International Journal of GEOMATE, Aug., 2018 Vol.15, Issue 48, pp.79-85
Geotec., Const. Mat. & Env., DOI: https://doi.org/10.21660/2018.48.7169
ISSN: 2186-2982 (Print), 2186-2990 (Online), Japan

MULTIPLE OBJECTIVE MANAGEMENT STRATEGIES FOR COASTAL AQUIFERS UTILIZING NEW SURROGATE MODELS

*Alvin Lal[1] and Bithin Datta[1]

[1]College of Science & Engineering, James Cook University, Australia

*Corresponding Author, Received: 15 Jan. 2018, Revised: 26 Feb. 2018, Accepted: 10 April 2018

ABSTRACT: Coastal aquifers are hydraulically connected to the sea and therefore susceptible to saltwater intrusion problems. This study proposes the utilization of a new surrogate model within coupled simulation-optimization (S/O) model for the management of coastal aquifers subjected to density-dependent saltwater intrusion processes. The simulation of the transient 3-dimensional density-dependent flow and transport model is based on the solution of an implemented numerical simulation model. Direct coupling of the numerical simulation model into the multi-objective genetic algorithm (MOGA) is computationally expensive. Hence, the solution of the numerical simulation model with random input variables are used to train and test the support vector machine regression (SVMR) surrogate models for approximately simulating the flow and transport processes. The performances of the new surrogate models are evaluated using various performance evaluation criteria. The resulting validated SVMR surrogate models are coupled to the MOGA and implemented for an illustrative coastal aquifer with an aim to develop efficient coastal aquifer management strategies. Based on the objective functions, execution of S/O model presented a set of optimal groundwater withdrawal rates from the simulated aquifer. It also ensured salinity levels at the designated monitoring wells are constrained within specified limits. The efficiency of the new SVMR surrogate models is also demonstrated. Evaluation results suggested that the projected S/O model is an effective way of developing feasible and reliable coastal aquifer management strategies. It also enhances the possibility of solving more realistic large-scale problems and developing regional-scale coastal aquifer management methodologies.

Keywords: Coastal aquifers, Density-dependent flow and transport model, Multi-objective, Surrogate models

1. INTRODUCTION

Saltwater intrusion (SWI) is imposing a long-term threat to the beneficial use of life-sustaining groundwater resources in coastal areas. Sustainable planning and management of coastal aquifers are decisive to ensure the future sustainability of these fragile resources. Simulation-optimization (S-O) models are one of the widely used methods for developing management methodologies for coastal aquifers subjected to SWI. The present study demonstrates an integration of a new yet efficient surrogate model into a coupled S-O framework for developing improved, and efficient optimal strategies for groundwater utilization from coastal aquifers.

The S-O approach has been used to develop different coastal aquifer management models [1-6]. In many cases, due to enormous computational requirements and time constraints, numerical simulation models in an S-O model are substituted by surrogate models. Some of the common surrogate models used in developing aquifer management strategies *via* an S/O model includes Radial Basis Function [7], Artificial Neural Network [8, 9], Modular Neural Network [10], and Genetic Programming [5, 11]. However, the reliability and accuracy of surrogate model incorporated S/O models are always questionable. The present study focuses on developing and implementing robust support vector machine regression (SVMR) surrogate model assisted S/O model for optimal coastal aquifer management.

In numerous surrogate model performance comparison studies, SVMR has been ranked as an accurate and most efficient predictive modeler [12-14]. Despite several successful applications and numerous benefits, SVMR surrogate model has never been used for predicting SWI in aquifers. It is for the first time, SVMR surrogate models are trained for emulating a numerical costal aquifer flow and transport simulation model FEMWATER [15] used for predicting the impact of variable groundwater pumping patterns. The surrogate model is then integrated into a multi-objective genetic algorithm (MOGA) optimization algorithm within the R2016a MATLAB environment, to develop a multi-objective coastal aquifer management strategy.

2. METHODS

2.1 Coastal Aquifer Management Model

The proposed coastal aquifer management model was designed to sustainably withdraw groundwater from installed production wells (PWs) and barrier wells (BWs). PWs were installed for withdrawing fresh groundwater for domestic utilization whereas BWs were installed near the shoreline to control SWI into the aquifer. Pumping from BWs induces a steeper hydraulic gradient towards the sea, averting encroachment of seawater into the aquifer [6]. Thus, two conflicting objectives i.e. maximizing total pumping from PWs and minimizing total pumping from BWs were considered. Monitoring wells (MWs) were installed for monitoring salinity level in the aquifer. Maintaining salinity levels at respective MWs within specified limits were incorporated as constraints in the SVMR-MOGA optimization framework. The mathematical expressions of the conflicting objective functions, constraints and bounds [16] are given by;

Maximize,

$$F_1(P) = \sum_{n=1}^{N} \cdot \sum_{t=1}^{T} P_n^t \qquad (1)$$

Minimize,

$$F_2(P) = \sum_{m=1}^{M} \cdot \sum_{t=1}^{T} p_m^t \qquad (2)$$

Constraints

$$c_i = \xi(P, p) \qquad (3)$$

$$c_i \le c_{\max,\, i}\ \forall i, t \qquad (4)$$

Bounds

$$P_{\min} \le P_n^t \le P_{\max} \qquad (5)$$

$$p_{\min} \le p_m^t \le p_{\max} \qquad (6)$$

P_n^t denotes pumping from *nth* PW at t^{th} time and p_m^t denotes pumping from the m^{th} BW at t^{th} time. c_i represents the saline concentration at the i^{th} monitoring well at the end of management time period. $\xi()$ symbolizes the surrogate model replacing the numerical FEMWATER model and constraint (3) denotes coupling of the surrogate model within the S-O framework. *M*, *N* and *T* are the total number of PW, BW and a total number of time steps in the management model. Inequality (4) represents the constraints imposed to keep salinity concentrations within specified limits at the respective MWs. Inequality (5) and (6) represents the upper and lower bounds of pumping from PWs and BWs respectively. Pumping bounds for both the PW and BW was set between 0 – 1300 m^3/day. The constraints imposed as permissible limits on concentrations (assumed to be a conservative pollutant) were $c_i \le c_{\max,\, i}$ of 425 mg/L at MW1, $c_i \le c_{\max,\, j}$ of 510 mg/L at MW2 and $c_i \le c_{\max,\, k}$ of 625 mg/L at MW3.

2.2 The Numerical Simulation Model

The 3D numerical simulation FEMWATER model was used for simulating pumping induced SI processes into an illustrative costal aquifer system. FEMWATER model (FM) allowed simulation of density-dependent coupled groundwater flow and transport processes in an aquifer system. An illustrative study area, similar to [16] containing of a portion of a multi-layered coastal aquifer was modelled using FEMWATER. The length of the coastline (sea side boundary) was 2.13 km and the other two boundaries were of 2.04 km (*Boundary A*) and 2.79 km (*Boundary B*) respectively. The 2.53 km^2 study area incorporated 5 BWs, 8 PWs and 3 MWs. The study area with specific good locations is presented in Fig. 1.

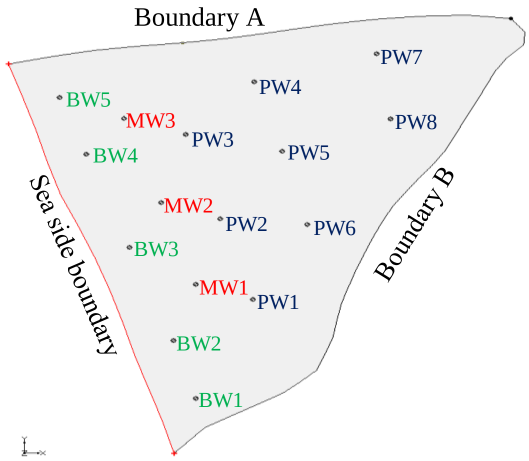

Fig.1 Illustrative study area with specific PW, BW and MW locations

The sea side boundary was assumed to be a constant head and constant concentration boundary having a concentration of 35000 mg/L. The other two boundaries of the study area were taken as no-flow boundaries. The modelled aquifer was discretized into triangular finite elements having an average element size of 150 m. The element size near the wells was set to 75 m. The total aquifer depth was 60 m, divided equally into 3 layers. A constant vertical groundwater recharge of 0.00054 m/d was specified over the entire study area. The screening interval of all the wells was taken from the second and third layers of the aquifer. The compressibility and velocity of water were taken as 6.69796 X 10^{-20} md^2/kg and 131.328 kg.md respectively. Other key parameters used for aquifer simulation are listed in Table 1.

The 3D transient simulation was instigated from an initially steady state condition of the aquifer, achieved via constant pumping of 300 m^3/day from 3 of the PWs for a period of 20 years only. The resulting heads and salinity concentrations after 20 years were used as initial conditions for the specified period of 4 years (4^{th}-time step) where pumping from all production and barrier wells were instigated.

Table 1 Key parameter values for model development

Properties		Values
Hydraulic Conductivity	*x* direction	15 m/d
	y direction	7.5 m/d
	z direction	1.5 m/d
Bulk density		1600 kg/m^3
Longitudinal dispersivity		50 m
Lateral dispersivity		25 m
Molecular dispersion coefficient		0.69 m^2/d
Density reference ratio		0.025
Soil porosity		0.46
Compressibility		8.5x10^{-15} md^2/kg

2.3 Development of Support Vector Machine Regression Surrogate Models

SVMR methodology is a statistical tool employed for numeric data prediction utilizing support vector machines (SVM). SVMs have been applied in various engineering fields because of its attractive features and encouraging empirical performance [17]. A comprehensive discussion on SVM is presented in [18-20] and only a brief theoretical background is given below. For a given training dataset (x_i, y_i) where x_i is the *i*th input pattern and y_i is the parallel target output and $y_i \in R$. The aim of the SVMR is to find a function $f(x)$ that has most ε deviation from the targets y_i for all training data, and also is at flat as possible [21]. A SVM takes advantage of the kernel function to map the input data onto a high-dimensional feature space. Later, linear regression is performed in the high-dimensional feature space. As a result, non-linear problems are addressed in a linear space through non-linear feature mapping. The final prediction function used by an SVM is:

$$f(x_i) = \sum_{m=1}^{M} \alpha_i y_i K(x_i, x_j) + b; \qquad (7)$$

where, $\mathbf{0 < \alpha_i < C}$

where, α_i is Lagrange multiplier, and x_i is a feature vector corresponding to a training object. The components of vector α_i and the constant b are optimized during training. C is the penalty factor, which regulates the trade-off between the flatness of f and the extent up to which deviations greater than ε can be accepted. The kernel function is one of the most important parts in the SVMR model. Commonly used kernels are linear, Gaussian, polynomial and sigmoid. The Gaussian function kernel is the most commonly used kernel because of its effectiveness and speed [17]. Mathematical expression of the Gaussian function is given by:

$$K(x_i, x_j) \exp(\gamma \left| x_i \cdot x_j \right|^2) \qquad (8)$$

where γ is the parameter of the kernel function with representing independent variables? Key parameters that control SVMR problems are the cost function C is the radius of the insensitive tube ε and the kernel parameter. These parameters are dependent on each other so altering the value of one parameter leads to a change in another parameter. For the present study, Gaussian kernel was used, with ε, C and γ having a value of 0.60, 6.49 and 0.001 respectively. Evaluating other values of these fundamental parameters were beyond the scope of this study, although, other parameter values may have provided more accurate results.

Five hundred data sets were generated out which 400 was used for training and 100 were used for testing the trained SVMR models. The training set was used to construct the SVMR models while the testing set assessed the prediction capabilities of the trained model. 500 transient pumping (inputs) were obtained from uniform sampling distribution using Latin Hypercube Sampling (LHS) having an upper bound of 1300 m^3/day and lower bound of 0 m^3/day. The resulting salt concentration at each monitoring well was obtained from FEMWATER simulation after each set of pumping from all production and barrier wells are fed to the model as inputs. Each FEMWATER simulation took approximately about 4-5 minutes to converge. 500 sets of pumping and resulting output concertation were assembled by running the simulation model 500 times. These input-output patterns were used for surrogate model training and testing purpose. The SVMR models were constructed to learn from the training data presented to them with the intent of capturing the functional relation between the pumping-concentration data sets. 3 SVMR surrogate models namely M^1, M^2 and M^3 were constructed for predicting salinity concentrations at MW1, MW2 and MW3 respectively.

2.4 Performance Analysis of Surrogate Models

Mathematical expressions of the performance evaluation indices (PEI's) are presented from Eq. (9) to (12).

$$RMSE = \sqrt{\frac{1}{K}\sum_{k=1}^{K}\left(c_k^o - c_k^p\right)^2} \tag{9}$$

$$MSE = \frac{1}{K}\sum_{k=1}^{K}\left(c_k^o - c_k^p\right)^2 \tag{10}$$

$$c_c = \frac{\sum_{k=1}^{K}\left(c_k^o - c^o\right)\left(c_k^p - c^p\right)}{\sqrt{\sum_{k=1}^{K}\left(c_k^o - c^o\right)^2}\sqrt{\sum_{k=1}^{K}\left(c_k^p - c^p\right)^2}} \tag{11}$$

$$NSE = 1 - \frac{\sum_{k=1}^{K}\left(c_k^o - c_k^p\right)^2}{\sqrt{\sum_{k=1}^{K}\left(c_k^o - c^o\right)^2}} \tag{12}$$

where, c_k^o and c_k^p are observed (from FM) and predicted values of saltwater concentrations respectively, c^o and c^p are observed and predicted saltwater concentration mean values respectively and K represents total data points.

2.5 Linked simulation-optimization approach

A flowchart for the linked surrogate assisted S/O framework in given in Fig. 2. The trained and tested SVMR models were used in the S/O model as a set of binding constraints for salinity concentration prediction purpose. Multi-objective Genetic algorithm (MOGA) was used as an optimization tool. MOGA has been used efficiently in solving multi-objective optimization problems [5, 16]. When using MOGA optimization tool, population size was set to 1500, 5200 generations, crossover fraction of 0.8 and mutation probability of 0.02.

3. RESULTS AND DISCUSSION

3.1 SVMR Model Prediction Capabilities

The error estimates during the training and testing phases are given in Table 2. The performances of the SVMR models at the training and testing phases for M^1, M^2 and M^3 in terms of the 4 evaluation criteria showed similar trends. RMSE, and MSE values for concentrations at M^1, M^2 and M^3 at both the stages were substantially smaller.

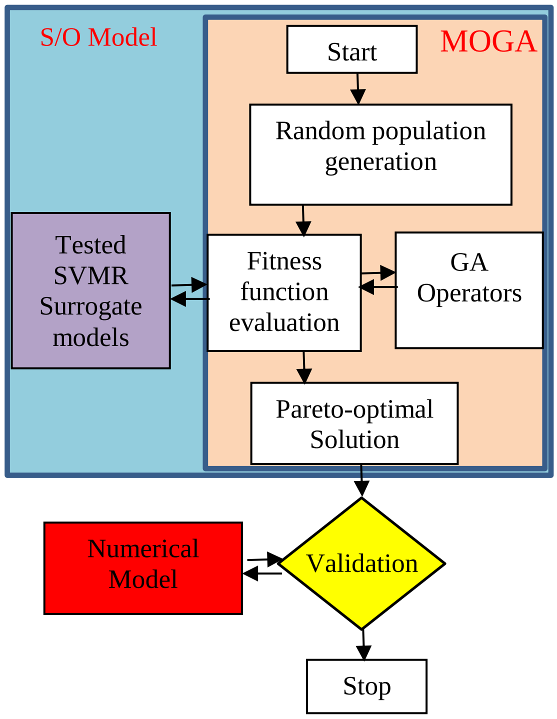

Fig. 2 Flowchart for the linked S/O model with optimal solution validation stage

Table 2 Performance evaluation results

Phase	PEI's	M^1	M^2	M^3
Training	MSE	0.164	0.057	0.184
	RMSE	0.405	0.238	0.428
	R^2	0.997	0.997	0.989
	NSE	0.990	0.990	1.00
Testing	MSE	0.202	0.105	0.187
	RMSE	0.449	0.323	0.432
	R^2	0.994	0.993	0.989
	NSE	0.990	1.000	1.000

During the training stage, SVMR model has a smallest RMSE value of 0.238 accomplished for M^2 and highest RMSE of 0.428 attained for M^3. Comparing RMSE values of SVMR models at the testing stage, the lowest 0.323 is obtained for M^2 and highest value of 0.449 is achieved for M^1. Accordingly, MSE values of the SVMR models are comparatively smaller. R^2 and NES values do not significantly differ for the three models and have values close to 1. NSE value of 1 presents a perfect estimate with no errors [22]. A model can be considered accurate if the calculated NSE value is greater than 0.8 [23]. The NSE values for M^1, M^2 and M^3 were greater than 0.8 indicating that it can reliably be used for SI prediction. A model's good performance in the testing phase is an indication of its accuracy in prediction and hence its practical utility [24]. Hence, the observed good performance of SVMR models at the testing stage establishes its credibility, suggesting accurate results when used for prediction purposes.

3.2 Optimal Groundwater Pumping Solution Set

After evaluating the prediction performances, the M^1, M^2 and M^3 replaced the original FEMWATER model in the S-O framework instituted for obtaining optimal groundwater pumping rates from the PWs and BWs installed in the simulated coastal aquifer. The execution of the SVMR-MOGA coastal aquifer management model presented a set of optimal pumping solution. The resulting Pareto-front from the executed S-O model is given in Fig. 3. Depending on the relative importance each of the two objective functions in the management model, a set of pumping solutions can be chosen and implemented for optimal and sustainable utilization of groundwater from the simulated aquifer.

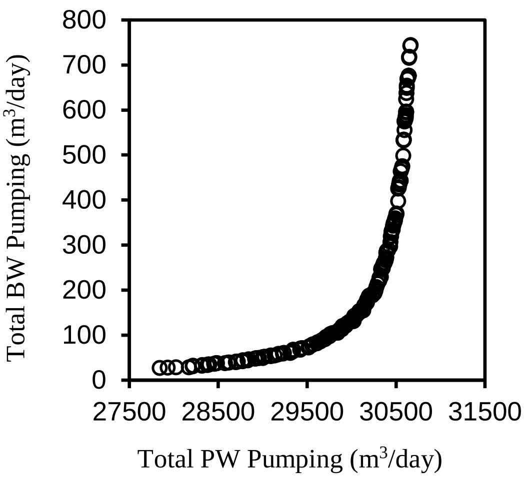

Fig. 3 Optimal Pareto-front from the executed SVMR-MOGA model

The selected optimal pumping solutions from the Pareto-front will ensure optimal pumping and maintaining salinity levels at the MWs within specified limits which were the major aim of the proposed management model. The Pareto optimal solutions also show the conflicting nature of the two objectives. For example, if the pumping for beneficial use increases beyond 30500 m^3/day, the marginal increase in a barrier well-pumping increases exponentially.

3.3 Validation of Optimal Pumping Solutions

To establish the validity of the resulting optimal solution, the salinity concentration obtained from the SVMR surrogate models were compared with the salinity concentration obtained from the original FM. Five random solutions from the optimal Pareto-front were chosen and the salinity comparison results are presented in Table 3.

Table 3 Optimal pumping solution validation results

$c_i \le c_{max,\ i}$ at MW1 ≤ 425 mg/L		$c_i \le c_{max,\ j}$ at MW2 ≤ 510 mg/L		$c_i \le c_{max,\ k}$ at MW3 ≤ 625 mg/L	
SVMR	FM	SVMR	FM	SVMR	FM
423.8	422.9	509.2	507.9	624.2	623.9
424.5	423.5	509.4	508.3	623.6	622.1
423.2	422.2	508.9	506.5	624.6	623.5
425.0	425.3	508.4	508.9	623.1	622.6
424.7	423.6	509.4	508.3	624.7	622.0

It was observed that the salinity concentration values from both the modeling tools were very close to each other. It was also observed that the optimal pumping values satisfied the specified constraints at monitoring locations. Also, the optimal solutions converged to the specified upper bound of the concentrations constraint. The validation results established that the SVMR surrogate models predicted the salinity concentrations at MWs accurately and can be utilized for developing coastal aquifer management strategies.

4. CONCLUSION

This paper presents an efficient and feasible SVMR model assisted coupled S/O model for the optimal sustainable management of coastal aquifers subjected to SWI. The S/O model maximizes pumping from PWs, and minimizes pumping from BWs while maintaining salinity concentrations at respective MWs within specified limits. Utilized SVMR surrogate models have effectively approximated density-dependent SWI processes in the simulated aquifer. The execution of the developed management model presented a set of optimal pumping patterns for both the PWs and BWs. An optimal pumping solution set can be chosen and implemented for the sustainable utilization of the modelled coastal aquifer. The implementation of the developed management model has the potential to aid in developing regional-scale coastal aquifer management strategies while ensuring computational efficiency, thus making it feasible to address much larger study areas. In future, it would be beneficial to compare the prediction performances of SVMR with other available surrogate modeling methods and also apply the established management model on a real case study area.

5. REFERENCES

[1] Javadi A., Hussain M., Sherif M., and R. Farmani., Multi-objective optimization of different management scenarios to control seawater intrusion in coastal aquifers, Water Resources Management, Vol. 29, 2015, pp. 1843-1857.

[2] Kourakos G. and Mantoglou A., Development of a multi-objective optimization algorithm using surrogate models for coastal aquifer management, Journal of Hydrology, Vol. 479, 2013, pp. 13-23.

[3] Ataie-Ashtiani B., Ketabchi H., and Rajabi M. M., Optimal management of a freshwater lens in a small island using surrogate models and evolutionary algorithms, Journal of Hydrologic Engineering, Vol. 19, 2013, pp. 339-354.

[4] Ataie-Ashtiani B. and Ketabchi H., Elitist continuous ant colony optimization algorithm for optimal management of coastal aquifers, Water resources management, Vol. 25, 2011, pp. 165-190.

[5] Sreekanth J. and Datta B., Multi-objective management of saltwater intrusion in coastal aquifers using genetic programming and modular neural network-based surrogate models, Journal of Hydrology, Vol. 393, 2010, pp. 245-256.

[6] Dhar A. and Datta B., Saltwater intrusion management of coastal aquifers. I: Linked simulation-optimization, Journal of hydrologic engineering, Vol. 14, 2009, pp. 1263-1272.

[7] Christelis V. and Mantoglou A., Pumping Optimization of Coastal Aquifers Assisted by Adaptive Metamodelling Methods and Radial Basis Functions, Water Resources Management, 2016, pp. 1-15.

[8] Bhattacharjya R. K., Datta B., and Satish M. G., Artificial neural networks approximation of density-dependent saltwater intrusion process in coastal aquifers, Journal of Hydrologic Engineering, Vol. 12, 2007, pp. 273-282.

[9] Bhattacharjya R. K. and Datta B., ANN-GA-based model for multiple objective management of coastal aquifers, Journal of Water Resources Planning and Management, Vol. 135, 2009, pp. 314-322.

[10] Kourakos G. and Mantoglou A., Pumping optimization of coastal aquifers based on evolutionary algorithms and surrogate modular neural network models, Advances in water resources, Vol. 32, 2009, pp. 507-521.

[11] Sreekanth J. and Datta B., Comparative evaluation of genetic programming and neural network as potential surrogate models for coastal aquifer management, Water resources management, Vol. 25, 2011, pp. 3201-3218.

[12] Khalifelu Z. A. and Gharehchopogh F. S., Comparison and evaluation of data mining techniques with algorithmic models in software cost estimation, Procedia Technology, Vol. 1, 2012, pp. 65-71.

[13] Li Y., Ng S. H., Xie M., and Goh T., A systematic comparison of metamodeling techniques for simulation optimization in decision support systems, Applied Soft Computing, Vol. 10, 2010, pp. 1257-1273.

[14] Mukherjee S., Osuna E., and Girosi F., Nonlinear prediction of chaotic time series using support vector machines, in Neural Networks for Signal Processing [1997] VII. Proceedings of the 1997 IEEE Workshop, 1997, pp. 511-520.

[15] Lin H.-C. J., Richards D. R., Yeh G.-T., Cheng J.-R., and Cheng H.-P., FEMWATER: A Three-Dimensional Finite Element Computer Model for Simulating Density-Dependent Flow and Transport in Variably Saturated Media, DTIC Document 1997.

[16] Sreekanth J. and Datta B., Coupled simulation ‐ optimization model for coastal aquifer management using genetic programming ‐ based ensemble surrogate models and multiple ‐ realization optimization, Water Resources Research, Vol. 47, Issue 4, 2011.

[17] Liu B., Hu J., Yan F., Turkson R. F., and Lin F., A novel optimal support vector machine ensemble model for NO_x emissions prediction of a diesel engine, Measurement, Vol. 92, 2016, pp. 183-192.

[18] Welling M., Support vector regression, Department of Computer Science, University of Toronto, Toronto (Canada), 2004.

[19] Smola A. J. and Schölkopf B., A tutorial on support vector regression, Statistics and Computing, Vol. 14, 2004, pp. 199-222.

[20] S. R. Gunn, Support vector machines for classification and regression, ISIS technical report, vol. 14, 1998.

[21] Betrie G. D., Sadiq R., Morin K. A., and Tesfamariam S., Uncertainty quantification and integration of machine learning techniques for predicting acid rock drainage chemistry: A probability bounds approach, Science of the Total Environment, Vol. 490, 2014, pp. 182-190.

[22] He Z., Wen X., Liu H., and Du J., A comparative study of artificial neural network,

adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region, Journal of Hydrology, Vol. 509, 2014, pp. 379-386.

[23] Shu C. and Ouarda T., Regional flood frequency analysis at ungauged sites using the adaptive neuro-fuzzy inference system, Journal of Hydrology, Vol. 349, 2008, pp. 31-43.

[24] Khalil A., Almasri M. N., McKee M., and Kaluarachchi J. J., Applicability of statistical learning algorithms in groundwater quality modeling, Water Resources Research, Vol. 41, Issue 5, 2005.