimization problem are sequentially solved in a feedback loop so that a much fewer number of experiments is required for the task of function approximation. The proposed approach adaptively utilizes the information obtained from function optimization, finds the regions where more data are needed, updates the training data set to fill the space and sequentially improves the accuracy of the meta-model.

In spite of extensive works carried out on this topic, comparison of different meta-modeling techniques especially in hydro-systems is limited. In this paper, Support Vector Machines (SVMs) and Polynomial models are adopted and tested as meta-models approximating MODSIM river basin simulation model. PSO (particle Swarm Optimization) algorithm is linked first to the MODSIM DSS resulting in PSO-MODSIM model to solving the Atrak river basin water allocation problem. Then the mentioned two surrogate models have replaced MODSIM in the PSO algorithm according to an adaptive sequentially space filling (ASSF) meta-modeling approach, in which the accuracy of the approximate model is sequentially improved during the course of optimization.

2. META-MODELING TECHNIQUES

2.1. Support Vector Machines (SVMs)

SVMs are learning systems that use a linear high dimensional hypothesis space called feature space. This method was introduced by Vapnik (1998) and his colleagues as a robust and significant learning tool, which uses a learning bias derived from statistical learning theory (SLT).

SVMs can be also employed for regression estimation, called as support vector regression (SVR) in which the real value functions are estimated. The goal of learning process is to find a function $f(x)$ as an approximation of the value $y(x)$ with minimum error, which is based on the available independent and identically distributed data, i.e:

$$(x_1, y_1), \dots, (x_m, y_m) \subseteq (X \subseteq R^n \times Y \subseteq R) \quad (1)$$

The approximate function is determined by a small subset of training samples called support vectors (SVs). A specific loss function called ϵ -insensitive loss is developed to make a sparseness property for SVR (Figure 1).

In order to learn the non-linear relations by linear machines, it is needed to select a set of non-linear features and to rewrite the data in the new representation; that is equivalent to applying a fixed non-linear mapping of the input space to a feature space in which the linear machine can be used. Therefore, the considering hypothesis set (i.e. linear functions in the feature space) will be functions of the type:

$$f(\mathbf{x}) = \langle W, \phi(\mathbf{x}) \rangle + b \quad W \in F, b \in R$$

Where $\phi: X \rightarrow F$ is a non-linear mapping from input space to some feature space. Finally, the following equation will be concluded;

$$f(\mathbf{x}) = \sum_{i=1}^m (\alpha_i^* - \alpha_i) \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}) \rangle + b \quad \mathbf{w} = \sum_{i=1}^m (\alpha_i^* - \alpha_i) \phi(\mathbf{x}_i) \tag{2}$$

If inner product $\langle \phi(x_i), \phi(x) \rangle$ is computed in the feature space as a function of the original input points in a direct way, the explicit mapping of data will not be required. This implicit computation method is called Kernel method. A Kernel function for all $x, z \in Z$ can be illustrated as follows:

$$K(\mathbf{x}, \mathbf{z}) = \langle \phi(\mathbf{x}), \phi(\mathbf{z}) \rangle \tag{3}$$

Linear, polynomial, sigmoid and radial basis functions (RBF) are some common Kernels. Eventually, by substituting the inner product $\langle \phi(x_i), \phi(x) \rangle$ with a suitable Kernel function the approximate function will be obtained as follows:

$$f(\mathbf{x}) = \sum_{i=1}^m (\alpha_i^* - \alpha_i) K(\mathbf{x}_i, \mathbf{x}) + b \tag{4}$$

2.2. Response surface model (Polynomial)

Response surface modeling techniques were originally developed to analyze the results of physical experiments to create empirically-based models of the observed response values. Response surface modeling can be written in the following form:

$$y(x) = f(x) + \epsilon \tag{5}$$

where $y(x)$ is the unknown function of interest, $f(x)$ is the polynomial approximation of x , and ϵ is random error that is assumed to be normally distributed with mean equal to zero and variance σ^2 . At each observation the error, ϵ_i , is assumed to be independent and identically distributed. The polynomial function, $f(x)$, used to approximate $y(x)$ is typically a low order polynomial, which is assumed to be either linear, e.g. (6), or quadratic, e.g. (7):

$$\hat{y} = \beta_0 + \sum_{i=1}^k \beta_i x_i \tag{6}$$

$$\hat{y} = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i=1}^k \sum_{j<i} \beta_{ij} x_i x_j \tag{7}$$

Parameters $\beta_0, \beta_i, \beta_{ii}$ and β_{ij} of the polynomials are determined using the least-squares regression method, which minimizes the sum of the squares of the deviations of predicted values, $\hat{y}(x)$, from the actual values, $y(x)$. In this study, Eq. (7) is used.

3. ADAPTIVE SEQUENTIALLY SPACE FEELING META-MODELING

The formulation of an optimization model, referred to as problem (1), objective function evaluation of which requires performance of a simulation model may be represented as below:

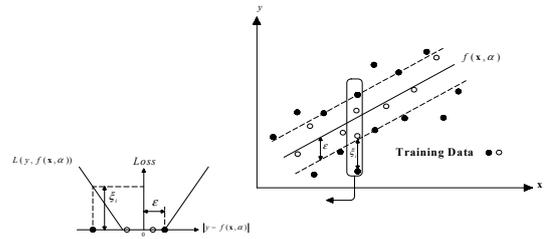


Figure 1. ϵ -insensitive Loss function

$$\text{Min } f(y) \tag{9}$$

Subject to:

$$h_i(x) = 0, \quad i = 1, \dots, m \quad g_j(x) \geq 0, \quad j = 1, \dots, n \quad y = \text{sim}(x)$$

In problem 1, $f(y)$ is the objective function to be minimized and $h_i(x)$ and $g_j(x)$ are the equality and inequality constraints of the problem. In order to evaluate $f(y)$, in many cases, there is a need to use a high-fidelity expensive simulation model. Moreover, the mathematical functions and relationships in the simulation model, $\text{sim}(x)$, may not be algebraic functions to be used in gradient-based optimization techniques. There are also many cases in which the simulation models are some commercial or public domain computer packages used as a black box model in the optimization procedure and the modeler may not have access to computer codes and routines (functions) of the model. It should be noted that function $f(y)$ may be multimodal with respect to decision vector x and also some of functions $h_i(x)$ or $g_j(x)$ may be nonsmooth or the feasible space of the problem may be nonconvex. Metaheuristic and evolutionary optimization techniques are promising in solving these types of optimization problems as they can be easily linked with any simulation model without the need to have access to computer codes or details of the function $\text{sim}(x)$. However, metaheuristic algorithms to converge to a good solution may require thousands of objective function ($f(y)$) evaluation each needs running the high-fidelity simulation model $\text{sim}(x)$. This makes solving problem (1) using a metaheuristic optimization algorithm computationally prohibitive. To address this difficulty, the incorporation of meta-models has been suggested. The purpose of using meta-models is to approximate the relationship between $f(y)$ and the vector of input decision variables x by using efficient mathematical models. Therefore, a function approximation technique may be used in solving problem (1) wherein approximate function $\tilde{f}(y)$ is optimized instead of optimizing exact function $f(y)$. The optimization problem in this case may be expressed as below, referred to as problem (2):

$$\text{Min } \tilde{f}(y) \tag{10}$$

Subject to:

$$h_i(x) = 0, \quad i = 1, \dots, m \quad g_j(x) \geq 0, \quad j = 1, \dots, n$$

In problem (2), $\tilde{f}(y)$ is used in lieu of the exact function $f(y)$. Hence, to solve main optimization problem (1), we need to determine the surrogate function $\tilde{f}(y)$. To have an accurate approximation the following condition should be satisfied (Mousavi et al. (2010)):

$$\left| \tilde{f} - f \right| \leq \varepsilon, \quad \forall x \in X \tag{11}$$

where ε is the accuracy parameter of approximation and X is the search space of problem (1). In order to obtain the required precision for function approximation used in the optimization problem (2), there is no other way but to design enough experiments to fill the search space uniformly. These experiments are used to construct the meta-model used as a function approximator. Therefore, one could say that D experiments are needed, which means that the function $f(y)$ has to be evaluated for D times. In other

words, in order to get the approximate function $\tilde{f}(y)$ in problem (2), a set of experiments known as $S_D = \{(x_k, f(\text{sim}(x_k)))\}$, $k = 1, \dots, D$, should be prepared which requires D times of evaluating function f . This problem is referred to as problem (3) which is the problem of design of experiments.

The procedure starts with generating a number of input-output data in the search space, just enough to construct an approximate surrogate model. The experiments are designed randomly and the function $f(\text{sim}(x))$ is evaluated for each experiment through performing the original simulation model (design of experiments and function evaluation, problem (3)). A meta-model is trained using the data which can be used for approximating function f (function approximation that is referred to as problem (2a)). Minimum

value of approximate function \tilde{f} is obtained using the PSO algorithm (approximate function optimization, problem (2)). This value will be almost the same as the minimum value of the exact function f , if the approximation procedure has been properly worked out. To verify this, one should evaluate the error of the approximation (ε) between \tilde{f} and f for the solution of problem (2). If ε is small enough, there is a high chance for the located minimum value of \tilde{f} in problem (2) to be a near-optimal solution for the main problem (1); Otherwise, the trained meta-model may have not been able to learn and approximate the exact function f accurately. If so, the solutions obtained for problem (2) are not taken as the solutions of problem (1) but just some representatives of the gaps in the search space of the problem where there is not enough number of experiments and thus more experiments around the solutions should be designed. These new experiments are added to the former set of experiments, SD , and the procedure is iterated until the whole search space of the problem and especially the regions that good solutions of the main problem (1) are in is covered with the required intensity of data.

The main feature of the proposed methodology is to consider the problems of design of experiments (function evaluation), function approximation and function optimization in a combined and related process. The relation between the different parts of the problem solving procedure is shown in Figure 2.

4. ATRAK RIVER BASIN WATER ALLOCATION PROBLEM

Optimum water allocation planning in Atrak river basin system in north-east of Iran has been considered as the case study problem in this research. To adequately simulate different complex institutional, hydrologic and socio-economic relationships in the river basin system under study, it is desirable to use a generalized river basin DSS as the simulator engine. MODSIM, a generalized river basin simulation model (Labadie, 1995) has been selected for this purpose.

Taking advantages of MODSIM's custom coding features, MODSIM has been already embedded in a PSO algorithm to solve a river basin system optimization problem (Shourian et al. 2008). Deciding on the size of water resource projects and how water resources should be allocated over time and space considering coordinated operation of the system components are some challenging issues in management of the basin, which could be dealt with by formulating a large scale optimization model for the Atrak river basin system. The schematic representation of the system in MODSIM's graphical user interface is shown in Figure 3.

As seen in Figure 3, there are 13 reservoirs (red triangles) in the basin among which Ghordanlu, Darband, Garmab, Amand, Chaily, Chandir and Sumbar reservoirs are not constructed yet; so they have to be sized. Also, relative priorities of objective reservoir storages are operational variables through which the network flow program in MODSIM decides whether water stored in reservoirs in a time period should be released to meet water demands of that time period or it should be kept in the reservoirs for future uses. These priorities indicate the significance of storing in or releasing water from reservoirs in every time step in comparison with the priorities of the water demands. Therefore, there are two main sets of decision variables which design variables include

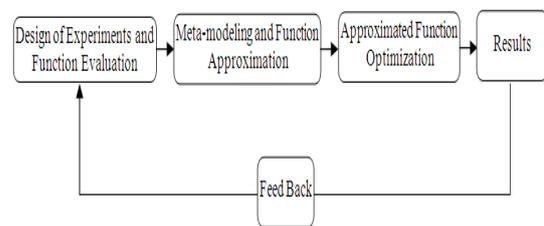


Figure 2. Relation between different parts of the proposed methodology for meta-modeling (Mousavi et al. (2010))

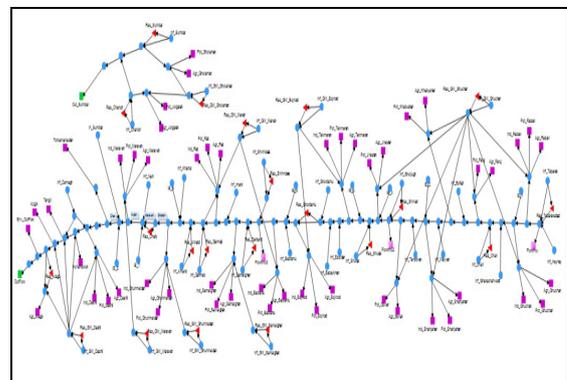


Figure 3. Representation of Atrak river basin in MODSIM's GUI

Table 1. minimum and maximum amount for design variables

Dam	Darband	Garmab	Ghordanlu
Maximum	10 MCM	10 MCM	50 MCM
Minimum	40 MCM	5 MCM	5 MCM

the sizes of unconstructed dams and operational variables contain the relative priorities of objective reservoir storages.

4.1. Results

The PSO-MODSIM model was performed first to solve the optimal water allocation problem of Atrak basin with parameter values as 10 particles, $w_{min} = 0.1$, $w_{max} = 1.2$ and $C_1 = 1.75$, $C_2 = 1.25$ and maximum iteration number of 50. Table 1 presents minimum and maximum values of design variables. Upper and lower bounds of priorities were also set as 10 and 1, respectively.

The problem was also solved by ASSF approach using SVM and Polynomial meta-models. The meta-models started with 200 initial random data. Table 2 reports the solutions found by the two meta-models using 30 particles and 3000 iterations. The results reported are the best solutions obtained among 10 runs.

As seen in Table 2 both PSO-MODSIM~SVM and PSO-MODSIM~POLYNOMIA models have been able to find the solutions which are the same as that of PSO-MODSIM model. The PSO-MODSIM~SVM has, however, required a less number of function evaluations compared to the PSO-MODSIM~POLYNOMIAL. This seems to

be reasonable as the accuracy of SVM with mean square error (MSE) of $5.37E^{-5}$ has been much better than that of Surface response model with MSE equal to 0.00188.

Figures 4 and 5 show prediction versus actual values of the objective function for PSO-MODSIM~SVM and PSO-MODSIM~POLYNOMIA models.

Table 2. Optimal water allocation at Atrak basin using Adaptive Sequentially Space Feeling meta-modeling technique

Method	PSO-MODSIM	PSO-MODSIM-SVM	PSO-MODSIM-POLYNOMIAL
Approximate optimum	-	137830	137577
Exact fitness value	137837	137837	137837
Error	-	5.37E-05	0.00188
number of exact function evaluation	2500	477	1667
Darband Dam (mcm)	5	5	5
Garmab Dam (mcm)	5	5	5
Gordanlu Dam (mcm)	40	40	40
Darband Priority	1	3	4
Garmab Priority	3	4	6
Ghordanlu Priority	10	3	1

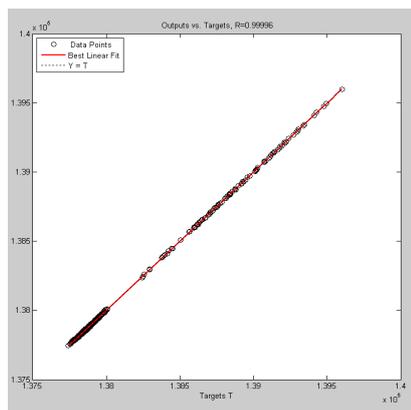


Figure 4. SVM-based prediction (y-axis) vs. actual values of the objective function (x-axis)

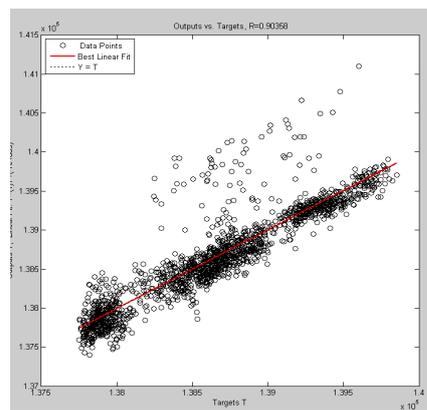


Figure 5. Polynomial-based prediction (y-axis) vs. actual values of the objective function (x-axis)

5. DISCUSSION AND CONCLUSIONS

In this study the PSO algorithm was linked first to the MODSIM DSS resulting in PSO-MODSIM model used in solving the problem of optimum water allocation of Atrak River Basin in north-east of Iran. Two surrogate models of Support Vector Machines (SVMs) and Polynomial response surface were then used within the adaptive sequentially space filling (ASSF) meta-modeling approach. The results showed that although both meta-models are efficient in reducing the PSO-MODSIM model's computational

burden, SVM performed better than the other one in terms of saving in the number of function evaluations required to solve the problem.

ACKNOWLEDGMENTS

The authors would like to thank Iran Water Resource Centre of Water Research Institute and Dr. Mojtaba Shourian for providing the required data of the case study.

REFERENCES

- Farhang-Mehr, A and Azarm, S. (2005). Bayesian meta-modeling of engineering design simulations: a sequential approach with adaptation to irregularities in the response behavior, *International Journal for Numerical Methods in Engineering*, (62): 2104-2126.
- Labadie, J. (1995), MODSIM: river basin network flow model for conjunctive stream-aquifer management, program user manual and documentation, Colorado State University.
- Liang, K. H., Yao, X. and Newton, C. (2000). Evolutionary search of approximated N-dimensional landscapes: *International Journal of Knowledge-Based Intelligent Engineering Systems*, 4(3): 172-183.
- Mousavi, S. J., and M. Shourian (2010). Adaptive sequentially space-filling metamodeling applied in optimal water quantity allocation at basin scale, *Water Resour. Res.*, 46, W03520, doi:10.1029/2008WR007076
- Nair, P. B. and Keane, A. J. (2001). Passive vibration suppression of flexible space structures via optimal geometric redesign, *AIAA*, 39(7): 1338-1346.
- Ratle, A. (2001). Kriging as a surrogate fitness landscape in evolutionary optimization. *Artificial Intelligence for Engineering Design Analysis and Manufacturing*, 15(1), 37-49.
- Robinson, G. M. and Keane, A. J (1999). A case for multi-level optimization in aeronautical design, *Aeronautical*, 103(1028): 481-485.
- Schölkopf, B., & Smola, A. J. (2002). *Learning with Kernels: Support vector machines, regularization, optimization, and beyond*. Cambridge, Mass: MIT Press.
- Shourian M., Mousavi S. J. and Tahershamsi A. (2008a). Basin-wide Water Allocation Planning by Integrating PSO Algorithm and MODSIM, *Water Resources Management*, 22, pp. 1347-1366.