

## A comparative assessment between artificial neural network, neuro-fuzzy, and support vector machine models in splash erosion modelling under simulation circumstances

Mahdi Borouhiani<sup>1</sup>, Somayeh Soltani<sup>2</sup>, Nafiseh Ghezelseflu<sup>3</sup>, Iman Pazhouhan<sup>4\*</sup>

<sup>1</sup>Center for Geosciences and Social Studies, Hakim Sabzevari University, Sabzevar, Iran

<sup>2</sup>Department of Water Sciences and Engineering,  
College of Agricultural and Natural Resources,  
Ardakan University, Ardakan, Iran

<sup>3</sup>Department of Watershed Management, Ardakan University, Ardakan, Iran

<sup>4</sup>Nature Engineering Department, Faculty of Natural Resource and Environment, Malayer University, Malayer, Iran

### Abstract

BOROUGHANI, M., SOLTANI, S., GHEZELSEFLU, N., PAZHOUHAN, I., 2022. A comparative assessment between artificial neural network, neuro-fuzzy, and support vector machine models in splash erosion modelling under simulation circumstances. *Folia Oecologica*, 49 (1): 23–34.

Splash erosion, as the first step of soil erosion, causes the movement of the soil particles and lumps and is considered an important process in soil erosion. Given the complexity of this process in nature, one way of identifying and modeling the process is to use a rainfall simulator and to study it under laboratory circumstances. For this purpose, transported material was measured with various rainfall intensities and different amounts of poly-acryl-amide. In the next step, artificial neural network (ANN), adaptive neuro-fuzzy inference system (ANFIS), and support vector machine (SVM) were used to model the transported materials. The results showed that among the three methods, the best values of evaluation criteria were related to SVM, and ANFIS respectively. Among the three studied durations, the experiment with a duration of 30 minutes received the best results. The results based on available data showed by increasing the number of membership functions, over-fitting happens in the ANFIS method. To reduce the complexity of the model and the likelihood of over-fitting, some rules were eliminated. The results showed that the performance of the model improved by eliminating some rules.

### Keywords

ANN, ANFIS, rainfall simulator, sediment load, soil particles, SVM

### Introduction

Soil erosion is a global problem that seriously threatens natural resources (TRIPATHI and SINGH, 2001). One of the most important environmental issues is agriculture and food production in the world that has intensified in recent years with increasing population and changing human

activities, such that each year 75 billion tons of soil is eroded (SOLTANI et al., 2019). Splash erosion is known as the first step in the process of erosion (LEGUEDOIS et al., 2005; KANE et al., 2008; SOUSA-VIEIRA et al., 2009; BOROUGHANI et al., 2011; ZUMR et al., 2020), which is a result of the bombardment of the soil surface by the rainfall. One of the important and reviewable variables

\*Corresponding author:

e-mail: imanpazhouhan@gmail.com

© 2022 Authors. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

in the occurrence of the erosion is the properties of soil. Removal of soil particles by splash process influenced by the adhesion affected by the soil characteristics such as organic matter and the soil texture (SOUSA-VIEIRA et al., 2009). One of the ways to reduce soil erosion is to use the soil stabilizers such as chemical polymers. Polymers, as soil improvers, increase the plant growth, reduce water and wind erosion and increase water retention by improving the soil structure. Absorbent polymers are from different types and are chemically classified to three categories of acrylate starch copolymers, poly-acrylamides and poly vinyl alcohols (BOROGHANI et al., 2012; SOLTANI et al., 2019). Poly-acrylamides are polymers with high molecular weight that are used to reduce soil erosion (WALLACE and WALLACE, 1986). Poly-acrylamides are wide ranges of molecular weights and diverse forms that converted to anions, cations and non-anions. Anionic poly-acrylamides soluble in water containing 150,000 monomers for each molecule used in controlling soil erosion and runoff, SOJKA et al. (2004). Compared with other polymers, poly-acrylamides are the best soil amendment because the amount of required poly-acrylamide to achieve the similar results and even better in soil conservation is 10 to 100 times lower (SOJKA, 2006). According to erosion damages, the use of methods that can increase the soil strength against corrosion is of great importance. Therefore, in this study the effect of different amounts of poly-acrylamide (0.2, 0.4 and 0.6 grams per square meter) on the levels of rainfall intensity in the rates of 60, 90 and 120 mm per hour in three different periods (10, 20 and 30 minutes) will be examined. In addition, the capabilities of three black-box models in the modeling of the amounts of the carried sediment in different intensities of rainfall and various amounts of poly-acrylamide are studied.

## Materials and methods

### Artificial neural networks (ANN)

One of the most common neural network used in the field of hydrology is multi-layered Perceptron Neural Network (ABRAHAM and NATH, 2001). This network has been composed of an input layer, an output layer and one or more layers between them that are not directly connected to input data and the output results. The units of the input layer are solely responsible for the distribution of the input values to the following layer, and the output layer also represents the response of output signal. In these two layers, the number of neurons is equal to the number of inputs and outputs and the hidden layer or layers have the responsibility of creating connection between the input and output layers. In an artificial neuron (Fig. 1), the scalar quantities of  $p$  and  $a$  are input and output of neuron, respectively. The scalar input of  $P$  before entering the main core of processor element is weighted, that is, multiplied by the  $w$  factor and produces  $wp$ . Other input into the neuron is the constant value of one that is multiplied by bias and then added to  $wp$  and  $n$  are obtained. At the last

stage, the value of  $n$  is put into activation function (which can be a function of step, sigmoid and etc type) and the final output value of the neuron, i.e.,  $a$ , is calculated (DEMUTH and BEALE, 1998). It is needed to calculate the values of neurons' weights in different layers in the neural network. To this end there are different methods such as the steepest gradient descent, conjugate gradient, quasi-Newton, Levenberg-Marquardt, etc., that in this study Levenberg-Marquardt method was used due to its high speed and the ability to optimize parameters.

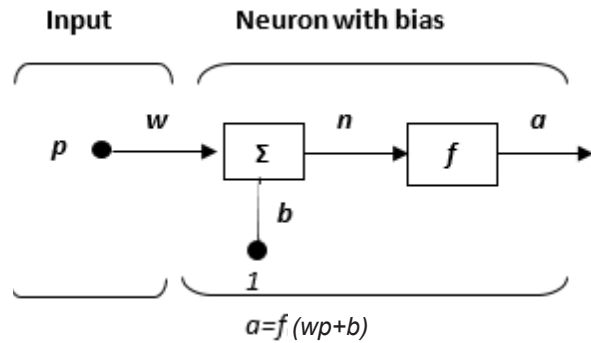


Fig. 1. A neuron with one input (RAJAEI et al., 2009).

ANFIS Fuzzy logic was first introduced in 1965 by Lotfi Askarzadeh. Fuzzy logic is a good solution for solving nonlinear systems that are less well known. Due to the problems in each of the two methods of fuzzy logic (lack of training, lack of a specific algorithm to determine the parameters, for example, the form of membership functions and problem of optimization) and neural network (problem of optimal design of the network, and the black box, and lack of direct and simple method for applying prior knowledge and deducing rules), the researchers sought to combine the neural network and fuzzy methods in order to use the advantages of both methods and avoid their disadvantages (ABRAHAM and NATH, 2001). Accordingly, one of the most common systems of neuro-fuzzy is ANFIS system that was first introduced by JANG, 1993. This model implements a Sugeno fuzzy system in a neural network structure and uses a combination of teaching methods of gradient descent and least squares method for training process assume that there is a fuzzy system with two inputs of  $x$  and  $y$ , and an output of  $z$ , by using Takagy-Sugeno system (Fig. 2), rules of 'if-then' are as follows:

Rule 1: if  $x$  is  $A_1$  and  $y$  is  $B_1$ , then,  $f_1 = p_1x + q_1y + r_1$ ,

Rule 2: if  $x$  is  $A_2$  and  $y$  is  $B_2$ , then,  $f_2 = p_2x + q_2y + r_2$ .

ANFIS structure of this example is shown in Figure 3. If the output of each layer of ANFIS model is shown as  $O_l$  and  $i$ , which is the output of  $i$ -th node in  $l$ -th layer, and then we can express the performance of different layers as following:

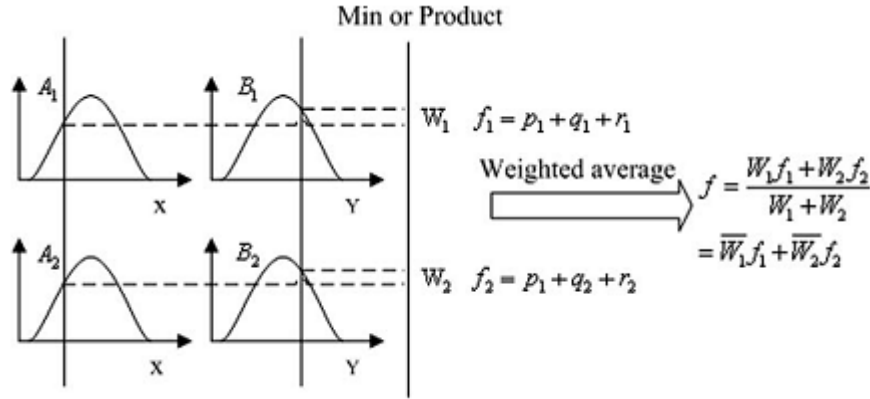


Fig. 2. Sugeno's fuzzy model with two inputs and two rules (ZHANG, 1993).

Layer 1: This layer is the input layer; each node in this layer is the equivalent of a fuzzy set. In this layer, the output of each node is equal to degree of membership of the input variable in this set (each node in this layer is a node adaptable with node functioning).

$$O_{1,i} = \mu A_i(x) \quad \text{for } i = 1 \text{ and } 2, \quad (1)$$

$$O_{1,i} = \mu B_{i-2}(y) \quad \text{for } i = 3 \text{ and } 4, \quad (2)$$

where  $x$  and  $y$  are the inputs of  $i$ -th node and  $A_i$  verbal description related to each node. Here, the output of each node is the membership degree of input variable. The parameters of each node determine the form of membership function equivalent to the fuzzy set of same node.

Layer 2: Each node in this layer is shown as  $\mu$  that in these nodes the input signals multiplied each other and the output is produced, for example:

$$w_i = \mu_{A_i}(x) \times \mu_{B_i}(y) \quad i = 1, 2. \quad (3)$$

The output of each node in this layer is the activity rate (firing strength) of a rule (in fact, the other T-norm

operators used as a node function in this layer).

Layer 3: Each node in this layer is shown as  $N$  that calculates the activity rate of  $i$ -th rule to the sum of activity rates of all rules (equation 7).

$$O_{3,i} = \bar{w}_i = \frac{w_i}{w_1 + w_2} \quad i = 1, 2, \quad (4)$$

where  $O_{3,i}$  is the output of this layer and is equal to the normalized firing strength of the  $i$ -th rule.

Layer 4: In this layer, each  $i$ - node calculates the share of  $i$ -th rule in the model output as a relation.

$$O_{4,i} = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i), \quad (5)$$

where  $\bar{w}_i$  is the output of the third layer, and  $p_i x + q_i y + r_i$  is the adaptive parameters of this layer. These parameters known as consequent parameters.

Layer 5: Each node in this layer is shown as  $\Sigma w_i$ , which calculates the final output as equation 6 (the number of nodes is equal to the number of outputs).

$$O_{5,i} = \sum_l \bar{w}_l f_l = \frac{\sum_i w_i f_i}{\sum_i w_i}. \quad (6)$$

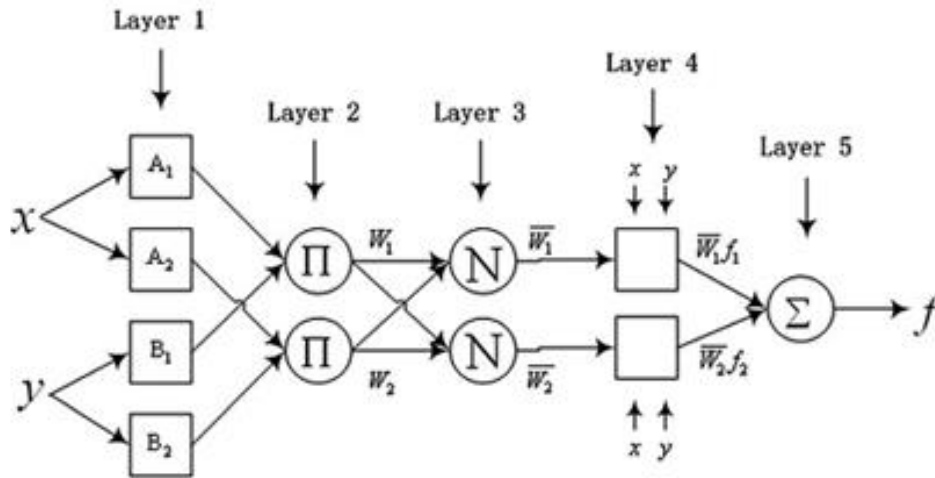


Fig. 3. The structure of ANFIS with two inputs (SCHILDT, 1998).

### Hybrid method used to train ANFIS

Each epoch of the hybrid method is composed of two forward and backward parts. In the forward path, the forward operating signals go ahead up to the layer 4, while the premise parameters are fixed; the consequent parameters (parameters of set) are determined using the method of least squares. In the backward path, the calculated error in the last layer is propagated toward the first layer and the premise parameters (parameters of set) are corrected using gradient descent method. In this part, the tally parameters are also fixed (ZHANG, 1993). Artificial neural networks (ANN), fuzzy logic (FL), and neuro-fuzzy (NF) methods have been used successfully by many researchers to estimate relationship between stream flow and sediment concentration (KISI, 2004a, b; CİĞİZOĞLU and KİŞİ, 2006; PARTAL and CİGİZOĞLU, 2008; HAMIDI and KAYAALP, 2008; RAJAEI et al., 2009; FIRAT and GUNGOR, 2010; KISI and SHIRI, 2012).

### Support vector machine (SVM)

SVM method is one of the most efficient methods for modeling the continuous and discrete data based on statistical training (WANG et al., 2009). In SVM method, the data that are non-linear at lower dimension are transferred to higher dimension using the kernel function, which are modeled as linear in this dimension. In fact, in SVM method the linear model fitted to the data in the higher dimension is equivalent to non-linear model in the lower dimension (YU et al., 2006). The components of a support vector machine are shown in Figure 4.

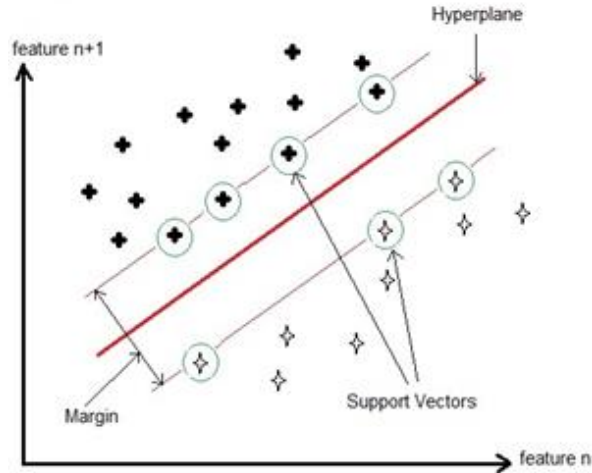


Fig. 4. The basis of the support vector machines (PARTAL and CİGİZOĞLU, 2008).

SVM method has been described in the literature for classification; therefore, in this study, just the regression support vector model is explained. It is assumed that there are  $N$  number of training data set, i.e.,  $\{(x_i, d_i)\}_i^N$ . Linear regression equation that in the feature space can be fitted with the higher data is in the form of equation (7).

$$y_i = f(x) = \langle w_i, \phi_i(x) \rangle + b, \quad (7)$$

where,  $w_i$  and  $b$  are the coefficients of relation in feature space at higher dimension and  $\phi_i(x)$  is a function that partitions the data from the input space to feature space. Like any other model, to estimate the parameters of the model we should consider a standard error. In this regard, different loss functions such as linear, quadratic, exponent, etc. are used. In this study, the Vapnik's- $\epsilon$  criteria that is presented as an equation is used to optimize the parameters of the above model (VAPNIK et al., 1997).

$$L_\epsilon(v, g(u)) = \begin{cases} 0 & \text{for } |v - g(u)| \leq \epsilon \\ |d - g(u)| - \epsilon & \text{otherwise,} \end{cases} \quad (8)$$

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*). \quad (9)$$

According to this equation, if the predicted values are within,  $\epsilon$  - tube loss equals to zero. Considering the loss function, the estimation of the regression equation parameters is changed to the optimization problem of convex that is expressed as follows.

$$\text{subject to } \begin{cases} v_i - (w \cdot \phi(u_i) + b) \leq \epsilon + \xi_i \\ (w \cdot \phi(u_i) + b) - v_i \leq \epsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, \quad i = 1, 2, \dots, n, \end{cases} \quad (10)$$

where  $\xi_i^*$  and  $\xi_i$  are slack variables (that determine the training error considering the error tolerance of high and low limits),  $C$  is regularization parameter and is it determines the compromise between the empirical risk and regularization term. Considering the Lagrange coefficients and Karush-Khan-Tuker conditions (ASSAD and GASS, 2011), the dual form of the above optimization problem changed to optimization problem of quadratic programming as the following equation.

$$\max -\frac{1}{2} \sum_{i,j=1}^n (\underline{a}_i - \bar{a}_i)(\underline{a}_j - \bar{a}_j) (\phi(u_i) \cdot \phi(u_j)) - \epsilon \sum_{i=1}^n (\underline{a}_i + \bar{a}_i) + \sum_{i=1}^n y_i (\underline{a}_i - \bar{a}_i), \quad (11)$$

subject to

$$\sum_{i=1}^n (\underline{a}_i - \bar{a}_i) = 0$$

$$0 \leq \underline{a}_i \leq C, i = 1, 2, \dots, n.$$

$$0 \leq \bar{a}_i \leq C, i = 1, 2, \dots, n.$$

By optimizing the equation (11), Lagrange coefficients,  $\underline{a}_i$  and  $\bar{a}_i$  are calculated and the vector of optimized weights of regression hyper plane is obtained as an equation.

$$w = \sum_{i=1}^n (\underline{a}_i - \bar{a}_i) K(\Phi(x_i), \Phi(x_j)). \quad (12)$$

The values of the input data for which the value of  $(\underline{a}_i - \bar{a}_i)$  is non-zero are considered as a support vector. With regard to this equation, the regression equation for support vectors expressed as an equation.

$$g(x) = \sum_{i=1}^n (\underline{a}_i - \bar{a}_i) K(\Phi(x_i), \Phi(x_j)) + b, \quad (13)$$

where  $n_l$  indicates the number of support vectors. In equations (12) and (13), the internal multiplication in feature space is done by kernel function. That is, if there are two vectors of  $\Phi(x_i)$  and  $\Phi(x_j)$  their internal multiplication is done by function as an equation. Every function which is true in Mercer conditions can be used as a kernel function (VAPNIK, 1995). Of common kernel functions, we can refer to the Gaussian radial basis function (equation (14)) and radial basic polynomials.

$$K_{xixj} = \exp(-\gamma \|x_i - x_j\|^2). \quad (14)$$

In this model, the parameters that affect the performance of the model and the need to be optimized include  $C$ , regularization parameter, the radius of the insensitive tube and the kernel function parameters.

### Research method

A FEL3 simulator model was used in the Soil Conservation and Watershed Management Institute for simulating rainfall, Tehran, Iran. FEL3 rainfall simulator is a battalion page simulator that creates a good distribution of the raindrop sizes with their kinetic energy (SOLTANI et al., 2019).

Tested soil samples were collected in the Taleghan watershed that is located 100 km from the Tehran. After sampling, all the samples were transferred to the Soil Conservation and Watershed Management Research Institute Tehran, Iran. Soil properties were studied in the soil laboratory of Soil Conservation and Watershed Management Research Institute Tehran, Iran. Soil texture according to a hydrometer method was determined as loamy and includes 18% clay, 33% sand and 49% silt (BOROGHANI et al., 2012). After the calibration of rainfall simulator and weighing the samples dried in the oven, and treating them with poly-acryl-amide material, the samples were placed in rainfall simulator under the rain with intensities of 60, 90 and 120 mm h<sup>-1</sup> for 10, 20 and 30 minutes. Each treatment was repeated 3 times. After completion of precipitation, again the samples were placed in an oven for 24 hours at temperature of 105 °C to calculate their secondary weights (BOROGHANI et al., 2012).

The method of calculating the splash rate. According to the definition of LOCKE and KITROMILIDES (1990), the rate of splash erosion per time and area units is called splash erosion rate and is calculated based on the equation 15 (KANE et al., 2008):

$$S = \frac{Dt_2 - Dt_1}{(t_2 - t_1) A}, \quad (15)$$

where  $S$  = splash rate in grams per minute per square meter,

$D_{t_1}$  and  $D_{t_2}$  sediment produced between times of  $t_1$  and  $t_2$  in grams,

$t_1$  and  $t_2$  = rainfall duration in minutes,

$A$  = area of splash bowl in square.

Splash erosion rate in each sample was calculated based on equation 15 and splash erosion rate for each treatment

was obtained from the average of three tested replicates for that treatment. After measuring the amount of the produced splash in various intensity values and percentage of different nanos, the modeling of produced sediment was applied using three methods of artificial neural networks, neuro-fuzzy and support vector machines. To implement these models, there are two main stages of pre-process. If the scale of the input data to network is different, each input has different effect such that in neural network, for example, some neurons may reach the firing limit and other neurons may not even be provoked. Therefore, it is required the data are placed at certain numeral range prior to entering them into model. For this purpose, in this study, using the following equation, all input data were set at the range of 0.05 to 0.95.

$$Y = 0.9 * \left( \frac{y - y_{min}}{y_{max} - y_{min}} \right) + 0.05, \quad (16)$$

where  $Y_{min}$  is the minimum observed data,  $Y_{max}$  is the maximum observed data and  $Y$  is the standardized value (HAN et al, 2007). After standardization, it is required that the training, validation and testing of data are done. Using the training patterns, model's parameters were estimated.

In this case, the training data must be chosen such that they include all patterns of observational data. Therefore, a greater percentage of observational data are usually allocated to training. In the training process with the introduction of training data to the model, the parameters of the model are estimated such that the error between observed and modeled data is minimized. This error is minimized with the progress of the training process. If training process lasted, the network begins to memorize training data and somehow an over fitting occurs. In this case, to prevent over fitting, a new data set (validation data) that is not directly used in the training process introduced to the model. In this case, as the error of validation data set begins to increase, the training process is stopped. The testing data are used to compare the performance of different models. Respect to available data (which included 36 data sets), the cross validation approach was used in this study. In this approach, all data were divided into six categories (each category included six data sets). In each model implementation, four categories for training, one category for validation and one category for testing were used. In this case, each data set at least used once as the test and once as validation. In the neural network, the structure of the network includes a number of hidden layers, a number of neurons of hidden layers, type of activation function and training algorithms. The more complex the structure of the model, the more the capability of the model in modeling more complex processes will be and on the other hand, by increasing the complexity of the model, the training process becomes longer and if the number of training data is low, the effectiveness estimation of parameters is not possible, and the possibility of trapping the model in local minimum increases. For this purpose, we need to determine the appropriate structure of the model considering the type of studied process and the observational data. HORNİK et al. (1989) stated that a network with one hidden layer and a sufficient number of

neurons in its hidden layer can determine the relationship between input and output data with any degree of accuracy. BERNARD et al. (1993) showed that a neural network with two hidden layers has had a weaker performance than a network with one hidden layer. So in this study for the artificial neural network, only one hidden layer was studied that its activation function was sigmoid and the number of its neurons was determined by trial and error. Also, for training the network Lonberg-Margotte method was used due to its high efficacy and speed (YU and WILAMOWSKI, 2011).

In ANFIS neuro-fuzzy method, the model parameters include the inputs, number, and form of membership functions for input and output; the type of membership functions (that in this study three Gaussian, triangular and bell membership functions were considered) and number of membership functions (that in this study range from 2 to 6) was determined by trial and error.

In method of support vector machine, model parameters values include the values of C, e and its kernel function parameters. There is no specific algorithm to determine these parameters. Therefore, these parameters are determined by trial and error method and grid search algorithm (TIEN BUI et al., 2012). In the grid search algorithm, because no directed optimization algorithm is used to optimize the parameters, it is necessary the model be implemented in several numbers. In order to solve this problem, the two-stage grid search algorithm used. In the first stage in this algorithm, the search for values of the parameters is done in a larger domain but with less precision. In the second stage, the search is done in more details and in a part of the parameters' space which had worked better in the previous stage (HSU et al., 2010; CHEN and YU, 2007). In SVM method, there are different kernel functions such as linear, radial base, polynomial, etc. In this connection, according to KAKAEI LAFDANI et al. (2014) the radial function has had better performance compared to the different kernel functions. In a research by DIBIKE et al. (2001), in investigating the types of kernel functions in modeling runoff rainfall, the radial function performed better than other functions. In other researches, such as KHAN and COULIBALY (2006), YU et al. (2006),

WANG et al. (2009), CIMEN (2008), HE et al. (2014), the radial base function has been used. Therefore, in this study we have used the radial base function in all models for each duration (10, 20 and 30 min) inputs were rainfall intensity and amounts of poly-acryl-amide.

In order to select the optimal structure of each model and compare different models in each stage, the evaluation criteria of correlation coefficient such as RMSE<sup>1</sup>, RSE, Nash-Sutcliff were used.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Q_i^o - Q_i^f)^2}, \quad (17)$$

$$RSE = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (Q_i^o - Q_i^f)^2}}{Q_{ave}^o}, \quad (18)$$

$$NS = 1 - \frac{\sum_{i=1}^n (Q_i^o - Q_i^f)^2}{\sum_{i=1}^n (Q_i^o - Q_i^p)^2}, \quad (19)$$

where  $Q_i^o$  is the observational value,  $Q_i^f$  is the estimated value by model, and  $Q_{ave}^o$  the value of observed data averages.

The disadvantages of the above criteria are that these criteria only consider the magnitude of the error and do not consider distribution of errors in assessing the different models. For this purpose, in addition to the above-mentioned criteria, the mean absolute relative error (MARE) was used (HE et al., 2014).

$$MARE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Q_i^o - Q_i^f}{Q_i^o} \right| * 100. \quad (20)$$

## Results and discussion

The results of the statistical analysis of different levels of splash rate in marl soil with rainfall intensities of 60, 90 and 120 mm per hour, and times of 10, 20 and 30 minutes indicate that there is a significant difference between various intensities of rainfall and the given times in terms of increasing the splash rate at 0.05 level, while the interactional effects of the intensity and duration on splash rate do not show a significant difference (Table 1).

Table 1. The results of two-way analysis of variance on two factors of the intensity and duration of precipitation

Sources of changes	Sum of squares	Degrees of freedom	Mean square	Sig	F
Rainfall intensity (mm s <sup>-1</sup> )	2,023.041	2	1,011.520	0.00**	60.268
Duration of rainfall (min)	1,945	2	174.360	0.01*	10.385
Rainfall intensity × rainfall duration	0.541	4	6.136	0.830 <sup>ns</sup>	0.366
Error	16.784	18	302.106	—	—

\*\*P ≤ 0.01, \*P ≤ 0.05, ns = not significant at a = 0.05.

Table 2. Statistical parameters of observation data for the splashes with different durations

	Duration of 10 min	Duration of 20 min	Duration of 30 min
Mean	47.96	50.88	57.12
Standard deviation	13.89	12.28	11.63
Minimum	16.49	23.12	32.84
Maximum	71.39	73.18	79.41

The results indicate that different amounts of poly-acryl-amide (0.2, 0.4 and 0.6 g m<sup>-2</sup>) in splash erosion control with the intensities of 60, 90 and 120 mm h<sup>-1</sup> and raining times of 10, 20 and 30 minutes show significant differences compared to control treatment, but the amount of 0.2 g m<sup>-2</sup> of this material in all three intensities and times of given rainfall does not show significant differences compared to control treatment. In general, the results showed that the amount of 0.6 g m<sup>-2</sup> of poly-acryl-amide had controlled better the splash rate in different intensities and rainfall durations (Table 2).

The statistical properties of observational data for splashes with different duration are given in Table 2. According to this table the highest amount of splash is related to the test with duration of 30 minutes.

### ANN Results

After pre-processing and entering data into the different models, the parameters of different models were estimated using training and validation data and were evaluated by testing the performance of different models. In the diagrams in Figure 5, the values of RMSE, MARE and correlation coefficient for different structures of the neural network are presented. According to this figure, it is observed that for the case that the duration of the test is 10 minutes, the network with two neurons in hidden layer, and for the test with duration of 20 minutes the network with two neurons in hidden layer and for the test with duration of 30 minutes the network with four neurons in hidden layer have had the best performance. According to the diagrams, it was found that the best performance of the neural network is related to tests with duration of 20 and 30 minutes.

ANFIS results for experiments with different durations, in the case when all rules made by ANFIS in modeling are applied, are shown in Figure 6. In the diagrams, the results of the best membership functions have been shown that for experiments with durations of 10 and 20, the Gaussian membership function and for experiments with duration of 30 minutes, the trapeze membership function has had better performances than other membership functions respectively. According to these figures and considering all experiments, the best results are related to the case when there are two membership functions and by increasing the number of membership functions the models have had worse performances. We can somehow state that by increasing the number of membership functions a kind of over-fitting has happened according to the data available. In artificial methods, over-fitting happens when the model has a complex structure or the training

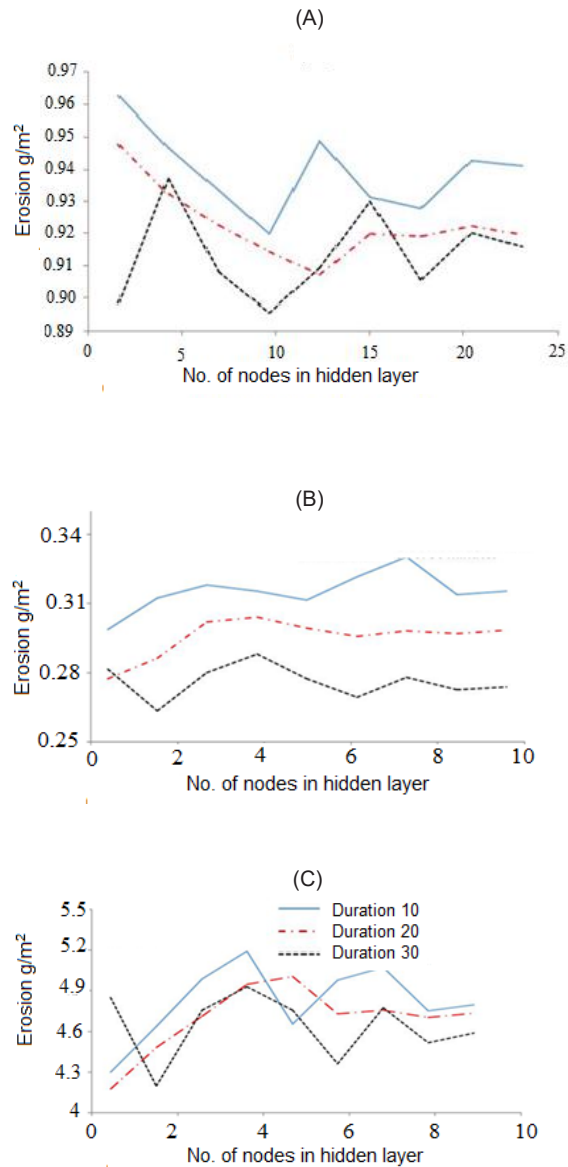


Fig. 5. The results of RMSE (A), MARE (B) and correlation coefficient (C) for different structures of ANN.

process is elongated. Since the validation data in training process were used in this study, we can say that the over-fitting has happened due to the complex structure of the model. Therefore, in order to more simplify the model, each of ANFIS structures with different membership functions was studied and determined as trial and error by

eliminating some rules. It should be noted that in this stage due to the better performance of the Gaussian membership function, only the Gaussian membership function was evaluated. ANFIS results when some of its rules have been eliminated are shown in Table 3. By studying the different structures in all three experiments and in all membership functions, the best results were related to the case that the number of rules was the smallest, that is, two rules. We can typically say that the simplest structures had the best performances. According to Table 3, it is observed that unlike the previous results (when the initial structure of ANFIS without eliminating the rules was applied to the model and the structure with two membership functions had the best performance), the best results are related to models with high numbers of membership functions.

### SVM method

In SVM method, the most important parameters include values of  $C$ ,  $\epsilon$ , and parameters of kernel functions.  $C$  is a regulating parameter that determines the amount of training error with regard to the band. The low values of  $C$  result in high training error and high values of  $C$  result to the over-fitting of the model. WANG et al. (2003) stated that the results of the model are less sensitive to the values of  $C$ . Similarly, for the parameter of  $\epsilon$  if its value is low, due to ignorance of many support vectors, the training error is low and the model does not have necessary complication for modeling the data, and by increasing its value the complexity of the model also increases and consequently the possibility of over-fitting is increased. In this regard in the present research, grid search algorithm was used to find some values of these parameters that give the best results. Finally, for the test with a duration of 10 minutes the optimal values of parameters  $C$ ,  $\epsilon$  and  $\gamma$  were 96.5, 0.09 and 8 respectively, for the test with duration of 20 minutes these amounts were 5, 0.0801 and 2 respectively and for the test with duration of 30 minutes these amounts were 537, 0.07 and 23 respectively. The results of different models for experiments with various durations are shown in Table 3. SHU and OUARDA (2008) stated that when the value of Nash-Sutcliff is more than 0.8, the estimation of the model is acceptable. In this connection, HE et al. (2014) found that the value of Nash-Sutcliff of black box models such as neural network, ANFIS and SVM has been greater than 0.82 in predicting monthly discharge amounts. In this study, according to the following table it is observed that the value of Nash-Sutcliff for all models and all experiments is greater than 0.85, which indicates the appropriate performance of these models. Also according to this table, result showed that the maximum value is Nash-Sutcliff and the minimum values are RMSE, RSE and MARE related to the SVM method. In this model, the values of MARE are close to zero which typically indicates that the overestimation and underestimation have happened much less and the minimum proportion of the error is related to the observations of this model.

According to Table 4 it is observed that in all models, based on the values of RMSE, RSE and MARE, the best results are related to the experiment with the duration of 30 minutes. SVM and ANFIS models calculated the test values with duration of 10 minutes better than the test values with duration of 20 minutes, while the neural network had better performance in modeling the values obtained with duration of 10 minutes. In all tests, the performance of SVM model in terms of all assessment criteria has been better than the other models. After the SVM model which has had the best performance among all models, ANFIS model is in second place. According to AKBARZADEH et al. (2009) in modeling of runoff and splash, the ANFIS model had a better performance than the neural network. According to HE et al. (2014), the SVM model had a better performance than ANFIS and neural network models to predict the stream flow of Pailugou River in China. In a study by WANG et al (2009), seven models were used to predict the monthly discharge of two rivers in China

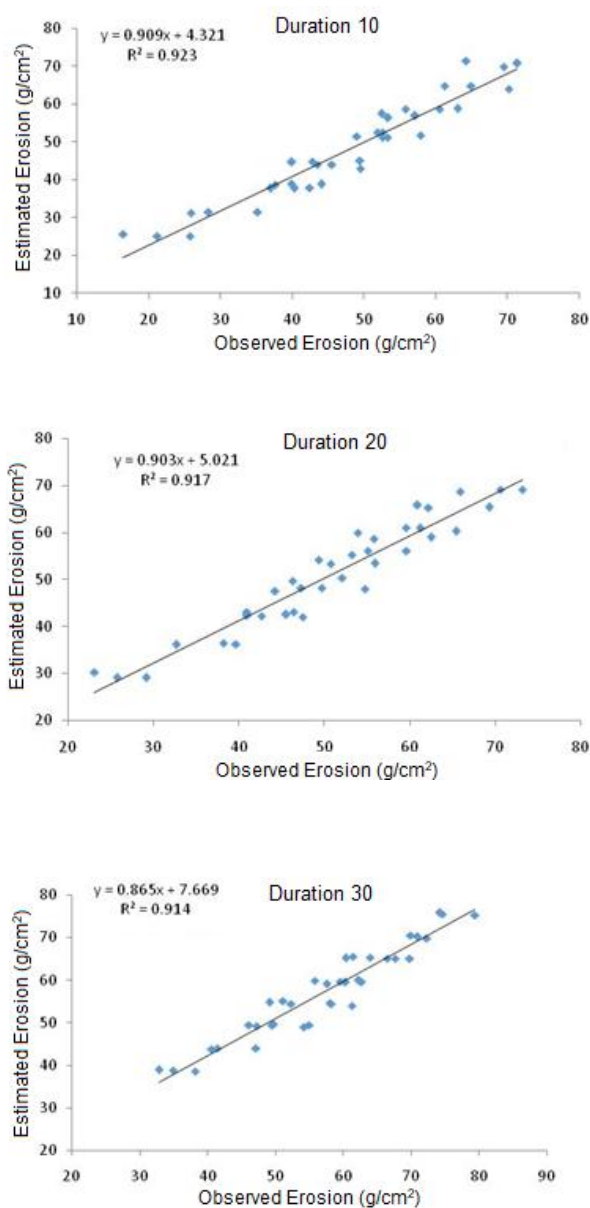


Fig. 6. The results of RMSE (A), correlation coefficient (C) and MARE (B) of ANFIS for the number of various membership functions in different experiments.

Table 3. The results of ANFIS when some of the rules eliminated

The number of membership functions	Testing with duration of 10 min				Testing with duration of 20 min				Testing with duration of 30 min			
	2	3	4	5	2	3	4	5	2	3	4	5
RMSE	4.266	4.269	3.944	3.861	4.383	4.532	4.076	3.959	3.970	3.728	4.063	3.908
RSE	0.089	0.089	0.082	0.081	0.086	0.089	0.080	0.078	0.070	0.065	0.071	0.068
R	0.951	0.951	0.958	0.960	0.933	0.929	0.942	0.945	0.938	0.946	0.935	0.941
MARE	0.088	0.297	0.296	0.287	0.287	0.290	0.280	0.274	0.248	0.239	0.257	0.246

that among them SVM model had a better performance than the two models of ANFIS and ANN. In a research by (KAKAIE LEFDANI et al., 2013), in the estimation of suspended sediment using neural network and SVM, the SVM model had a better performance than the neural network. CIMEN (2008) found that in predicting the amount of suspended sediment, the SVM method had a better performance than the methods of fuzzy-inference systems and Artificial Neural Networks. In Figure 7, diagram of the distribution of SVM results have been presented. In the previous sections, it was observed that while the values of RMSE, RSE and MARE for the test with duration of 10 minutes were lower compared to the test with the duration of 30 minutes, but the correlation coefficient of the test was greater for the test with duration of 10 minutes this issue can be observed in the pattern of distribution of observed and predicted data in Figure 7. According to this figure, it is observed that in the test with duration of 10 minutes there is a kind of symmetry between the underestimated and overestimated data, but the distance of some points to the fitted line is greater this problem has caused while the correlation coefficient between the observed and predicted data and Nash-Sutcliffe criteria for these data is high, but the amounts of other error criteria for these data are not suitable.

The modeled observational values by all three methods have been presented in Figure 8. According to diagrams, although SVM model has had a better performance than ANFIS and neural network models in terms of evaluation criteria at all three experiments, but in some observational values, the ANFIS and neural network models have had a better performance than SVM model. For example, in the experiment with duration of 20 minutes, neural network model in 15 cases and ANFIS in 13 cases have had a better performance than SVM model. However, in comparison

to the other two methods, SVM model predicted wrong values are less that this issue can also be observed in the calculated values of MARE.

### Conclusions

In this study, using three approaches of artificial intelligence including artificial neural network, ANFIS and SVM and taking into account the two inputs of rainfall intensity and amount of poly-acryl-amide, the amount of the produced splash was modeled. The results showed that in ANFIS model, according to low number of the observed data, a kind of over-fitting has happened by increasing the number of membership functions and the best results are related to the case that the membership functions are less. To solve the problem of over fitting in ANFIS, some rules were eliminated and the model was re-run. The results showed that by making the structure of ANFIS model simpler, the performance becomes better. WANG et al. (2009) observed that the ANFIS model had a better performance compared to the models of the neural network, SVM, genetic programming, and ARMA in training, but in testing performances, the other models have been better which indicates that in ANFIS model the probability of over-fitting is more than other models. In a research by HE et al. (2014), also the best performance is related to the case that the number of membership functions was 2 and 3. Among all studied models, the best results were obtained from SVM model, followed by ANFIS model. Based on the data used, the results showed that the amount of 0.6 g m<sup>-2</sup> of poly-acryl-amide had the greatest effect in reducing splash erosion in three intensities and various durations of rainfall in comparison

Table 4. The results of the different models

	ANN			ANFIS			SVM		
	Duration (10 min)	Duration (20 min)	Duration (30 min)	Duration (10 min)	Duration (20 min)	Duration (30 min)	Duration (10 min)	Duration (20 min)	Duration (30 min)
RMSE	4.195	4.038	4.068	3.861	3.959	3.728	3.796	3.476	3.411
RSE	0.087	0.084	0.071	0.081	0.078	0.065	0.079	0.068	0.060
R	0.955	0.943	0.935	0.960	0.945	0.946	0.961	0.958	0.956
MARE	0.299	0.271	0.253	0.287	0.274	0.239	0.279	0.254	0.232
NS	0.906	0.889	0.874	0.921	0.893	0.894	0.923	0.918	0.912

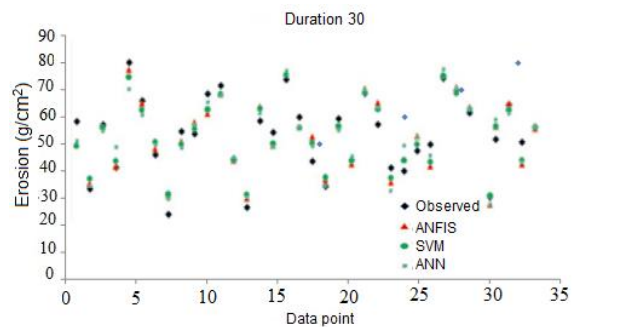
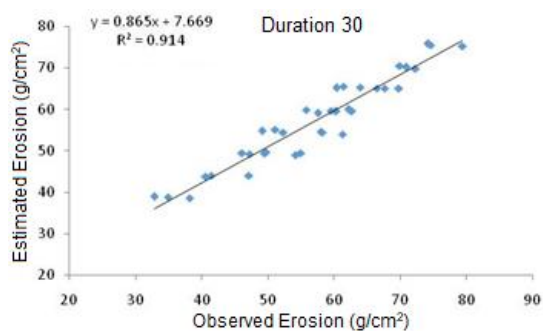
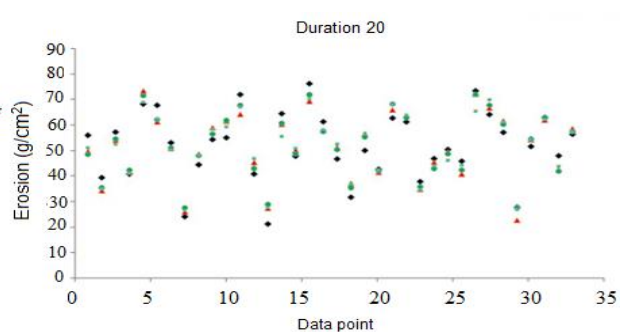
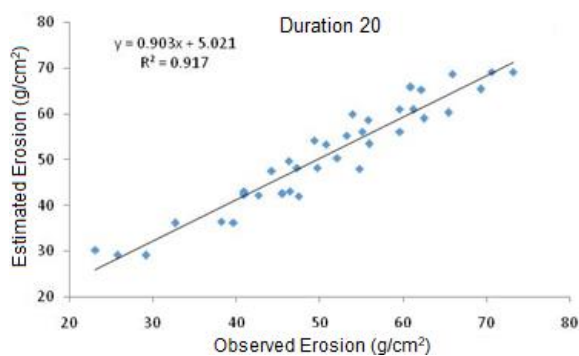
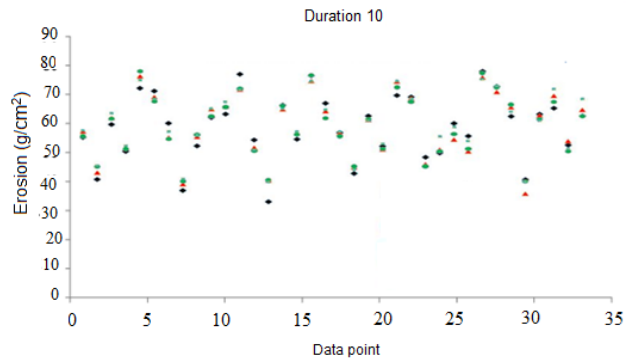
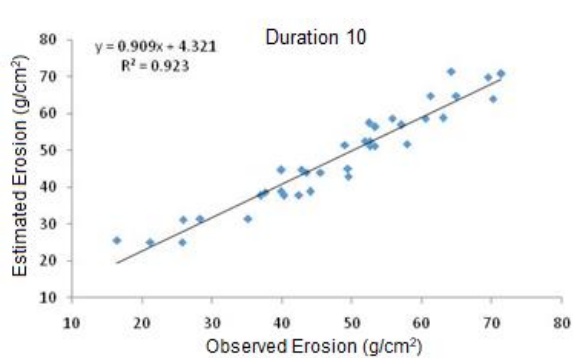


Fig. 7. Diagram of the distribution of observed and estimated values by SVM model.

Fig. 8. Observed and estimated values using three methods used.

to control treatment that is in agreement with the results obtained by BOROGHANI et al. (2012). In this study, the best results were obtained when the duration of the test was 10 minutes. SVM method was capable of modeling soil erosion such that there is the least difference between the observed and estimated data. In this method, the most important issue is the determination of the optimal values of the C, and parameters such that there is the smallest difference between the observed and predicted values. For this purpose, the grid search algorithm was used to find the optimal parameters in this study. In this method, all spaces of parameters searched to find the optimal parameters. According to the dimensions of the issue, one of the problems of this method is its consumption of time. Using the optimization methods such as genetic algorithm, we can reduce this time by implementing a directed search and obtain the values of parameters with higher accuracy.

## References

- ABRAHAM, A., NATH, B., 2001. *Hybrid intelligent systems design – A review of a decade of research*. Technical Report, 5/2000. Gippsland School of Computing and Information Technology, Monash University, Australia.
- AKBARZADEH, A., TAGHIZADEH MEHRJARDI, R., ROUHIPOUR, H., GORJI, M., REFAH, H.G. 2009. Estimating of soil erosion covered with rolled erosion control systems using rainfall simulator (neuro-fuzzy and artificial neural network approaches). *Journal of Applied Sciences Research*, 5: 505–514. (In Persian).
- ASSAD, A.A., GASS, S.I. (eds), 2011. *Profiles in operations research: Pioneers and innovators*. International Series in Operations Research et Management Science, 147. Boston: Springer Science & Business Media. 824 p.
- BERNARD, V., THOMAS, J., ABARBANELL, J., 1993. How sophisticated is the market in interpreting earnings

- new? *Applied Corporate Finance*, 6: 54–63. <https://doi.org/10.1111/j.1745-6622.1993.tb00383.x>
- BOROGHANI, M., HAYAVI, F., NOOR, H., 2012. Affectability of splash erosion by poly-acryl-amide application and rainfall intensity. *Soil and Water Research*. (4): 159–165. <https://doi.org/10.17221/45/2011-SWR>
- BOROGHANI, M., MIRNIA, S. K., VAHHABI, J., AHMADI, S. J., CHARKHI, A., 2011. Nanozeolite synthesis and the effect of on the runoff and erosion control under rainfall simulator. *Australian Journal of Basic and Applied Sciences*, 5 (12): 1156–1164.
- CHEN, ST., YU, PS., 2007. Pruning of support vector networks on flood forecasting. *Journal of Hydrology*, 347: 67–78. <https://doi.org/10.1016/j.jhydrol.2007.08.029>
- CIĞIZOĞLU, HK., KIŞI, Ö., 2006. Methods to improve the neural network performance in suspended sediment estimation. *Journal of Hydrology*, 317: 221–238. <https://doi.org/10.1016/j.jhydrol.2005.05.019>
- CIMEN, M., 2008. Estimation of daily-suspended sediments using support vector machines. *Hydrological Sciences Journal*, 53: 656–666. <https://doi.org/10.1623/hysj.53.3.656>
- DEMUTH, H., BEALE, M., 1998. *Neural Network Toolbox for Use with MATLAB. User's Guide. Version 3*. Natick: The MathWorks Inc.
- DIBIKE, YB., VELICKOV, S., SOLOLATINE, D.P., ABBOTT, MB., 2001. Model induction with support vector machine: introduction and application. *Journal of Computing in Civil Engineering*, 15: 208–216. [https://doi.org/10.1061/\(ASCE\)0887-3801\(2001\)15:3\(208\)](https://doi.org/10.1061/(ASCE)0887-3801(2001)15:3(208))
- FIRAT, M., GUNGOR, M., 2010. Monthly total sediment forecasting using adaptive neuro fuzzy inference system. *Stochastic Environmental Research and Risk Assessment*, 24: 259–273. <https://doi.org/10.1007/s00477-009-0315-1>
- HAMIDI, N., KAYAALP, N., 2008. Estimation of the amount of suspended sediment in the Tigris River using artificial neural networks. *Clean*, 4: 380–386. <https://doi.org/10.1002/clen.200700094>
- HAN, D., CHAN, L., ZHU, N., 2007. Flood forecasting using support vector machines. *Journal of Hydroinformatics*, 4: 267–276. <https://doi.org/10.2166/hydro.2007.027>
- HE, Z., WEN, X., LIU, H., DU, J., 2014. 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*, 509: 379–386. <https://doi.org/10.1016/j.jhydrol.2013.11.054>
- HORNİK, K., STINCHCOMBE, M., WHITE, H., 1989. Multilayer feed forward networks are universal approximates. *Neural Networks*, 5: 359–366. [https://doi.org/10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8)
- HSU, R.J., LIN, C.Y., HOI, HS, ZHENG, SK, LIN, CC, TSAI, HJ, 2010. Novel intronic microRNA represses zebrafish myf5 promoter activity through silencing dickkopf-3 gene. *Nucleic Acids Research*, 38 (13): 4384–4393. <https://doi.org/10.1093/nar/gkq148>
- JANG, J.S.R., 1993. ANFIS: Adaptive-network-based fuzzy inference systems. *IEEE Transactions on Systems, Man, and Cybernet*, 23: 665–685. <https://doi.org/10.1109/21.256541>
- KAKAEI LAFDANI, E., MOGHADDAMNIA, A., AHMADI, A., EBRAHIMI, H., 2014. Assessing the impact of input variables preprocessing into support vector machine through gamma test method for suspended sediment volume prediction. *Journal of Range and Watershed Management*, 67 (2): 289–303. (In Persian). DOI: 10.22059/jrwm.2014.51833
- KANE, E.S., BETTS, E.F., BURGİN, A.J., CLILVERD, H.M., CRENSHAW, C.L., FELLMAN, J.B., MYERS-SMITH, I.H., O'DONNELL, J.A., SOBOTA, D.J., VERSEVELD, W.J., JONES, J.B., 2008. Precipitation control over inorganic nitrogen import-export budgets across watersheds: A synthesis of long-term ecological research. *Ecohydrology*, 1: 105–117. <https://doi.org/10.1002/eco.10>
- KANE, J.M., LAURIELLO, J., LASKA, E., DI MARINO, M., WOLFGANG, C.D., 2008. Long-term efficacy and safety of iloperidone: Results from 3 clinical trials for the treatment of schizophrenia. *Journal of Clinical Psychopharmacology*, 28 (2): S29–S35. DOI: 10.1097/JCP.0b013e318169cca7
- KHAN, M.S., COULIBALY, P., 2006. Application of support vector machine in lake water level prediction. *Journal of Hydrologic Engineering*, 3: 199–205. [https://doi.org/10.1061/\(ASCE\)1084-0699\(2006\)11:3\(199\)](https://doi.org/10.1061/(ASCE)1084-0699(2006)11:3(199))
- KIŞI, Ö., 2004a. Daily suspended sediment modelling using a fuzzy differential evolution approach. *Hydrological Sciences Journal*, 1: 183–197. <https://doi.org/10.1623/hysj.49.1.183.54001>
- KIŞI, Ö., 2004b. Multi-layer perceptrons with Levenberg-Marquardt training algorithm for suspended sediment prediction and estimation. *Hydrological Sciences Journal*, 6: 1025–1040. <https://doi.org/10.1623/hysj.49.6.1025.55720>
- KISI, O., SHIRI, J., 2012. River suspended sediment estimation by climatic variables implication: Comparative study among soft computing techniques. *Computers and Geosciences*, 43: 73–8. <https://doi.org/10.1016/j.cageo.2012.02.007>
- LEGUEDOIS, S., PLANNCHON, O., LEGOUT, C., BISSONNAIS, Y.L., 2005. Splash projection distance for aggregated soils: Theory and experiment. *Soil Science Society of America Journal*, 69: 30–37. <https://doi.org/10.2136/sssaj2005.0030>
- LOCKE, J., KITROMILIDES, P., 1990. *Deuterē pragmateia peri kyvernēseōs: dokimio me thema tēn alēthinē archē, ektasē kai skopo tēs politikēs exousias* [Essay concerning the true original extent and end of civil government]. Athēna: Ekdoseis “Gnōsē”. 302 p.
- PARTAL, T., CIĞIZOĞLU, HK., 2008. Estimation and forecasting of daily suspended sediment data using wavelet-neural networks. *Journal of Hydrology*, 358: 317–331. <https://doi.org/10.1016/j.jhydrol.2008.06.013>
- RAJAEI, T., MIRBAGHERI, S.A., ZOUNEMAT-KERMANI, M., NOURANI, V., 2009. Daily suspended sediment concentration simulation using ANN and neuro-fuzzy models. *Science of the Total Environment*, 407: 4916–4927. <https://doi.org/10.1016/j.scitotenv.2009.05.016>
- SCHILDT, G., 1998. *A distributed user adaptive neuro-fuzzy controller application for home automation*. Foundation

- research project. Vienna University of Technology, Institute of Automation, Austria.
- SHU, C., OUARDA, T.B.M.J., 2008. Regional flood frequency analysis at ungauged sites using the adaptive neuro-fuzzy inference system. *Journal of Hydrology*, 349: 31–43. <https://doi.org/10.1016/j.jhydrol.2007.10.050>
- SOJKA, R.E., 2006. *PAM Research Project*. [cit. 2021-05-25]. <http://kimberly.ars.usda.gov/pampage.shtml>
- SOJKA, R.E., ORTS, W.J., ENTRY, J.A., 2004. Soil physics and hydrology: conditioners. In HILLEL, D. (ed.). *Encyclopedia of soil science*. Oxford: Elsevier, p. 301–306. <https://doi.org/10.1016/B0-12-348530-4/00501-4>
- SOLTANI, S., GHEZELSOFLU, N., BOROUGHANI, M., 2019. Using polyacrylamide to control soil splash erosion in rainfalls with variable intensity and duration. *Spanish Journal of Soil Science*, 9 (3): 213–222. <https://doi.org/10.3232/SJSS.2019.V9.N3.05>
- SOUSA-VIEIRA, M. E., SUAREZ-GONZALEZ, A., LOPEZ-ARDAO, J. C., LOPEZ-GARCIA, C., 2009. Study and M/G/∞-based modeling of the correlation characteristics of H. 264/AVC video traffic and the SVC extension. In *2009 IEEE International symposium on broadband multimedia systems and broadcasting : Broadband multimedia symposium 2009. 13-15 May 2009, Euskalduna Conference Centre, Bilbao, Spain*. Piscataway, NJ: IEEE, p. 1–5.
- TIEN BUI, D., PRADHAN, B., LOFMAN, O., REVHAUG, I., 2012. Landslide susceptibility assessment in Vietnam using support vector machines, decision tree, and Naive Bayes models. *Mathematical Problems in Engineering*, 2012: article ID 974638, 26 p.. <https://doi.org/10.1155/2012/974638>
- TRIPATHI, R.P., SINGH, H.P., 2001. *Soil erosion and conservation*. New Delhi: New Age International Limited Publication. 210 p.
- VAPNIK, V., 1995. *The nature of statistical learning theory*. New York: Springer. 189 p. <https://doi.org/10.1007/978-1-4757-3264-1-8>
- VAPNIK, V., GOLOWICH, S.E., SMOLA, A., 1997. Support vector method for function approximation, regression estimation, and signal processing. In *NIPS'96: Proceedings of the 9th International conference on neural information processing systems*. Cambridge, MA, USA: M.I.T. Press, p. 281–287. <https://doi.org/10.1007/978-1-4757-2440-0>
- WALLACE, A., WALLACE, G.A., 1986. Effects of soil conditioners on emergence and growth of tomato, cotton, and lettuce seedlings. *Soil Science*, 141: 313–316. <https://doi.org/10.1097/00010694-198605000-00002>
- WANG, W.C., CHAU, K.W., CHENG, C.T., QIU, L., 2009. A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. *Journal of Hydrology*, 374: 294–306. <https://doi.org/10.1016/j.jhydrol.2009.06.019>
- WANG, W.J., XU, Z.B., LU, W.Z., ZHANG, X.Y., 2003. Determination of the spread parameter in the Gaussian kernel for classification and regression. *Neurocomputing*, 55: 643–663. [https://doi.org/10.1016/S0925-2312\(02\)00632-X](https://doi.org/10.1016/S0925-2312(02)00632-X)
- YU, H., WILAMOWSKI, B.M., 2011. Levenberg-Marquardt training. In WILAMOWSKI, B.M., IRWIN, J.D. *The industrial electronics handbook. Intelligent Systems. Vol. 5*. Boca Raton: CRC Press, Chapter 12, p. 12-1–12-15. <https://doi.org/10.1201/9781315218427-12>
- YU, P.S., CHEN, S.T., CHANG, I.F., 2006. Support vector regression for real-time flood stage forecasting. *Journal of Hydrology*, 328: 704–716. <https://doi.org/10.1016/j.jhydrol.2006.01.021>
- ZHANG, Y., 1993. A modified effective binary diffusion model. *Journal of Geophysical Research-Solid Earth*, 98 (7): 11901–11920. <https://doi.org/10.1029/93JB00422>
- ZUMR, D., MÜTZENBERG, D.V., NEUMANN, M., JEŘÁBEK, J., LABURDA, T., KAVKA, P., JOHANNSEN, L.L., ZAMBON, N., KLIK, A., STRAUSS, P., DOSTÁL, T., 2020. Experimental setup for splash erosion monitoring—Study of silty loam splash characteristics. *Sustainability*, 12 (1): 157. <https://doi.org/10.3390/su12010157>

Received April 24, 2021  
Accepted October 12, 2021