



Artificial Neural Network and Support Vector Machine Models for Inflow Prediction of Dam Reservoir (Case Study: Zayandehroud Dam Reservoir)

Mohammad Babaei¹ · Ramtin Moeini¹ · Eghbal Ehsanzadeh²

Received: 12 September 2018 / Accepted: 25 March 2019 /

Published online: 3 April 2019

© Springer Nature B.V. 2019

Abstract

Inflow prediction of reservoirs is of considerable importance due to its application in water resources management related to downstream water release planning and flood protection. Therefore, in this research, different new input patterns for predicting inflow to Zayandehroud dam reservoir is proposed employing artificial neural network (ANN) and support vector machine (SVM) models. Nine different models with different patterns of input data such as inflow to the dam reservoir considering time duration lags, time index, and monthly rainfall of Ghaleh-Shahrokh station have been proposed to predict the inflow to the dam reservoir. Comparison of the results indicates that the ninth proposed model has the least error for inflow prediction in which the results of SVM model outperform those of ANN model. That is, the least error has been obtained using the ninth SVM (ANN) model with correlation coefficient (R) values of 0.8962 (0.89296), 0.9303 (0.92983) and 0.9622 (0.95333) and root mean squared error (RMSE) values of 47.9346 (48.5441), 42.69093 (43.748) and 23.56193 (28.5125) for training, validation and test data, respectively.

Keywords Inflow discharge prediction · Artificial neural network · Support vector machine · Input data pattern · Zayandehroud dam

✉ Ramtin Moeini
r.moeini@eng.ui.ac.ir

Mohammad Babaei
mohammad.babaei@eng.ui.ac.ir

Eghbal Ehsanzadeh
eehsanzadeh@gmail.com

¹ Department of Civil Engineering, Faculty of Civil Engineering and Transportation, University of Isfahan, Isfahan 81746-73441, Iran

² Department of Water Engineering, Faculty of Agriculture, Ilam University, Ilam, Iran

1 Introduction

Prediction of water resources quantity with acceptable accuracy is one of the most important parts of water resources management. In other words, determining the actual and optimal amounts of water resources including water release from reservoirs provides valuable information that helps planners for managing and allocating optimal water resources. The optimal values of water release from the reservoir is determined by different parameters particularly inflow into the reservoir and downstream demands. Due to the uncertain values of inflow into the reservoir and in order to achieve optimal management of surface water resources, it is necessary to predict inflow discharges into the reservoirs. For this purpose, various methods such as autoregressive moving average (ARMA), autoregressive integrated moving average (ARIMA), artificial neural network (ANN), and adaptive neuro fuzzy inference system (ANFIS) have been proposed which can be generally categorized into two classes of data-driven and conceptual methods. The main difference between these two classes is the degree of dependency on the amount of input data. Due to less dependency on the amount of input data and the less complexity, the data-driven methods are recently receiving much more attention (Lima et al. 2014). One of the most usual data-driven methods is the artificial neural network (ANN) model which is based on the structure of human nervous system. Due to high predictive power and flexibility, the ANN model has been extremely used recently.

Reviewing the literature shows that ANN model has been used as a powerful tool for solving many engineering problems including water resources engineering problems (Noori et al. 2010). This is due to the ability of this model to simulate and estimate nonlinear functions with proper precision (Menhaj 2009). Another data-driven method is support vector machine (SVM) model which has many advantages over ANN models (Liu et al. 2016). The SVM model is a machine learning method and is usually used to simplify complex processes, simulate nonlinear behaviors, and obtain the necessary information to make an appropriate decision. In other words, SVM model is a powerful method for describing various aspects of uncertainty in a prediction problem (Wenjian et al. 2008). The SVM model was originally developed by Vapnik (1995) and later Dibike et al. (2001) proved the applicability of this model for classification and regression problems.

Numerous research works have been done on the use of data-driven models for inflow predicting into dam reservoirs. For example, Coulibaly et al. (2000) used feed forward neural network (FFNN) to predict daily water inflow into a reservoir in Quebec, Canada. They used an approach known as stopped training approach to improve the ANN model performance. The results showed that using proposed approach increased the prediction accuracy of water inflow to the reservoir. Asefa et al. (2006) used a SVM model to predict water discharge of the Sevier River for short and long time conditions. In this study, different combinations of data were considered as input data for both conditions. The results showed the capability of SVM model to accurate prediction of water discharge. In addition, using the SVM model led to much simpler prediction specifically for predicting extreme seasonal water discharge in comparison with transfer function noise (TFN) model. Lin et al. (2006) predicted the monthly discharge of a river in China using SVM model considering Gaussian kernel function. The parameters of the model were determined using an evolutionary algorithm. In addition, comparing the results of SVM, ANN, and ARMA models showed that using SVM model led to better results. Lohani et al. (2012) investigated the performance of ANN, ANFIS, and regression methods to predict monthly inflow into the reservoir of Bhakra dam constructed on Sutlej River in India. The results indicated the accuracy of the ANFIS model in comparison

with all other models in which making use of periodic time period for the monthly inflow improved the prediction ability of the ANFIS model. Valipour et al. (2013) compared the performance of ARIMA, ARMA, and static and dynamic ANN models for predicting monthly inflow into the Dez dam reservoir. The results showed that using dynamic ANN model with sigmoid transfer function and using 17 neurons in the hidden layer led to the best results. Awchi (2014) predicted the water discharge of two branches of the Tigris River in northern Iraq using FFNN, generalized regression neural network (GRNN), and radial basis function (RBF) of ANN models. Comparison of the results showed that FFNN model had better results in comparison with other models in which adding the time index as an model input led to better prediction results. In addition, the performance of the ANN model was studied for predicting high peaks and low flows of water discharge using other river data. The obtained results indicated the capability of this model to predict high and low flows into the reservoir. He et al. (2014) predicted the river water discharge of semi-arid mountainous regions of northwestern China using ANN, ANFIS, and SVM models considering different combinations of input data. Comparison of the results showed that the model with three time lags for river discharge was the best model. In addition, the results of SVM model were better compared to other employed models. Liu et al. (2014) used a model named Bayesian wavelet–support vector regression model (BWS) to predict monthly and daily water discharges of two rivers in Dongjiang region of southern China. Comparing the results of this model with ANFIS model indicated that using the proposed model led to better results. In addition, the BWS model provided more detailed information on the uncertainty of the data used in the study. Kalteh (2015) used support vector machine wavelet regression model to predict monthly water discharge of two rivers in northern Iran and compared the results with SVM model. In this study, genetic algorithm was used to determine the optimal parameters of SVM model. Comparison of the results showed that using support vector machine wavelet regression model led to more accurate results. Finally, Zhu et al. (2016) combined the SVM model with discrete wavelet transform (DWT) and empirical mode decomposition (EMD) models to predict monthly water discharge of Jinsha River in China. The results of this study showed that using both DWT and EMD methods improved the accuracy of prediction in which the DWT method was superior for water discharge prediction.

Generally, inflow into reservoirs is very important for managing and planning water supply. However, the inflow values in the future are uncertain and therefore it is necessary to obtain as much information as possible about the possible variations of this variable in the future. This is important for decision making and optimal operation of water related infrastructures. Reviewing the literature indicates the wide application of data-driven methods in field of water resources management and therefore it is an attractive research topic for researchers. In current study, due to the unique capabilities of ANN and SVM models for inflow prediction, these models are used to predict inflow into the Zayandehroud dam reservoir. In order to improve the accuracy of the obtained results, here, different patterns are proposed by combining different sets of input data to increase the accuracy of prediction as well as adding time index (T) as input data. The structure of the present research is as follows. The literature review and the novelty and necessity of this research are presented in the Section 1. The detailed information on the study area (Zayandehroud dam in Isfahan) and data used in the study are presented in Section 2. In Section 3, different ANN and SVM models for predicting the inflow into the reservoir are described in detail. The results of prediction of inflow into the Zayandehroud dam reservoir using the proposed models are presented in Section 4. Finally, in Section 5 the concluding remarks and recommendations for further work are presented.

2 Study Area and Data

The study area of this research is Zayandehroud basin, located in the central plateau of Iran and between latitudes $31^{\circ} 15'$ and $33^{\circ} 45'$ north and longitudes $50^{\circ} 02'$ and $53^{\circ} 45'$ east. This basin is located in a warm and arid region of Iran. The total area of this basin is $41,524 \text{ km}^2$ which covers 2.5% of Iran's total area. About $16,670 \text{ km}^2$ of this basin is mountainous and $24,854 \text{ km}^2$ consists of plains and foothills. The basin is limited from north to the Salt Lake catchment, from east to Dagh-Sorkh Playa and Siyah-Kooh mountains, from south to Abarqoo Desert basin, and from west and southwest to the Karun River basin. The total precipitation of the basin varies from 1500 mm in the west to 50 mm in the east with average annual value of 140 mm (Jamali et al. 2007).

Zayandehroud basin has several hydrometric stations including Ghaleh-Shahrokh, Skandari, Zayandehroud regulatory dam, Zamankhan Bridge, Kale Bridge, Varzaneh, Nekoabad (Lenj), Chom Bridge, Mosian, Diziche, Chelgerd and Dare stations. Among these hydrometric stations only Ghaleh-Shahrokh station has quality data with sufficient statistically significant duration in comparison with all other stations. Therefore, Ghaleh-Shahrokh station is considered to be the best station for the purpose of this study due to both the location and the proximity of this station to a climatologically station with the same name. Therefore, the monthly precipitation of this station is used for prediction of inflow to Zayandehroud dam reservoir in this study. A 44-year time span (1971 to 2015) of monthly rainfall data of this station (Fig. 1) has been used as input data to different models employed in this study. The statistical properties of monthly precipitation data are presented in Table 1. The Ghaleh-Shahrokh hydrometric station is located upstream of the Zayandehroud dam reservoir and detailed information of this station are presented in Table 2 (Zayandab Consulting Engineers 2008).

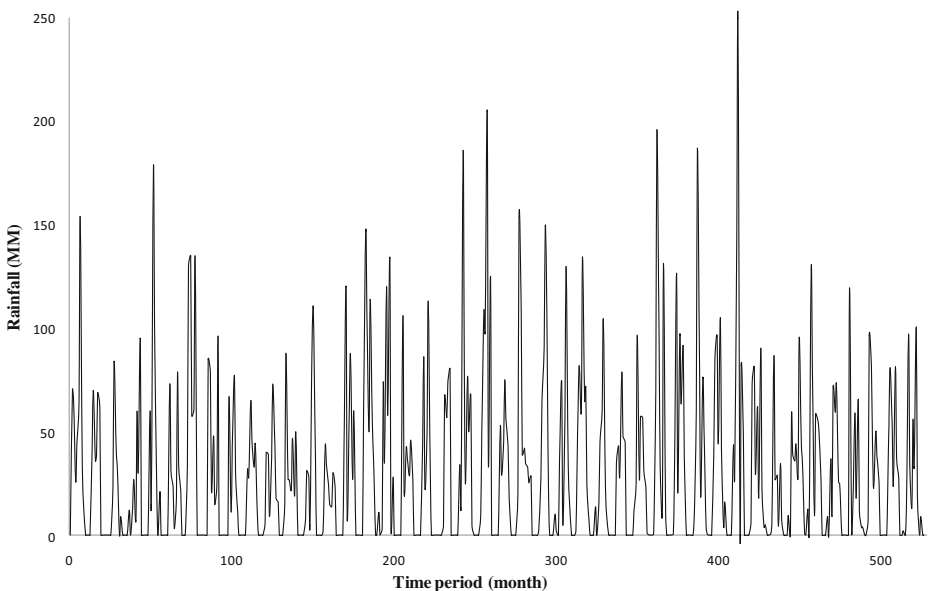


Fig. 1 Monthly rainfalls of Ghaleh-Shahrokh hydrometric stations for 44-year time period (1971 to 2015)

Table 1 Statistical properties of inflow data sets and rainfall (Zayandehroud basin)

Month	Inflow (MCM)		Rainfall (mm)	
	Mean	Standard deviation	Mean	Standard deviation
Oct	36.4	15.2832	4.2	8.3187
Nov	54.6	25.3965	50.1	40.1694
Dec	62.9	33.2028	62.3	46.4860
Jan	58.7	24.0407	54.6	43.3797
Feb	70.0	35.9520	54.8	40.8977
Mar	138.1	67.4761	61.9	43.2258
Apr	297.2	114.3445	54.9	37.7035
May	281.6	110.6383	29.5	26.6135
Jun	179.2	66.6396	3.9	8.2815
Jul	113.0	49.6230	0.4	1.6303
Aug	64.8	32.4250	0.3	1.6786
Sep	40.3	20.8265	0.9	3.0089

Zayandehroud dam is a concrete two-curved dam located on Zayandehroud River about 110 km west of Isfahan city. This dam is one of the main water supply facilities of the Zayandehroud basin which includes main reservoir, hydroelectric power plant, and regulatory dam located 4 km downstream of Zayandehroud dam (Jamali et al. 2007). Similar to the precipitation data, a 44-year time period (1971 to 2015) inflow into Zayandehroud dam reservoir has been used for modeling purposes of this study (Fig. 2). It should be noted that the total cumulative water inflow during this time period is 61,456.5 million cubic meter (MCM). The statistical properties of monthly inflow data are presented in Table 1.

3 Methodology

In this paper, ANN and SVM models have been used to predict inflow to Zayandehroud dam reservoir. By changing the combination of input data, different input patterns have been created and proposed for ANN and SVM models leading to different models. Therefore, nine different models have been proposed making use of different combinations of different variables such as water inflow to the reservoir with different lag times and rainfall data (Yazdani et al. 2009; Sattari et al. 2012; Lohani et al. 2012; He et al. 2014; Hassan et al. 2015). In the following subsections, brief descriptions of the basics of the proposed methodologies are presented.

Table 2 The specifications of Ghaleh-Shahrokh hydrometric station (Zayandab Consulting Engineers 2008)

Station code	Station name	Geographical attributes			Established year
		Longitude (degree-minute)	Latitude (degree-minute)	Elevation (m)	
42-003	Ghaleh-Shahrokh	50°28'	32°39'	2080	1349

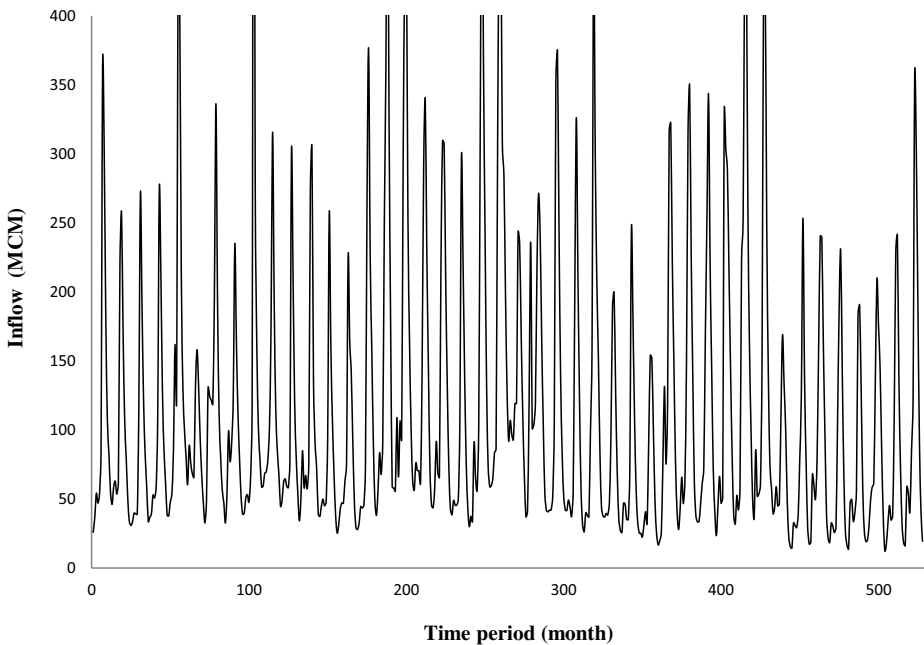


Fig. 2 Monthly inflow into Zayandehroud dam reservoir over 44-year time period (1971 to 2015)

3.1 Artificial Neural Network (ANN) Model

ANN model is an intelligent dynamic system which is based on human's brain neural network (Menhaj 2009). In recent decades, due to the ability of this model to simulate and estimate nonlinear functions with proper accuracy, ANN models have been widely used to solve a number of engineering problems including water resources engineering problems.

In general, the number of neurons, layers (consisting input, hidden and output layers), and weights should be defined to create the structure of an ANN model. The number of neurons in input and output layers is defined by the number of input and output variables, respectively. Further, the number of hidden layers and the number of neurons in each hidden layer should be determined using experimental and try and error methods to avoid under and/or over fitting problems (Valipour et al. 2013).

In this study, based on different combinations of input data, overall nine ANN models are employed to predict inflow to Zayandehroud dam reservoir. The input data of the first to the seventh models are monthly inflow of Zayandehroud dam reservoir with 1, 2, 3, 4, 5, 6 and 12 months time delays, respectively. In the eighth model, however, another input parameter called the time index (T), related the studied month, is added to predict inflow to the reservoir. The value of this parameter varies from 1 to 12 with respect to the studied month (Kisi 2008; Awchi 2014). Finally, monthly rainfalls of Ghaleh-Ghahrokh station considering 1, 2, 3, 4, 5, 6, and 12 month time lags are also added to the input data sets for the ninth ANN model. In all proposed models, the inflow discharge of the target (current) month is considered as the output data. Table 3 shows the details of proposed models for inflow prediction of Zayandehroud dam reservoir in which $Q(t)$ is discharge of month t (target month), $Q(t-1), \dots, Q(t-6), Q(t-12)$ are water discharges of the months with 1 to 6 and 12 month time lags respectively, $P(t)$ is the

Table 3 Details of proposed models for water inflow prediction of the Zayandehroud dam reservoir

Model	Input data	Output data
One	$Q(t-1)$	$Q(t)$
Two	$Q(t-1), Q(t-2)$	
Three	$Q(t-1), Q(t-2), Q(t-3)$	
Four	$Q(t-1), Q(t-2), Q(t-3), Q(t-4)$	
Five	$Q(t-1), Q(t-2), Q(t-3), Q(t-4), Q(t-5)$	
Six	$Q(t-1), Q(t-2), Q(t-3), Q(t-4), Q(t-5), Q(t-6)$	
Seven	$Q(t-1), Q(t-2), Q(t-3), Q(t-4), Q(t-5), Q(t-6), Q(t-12)$	
Eight	$Q(t-1), Q(t-2), Q(t-3), Q(t-4), Q(t-5), Q(t-6), Q(t-12), T$	
Nine	$Q(t-1), Q(t-2), Q(t-3), Q(t-4), Q(t-5), Q(t-6), Q(t-12), P(t), P(t-1), P(t-2), P(t-3), P(t-4), P(t-5), P(t-6), P(t-12), T$	

rainfall of month t (target month), $P(t-1), \dots, P(t-6), P(t-12)$ are rainfalls of months with 1 to 6 and 12 month time lags respectively, and T is the time index.

Generally, different learning algorithms have been proposed for ANN models. In ANN models, the purpose of learning is to find the best relationship between input and output parameters and therefore the ANN model should be trained using a proper learning algorithm. The Levenberg–Marquardt algorithm due to a high convergence rate has been widely used to train ANN networks (Awchi 2014) and as such was adopted to train the ANN models in this research.

In this study, a multi-layer perceptron network with an error-back propagation algorithm has been used with one input layer, one hidden layer, and one output layer. The reason for choosing this three-layer structure is the high capability of this network to estimate complex relationships (Coulibali 1999). In addition, the transition functions for hidden and output layers are sigmoid tangent and linear, respectively. Figure 3 shows the general structure of an ANN model (ninth model).

To implement the proposed models, all study data were divided into three sets of training, validation, and test data sets. Generally, the “training data” refers to the data used for model training and the model parameters are adjusted according to the errors obtained from these data. In addition, training data are used to measure model generality. It is worth noting that the evaluation error during the training process should be reduced. However, the evaluation error increases when the network tries to overlap with the observational data. When the evaluation error increases for a certain number of iterations, the training process is stopped and the weights and bias values are matched with the values at time of minimal error. Finally, the third subset of data (test data) is used to evaluate and comparing different models (Coulibaly et al. 2000). Here, 70% of the data (362 months) have been considered as training data sets, 15% (77 months) have been considered as validation data set, and finally 15% of the remaining data (77 months) have been considered as the test data. It should be noted that all data should be normalized (between 0 and 1) to standardize different values of input and output data sets (Noori et al. 2011b).

Two evaluation criteria including root mean squared error (RMSE) and correlation coefficient (R) are commonly used to evaluate and compare the performance of different models and as such are used to select the best model for inflow prediction in this study. It is worth noting that a smaller value of RMSE (close to zero) and a larger value of R (close to one) indicate better performance of the employed model.

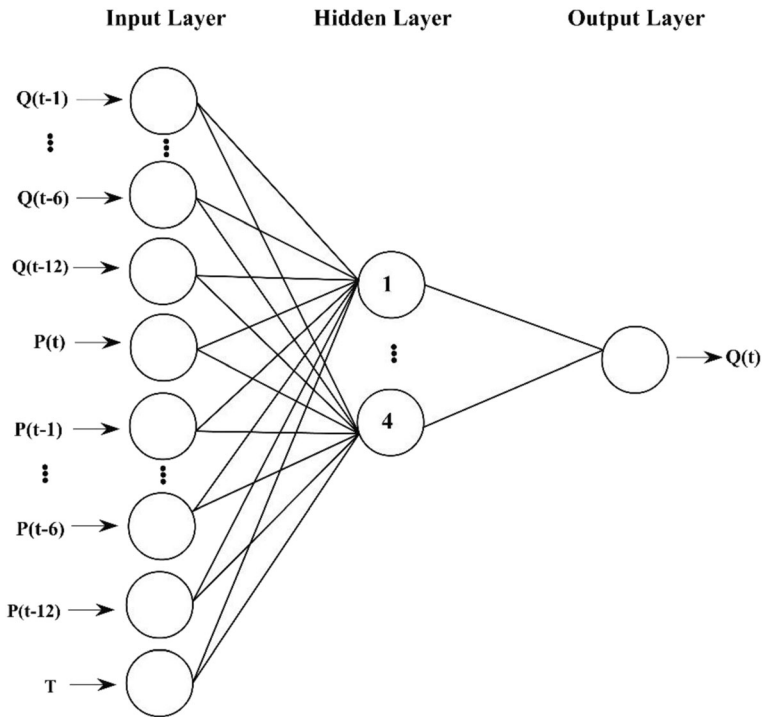


Fig. 3 General structure of the ANN model (the ninth model)

3.2 Support Vector Machine (SVM) Model

Support vector machine (SVM) is an intelligent model that is based on statistical learning theory. Generally, it is a well-suited method for regression, classification, and time series prediction. The performance of this model has been found to be acceptable in hydrological studies (Vapnik 1995, 1998).

In general, this model creates the learning function for inputs and outputs $((x_1, y_1), \dots, (x_l, y_l))$. In other words, in SVM model, it is necessary to estimate the dependence function of the dependent variable (y) based on the independent variable (x). In basic SVM model, the linear function has been defined to solve a problem in a linear regression case and therefore a convex optimization problem can be defined (Smola and Scholkopf 1998). In other words, approximated function $f(x)$ can be calculated using all pairs (x_i, y_i) with the minimum precision (ε) and loss function of Vapnik that is named ε -insensitive loss function (Vapnik 1995).

This proposed formulation is solved using Lagrange method and weight vector (w) can be defined as a combination of weighted training data of x_i . Finally, the multipliers value can be determined solving the optimization problem and therefore the estimated function can be defined as follows:

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) \langle x_i, x \rangle + b \quad (1)$$

Where, \langle, \rangle denotes the dot product of the two vectors, x is the problem variables vector, b is the additive noise (bias), α_i and α_i^* are dual variables, and other parameters are as defined before.

This equation is the so-called support vector expansion that is used for linear regression. In this model, the proposed method is completely defined by the internal multiplication of the input vectors. Therefore, computing function $f(x)$ does not require the explicit computation of W .

Generally, most of the engineering problems are nonlinear and therefore using linear regression is not suitable for solving such problems. Therefore, the proposed method should be extended to solve nonlinear problems using kernel functions. Generally, different kernel functions, such as linear kernel, polynomial kernel, hyperbolic tangent kernel, and Gaussian radial base function (RBF) have been proposed and used in SVM nonlinear modeling studies (Lin et al. 2006; Dibike et al. 2001; Yu et al. 2006). The commonly used kernel functions are listed in Table 4.

Generally, choosing the proper kernel function is an important issue for solving a problem using a SVM model (Dibike et al. 2001; Asefa et al. 2004; Liong and Sivapragasam 2002; Yu et al. 2006; Kisi and Cimen 2011). Because of the better performance of RBF kernel function in comparison with other kernel functions for hydrological predictions (Yu et al. 2006; Kisi and Cimen 2011), this function has been used in this research which is defined as follows:

$$K(x, x_i) = \exp\left(-\gamma\|x-x_i\|^2\right) \tag{2}$$

Where, γ is Gaussian kernel function parameter and other parameters were defined before. It is worth noting that changing γ has a significant effect on model training and the error rate of the problem and therefore the value of this parameter should be determined carefully for different conditions of the problem. Finally, by replacing the kernel function, the $f(x)$ function can be defined as follows:

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(x_i, x) + b \tag{3}$$

Accuracy in estimating C , ε , and γ parameters is very important and has significant effect on reducing the error estimation of the problem (Smola and Scholkopf 1998). Therefore, determining optimal parameters for SVM model is an important step of designing this model which significantly affects the model performance and therefore it is essential for predicting model output results.

In recent years, a variety of software has been developed to implement the SVM model. In this study, the LIBSVM software has been used due to the ease of application and making use of new and up-to-date methods for implementation. Two types of SVM models have been defined in this software named ν -SVM (type one) and ε -SVM (type two) model (Noori et al. 2011a). Generally, ε -SVM is more applicable and therefore it has been used in this study. Here, according to the type of kernel function used in this research, the proper values of C , ε and γ have been determined using a trial and error method. Using this procedure, the best values of C , ε and γ have been found to be 15, 0.039 and 0.05, respectively.

Table 4 The commonly used kernel functions

Kernel type	Function	Parameters
Linear	$\langle x, x_i \rangle$	—
Polynomial	$(\langle x, x_i \rangle + 1)^d$	d
Gaussian radial base function	$\exp(-\ x - x_i\ ^2/2\sigma^2)$	σ
Hyperbolic tangent	$\tanh(b\langle x, x_i \rangle + c)$	b, c

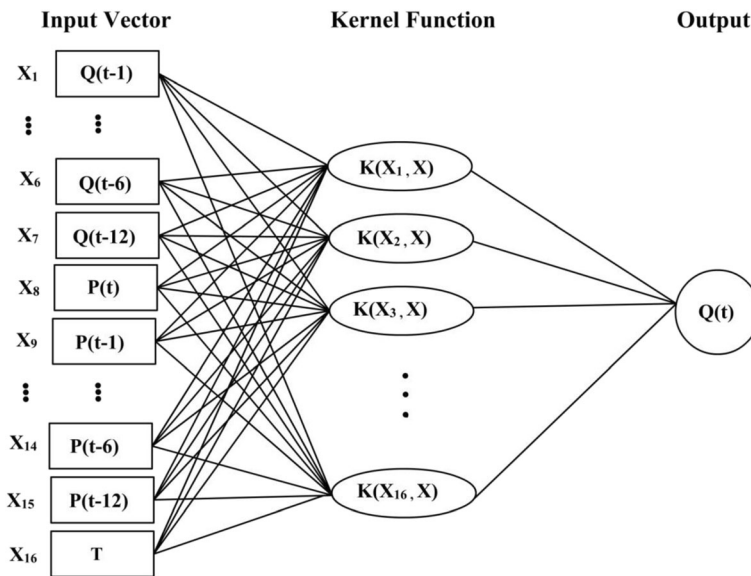


Fig. 4 General structure of the SVM model

Here, the input data of ANN model have been used to create different SVM models. Based on the obtained results, using the input data of the best ANN model (ninth model) leads to better results and therefore this input data have been used to create a SVM model in this paper. In other words, the monthly inflow of Zayandehroud dam reservoir with 1, 2, 3, 4, 5, 6 and 12 months time delays, the time index (T), and the monthly rainfalls of Ghaleh-Ghahrokh station with 1, 2, 3, 4, 5, 6, and 12 months time lags have been considered as input data for the SVM model. In addition, the inflow discharge of the target (current) month has been considered as the output data. In this model, 70% (362 months), 15% (77 months), and 15% (77 months) of data have been considered as training, validation, and test data sets. It is worth noting that in SVM the model parameters are selected in such a way that the training error rate is minimized and therefore the trial and error method has been used here for this purpose. Furthermore, to reduce the complexity of the problem and standardize different values of data sets, the input and output data should be normalized (between 0 and 1). Figure 4 shows the general structure of an SVM model.

Table 5 The best results of proposed ANN models for inflow prediction of the Zayandehroud dam reservoir

Model	Number of hidden layer neurons	Training process		Validation process		Test process	
		RMSE	R	RMSE	R	RMSE	R
One	15	74.9521	0.71325	78.3471	0.73508	49.9695	0.78413
Two	20	66.9964	0.79504	66.1758	0.82716	43.8354	0.88629
Three	13	63.937	0.80151	65.46.04	0.82273	40.3311	0.89045
Four	7	63.258	0.80688	63.4783	0.83489	37.632	0.8969
Five	13	62.5162	0.8111	62.2974	0.84402	36.3041	0.89811
Six	19	59.8496	0.82944	60.5514	0.85561	35.5529	0.91814
Seven	19	52.1017	0.87323	56.3514	0.87292	29.089	0.93733
Eight	5	50.3716	0.88197	49.8228	0.90262	28.8021	0.95013
Nine	4	48.5441	0.89269	43.748	0.92983	28.5125	0.95333

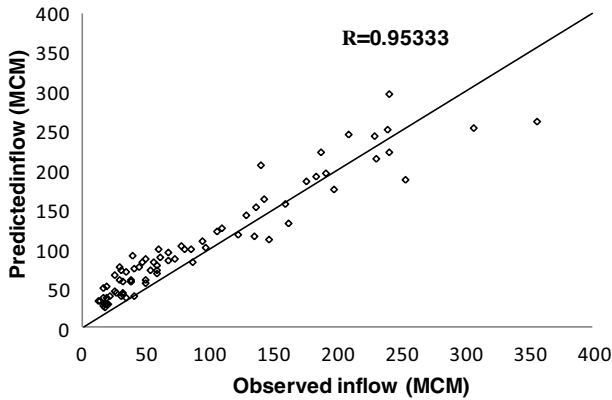


Fig. 5 Results of ninth ANN model for test data

4 Results and Discussion

The performance of nine proposed models has been studied for predicting inflow of Zayandehroud dam reservoir using ANN and SVM methods. For this purpose, first, the nine proposed ANN models have been evaluated and then the best pattern has been selected to be used in the SVM model.

As the first step, proposed ANN models have been used to predict the inflow to Zayandehroud dam reservoir and the results have been presented and compared. Table 5 shows the best results of various proposed models along with the corresponding R and RMSE values for training, validation, and test data sets. This table also presents the best number of hidden layer neurons in each case. Comparison of the results shows that, in the first seven models, by adding each time duration lag of inflow as input data, the R and RSME values of training, validation and test processes are improved. Due to this fact, using the seventh model leads to the best results in comparison with the first to sixth models. Further, adding time index

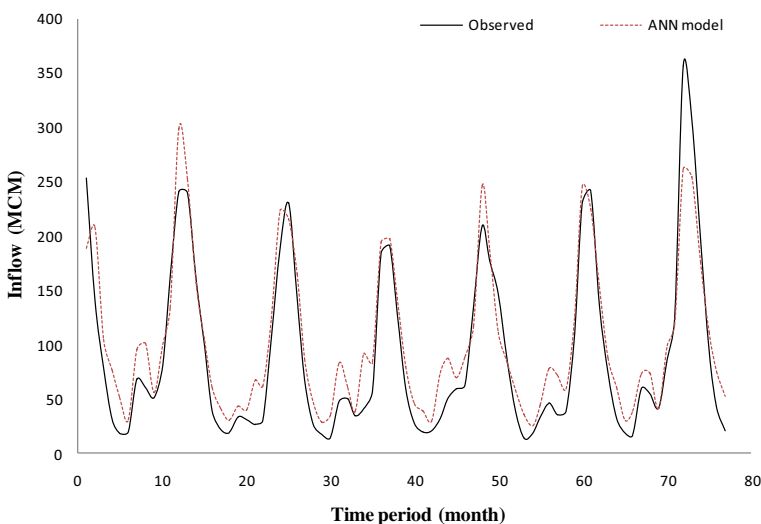


Fig. 6 Comparison of the best results of inflow prediction using the ninth ANN model with observed data

Table 6 The best results of proposed SVM model for inflow prediction of the Zayandehroud dam reservoir

Model	Training process		Validation process		Test process	
	RMSE	R	RMSE	R	RMSE	R
Nine	47.9346	0.8962	42.69093	0.9303	23.56193	0.9622

(T) in the eighth model has led to a significant increase in R and considerable decrease in RSME values in comparison with the seventh model in all training, validation, and test processes. Finally, in the ninth model, by adding the monthly rainfall of Ghaleh-Ghahrokh station with different time lags as input data, the best results have been obtained in all training, validation and test processes in comparison with other proposed models (Table 5). In other words, the minimum error has been obtained using the ninth model with R values of 0.89296, 0.92983 and 0.95333 and RSME values of 48.5441, 43.748 and 28.5125 for training, validation, and test processes, respectively. In other words, the RMSE value of ninth ANN model in test process has decreased by 42.94, 34.96, 29.31, 24.23, 21.46, 19.80, 1.98, and 1.01% compared to the results of the first, second, third, fourth, fifth, sixth, seventh, and the eighth ANN models, respectively. In addition, the R value of ninth ANN model in test process has increased by 21.58, 7.56, 7.06, 6.29, 6.15, 3.83, 1.71, and 0.34% compared to the results of the first, second, third, fourth, fifth, sixth, seventh, and the eighth ANN models, respectively.

The best results of the ninth ANN model and the corresponding R values in test process have been presented in Fig. 5. In addition, the predicted values of inflow into the Zayandehroud dam reservoir using the ninth ANN model in the test process along with actual observed inflow have been presented in Fig. 6. Comparison of obtained results of all nine proposed ANN models show that using the ninth proposed ANN model leads to the best results. In other words, the least error for predicting inflow to the reservoir of Zayandehroud dam has been obtained using the ninth ANN model. Since the best results have been obtained by the ninth ANN model, the input data pattern of this model has been used in the SVM model to predict inflow to the Zayandehroud dam reservoir. Table 6 shows the best results of the ninth proposed SVM model along with the corresponding R and RMSE values for training, validation and test data sets. Comparison of the results shows that acceptable values have been obtained using the proposed SVM Model. Finally, the best obtained results of ANN and SVM models to predict inflow into the Zayandehroud dam reservoir have been presented and compared in Table 7. These results indicate that the obtained results of SVM models are better than all proposed ANN models in all training, validation, and test processes. In other words, the results of SVM model have less error than all ANN models in all training, validation and test processes. That is, the RMSE values of SVM model have reduced by 1.26, 2.42, and 17.36% in comparison with the results of the ninth ANN model at training, validation and test

Table 7 Comparison the best results of ANN and SVM models for inflow prediction of the Zayandehroud dam reservoir

Model	Training process		Validation process		Test process	
	RMSE	R	RMSE	R	RMSE	R
SVM	47.9346	0.8962	42.69093	0.9303	23.56193	0.9622
ANN	48.5441	0.89269	43.748	0.92983	28.5125	0.95333

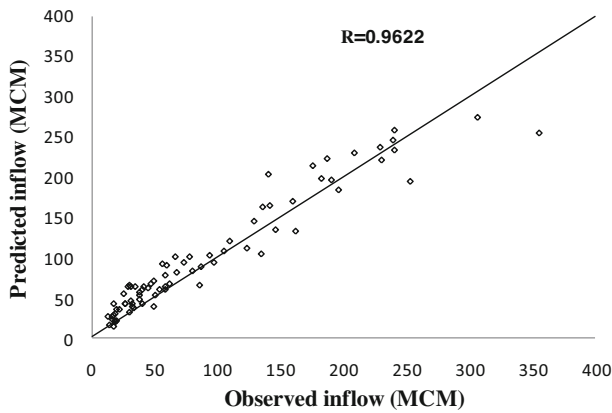


Fig. 7 Results of the SVM model for test data

processes, respectively. In addition, the R value of SVM model has increased by 0.39, 0.05, and 0.93% in comparison with the results of the ninth ANN model at training, validation, and test processes, respectively. These results indicate the capability of SVM model to predict the inflow to Zayandehroud dam reservoir with acceptable accuracy. In other words, the performance of the SVM model for inflow prediction is superior to that of ANN model. Based on this finding, one may conclude that the SVM model might be a useful tool to be used in prediction problems in water resources engineering as well as other fields of science and engineering.

Finally, the best results of the proposed SVM model for inflow prediction and the related R values in test processes have been presented in Fig. 7. In addition, the predicted values of inflow into the Zayandehroud dam reservoir using the SVM model in the test process in comparison with the actual observed inflow have been presented in Fig. 8. Further, the best predicted values of inflow into the reservoir using the ANN and the SVM models and the respected observed inflow have been presented and compared in Fig. 9. Comparison of the values presented in this figure also indicates the superior performance of the SVM model compared to the ANN model.

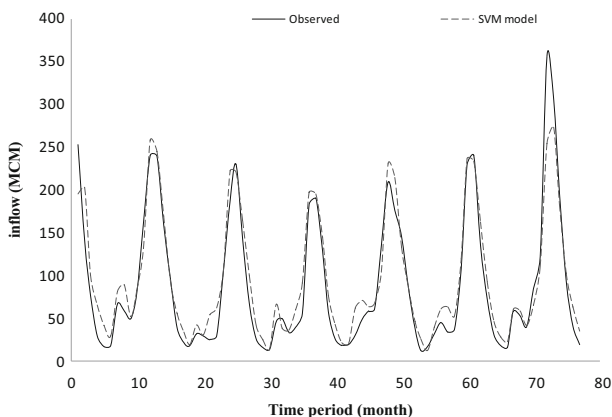


Fig. 8 Comparison of the best results of inflow prediction using the SVM model with the observed data

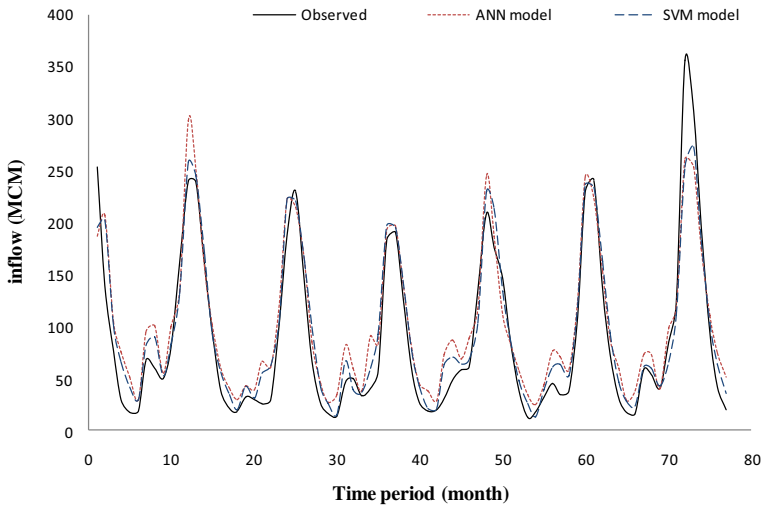


Fig. 9 Comparison of the best results of inflow prediction using the ninth ANN and SVM model with the observed data

5 Concluding Remarks

In this research, ANN and SVM data-driven models were used for inflow prediction of Zayandehroud dam reservoir employing nine different input data patterns. At first, nine ANN models were utilized for inflow prediction using proposed input data patterns. In the first to the seventh proposed models, the monthly inflow into the reservoir were taken as input data considering 1, 2, 3, 4, 5, 6 and 12 monthly time lags, respectively. However, in the eighth model, another input parameter called the time index (T) was also added as the supplementary input data to the model. Finally, in the ninth model, the monthly rainfall of Ghaleh-Ghahrokh station were also added to the input data considering 1, 2, 3, 4, 5, 6, and 12 monthly time lags. Comparison of the results showed that feeding the ANN model with the ninth proposed input data pattern led to the least error for inflow prediction of Zayandehroud dam reservoir. Further, the input data pattern of the best ANN model (ninth model) were used to feed the SVM model for inflow prediction of Zayandehroud dam reservoir and the obtained results were compared to those of ANN models. Comparison of the results showed that the proposed SVM model has more accurate results compared to that of ANN model for inflow prediction of Zayandehroud dam reservoir and therefore it can be considered as a superior alternative for prediction purposes in water resources engineering as well as other engineering and science fields.

Compliance with Ethical Standards

Conflict of Interest None.

References

- Asefa T, Kemblowski MW, Urroz G, McKee M, Khalil A (2004) Support vectors-based groundwater head observation networks design. *Water Resour Res* 40(11):W11509

- Asefa T, Kemblowski MW, McKee M, Khalil A (2006) Multi-time scale stream flow predictions: the support vector machines approach. *J Hydrol* 318:7–16
- Awchi TA (2014) River discharges forecasting in northern Iraq using different ANN techniques. *Water Resour Manag* 28(3):801–814
- Coulibali CG (1999) Daily stream flow forecasting: application of ANN. *J Hydrol* 3:123–128
- Coulibaly P, Ancitl F, Bobee B (2000) Daily reservoir inflow forecasting using artificial neural networks with stopped training approach. *J Hydrol* 230(3):244–257
- Dibike YB, Yelickov S, Solomatine DP, Abbott MB (2001) Model induction with support vector machines: introduction and application. *Journal of Computing in Civil Engineering Management* 15(3):208–216
- Hassan M, Shamim MA, Hashmi HN, Ashiq SZ, Ahmed I, Pasha GA, Naeem UA, Ghumman AR, Han D (2015) Predicting streamflows to a multipurpose reservoir using artificial neural networks and regression techniques. *Earth Sci Inf* 8(2):337–352
- 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. *J Hydrol* 509:379–386
- Jamali S, Abrishamchi A, Tajrishi M (2007) River stream-flow and Zayanderoud reservoir operation modeling using the fuzzy inference system. *Journal of Water and Wastewater* 18(4):25–34 (in Persian)
- Kalteh AM (2015) Wavelet genetic algorithm-support vector regression (wavelet GA-SVR) for monthly flow forecasting. *Water Resour Manag* 29(4):1283–1293
- Kisi O (2008) River flow forecasting and estimation using different artificial neural network techniques. *Hydrol Res* 39(1):27–40
- Kisi O, Cimen M (2011) A wavelet-support vector machine conjunction model for monthly streamflow forecasting. *J Hydrol* 399:132–140
- Lima LM, Popova E, Damien P (2014) Modeling and forecasting of Brazilian reservoir inflows via dynamic linear models. *Int J Forecast* 30(3):464–476
- Lin J-Y, Cheng C-T, Chau K-W (2006) Using support vector machines for long-term discharge prediction. *Hydrol Sci J* 51(4):599–612
- Liong S, Sivapragasam C (2002) Flood stage forecasting with support vector machines. *J Am Water Resour Assoc* 38:173–186
- Liu Z, Zhou P, Zhang Y (2014) A probabilistic wavelet-support vector regression model for stream flow forecasting with rainfall and climate information input. *J Hydrometeorol* 16:2209–2229
- Liu J, Yan K, Zhao X, Hu Y (2016) Prediction of autogenous shrinkage of concretes by support vector machine. *International Journal of Pavement Research and Technology* 9(3):169–177
- Lohani AK, Kumar R, Singh RD (2012) Hydrological time series modeling: a comparison between adaptive neuro-fuzzy, neural network and autoregressive techniques. *J Hydrol* 442:443:23–35
- Menhaj MB (2009) Computational intelligence. AmirKabir University of Technology (in Persian)
- Noori R, Khakpour AR, Omidvar B, Farokhnia A (2010) Comparison of ANN and principal component analysis-multivariate linear regression models for predicting the river flow based on developed discrepancy ratio statistic. *Expert Syst Appl* 37(8):5856–5862
- Noori R, Karbassi AR, Moghaddamnia A, Han D, Zokaei-Ashtiani MH, Farokhnia A, Ghafari Gousheh M (2011a) Assessment of input variables determination on the SVM model performance using PCA, gamma test and forward selection techniques for monthly stream flow prediction. *J Hydrol* 401:177–189
- Noori R, Karbassi AR, Mehdizadeh HM, Vesali-Nasch M, Sabahi MS (2011b) A framework development for predicting the longitudinal dispersion coefficient in natural streams using an artificial neural network. *Environ Prog Sustain Energy* 30(3):439–449
- Sattari MT, Yurekli K, Pal M (2012) Performance evaluation of artificial neural network approaches in forecasting reservoir inflow. *Appl Math Model* 36(6):2649–2657
- Smola AJ, Scholkopf B (1998) A tutorial on support vector regression. *NeuroCOLT2 Technical Report Series. NC2-TR-1998-030*
- Valipour M, Banihabib ME, Behbahani SMR (2013) Comparison of the ARMA, ARIMA, and the autoregressive artificial neural. *J Hydrol* 476:433–441
- Vapnik V (1995) The nature of statistical learning theory. Springer-Verlag, New York
- Vapnik V (1998) Statistical learning theory. Wiley, New York
- Wenjian W, Changqian M, Weizhen L (2008) Online prediction model based on support vector machine. *Neurocomputing* 71(5):550–558
- Yazdani MR, Saghafian B, Mahdian MH, Soltani S (2009) Monthly runoff estimation using artificial neural networks. *J Agric Sci Technol* 11(3):355–362
- Yu PS, Chen ST, Chang IF (2006) Support vector regression for real-time flood stage forecasting. *J Hydrol* 328: 704–716

- ZayandAb Consulting Engineers (2008) Studies on resources and water consumption for Zayandehroud Basin. Isfahan Regional Water Company
- Zhu S, Zhou J, Ye L, Meng C (2016) Stream flow estimation by support vector machine coupled with different methods of time series decomposition in the upper reaches of Yangtze River, China. *Environ Earth Sci* 75(531):1–12

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.