J. Hydrol. Hydromech., *60*, 2012, 1, 16–32
DOI: 10.2478/v10098-012-0002-7

COMPARISON BETWEEN ACTIVE LEARNING METHOD AND SUPPORT VECTOR MACHINE FOR RUNOFF MODELING

HAMID TAHERI SHAHRAIYNI[1)], MOHAMMAD REZA GHAFOURI[2)], SAEED BAGHERI SHOURAKI[3)], BAHRAM SAGHAFIAN[4)], MOHSEN NASSERI[5)]

[1)]Faculty of Civil and Environmental Engineering, Tarbiat Modares Univ., Tehran, Iran; Mailto: hamid.taheri@modares.ac.ir;
[2)]Shahrood University of Technology, Shahrood, Iran; Mailto: mrgh.ghafouri@gmail.com;
[3)]Department of Electrical Eng., Sharif Univ. of Tech., Tehran, Iran; Mailto: bagheri-s@sharif.edu;
[4)]Soil Conservation & Watershed Management Research Institute, Ministry of Jihad Agriculture, Tehran, Iran;
Mailto: saghafian@scwmri.ac.ir;
[5)]School of Civil Engineering, College of Engineering, University of Tehran, Tehran, Iran; Mailto: mnasseri@ut.ac.ir

In this study Active Learning Method (ALM) as a novel fuzzy modeling approach is compared with optimized Support Vector Machine (SVM) using simple Genetic Algorithm (GA), as a well known data-driven model for long term simulation of daily streamflow in Karoon River. The daily discharge data from 1991 to 1996 and from 1996 to 1999 were utilized for training and testing of the models, respectively. Values of the Nash-Sutcliffe, Bias, R^2, MPAE and PTVE of ALM model with 16 fuzzy rules were 0.81, 5.5 $m^3 s^{-1}$, 0.81, 12.9%, and 1.9%, respectively. Following the same order of parameters, these criteria for optimized SVM model were 0.8, –10.7 $m^3 s^{-1}$, 0.81, 7.3%, and –3.6%, respectively. The results show appropriate and acceptable simulation by ALM and optimized SVM. Optimized SVM is a well-known method for runoff simulation and its capabilities have been demonstrated. Therefore, the similarity between ALM and optimized SVM results imply the ability of ALM for runoff modeling. In addition, ALM training is easier and more straightforward than the training of many other data driven models such as optimized SVM and it is able to identify and rank the effective input variables for the runoff modeling. According to the results of ALM simulation and its abilities and properties, it has merit to be introduced as a new modeling method for the runoff modeling.

KEY WORDS: Runoff Modeling, Active Learning Method (ALM), Support Vector Machine (SVM), Fuzzy Modeling, Genetic Algorithm, Karoon River Basin.

Hamid Taheri Shahraiyni, Mohammad Reza Ghafouri, Saeed Bagheri Shouraki, Bahram Saghafian, Mohsen Nasseri: POROVNANIE METÓDY AKTÍVNEHO UČENIA S METÓDOU VEKTORMI PODPORENÝCH STROJOV PRI MODELOVANÍ ODTOKU. J. Hydrol. Hydromech., *60*, 2012, 1; 66 lit., 9 obr., 3 tab.

Cieľom štúdie bolo porovnať možnosti dlhodobej simulácie denných prietokov v rieke Karoon pomocou novovyvinutej fuzzy metódy aktívneho učenia (Active Learning Method – ALM) a známej metódy vektormi podporených strojov (Support Vector Machine – SVM), optimalizovanej genetickým algoritmom (GA). Na tréning a testovanie modelov boli použité časové rady denných prietokov za obdobie rokov 1991 až 1996 a 1996 až 1999. Hodnoty parametrov Nash-Sutcliffe, Bias, R^2, MPAE a PTVE pre model ALM boli 0,81; 5,5 $m^3 s^{-1}$; 0,81; 12,9% a 1,9%. Parametre v tom istom poradí pre model SVM boli 0,8 –10,7 $m^3 s^{-1}$, 0,81; 7,3%; a –3,6%. Z výsledkov simulácií vyplýva, že aplikáciou metód ALM a SVM možno získať porovnateľné a akceptovateľné výsledky. Podobnosť výsledkov medzi ALM a SVM implikuje vhodnosť novovyvinutej metódy ALM pre simuláciu odtoku. Tréning ALM je ľahší a jednoduchší ako je tréning ďalších dátami riadených modelov podobného typu. Navyše algoritmus ALM je schopný identifikovať a zoradiť efektívne vstupné premenné pre modelovanie odtoku. Na základe dosiahnutých výsledkov možno metódu ALM zaradiť medzi nové, alternatívne metódy modelovania odtoku.

KĽÚČOVÉ SLOVÁ: modelovanie odtoku, metóda aktívneho učenia (ALM), metóda vektormi podporených strojov (SVM), fuzzy modelovanie, genetický algoritmus, povodie rieky Karoon.

1. Introduction

Estimation of streamflow has a significant economic implication in agricultural water management, hydropower generation and flood - drought control. Many techniques are currently used for modeling of hydrological processes and generating of synthetic streamflow. One of these techniques are physically based (conceptual) methods which are specifically designed to simulate the subprocesses and physical mechanisms, related to the hydrological cycle. Implementation and calibration of these models can typically present various complexities (*Duan* et al., 1992), requiring sophisticated mathematical tools (*Sorooshian* et al., 1993), significant amount of calibration data (*Yapo* et al., 1996), and some degree of expertise and experience (*Hsu* et al., 1995). For a case study which has insufficient or no measured data of watershed characteristics, data-driven models (non-physically models) are often used for runoff simulation (*Wang*, 2006). These models are useful because they can be applied easily by avoiding of mathematical complexities. Most frequently used data-driven models are regression based, time series, artificial neural network (ANN) and fuzzy logic (FL) (e.g. *Hsu* et al., 1995; *Smith* and *Eli*, 1995; *Saad* et al., 1996; *Shamseldin*, 1997; *Markus*, 1997; *Maier* and *Dandy*, 1998; *Tokar* and *Johnson*, 1999; *Zealand* et al., 1999; *Jain* et al., 1999; *Chang* and *Chen*, 2001; *Cheng* et al., 2002; *Sivakumar* et al., 2002; *Chau* et al., 2005, *Kisi*, 2005; *Lin* et al., 2006; *Zounemat-Kermani* and *Teshnehlab*, 2007; *Anvari Tafti*, 2008). In addition, in the recent years, new data driven models have been frequently used for hydrological modeling and forecasting. For example *Elshafie* et al. (2007) compared ANFIS with ANN in the flow forecasting in the Nile River and results demonstrated that ANFIS has more capability from ANN for flow forecasting. *Firat* (2008) applied the ANFIS, ANN and AR (Auto Regressive) models for the forecasting of daily river flow in Seyhan and Cine rivers. The results exhibited that ANFIS is better other models.

Support Vector Machine (SVM) as a new data-driven model have remarkable successes in various fields and its ability has been demonstrated in hydrological prediction and runoff modeling (*Dibike* et al., 2001; *Smola* and *Schlkopf*, 2004; *Asefa* et al., 2006; *Yu* et al., 2006; *Behzad* et al., 2009). *Wang* et al. (2009) utilized of ARMA, ANN, ANFIS, SVM and GP (Genetic Programming) for the simulation of monthly flow discharge in two rivers (Lancang and Wujiang Rivers in China). The results demonstrated that ANFIS, GP and SVM are the best models. *Wu* et al. (2009) studied on the application of ARMA, ANN and SVM for monthly runoff forecasting in the Xiangjiabe (1, 3, 6 and 12 month ahead forecasting). The results showed that SVM outperformed than the other models.

Comparison between the ANN and SVM for one day ahead forecasting in the Bakhtiyari River-Iran demonstrated that SVM has better performance than ANN (*Behzad* et al., 2009).

According to the literature, the ANFIS and SVM are promising method for appropriate runoff simulation and forecasting. But the training of these methods is time consuming and need to expertise.

This subject leads to finding and utilizing of an artificial intelligence method with straightforward and easy training construction.

Early concepts on principles of fuzzy logic were proposed by *Zadeh* (1965). Although in the beginning, fuzzy logic was thought not to comply with scientific principles, but its capability was demonstrated by an application carried out by *Mamdani* and *Assilian* (1975). A fuzzy logic system can model human's knowledge qualitatively by avoiding delicate and quantitative analyses. Today, fuzzy logic is applied to most engineering fields. Several studies have been carried out using fuzzy logic in hydrology and water resources planning (e.g. *Liong* et al., 2006; *Mahabir* et al., 2000; *Chang* and *Chen*, 2001; *Nayak* et al., 2004a; *Sen* and *Altunkaynak*, 2006; *Tayfur* and *Singh*, 2006). *Bagheri Shouraki* and *Honda* (1997) suggested a new fuzzy modeling technique similar to the human modeling method that its training is very easy and straightforward. This method, entitled the active learning method (ALM) has a simple algorithm and avoids mathematical complexity. *Taheri Shahraiyni* (2007) developed a new heuristic search, fuzzification and defuzzification methods for ALM algorithm that resulted in a modified ALM.

Up to now, no research has been performed using ALM as a novel fuzzy method for the streamflow modeling. In this study, for the evaluation of ALM in runoff modeling, it is compared with optimized SVM via Genetic Algorithm (GA) as a well-known model for the simulation of daily runoff in Karoon III River.

2. Case study

Karoon III basin (a subbasin of the Large Karoon) is located in the southwest of Iran and drains into the Persian Gulf. The basin lies within 49° 30' to 52° E longitude and 30° to 32° 30' N latitude with an area of approximately 24,200 km^2. Some 30 reliable climatology and synoptic gauges are operated in the basin. The elevation ranges from 700m at the Pol-e-Shaloo hydrometric station (outlet of the Karoon III basin and just upstream of Karoon III dam) to 4500 m in Kouhrang and Dena Mountains. The digital elevation model (DEM) and major drainage system of the basin is shown in Fig. 1. About 50% of the area is higher than 2500m from MSL (Mean Sea Level). Average annual precipitation of watershed is about 760 mm. 55% of the precipitation is as snowfall. Average daily discharge flow of Karoon III basin is about 384 m^3 s^{-1}.

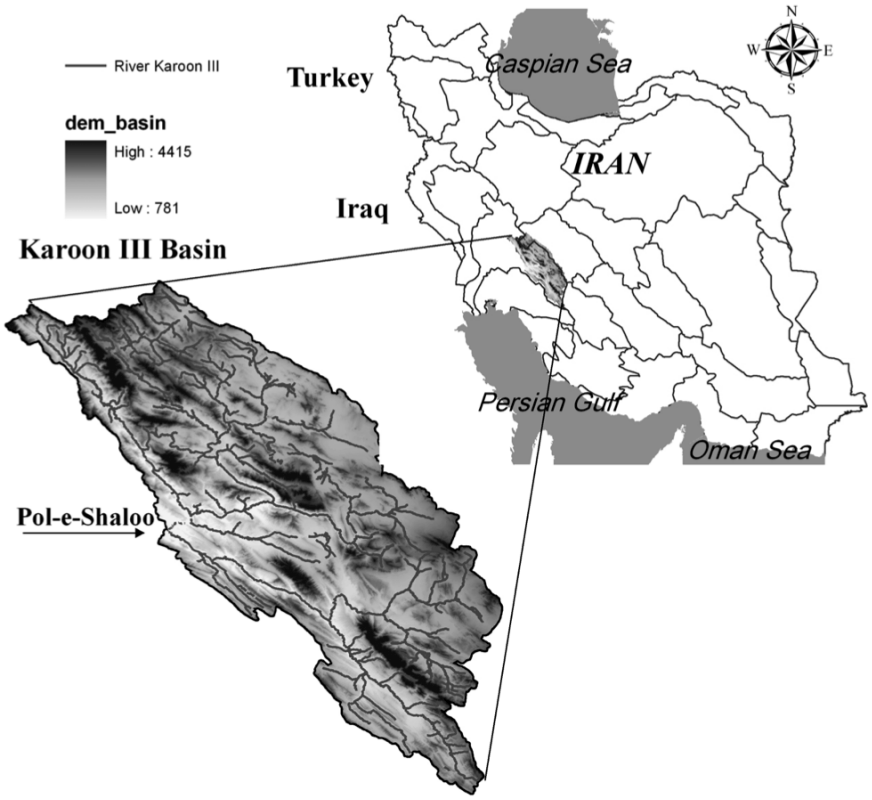


Fig. 1. DEM and major drainage network of Karoon III basin (subbasin of large Karoon).

3. Support Vector Machine (SVM) Theory

SVM principals have been developed by *Vapnik* and *Cortes* (1995). SVM is a well known modeling method, hence it is explained briefly. It is a new generation of statistical learning methods which aim to recognize the data structures. One of the SVM utilities in detecting the data structure is transformation of original data from input space to a new space (feature space) with a new mathematical paradigm entitled Kernel function which has been developed by *Boser* et al. (1992). For this purpose, a non-linear transformation function $\phi(.)$ is defined to map the input space to a higher dimension feature space. According to Cover's theorem, a linear function $f(.)$ can be formulated in the high dimensional feature space to represent a non-linear relation between the inputs (x_i) and the outputs (y_i) as follows (*Vapnik* and *Cortes* 1995):

$$y_i = f(x_i) = \langle w, \phi(x_i) \rangle + b. \quad (1)$$

where w and b are the model parameters, which are solved mathematically. SVM can be used for both regression as Support Vector Regression (SVR), and classification as Support Vector Classification (SVC). In this study, the SVR structure will be used for runoff simulation. SVR was developed using more sophisticated error functions (*Vapnik*, 1998).

3.1. Feature selection

Feature selection is a general procedure of selecting a suitable subset of the pool of original feature spaces according to discrimination capability to improve the quality of data, and performance of simulation technique. Feature selection techniques can be categorized into three main branches (*Tan* et al., 2006); Embedded approaches, Wrapper approaches, and Filter approaches.

Embedded approaches are preventive and they have been developed for particular (not general) classification algorithm.

In the wrapper methods, the objective function is usually a pattern classifier or a mathematical regression model which evaluates feature subsets by their predictive accuracy using statistical resampling or cross-validation approach. The most important weakness of wrapper method is its computation cost and it is not recommended for large feature sets.

Filter approach utilizes a statistical criterion to find the dependency between the input candidates and output variable(s). This criterion acts as a statistical benchmark for reaching the suitable input variable dataset. The three famous filter approaches are linear correlation coefficient, Chi-square criterion and Mutual Information (MI). The linear correlation coefficient investigates the dependency or correlation between input and output variables (*Battiti*, 1994). In spite of popularity and simplicity of linear correlation coefficient, this approach has shown inappropriate results for the feature selection in the real non-linear systems.

Chi-square criterion is considered for evaluation of goodness of fit and is based on nonlinearity of data distribution and known as a classical nonlinear data dependency criterion (*Manning* et al., 2008).

Mutual Information (MI), as another filtering method, describes the reduction amount of uncertainty in estimation of one parameter when another is available (*Liu* et al., 2009). It has been widely used for feature selection which is nonlinear and can effectively represent the dependencies of features (*Liu* et al., 2009). This method as a non linear filter method has recently been found to be a more suitable statistical criterion in feature selection. It has also been found to be robust due to its insensitivity to noise and data transformations and also has no pre-assumption in correlation of input and output variables (*Battiti*, 1994; *Bowden* et al. 2002; *Bowden* et al. 2005a, 2005b; *May* et al. 2008a, 2008b).

Mutual Information (MI) index developed in two form of parameter types, continues and discrete parameters. In realm of discrete parameters, as the current case, MI index could be estimated for two variables X and Y as follows,

$$I(X,Y)=\sum_{y\in Y}\sum_{x\in X}p(y,x)L\log\left(\frac{p(y,x)}{p(x)p(y)}\right). \qquad (2)$$

In this equation $p(x,y)$, $p(x)$ and $p(y)$ are joint probability and marginal probability of two parameter x and y respectively, and $I(X,Y)$ is the MI of X and Y.

4. The Active Learning Method (ALM)

4.1 ALM Algorithm

The ALM algorithm has been presented in Fig. 2.

For the purpose of explaining the ALM algorithm, the *Sugeno* and *Yasukawa* (1993) dummy non-linear static problem (Eq. (3)) with two input variables (x_1 and x_2) and one output (y) is solved by this method.

$$y=\left(1+x_1^{-2}+x_2^{-1.5}\right)^2, \quad 1\le x_1,x_2\le 5. \qquad (3)$$

First, some data are extracted from Eq. (3) (step 1). Then the data are projected on x–y plane (Figs. 3a and 3b) (step 2).

Step 3: The heart of calculation in ALM is a fuzzy interpolation and curve fitting method which is entitled IDS (Ink Drop Spread). The IDS searches fuzzily for continuous possible paths on data planes. Assume that each data point on each x–y plane is a light source with a cone or pyramid shape illumination pattern. Therefore, with increase of distance of each data point, the intensity of light source decreases and goes toward zero. Also the illuminated pattern of different data points on each x–y plane interfere together and new bright areas are formed. The IDS is exerted to each data point (pixel) on the normalized and discretized x–y planes. The radius of the base of cone or pyramid shape illumination pattern in each x–y plane is related to the position of data in it. The radius increases until the all of the domain of variable in x–y plane be illuminated. Figs. 3c and 3d show the created illumination pattern (IL values) after interfere of the illumination pattern of different points in x_1 – y and x_2 – y planes, respectively. Here, pyramid shape illumination pattern has been used.

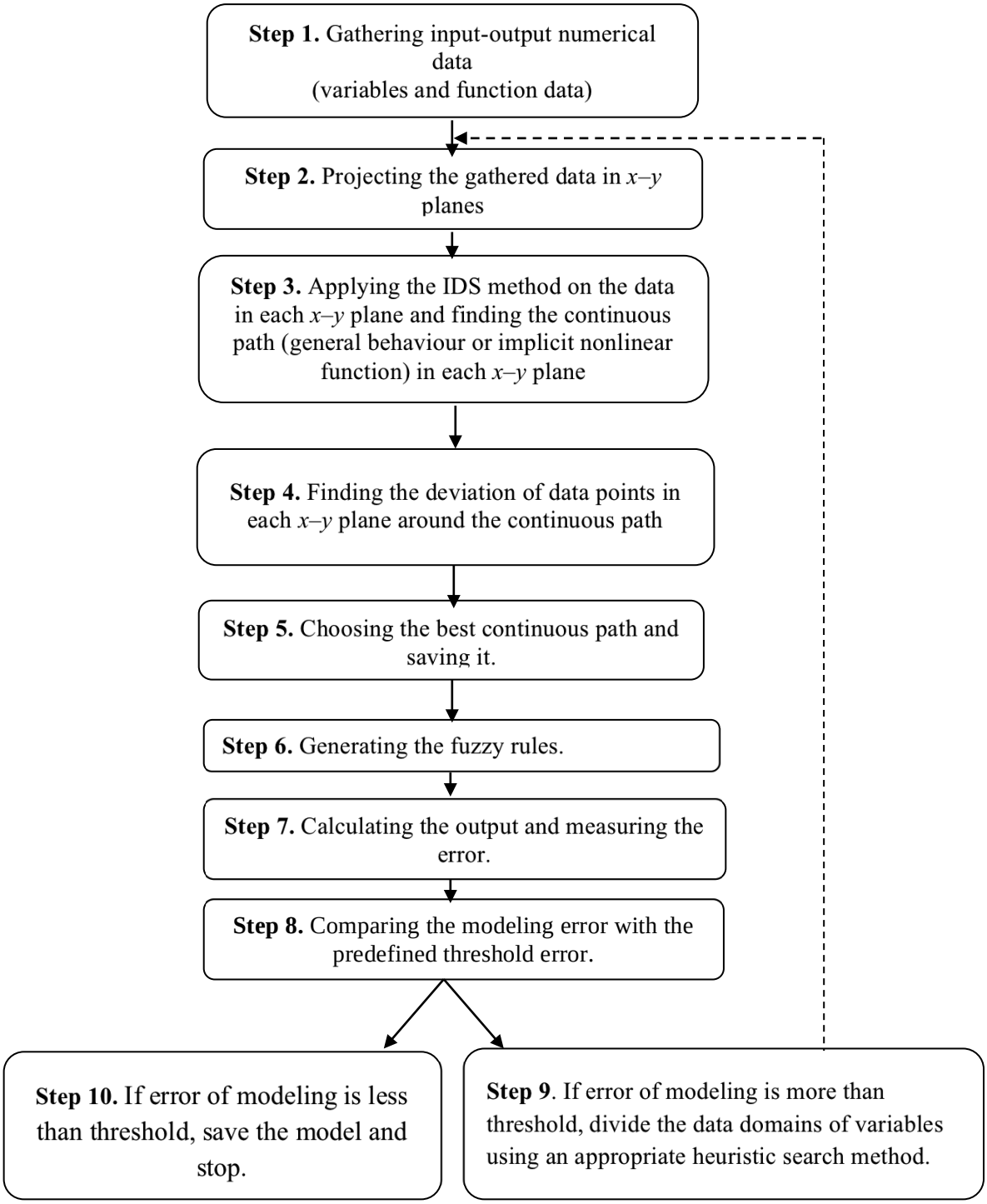


Fig. 2. Proposed algorithm for Active Learning Method.

Now, the paths or general behavior, or implicit nonlinear functions are determined by applying the center of gravity on y direction. The center of gravity is calculated using this equation:

$$y(x_i) = \frac{\sum_{j=1}^{M} (y_j \times IL(x_i, y_j)}{\sum_{j=1}^{M} y_j}, \tag{4}$$

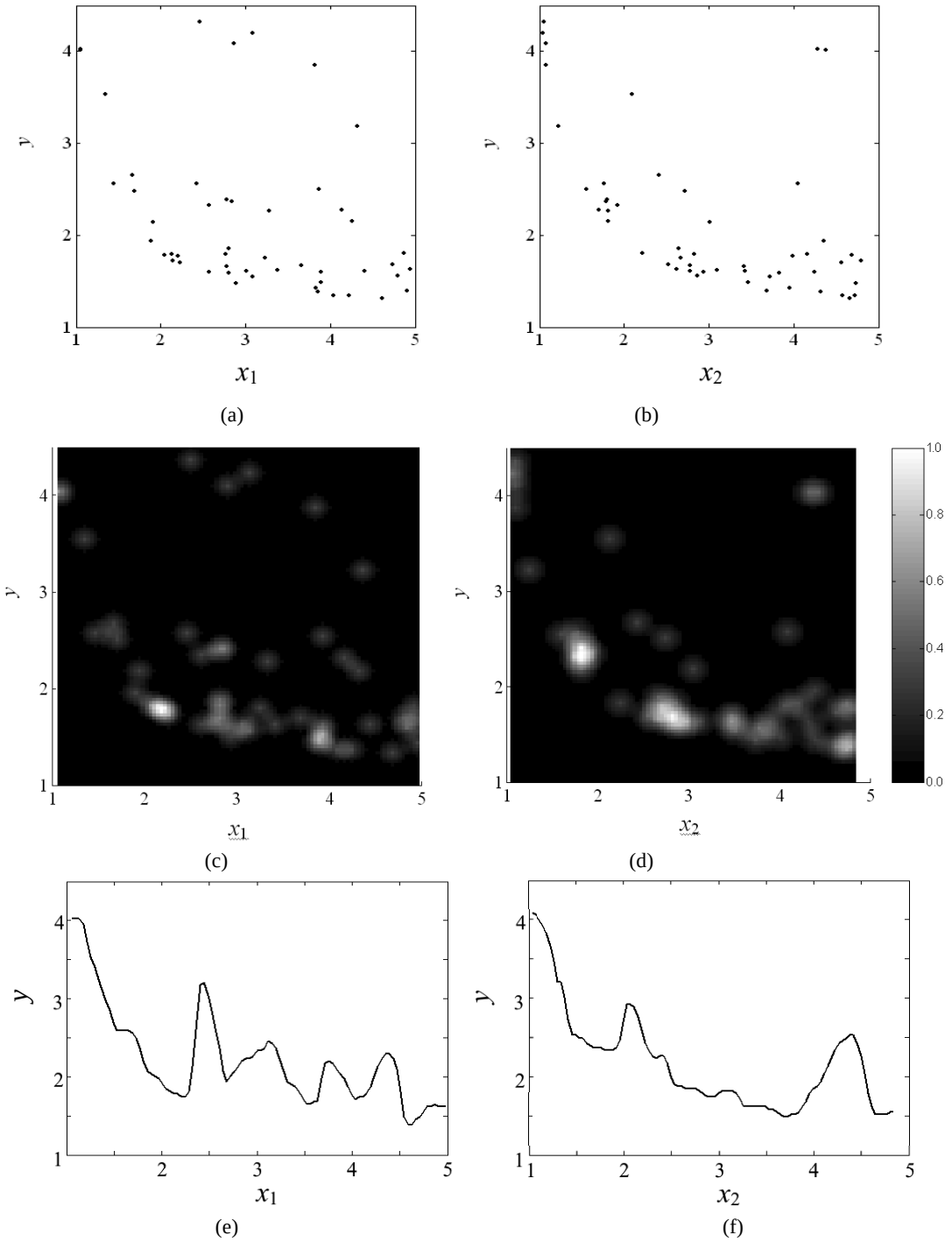


Fig. 3. (a) Projected data on x1–y plane; (b) projected data on x2–y plane; (c) results of applying IDS method on the data points in x1–y plane; (d) results of applying IDS on the data points in x2–y plane; (e) extracted continuous path by applying center of gravity method on Fig. 3c; (f) extracted continuous path by applying center of gravity method on Fig. 3d.

where j:1…M, M is the resolution of y domain, y_j – the output value in j-th position, $IL(x_i,y_j)$ – the illumination value on $x - y$ plane at the (x_i, y_j) point or pixel, and $y(x_i)$ is the corresponding function (path) value to x_i .

Hence, by applying the centre of gravity method on Figs. 3c and 3d, continuous paths are extracted (Figs. 3e and 3f).

Subsequently, the deviation of data points around each continuous path can be calculated by various

methods such as coefficient of determination (R^2), Root Mean Square Error (RMSE) or Percent of Absolute Error (PAE). The PAE values of continuous paths on x_1–y and x_2 – y planes (Figs. 3e and 3f) are 20.4 and 13.5%, respectively (Step 4).

The results show that the path of Fig. 3f is better than the path of Fig. 3e. The selected paths should be saved because these are implicit non-linear functions. The path can be saved as a look-up table, heteroassociative neural network memory (*Fausset*, 1994) or fuzzy curve expressions such as Takagi and Sugeno method (TSM) (*Takagi* and *Sugeno*, 1985). Look up tables are most convenient method and it is used for path saving in this example (Step 5).

We have no rules in the first iteration of ALM algorithm, hence we go to step 7.

The PAE of chosen path is more than a predefined threshold PAE value (5%). Hence, the error is more than predefined error (Steps 7 and 8) and we divide each space in two by using only one variable (Step 9) and go to the step 2 of Fig. 2. Dividing can be performed crisply or fuzzily, but for simplicity, a crisp dividing method is used here and the fuzzy dividing will be illustrated later. The results of ALM modeling after crisp division of space to four subspaces using a heuristic search method is presented in Fig. 4. From Fig. 4, the following rules are generated (Step 6):

If ($x_2 \geq 1.0$ & $x_2 < 1.9$) then $y = f(x_2)$

If ($x_2 \geq 1.9$ & $x_2 < 2.9$) then $y = g(x_1)$

If ($x_2 \geq 2.9$ & $x_1 < 2.9$) then $y = h(x_1)$

If ($x_2 \geq 2.9$ & $x_1 > 2.9$) then $y = u(x_2)$.

Whenever the PAE value of the above rules is less than the threshold of 5%, the procedure of ALM modeling is stopped. Here, using four rules, a PAE of 3.8% is achieved.

Then, the modeling error (e_{11}) is calculated for the above rules. Similarly, the domain of other variables are divided and their modeling errors are calculated and a set of k errors (e_{11}, e_{12},...,e_{1k}) are generated. For example, e_{1k} shows the minimum modeling error after dividing the domain of k-th variable in the first step of dividing. The variable corresponding to the minimum error is the best one for dividing of space. Suppose e_{1s} is the minimum error and it is correspond to x_s, then, the x_s domain is divided into small and big values. If e_{1s} is more than the threshold error, the dividing algorithm should continue.

Step 2. Consider all possible combinations of x_s – x_j (j = 1,2,…,k) for each part of x_s and then divide the domain of x_j again into two parts. Thus, 2k combinations are generated (k combinations of $x_{s(small)} - x_j$ and k combinations of $x_{s(big)} - x_j$) where each combination has two parts. For example, $x_{s(big)} - x_j$ means that when x_s has a big value, the domain of x_j is divided into small and big parts. Similarly, the ALM algorithm is applied to each part and the minimum modeling error is calculated for each k–combination. Suppose these are e_{2m} and e'_{2n} which mean the minimum modeling errors in the second step of dividing the space of variables is related to m-th and n-th variables for the small and big parts of x_s, respectively. Based on minimum errors, x_m and x_n are divided and the rules for modeling after dividing are:

If (x_s is small & x_m is small) then …
If (x_s is small & x_m is big) then …
If (x_s is big & x_n is small) then …
If (x_s is big & x_n is big) then …

e_{2m} and e'_{2n} are the local minimum errors. The appropriate global error (e_2) can be calculated using minimum local errors (e_{2m} and e'_{2n}). Dividing continues until the global error is less than the threshold error. In this heuristic search method, the global error is decreased simultaneously by decreasing the local errors.

Fig. 5 depicts the next step of dividing algorithm which is step 3.

This heuristic search method uses an appropriate criterion to select a variable for dividing and the median of data is used as the boundary for crisp dividing. Hence, the numbers of data points in the subspaces are equal.

4.2 Fuzzy dividing

Although, ALM implements crisp or fuzzy dividing methods, but fuzzy dividing and modeling methods can improve the ALM performance (*Taheri* et al., 2009).

Fuzzy dividing is similar to crisp dividing. In crisp dividing, the dividing point of a variable is the median as shown in Fig. 6a. But in fuzzy dividing, the boundary of small values of a variable is bigger than the median (Fig. 6b) and vice versa (Fig. 6c). Hence, the regions of small and big values of a variable can overlap.

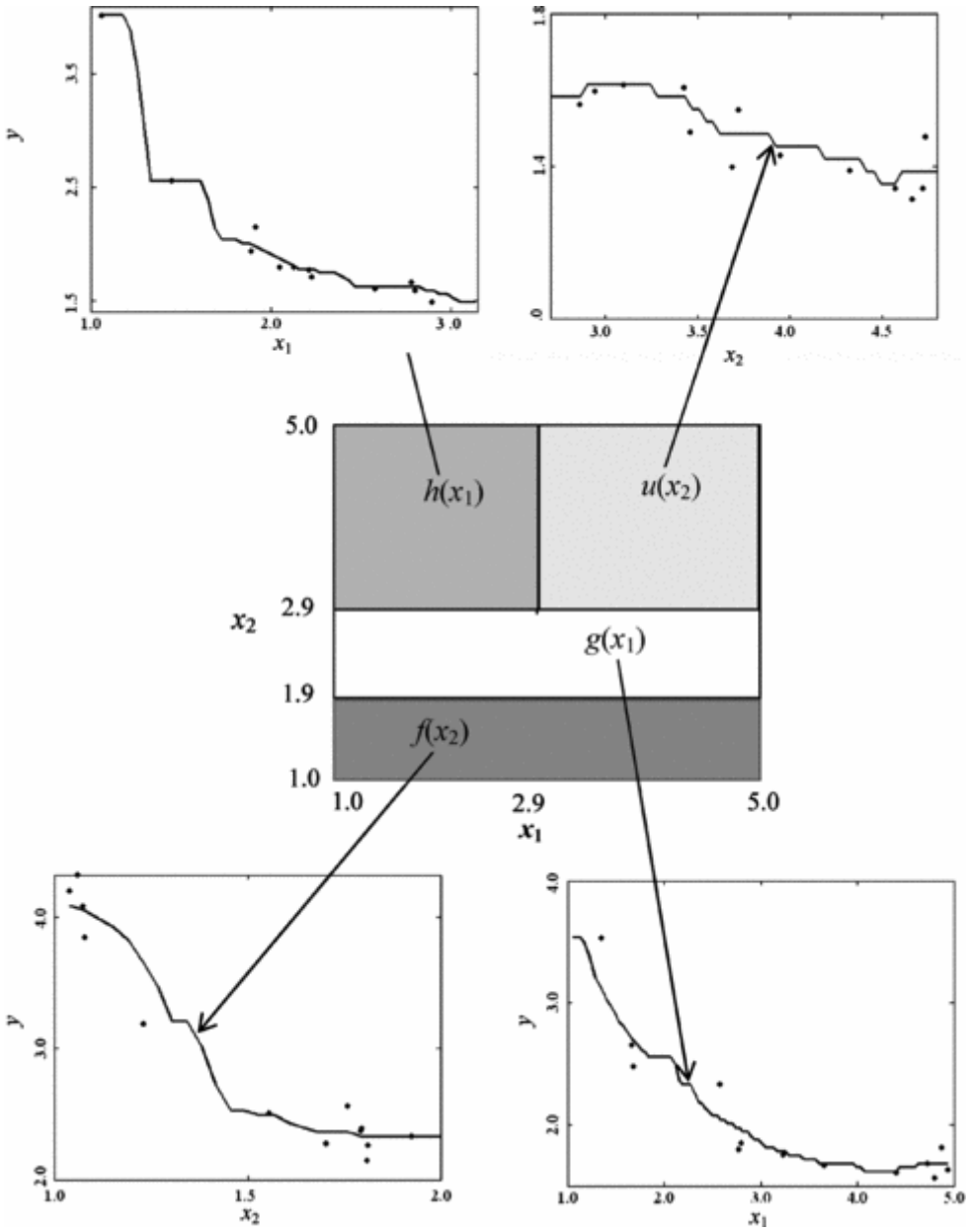


Fig. 4. Divided entire space to four subspaces using the heuristic search method and the best continuous path (implicit non-linear function), extracted for each subspace (the data points in each subspace have been shown by black circles).

The fuzzy systems are not too sensitive to the dividing points. Therefore, the appropriate points for fuzzy dividing can be calculated by investigating various alternatives to select the most appropriate one.

4.3 Fuzzy modeling in ALM

Since the presented new heuristic method utilizes a complicated dividing method, the typical fuzzification methods are not compatible with it. Here, a new simple fuzzy modeling method is presented which is attuned to the heuristic search method. This fuzzy modeling method has been developed by *Taheri Shahraiyni* (2007).

We denote the membership function of a fuzzy set as $A_{ij}^{ks}\left(x_k^m\right)$ in which i is the dividing step, j – the number of dividing in each i which has a value between 1 and 2^{i-1}, s – the membership function that is related to small (s = 1) and big parts (s = 2) of a variable domain, k denotes the divided variable number and x_k^m is the m-th member of the k-th

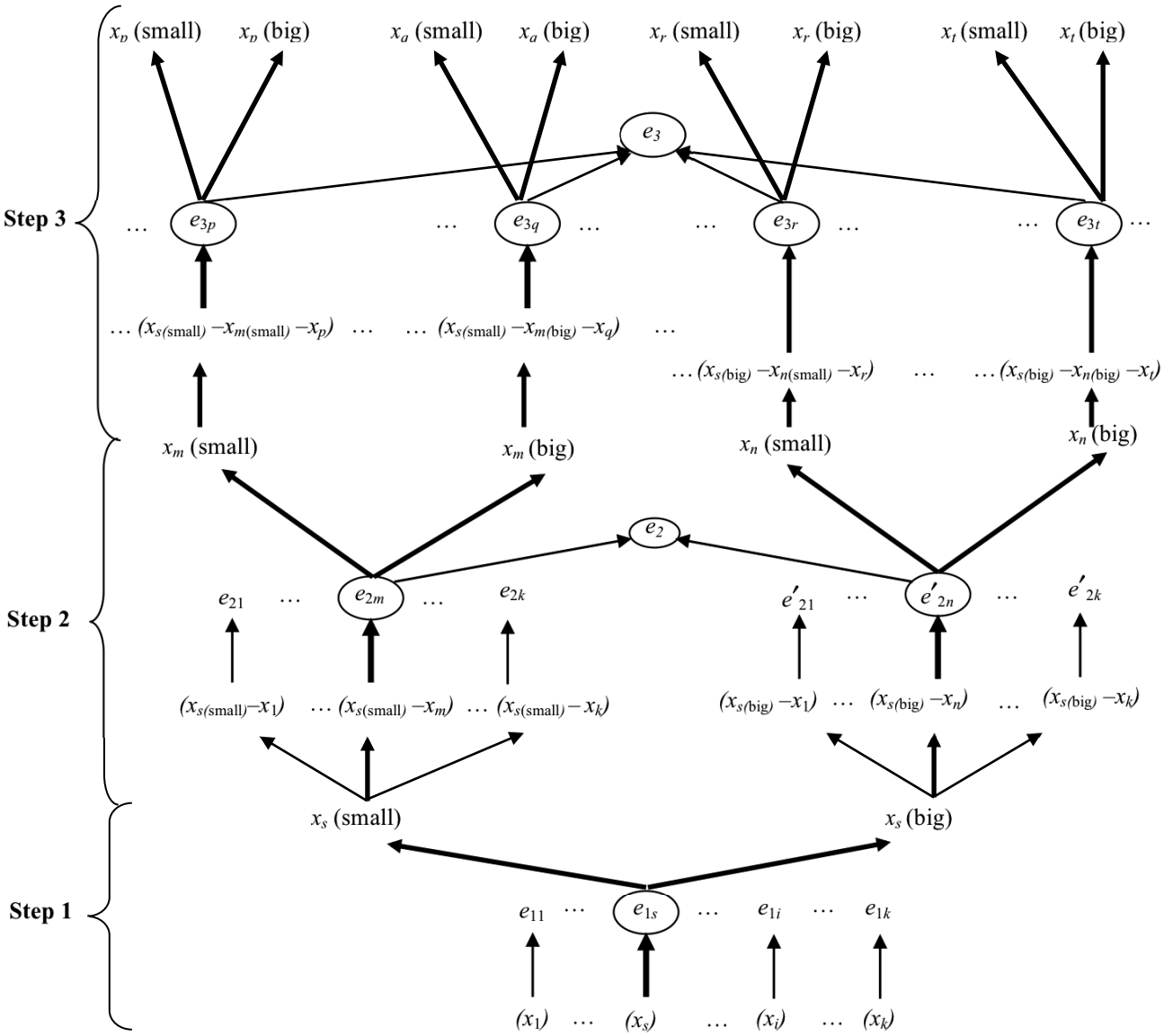
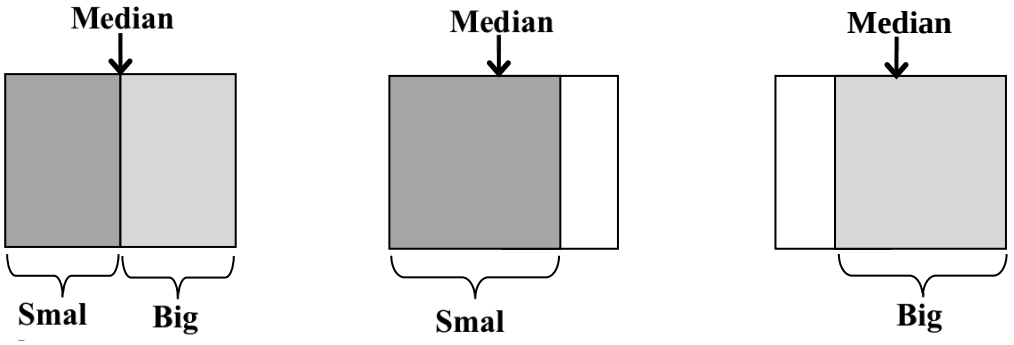


Fig. 5. Algorithm of the new heuristic search method for dividing the space.



(a) (b) (c)

Fig. 6. Schematic view of different dividing methods; (a) crisp dividing; (b) small part of variable domain in fuzzy dividing; (c) big part of variable domain in fuzzy dividing.

variable (X_k) ($x_k^m \in X_k$) and $X_k \in X$, $X = \{X_1, ..., X_n\}$ is a set of *n* variables. ALM can be implemented by fuzzy modeling with miscellaneous shapes of membership functions and the performance of ALM as a fuzzy modeling method is not sensitive to the shape of membership function. Trapezoidal membership functions are one of the most used membership functions. In addition, implementation of a fuzzy modeling method using trapezoidal membership functions is very straightforward. Hence, trapezoidal membership functions are applied here.

The truth value of a proposition is calculated by a combination of membership degrees. For example, the truth value of 'x_1^1 is A_{11}^{11} and x_2^1 is A_{21}^{22}' and is expressed as:

$$\left(x_1^1 \text{ is } A_{11}^{11} \text{ and } x_2^1 \text{ is } A_{21}^{22}\right) = \left(A_{11}^{11}\left(x_1^1\right) \wedge A_{21}^{22}\left(x_2^1\right)\right) = \left(A_{11}^{11}\left(x_1^1\right) \times A_{21}^{22}\left(x_2^1\right)\right).$$

In this fuzzy method, the general fuzzy rules are defined as below:

R^p: If ($x_{k_1}^m$ is $A_{1j_1}^{k_1 s_1}$ & $x_{k_2}^m$ is $A_{2j_2}^{k_2 s_2}$ & …

then $y_p^m = f_p(x_{k_3}^m)$,

where *p* is the rule number and has a value between 1 and *h* (*h* is total number of fuzzy rules), R^p – the *p*-th rule and f_p is the *p*-th one–variable non-linear function for the *p*-th subspace (*p*-th rule).

$1/P(f_p)$ is considered as the weight of the *p*-th rule (W_{rp}) where $P(f_p)$ is PAE of f_p (continuous path in the *p*-th rule). Fire strength or membership degree of the *p*-th rule, W_{fp}^m is equal to the truth value of the proposition which is:

$$W_{fp}^m = A_{1j_1}^{k_1 s_1}\left(x_{k_1}^m\right) \times A_{2j_2}^{k_2 s_2}\left(x_{k_2}^m\right) \times \ldots \quad (5)$$

Obviously, the summation of truth values of all of the propositions should be equal to 1 ($\sum_{p=1}^{h} W_{fp}^m = 1$).

Finally, the corresponding output (y^m) to *m*-th set of input dataset ($x_1^m, \ldots x_k^m, \ldots x_n^m$) is calculated as:

$$y^m = \frac{\sum_{p=1}^{h} (y_p^m \times W_{fp}^m \times W_{rp})}{\sum_{p=1}^{h} (W_{fp}^m \times W_{rp})}. \quad (6)$$

5. Modeling procedures

5.1 Statistical Evaluation Index

Statistical goodness-of-fit indices such as mean percent of absolute error (MPAE), coefficient of determination (R^2), mean bias, Nash-Sutcliffe efficiency (NS), root mean square error (RMSE), percent of total volume error (PTVE) and peak-weighted root mean square error (PW-RMSE) was employed for comprehensive evaluation of boss models. Mathematical equations of these utilized indices are presented in Tab. 1. In addition, graphical goodness-of-fit criteria such as quantile-quantile (Q-Q) diagram, scatter plot, hydrographs and time series of residuals were used for comprehensive evaluation of simulation results.

5.2. Support Vector Machine Modeling

Daily discharge of Karoon River at Pol-e-Shaloo hydrometric station (Fig. 1) from 23 Sep. 1991 to 22 Sep. 1999 were used for training and testing of SVM model. The first five hydrological years (23 Sep 1991 to 22. Sep 1996) were used for the training of model and the residual data were used for testing phase (three hydrological years).

For selection of the appropriate input data for the SVM, the meteorological (daily precipitation, temperature, relative humidity and vapor pressure) and hydrometric data (daily discharge data) of Karoon III basin were gathered. Then AMI (Average Mutual Information) index has been utilized for the determination of useful input variables for modeling. Tab. 2 exhibits the AMI values of hydrometric and meteorological data with different lags. The more important variable has higher AMI value. According to the Tab. 2, the discharges from 1 to 5 days lag were selected as appropriate input variables for SVM modeling in this study. Hence, SVR modeling will be performed using five lags of runoff as the input variables.

Then, SVR (regression based of SVM) has been implemented and its parameters have been tuned using simple Genetic Algorithm (GA). NS indicator has been selected as fitness function in simulating

this optimization procedure. Statistical learning paradigm and mathematical kernel function have been Epsilon-SVR and Radial Basis Function (RBF), respectively, and seven parameters including scaling factor and constant parameter in RBF, cost parameter in epsilon-SVR, value of epsilon and suitable tolerance of statistical learning have been optimized via simple GA.

According to previous section, about 60% of the whole dataset is used for training and the remained data is applied for the test of trained SVR model. The polynomial kernel in order 3 used as appropriate transformation in the raw dataset (with 0.1 as coefficient with zero intercept). LIBSVM 3.1 and simple GA in MATLAB media have been implemented for the runoff simulation in this study (*Chang* and *Lin*, 2011).

T a b l e 1. Mathematical equations of utilized Goodness of fit Indices.

Goodness of fit indices (statistical criteria)	Equations
Nash-Sutcliffe (NS) (Model efficiency)	$1-\left[\frac{\sum_{i=1}^{n}(O_i-S_i)^2}{\sum_{i=1}^{n}(O_i-\hat{O})^2}\right]$
Bias	$\frac{1}{n}\sum_{i=1}^{n}(O_i-S_i)$
Coefficient of determination (R^2)	$\left\{Cov(O_i,S_i)/\sqrt{Cov(O_i,O_i).Cov(S_i,S_i)}\right\}^2$
Percent of Total Volume Error (PTVE)	$\frac{\sum_{i=1}^{n}O_i-\sum_{i=1}^{n}S_i}{\sum_{i=1}^{n}O_i}\times 100$
Mean Percent of Absolute Error (MPAE)	$\frac{1}{n}\sum_{i=1}^{n}\frac{\lvert O_i-S_i\rvert}{O_i}\times 100$
Root Mean Square Error (RMSE)	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(O_i-S_i)^2}$
Peak Weighted RMSE (PW-RMSE) (USACE, 2000)	$\left\{\frac{1}{n}\left[\sum_{i=1}^{n}(O_i-S_i)^2\left(\frac{O_i+\hat{O}}{2\hat{O}}\right)\right]\right\}^{0.5}$

*In the above table, n – number of discharge data, O_i and S_i – observed and simulated discharge data in *i*-th time step, $\hat{O}$ – average of observed discharge and Cov – covariance of data.

T a b l e 2. AMI of various hydrological parameters for input selection in optimized SVR model.

Parameter	No. of lags					
	0	1	2	3	4	5
Precipitation	0.0244	0.0345	0.0277	0.0198	0.0177	–
Temperature	0.0394	0.0409	0.0424	0.0436	0.0446	–
Humidity	0.0393	0.0429	0.0408	0.0409	0.0412	–
Vapor	0.0160	0.0185	0.0173	0.0154	0.0132	–
Runoff	–	0.3067	0.2638	0.2367	0.2169	0.2050

5.3 ALM modeling

Similarly, about five years data (23 Sep 1991 to 22 Sep 1996) were used for the training and the remained data were used for the testing of the ALM model. Daily discharge data with 1 to 5 days time lags were used as a set of input data for ALM modeling. Contrary to many other modeling methods (e.g. ANNs), the ALM does not need initial parameters to start the training and thus it does not repeat the training. Hence the ALM training is easy, straightforward, and time efficient (*Taheri Shahraiyni* et al., 2009; *Taheri Shahraiyni*, 2010). When we divided the domain of a variable fuzzily, some of the data were shared in small and large parts of the variable domain. The percent of common data in small and large parts is related to the fuzzy dividing points. The fuzzy systems are not too sensitive to the dividing points. Therefore, the appropriate points for fuzzy dividing can be calculated by in-

vestigating various alternatives to select the most appropriate one. *Bagheri Shouraki* and *Honda* (1999) and *Taheri Shahraiyni* et al. (2009) showed that the first and the third quantiles of data are the best dividing points. In this study, firstly the ALM was applied to input set and appropriate fuzzy dividing points were determined. Variety of statistical objective functions and graphical goodness of fit indices were used for the evaluation of the ALM and optimized SVR modeling results.

6. Results and discussion

In the ALM modeling, for determination of appropriate fuzzy dividing points, the ALM model was executed with several fuzzy dividing alternatives (20%, 40%, 50%, 60% and 80% of data shared or common in small and big parts) using daily discharge data with 1 to 5 time lags as input data (Fig. 7). The results showed that ALM is not so sensitive to the location of fuzzy points. Similar findings have been presented by *Bagheri Shouraki* and *Honda* (1999). *Taheri Shahraiyni* et al. (2009) demonstrated that the optimum points for fuzzy dividing are the first and third quarters of data hence according to Fig. 7 and the results of *Taheri Shahraiyni* et al. (2009), the first and third quarters of data were selected as fuzzy dividing points in this study.

Then, the normalized daily discharge data with 1 to 5day lags were used as input set for ALM and SVR modeling. The statistical results of the simulated flow data for the training and testing phases with different number of fuzzy rules (for ALM) are presented in Tab. 3.

In the ALM model with increasing of number of fuzzy rules to more than 16 rules doesn't improve the modeling results. Thus, streamflow modeling by 16 fuzzy rules is considered to be the best model. According to the Tab. 2, Nash-Sutcliffe efficiency coefficients are more than 0.8 in training and testing phases of the modeling. Nash-Sutcliffe efficiency coefficient values of less than 0.5 are considered as unacceptable, while values greater than 0.6 are considered as good and greater than 0.8 are considered excellent results (*Garcia* et al., 2008). Therefore, ALM and optimized SVR have been presented excellent Nash-Sutcliffe values. Bias statistic for ALM and optimized SVR models in the testing period is equal to 5.5 and –11.7 $m^3 s^{-1}$ and percent of total volume error (PTVE) is equal to 1.9 and –3.9%, respectively. These results imply that the ALM slightly overestimates and optimized SVR underestimated the streamflow. Although ALM has better Bias and PVTE than optimized SVR, but optimized SVR has smaller MPAE than ALM. In addition, other statistical goodness of fit indices like to R^2, RMSE and PWRMSE express the similar and acceptable results in the simulation of streamflow in the both models.

Fig. 8 shows scatter plot and Q-Q diagram of ALM (left) and optimized SVR (right) models. Q-Q diagrams are often used to determine whether the model could extract the behavior of observed data (*Chambers* et al., 1983). As can be seen from the scatter plot and Q-Q diagram in Fig. 8, results of the models in the simulation of low to mid flows are good. ALM and optimized SVR present poor performance in the peak flow estimation. The observed hydrographs and the time series of residuals of ALM and optimized SVR have been presented in Figs. 9. Hydrographs and time series of residuals exhibit the acceptable results in the runoff modeling in the both models. Weak simulation of intense peaks is obvious in the time series of residuals for

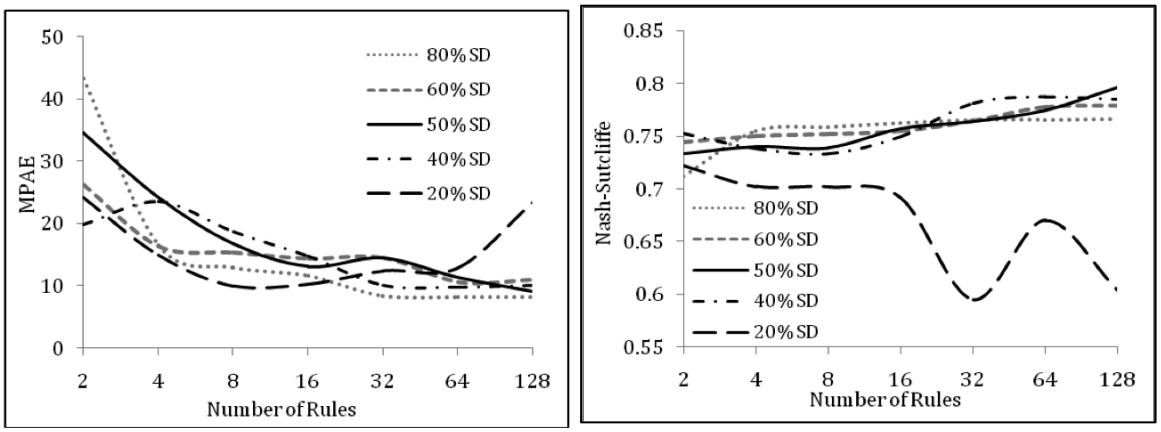


Fig. 7. The effect of changing fuzzy points at different fuzzy rules in the testing phase (SD: shared data).

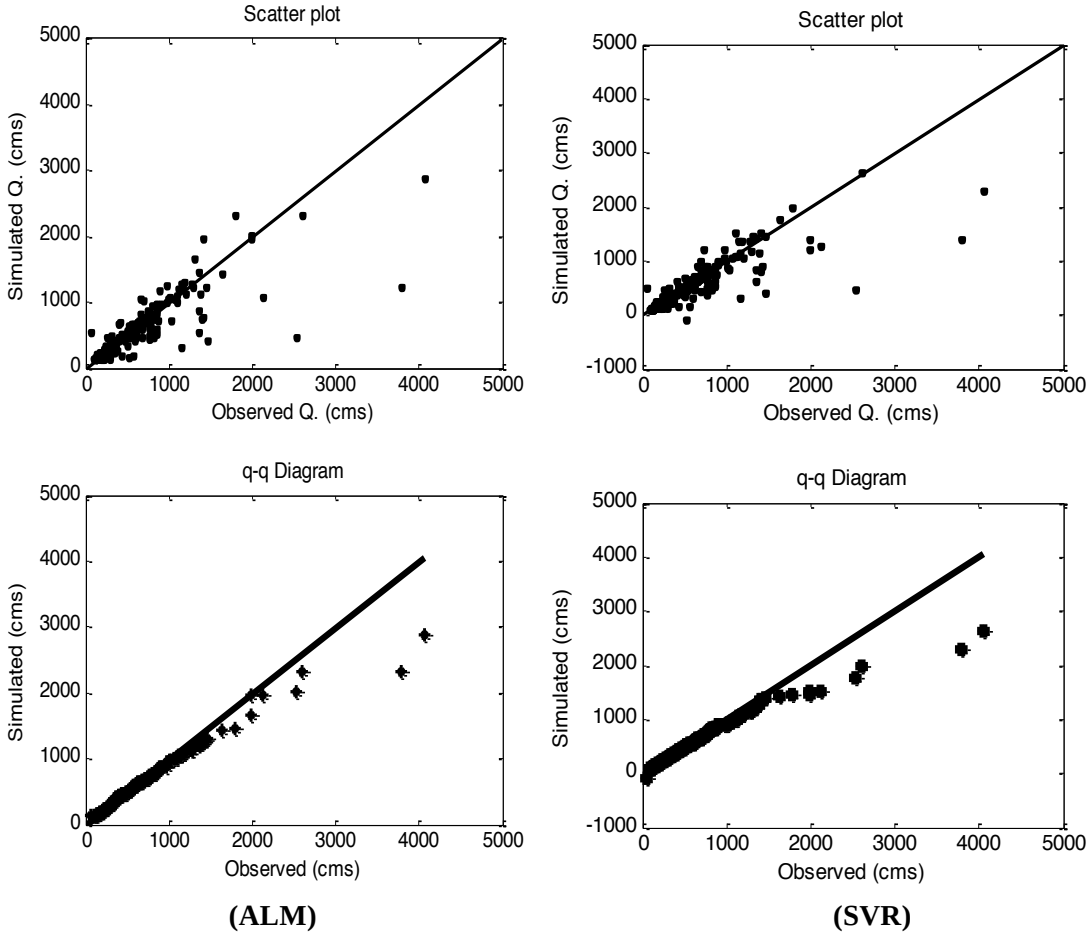


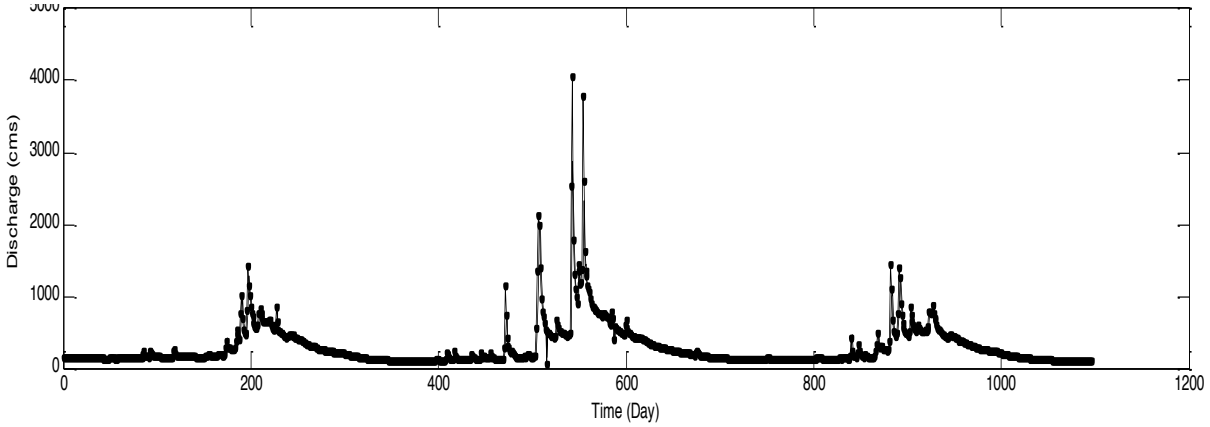
Fig. 8. Scatterplot and Q-Q diagram of simulated hydrographs using ALM with 16 fuzzy rules (left) and optimized SVR (right) models.

both models. The weak performance of ALM and SVM in the intense peak flows is a consequence of small number of intense extreme flows. This is highly related to the hydrological regime of Karoon River, which has low flows in the most of the time and it has only a few number of intense peak. In these cases, the learning algorithm of SVM and ALM has tendency to be adapted to the low and average flows. Therefore the generalization of SVM and ALM model reduces for the high flows. To overcome this problem, input dataset with high number of peak flows would be available for the model training. The poor performance in the simulation of high flow has been achieved in the other similar studies e.g. *Firat* and *Gungor* (2006); *Pulido-Calvo* and *Portela* (2007); *Firat* (2008) and *Behzad* et al. (2009). The PW-RMSE is an implicit expression for the evaluation of model performance in the peak flow simulation. PW-RMSE values higher than RMSE in Tab. 3 show that peak flow estimation is worse than the other parts of the hydrograph (*USACE*, 2000).

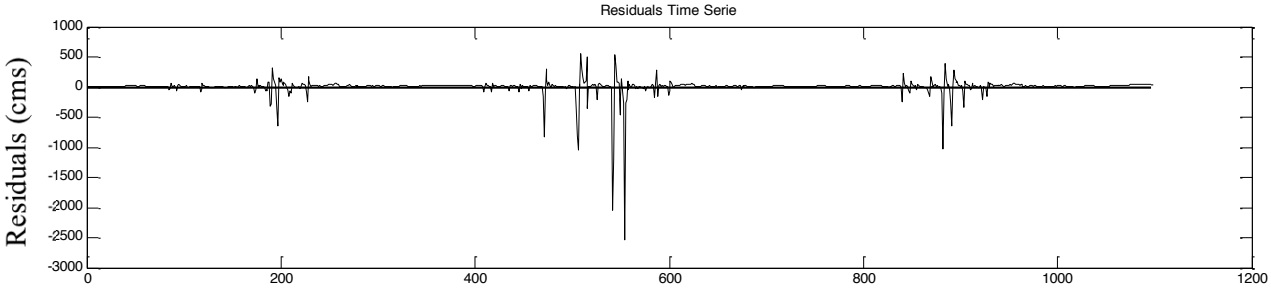
In general, comparison between statistical and graphical results of ALM and optimized SVR models shows that the ALM could simulate the streamflow as good as the optimized SVR model.

The parameters of SVR should be tuned manually or by a tuning method during the training phase of modeling, but ALM does not need to any method for parameter tuning and its training is very easy and straightforward.

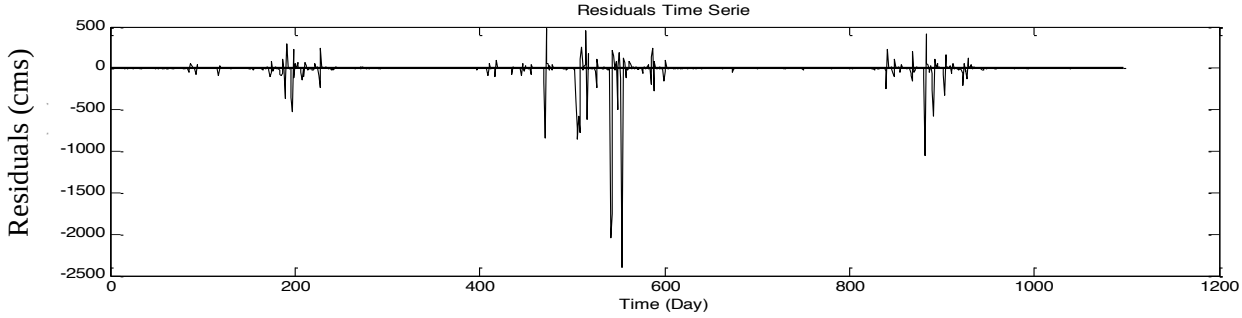
ALM is able to understand and find the important variables for different subspaces of the space of variables. Also it is able to find the divided variables and one-variable functions in each step of modeling. Therefore, the ranking of variables and their shares in modeling can be performed by ALM. The results of ALM simulation using 16 fuzzy rules showed that it only has utilized of runoff with time lag 1 in the one-variable functions in the simulation. According to the Tab. 2, discharge with lag 1



(a)



(b)



(c)

Fig. 9 a) Observed hydrographs in the test period; b) the residuals of simulated hydrograph using ALM with 16 fuzzy rules and c) the residuals of simulated hydrograph using optimized SVR model.

has the highest AMI value with the discharge with lag 0 (output), therefore ALM has selected the most appropriate variable for the modeling.

Similarly the role of different variables in dividing the space of variables can be calculated. ALM with 16 fuzzy rules has utilized of discharge with lag 1 more than the discharge with other lags for dividing the space of variables. Hence, the discharge with lag 1 is the most important variable for dividing the space of variables. Consequently, one of the most important properties of ALM is the ability of finding the important variables and ranking the effective variables in a system. The variables which have no role in modeling are recognized as excess variables and ALM model removes them from the input dataset.

T a b l e 3. Statistical results of ALM with different fuzzy rules and SVR models.

Goodness of fit indices		Nash-Sutcliffe		Bias [$m^3 s^{-1}$]		R^2		MPAE [%]		PTVE [%]		RMSE [$m^3 s^{-1}$]		PWRMSE [$m^3 s^{-1}$]	
	Rules	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
ALM	2	0.85	0.76	–10.7	–3.6	0.85	0.77	10.5	13.0	–2.5	–1.2	147.2	154.3	241.2	350.9
	4	0.85	0.76	–4.1	1.5	0.86	0.77	10.3	12.9	–0.9	0.5	145.5	153.1	238.6	349.0
	8	0.86	0.80	1.0	5.5	0.87	0.81	10.6	13.2	0.2	1.9	140.4	139.6	227.1	306.7
	16	*0.87*	*0.81*	*1.9*	*5.5*	*0.87*	*0.81*	*10.4*	*12.9*	*0.4*	*1.9*	*139.9*	*137.2*	*227.2*	*297.2*
	32	0.87	0.81	3.0	5.9	0.87	0.81	10.3	13.1	0.7	2.0	139.3	137.5	226.3	297.3
	64	0.87	0.81	4.6	6.9	0.87	0.81	10.4	13.4	1.1	2.3	137.6	137.6	223.2	296.8
	128	0.87	0.81	5.7	7.5	0.87	0.81	10.3	14.0	1.3	2.5	136.2	138.1	221.2	297.5
SVR		*0.89*	*0.80*	*–15.8*	*–10.7*	*0.89*	*0.81*	*4.7*	*7.0*	*–3.6*	*–3.6*	*126.7*	*141.6*	*211.5*	*307.1*

7. Conclusions

In this study, active learning method (ALM) as a novel fuzzy modeling method was used for simulation of daily streamflow of Karoon river. Also, optimized support vector machine (SVR type) was selected as a well known data-driven model for the evaluation of ALM results and comparison with it. ALM simulated the river flow as good as the optimized SVR model. The results of test of ALM model showed that the best model is the ALM model with 16 fuzzy rules and its Nash-Sutcliffe, Bias, R^2, MPAE and PTVE were 0.81, 5.5 $m^3 s^{-1}$, 0.81, 12.9%, and 1.9%, respectively. Similarly, these criteria for optimized SVR model were 0.80, – 11.7 $m^3 s^{-1}$, 0.81, 5.3%, and –3.9% respectively. Results of this study demonstrated acceptable streamflow simulation by ALM and optimized SVR models for the continuous streamflow simulation. In addition, training of the ALM is easier and more straightforward than the training of other data driven models such as optimized SVR. Also, ALM was able to identify and rank the effective variables of the system under investigation.

In general, according to the ALM abilities and properties and its similar results to optimized SVR, it has merit to be introduced as a new appropriate modeling method for the runoff simulation.

REFERENCES

ABBASS H. A., SARKER R. A., and NEWTON C. S., 2002: Data mining: a heuristic approach. Idea Group Publishing.

ANVARI TAFTI S., 2008: Upgrading intelligent models for streamflow forecasting using distributed climatic data and snow cover. MS thesis, Faculty of Agriculture, Tarbiat Modares University, Iran.

ASEFA T., KEMBLOWSKI M. W., MCKEE M., and KHALIL A., 2006: Multi-time scale stream flow prediction: The support vector machines approach. J. Hydrol., 318, 7–16.

BAGHERI SHOURAKI S. and HONDA N., 1997: A new method for establishing and saving fuzzy membership functions. Proc., Int. 13th Fuzzy Symposium, 4–6 June, Toyama, Japan, pp. 91–94.

BAGHERI SHOURAKI S., and HONDA N., 1999: Recursive fuzzy modeling based on fuzzy interpolation. J. Advanced Computational Intelligence, 3, 114–125.

BATTITI R., 1994: Using mutual information for selecting features in supervised neural net learning. IEEE Transactions on Neural Networks, *5*, 4, 537–550.

BEHZAD M., ASGHARI K., EAZI M., PALHANG M., 2009: Generalization performance of support vector machines and neural networks in runoff modeling, Expert Systems with Applications, 36, 7624–7629.

BOWDEN G. J., MAIER H. R., and DANDY G. C., 2002: Optimal division of data for neural network models in water resources applications. Water Resour. Res., *38*, 2, 1010, doi: 10.1029/2001WR000266. (http://www.agu.org/pubs/crossref/2002/2001WR000266.shtml)

BOWDEN G. J. DANDY G. C. & MAIER, H. R., 2005a: Input determination for neural network models in water resources applications, Part 1–background and methodology. J. Hydrol., *301*, 75–92.

BOWDEN, G. J. DANDY G. C. & MAIER, H. R. (2005b): Input determination for neural network models in water resources applications, Part 2. Case study: forecasting salinity in a river. J. Hydrol., 301, 93–107.

BOSER B. E., GUYON I. M., VAPNIK V. N., 1992: A training algorithm for optimal margin classifiers. In D. Haussler, editor, 5th Annual ACM Workshop on COLT, 144–152, Pittsburgh, PA, ACM Press.

CHAMBERS J. M., CLEVELAND W. S., KLEINER B., TUKEY P. A., 1983: Graphical methods for data analysis. Wadworth, Belmont, California, pp. 395.

CHANG C. C and LIN C. J., 2011: LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1–27:27. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

CHANG F. J. and CHANG Y. T., 2006: Adaptive neuro-fuzzy inference system for prediction of water level in reservoir. Adv. Water Res, *29*, 1–10.

CHANG F. J. and CHEN Y. C., 2001: Counter propagation fuzzy-neural network modeling approach to real time streamflow prediction. J. Hydrol., *245*, 153–164.

CHAU K. W., WU C. L. and LI Y. S., 2005: Comparison of several flood forecasting models in Yangtze River. J. Hydrologic Engng., ASCE, *10*, 6, 485–491.

CHEN S. H., LIN Y. H., CHANG L. C., and CHANG F. J., 2006: The strategy of building a flood forecast model by neuro-fuzzy network. Hydrol. Process., 20, 1525–1540.

CHENG C. T., OU C. P. and CHAU K. W., 2002: Combining a fuzzy optimal model with a genetic algorithm to solve multiobjective rainfall-runoff model calibration. J. Hydrol., 268, 72–86.

DIBIKE Y. B., VELICKOV S., SOLOMATINE D., and ABBOTT M. B., 2001: Model Induction with support vector machines: Introduction and Application. J. Computing in Civil Engng., *15*, 3, 208–216.

DUAN D., FERMINI B., and NATTELI S., 1992: Sustained Outward Current Observed after Ito1 Inactivation in Rabbit Atrial Myocytes Is a Novel Cl^- current". Am. J. Physiol., *263*, H1967–1971.

FAUSSET L., 1994: Fundamentals of neural networks: Architectures, algorithms and applications. Prentice Hall, 461p.

FERNANDO T. M. K. G., MAIER H. R., DANDY G. C., 2009: Selection of input variables for data driven models: An average shifted histogram partial mutual information estimator approach. J. Hydrol., *367*, 3-4, 165–176.

FIRAT M., and GUNGOR M., 2006: River flow estimation using adaptive neuro fuzzy inference system. Math. Compute. Simulate., *75*, 3-4, 87–96.

FIRAT M., 2008: Comparison of artificial intelligence techniques for river flow forecasting. Hydrol. Earth Syst. Sci., *12*, 123–139.

HSU K., GUPTA H. V., and SOROOSHIAN S., 1995: Artificial neural network modeling of rainfall-runoff process. Water Resour. Res., 31, 2517–2530.

JAIN S. K., DAS D., and SRIVASTAVA D. K., 1999: Application of ANN for reservoir inflow prediction and operation. J. Water Resources Planning and Management, ASCE, *125*, 263–271.

KISI O., 2005: Daily river flow forecasting using artificial neural networks and auto-regressive models. Turkish J. Eng. Env. Sci., *29*, 9–20.

LIN J. Y., CHENG C. T. and CHAU K. W., 2006: Using support vector machines for long-term discharge prediction. Hydrological Sciences J., *51*, 4, 599–612.

LIONG S. Y., LIM W. H., KOJIRI T., and HORI H., 2006: Advance flood forecasting for flood stricken Bangladesh with a fuzzy reasoning method. Hydrol. Process, *14*, 431––448.

LIU H., SUN J., LIU L. and ZHANG H., 2009: Feature selection with dynamic mutual information. Pattern Recognition, *42*, 7, 1330–1339.

MAHABIR C., HICKS F. E., and FAYEK A. R., 2000: Application of fuzzy logic to the seasonal runoff. Hydrol. Process, *17*, 3749–3762.

MAIER H.R. AND DANDY G. C., 1998: Neural networks for the prediction and forecasting of water resources variables: a review of modeling issues and applications. Environmental Modeling and Software, *15*, 101–124.

MAMDANI E. A., 1976: Advances in the linguistic synthesis of fuzzy controllers. Int. J. Man–Machine Studies, *8*, 669––678.

MAMDANI E. A. and ASSILIAN S., 1975: An experiment in linguistic synthesis with a fuzzy logic controller. Int. J. Man–Machine Studies, *7*, 1–3.

MANNING, C. D., RAGHAVAN P. and SCHÜTZE H., 2008: Introduction to Information Retrieval. Cambridge University Press.

MARKUS M., 1997: Application of neural networks in streamflow forecasting. PhD thesis, Department of Civil Engineering. Colorado State University. Fort Collins. Colorado.

MAY R. J., DANDY G. C., MAIER H. R., and NIXON J. B., 2008a: Application of partial mutual information variable selection to ANN forecasting of water quality in water distribution systems. Environmental Modeling and Software, 23, 1289–1299.

MAY R. J., MAIER H. R., DANDY G. C., and FERNANDO T. G., 2008b: Non-linear variable selection for artificial neural networks using partial mutual information. Environmental Modeling and Software, *23*, 1312–1326.

NAYAK P. C., SUDHEER K. P. and RAMASASTRI K. S., 2004a: Fuzzy computing based rainfall-runoff model for real time flood forecasting. Hydrol. Process, *17*, 3749–3762.

PULIDO-CALVO I. and PORTELA, M. M., 2007: Application of neural approaches to one-step daily flow forecasting in Portuguese watersheds. J. Hydrol., 332, 1–15.

SAAD M., BIGRAS P., TURGEON A., and DUQUETTE R., 1996: Fuzzy learning decomposition for the scheduling of hydroelectric power systems. Water Resour. Res., *32*, 179––186.

SEN Z. and ALTUNKAYNAK A., 2006: A comparative fuzzy logic approach to runoff coefficient and runoff estimation. Hydrol. Process, *20*, 1993–2009.

SHAMSELDIN A. Y., 1997: Application of a neural network technique to rainfall-runoff modeling. J. Hydrol., *199*, 272––294.

SIVAKUMAR B., JAYAWARDENA A. W. and FERNANDO T. M. K. G., 2002: River flow forecasting: Use of phase space reconstruction and artificial neural networks approaches. J. Hydrol., *265*, 225–245.

SMITH J., and ELI R. N., 1995: Neural network models of rainfall-runoff process. J. Water Resources Planning and Management, 499–508.

SMOLA A. J., SCHLKOPF B., 2004: A tutorial on support vector regression. Statistics and Computing, *14*, 199–222.

SOROOSHIAN S., DAUN Q., and GUPTA V. K., 1993: Calibration of Rainfall-Runoff Models; Application of Global Optimization to the Sacramento Soil Moisture Accounting Model. Water Resour. Res., *29*, 1185–1194.

SUGENO M,. and YASUKAWA T., 1993: A fuzzy-logic based approach to qualitative modeling. IEEE Transactions on Fuzzy Systems, *1*, 7–31.

TAHERI SHAHRAIYNI H., 2007: Extraction of water quality parameters using the satellite data and inverse modeling of radiative transfer equation by Active Learning Method. PhD Thesis, School of Civil Engineering, Sharif University of Technology, Iran.

TAHERI SHAHRAIYNI H., BAGHERI SHOURAKI S., FELL F., SCHAALE M., FISCHER J., TAVAKOLI A., PREUSKER R., TAJRISHY M., VATANDOUST M., and KHODAPARAST H., 2009: Application of the active learning method to the retrieval of pigment from spectral remote sensing reflectance data. Int. J Remote Sensing, *30*, 1045––1065.

TAHERI SHAHRAIYNI H., 2010: Fuzzy modeling by active learning method. Advanced Technologies, 1st Edn., In-Tech Publisher.

TAKAGI T., and SUGENO M., 1985: Fuzzy identification of systems and its applications to modeling and control., IEEE Transactions on Systems. Man. and Cybernetics, *SMC15*, 116–132.

TAN P. N., STEINBACH M. and KUMAR V., 2006: Introduction to data mining. Addison Wesley.

TOKAR A. S. and JOHNSON P. A., 1999: Rainfall-runoff modeling using artificial neural networks. J. Hydrol. Eng., ASCE, *4*, 232–239.

TAYFUR, G. and SINGH, V.P., 2006: ANN and Fuzzy Logic Models for Simulating Event-Based Rainfall-Runoff. J. Hydraul. Engng, *132*, 12, 1321–1330.

US ARMY CORPS of ENGINEERS (USACE)., 2000: HEC-HMS Technical Reference Manual. Davis, C.A, http://www.hec.usace.army.mil/.

VAPNIK V. N., CORTES C., 1995: Support vector networks. Machine Learning, *20*, 273–297.

VAPNIK V. N., 1998: Statistical Learning Theory. Wiley, New York.

WANG W., 2006: Stochasticity, nonlinearity and forecasting of streamflow processes. PhD Thesis, Hohai University.

WANG W. C., CHAU K. W., CHENG C. T. and QIU L., 2009: A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. J. Hydrol., *374*, 3-4, 294–306.

WU C.L., CHAU K. W. and LI Y. S., 2009: Predicting monthly streamflow using data-driven models coupled with data-preprocessing techniques. Water Resour. Res., *45*, W08432, doi:10.1029/2007WR006737.

YAPO P., GUPTA V. K., and SOROOSHIAN S., 1996: Calibration of Conceptual Rainfall-Runoff Models: Sensitivity to Calibration Data. J. Hydrol., *181*, 23–48.

YU P.S., CHEN S.T., and CHANG I.F., 2006: Support vector regression for real-time flood stage forecasting. J. Hydrol., *328*, 704–716.

ZADEH L. A., 1965: Fuzzy sets. Information and Control. 338–353.

ZEALAND C. M., BURN D. H., and SIMONOVIC S. P., 1999: Short-term streamflow forecasting using artificial neural networks. J. Hydrol., 214, 32–48.

ZOUNEMAT-KERMANI M., TESHNEHLAB M., 2007: Using adaptive neuro-fuzzy inference system for hydrological time series prediction. Appl. Soft Comput. J., doi:10.1016/j.asoc.2007.07.011.

Received 29 January 2011
Accepted 19 October 2011