Soft Computing
https://doi.org/10.1007/s00500-018-3528-8

METHODOLOGIES AND APPLICATION

Prediction of the intermediate block displacement of the dam crest using artificial neural network and support vector regression models

Mahmoud Mohammad Rezapour Tabari[1] · Hamed Reza Zarif Sanayei[1]

Abstract

Concrete arch dams are three-dimensional structures which are statically indeterminate due to integrity and arching performance. Hence, the spatial and temporal temperature gradients in concrete arch dams affect the volume of the structures and generated internal stresses threaten stability of the structures. Accordingly, estimation of long-term thermal behavior of these structures for proper serviceability with considering dam crest displacement is necessary, and this issue requires the application of appropriate prediction models. The goal of this study is to implement the support vector regression (SVR) and artificial neural network (ANN) models for prediction of the intermediate block displacement of the dam crest. For this purpose, displacement of dam crest is investigated with ABAQUS simulation model over a period of 8 years, and then, the results of the simulation are used in the soft models (SVR and ANN) as the input data. The analysis of the results of two models with five error indicators shows that the error reduction in the SVR model is about 32% less than the ANN model in the testing stage. Also, investigation of the normal cumulative probability distribution related to the outputs of two models indicates high degree of deviation on cumulative probability distribution of the ANN model. This is due to the fact that the ANN model ignores fundamental errors in the training process. Therefore, based on the SVR model one can predict the dam stability in an acceptable accuracy range, only by measuring two different parameters including reservoir water level and the air temperature.

Keywords Concrete arch dams · Predicting displacement behavior · Dam crest · Support vector regression · Artificial neural network

1 Introduction

Concrete arch dams are three-dimensional structures which are statically indeterminate due to integrity and arching performance. Hence, the spatial and temporal temperature gradients in concrete arch dams affect the volume of the structures and generated internal stresses threaten stability of the structures. As a result, stresses greater than tensile strength of concrete induce cracks in the external surface of the dam. In concrete arch dams, these cracks usually have direct effects on the sliding and overturning stabilities of dam.

Communicated by V. Loia.

✉ Mahmoud Mohammad Rezapour Tabari
mrtabari@sku.ac.ir

Hamed Reza Zarif Sanayei
zarif@sku.ac.ir

1 Department of Engineering, Shahrekord University, Shahrekord, Iran

Increasing of cracks can be critical when thermal loads are combined with hydrodynamical and hydrostatical loads of water and thawing–freezing cycle. Therefore, it endangers the overall stability of the dam (Agullo et al. 1991). This subject shows the importance of determination of the thermal stresses for safety evaluation of the structures. Since the effect of temperature variations on the structures is observed in all design, construction, and operation phases for dam safety evaluation, the analysis of arch dams is of great importance (US Army Corps of Engineering 1994). Due to permanent behavior of thermal reactions, it has significant effects on the useful life and thermal behavior of the dam at the time of operation.

In concrete arch dams, the reservoir water level is low in relatively long time period. Consequently, more area of external surface of dam is affected by direct solar radiation, and thus, the time phase difference between internal temperature of concrete and the air temperature would decrease. So, irregular behavior of dam is expected due to these phenom-

Published online: 12 September 2018

ena (Bofang 2014). With regard to the drought and water shortage conditions in Iran in recent years, the water level in many dams is sometimes lower than the minimum operational level and the thermal cracking resulted from water drawdown in dam body can be observed. Accordingly, the necessity of management and assessment of arch dam thermal behavior for long-term operation considering concrete dam crest displacement is clear. Therefore, application of accurate and reliable predictive models seems essential in this field.

In recent years, in information processing of problems without any explicit solution, implementation of intelligent systems, such as artificial neural network (ANN) and support vector regression (SVR), is widely increased. These soft models have been applied to solve miscellaneous water resource issues such as forecasting water levels and water quality in rivers, dams, and groundwater (Yoon et al. 2011; Hipni et al. 2013; He et al. 2014; Seo et al. 2016; Hosseini and Mahjouri 2016; Humphrey et al. 2016; Barzegar et al. 2017; Csábrági et al. 2017; Ebrahimi and Rajaee 2017).

In water resources management studies, the ANN was used since French et al. (1992) and then these soft models were used increasingly in water sciences.

The SVR is a relatively new and powerful tool, providing solution for classification and regression based on statistical learning theory. This model was first introduced by Vapnik (1995). Researches show that the SVR has the capability and good performance, even when little data are available (Vapnik 1998). Although many applications of ANN and SVR models were presented in various fields, only a few applications were observed in design of hydraulic structures.

The soft models have been applied in different subjects, such as dam engineering. Some studies related to this topic can be referred to as follows: Gaziev (2000), Fedele et al. (2006), Cao et al. (2009), Karimi et al. (2010), Guang-yong et al. (2011), Popescu (2012), Zhou et al. (2015), Salazar et al. (2015), Fisher et al. (2016), Stojanovic et al. (2016), and Saqib and Ansari (2017).

Mata (2011) evaluated the performance of multiple linear regression and artificial neural network models for safety control of concrete dams under various environmental loads (hydrostatic pressure and temperature gradients). Kao and Loh (2013) used the static neural networks (SNNs), the nonlinear autoregressive with exogenous neural network (NARXNN), the principal component analysis (PCA), and the auto-associate network (AAN) models for health monitoring of Fei-Tsui dam. Behnia et al. (2013) estimated the subsidence of dams (crest settlement) at the design stage using adaptive neuro-fuzzy inference system and gene expression programming intelligent methods.

Rankovic et al. (2014) predicted tangential displacement of a concrete dam based on support vector regression (SVR). The results showed that the SVR model provides more accurate results than the other models.

The literature review reveals that the hybrid of ABAQUS simulation model and soft-computing tools is not considered in predicting the displacement of intermediate block of the dam crest. Due to the complexity of the parameters that need to be adjusted in the ABAQUS model, it is not possible to directly communicate the results of the ABAQUS model with management models of the dam operation. Therefore, according to high efficiency of black-box simulation models in predicting the structural response to thermal stresses of dams, the development of these models is essential as an interface between the ABAQUS model and the optimal operation models of the dam reservoirs. It should be noted that in this study, these interface models as thermal behavior simulator of the dam determine the displacement of intermediate block of the dam crest under different hydrological conditions. Under these conditions, the dam reservoir is placed under different thermal stress and water level due to the amount of water stored in it.

Hence, the objective of this study is to evaluate the possibility of using ANN and SVR models as a substitution to the finite element method for predicting the displacement of intermediate block of the dam crest. These soft models are able to establish complex nonlinear relationships between input and output parameters. In order to apply the proposed approach and study the effects of using these soft models, Karaj dam in Iran is considered as a case study for prediction of the influences of variations in water level and temperature conditions of dam reservoir on the displacement of intermediate block of the dam crest. The major contributions of this study are as follows:

- Development of a two data-based forecasting models for predicting short-term nonlinear fluctuations of intermediate block displacement of the dam crest in a strategic dam reservoir due to variations in water level and temperature conditions using measured data.
- Comparison of the efficiency and accuracy of the ANN (with different training algorithms) and SVR (with miscellaneous kernel function) models with different error criteria in the forecasting performance.
- Proposal of best data-driven model as an interface model for forecasting intermediate block displacement of the dam crest for using in water resource management optimization models.

2 Study area and data

Karaj dam, located on 63 km northwest of Tehran, with watershed area of 764 km^2 was constructed on Karaj river. The average annual inflow to Karaj dam is 472 MCM (million

Table 1 General characteristics of Karaj concrete arch dam

Maximum height above foundation	168 m
Crest length	384 m
Foundation altitude	1606 masl
Crest altitude	1768 masl
Buttress width	32 m
Crest width	7.85 m
Reservoir normal capacity	203 MCM
Reservoir minimum capacity	33 MCM
Normal water elevation	1765 masl
Minimum water elevation	1692 masl

cubic meter). The purposes of Karaj dam construction were as follows: controlling spring flood and preventing damage caused by flood, annual supply of Tehran drinking water up to 340 MCM, regulating water for irrigation demands, agriculture development in farm lands near Karaj to 130 MCM, and producing hydroelectric power to help the national electricity network, especially during consumption peak hours to an annual rate of 150,000 MWh. According to the importance of this dam for water supply of capital of Iran, Karaj dam is used as a case study to investigate structural response of the dam body to exerted thermal stresses and to predict the displacement of intermediate block of the dam crest. Table 1 shows general characteristics of Karaj dam.

Karaj concrete dam is a double curvature dam, whose horizontal arches are single-centered circular (without any filet with constant thickness) and vertical arches are parabolic. The centers of the internal and the external arches of dam are coincident, and dam is perfectly symmetrical with regard to the axis passing through the center of the arches. This symmetry is proportional to the design, calculation, and construction of the radius of curvature and is presented as one of the special characteristics of the dam.

In this study, the ABAQUS software is used for modeling the dam, valley, and foundation. The first step of modeling is identifying the required data. Since thermal gradient should be considered in modeling climate effects, the developed model in ABAQUS is based on the radiation and the convection. The required data include: the value of the reflection coefficient of the surrounding land, the water level reflection coefficient, annual mean of wind speed, Boltzmann constant, absolute temperature, and the convection coefficient. Since the arch dams, in terms of geometrical shape, are considered as large shell structures, the use of solid elements in modeling is recommended (Labibzadeh and Khajehdezfuly 2010). In order to define thickness of the dam, especially in the foundation level, the elements with solid thickness are used. The block elements are used for modeling the body of dam because of similarity of foundation and body elements, and increase in the thickness of the dam body at levels near the foundation. Four-sided cubic elements, available in the program, are used. The attenuation model of the concrete is considered using the concrete damage plasticity (CDP) criteria with coefficient of 3%. This criterion, for the concrete material, consists of three parts: the required parameters of the plasticity behavior of the concrete, the pressure–impairment parameters of the concrete and the tensile–damage parameters of the concrete. The pressure–impairment and the tensile–damage parameters have important roles in analyzing the real behavior of concrete. Therefore, determining their value is of paramount importance. The impairment of the concrete begins with the nonlinear behavior of the materials. If there is not any impairment, the values of pressure–impairment and tensile–damage parameters are zero. For the case of full impairment, the values of these parameters reach to 1 or 100%. By assuming the value of 0.2 for Poisson coefficient, elasticity properties are defined as well. The total number of solid elements in the model of the dam body is 364. The solid element has 20 nodes, and each node has six degrees of freedom including freedom of the motion in three direction of x, y, and z as well as the freedom of rotation around axes of x, y, and z. This element is capable of accepting large deformations. The interaction can be modeled by this element too. This element is used for the static (the forces resulted from water), the dynamic (the effect of earthquake), and the complex linear and nonlinear analysis. It can be also used when the model consists of fluid. By the implementation of the ABAQUS simulation model for different modes of temperature and reservoir water level, the displacement of intermediate block of the dam crest is predicted. Input and output time series data used in this study are presented in Figs. 1 and 2. Also, Table 2 shows the statistical properties of each of input (temperature and reservoir water level) and output (displacement of dam crest) parameters.

3 Methodology

Long-term thermal behavior of arch dams for proper serviceability with considering dam crest displacement is necessary. This issue requires the application of appropriate prediction models. In this study, two soft models (artificial neural networks and support vector regression) are used for predicting displacement of the intermediate block of dam crest. For this purpose, displacement of the dam crest is simulated over a period of 8 years (2002–2009) using ABAQUS simulation model under different water levels of the reservoir using monthly temperature average as input data. Data generated by ABAQUS simulation model are considered as soft models output. Then, based on the produced input–output data set, different structures are considered for each of the soft models and their performance is evaluated.

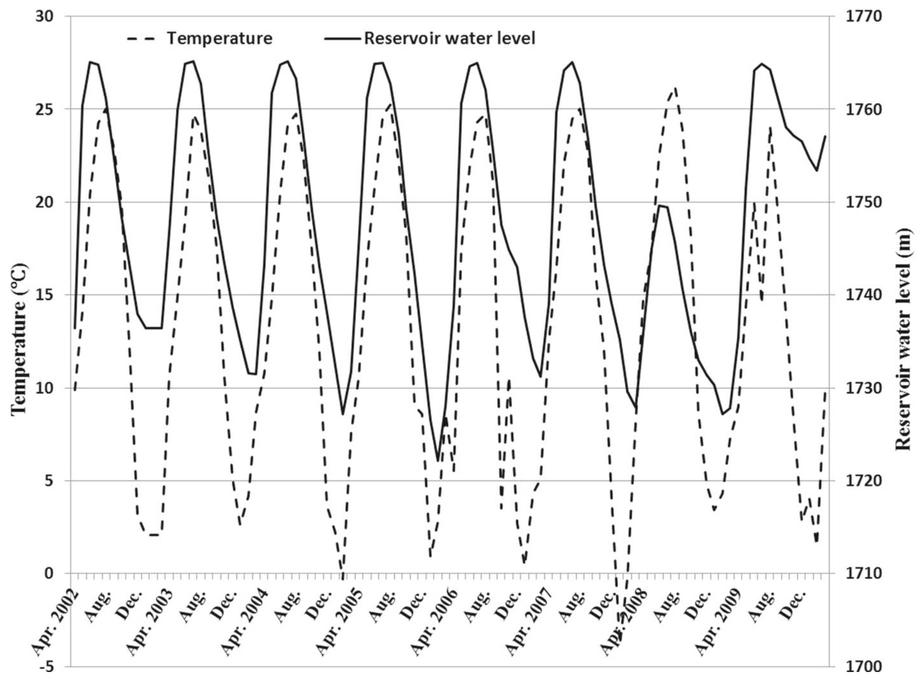


Fig. 1 Input time series data during the study period

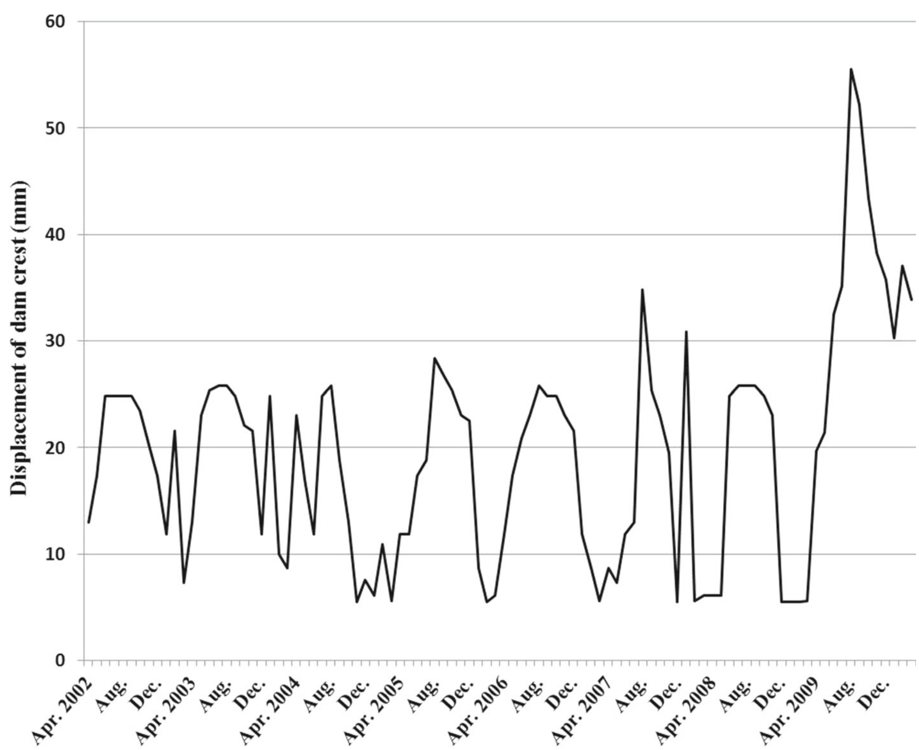


Fig. 2 Output time series data during the study period

3.1 Artificial neural networks

ANNs are developed using the function of the brain, the nervous system, and mathematical methods for determination of pattern of connection between input and output nodes (Senthil kumar et al. 2013; Tayfur et al. 2014). This soft model is an effective tool for simulating, predicting, and forecasting water resource variables. In ANNs, the multilayer perceptron (MLP) is the most widely used and is defined by three types of layers including input neurons, output neurons, and hidden neurons. The back-propagation algorithm (BPA) can effectively train the network for complex and nonlinear systems in the training stage. The BPA is basically a gradient descent-based optimization method described by Rumelhart and McClelland (1986). This method was used in many researches such as Mohanty et al. (2010); Khashei-

Table 2 The statistical parameters of input and output data used in simulation models

Parameter	The statistical characteristics	Year							
		2002	2003	2004	2005	2006	2007	2008	2009
Temperature (°C)	Max.	25.02	24.74	24.76	25.23	24.74	25	26.21	24.1
	Min.	2.1	2.6	− 0.28	0.9	0.39	− 3.78	3.44	1.48
	SD	9.22	7.75	8.76	8.36	9.38	9.68	8.72	7.18
Reservoir water level (m)	Max.	1765.1	1765.2	1765.1	1764.9	1764.9	1765.1	1749.6	1764.9
	Min.	1736.4	1731.5	1727.2	1722.2	1731.2	1727.9	1727.2	1735.5
	SD	12.1	12.8	14	15.9	12.44	13.73	8.2	7.97
Displacement of dam crest (mm)	Max.	24.88	25.83	25.83	28.4	25.83	34.83	25.83	55.52
	Min.	7.3	8.7	5.5	5.5	5.6	5.5	5.5	19.68
	SD	5.98	6.76	7.59	8.2	7.1	10.42	10.13	10.58

Siuki and Sarbazi (2013); Emamgholizadeh et al. (2015). The mathematical expression of the MLP with one output is as follows:

$$z_j = f\left(\sum_{i=1}^{N} W_{ij}X_i + b_j\right), \quad j = 1, 2, 3, \ldots, M \tag{1}$$

$$y = g\left(\sum_{j=1}^{M} W'_j z_j + b'\right) \tag{2}$$

where X_i is the input value to the ith neuron of input layer, z_j is the output value of jth neuron in the hidden layer, b_j is the bias of the jth neuron in the hidden layer, b' is the bias for output layer, W_{ji} is weight of jth neuron of hidden layer connected to the ith neuron of the input layer, W'_j is the weight of the output layer (y) connected to the jth neuron of the hidden layer, N and M are number of neurons in the input and hidden layers, f and g are the activation function in the hidden and output layers. As shown in Eq. (1), the initial values of the weight factor and bias vectors could play a crucial role in the creation of appropriate structures to predict the output parameters. In order to prevent the solution from being captured in local optimum value, the ANN structure is trained using 30 sets of random initial weight factors and bias to select the best initial weights and bias vectors. The schematic diagram of the ANN used in this study is shown in Fig. 3.

In this study, the performance of the proposed architecture of the network [gradient descent back-propagation (GD), gradient descent with momentum back-propagation (GDM), gradient descent with momentum and adaptive learning rate back-propagation (GDX), and Levenberg–Marquardt (LM)] is investigated using training algorithms.

Levenberg–Marquardt algorithm is one of the fastest, efficient, and more robust algorithms than the Gauss–Newton algorithm for MLP–FNN training. This algorithm was designed to approach second-order training speed and accuracy without having to compute the Hessian matrix. In fact, this method interpolates between the Gauss–Newton algorithm and the method of gradient descent to finding an optimal solution in a minimization problem. The Levenberg–Marquardt algorithm uses an approximate to the Hessian matrix in the following Newton-like weights update (Mohanty et al. 2010; Sahoo and Jha 2013):

$$\Delta w_t = [H(w_t) + \mu I]^{-1} J^{\mathrm{T}} e(w_t) \tag{3}$$

where w_t is weight in iteration tth, J Jacobian matrix that contains first derivatives of the network errors with respect to the weights and biases, J^{T} transpose matrix of J, $H = J^{\mathrm{T}}J$ Hessian matrix, μ learning parameter, I identity matrix and e a vector of network errors.

GDX algorithm combines adaptive learning rate with momentum training. In this method, the back-propagation is used to calculate derivatives of performance (g_t) with respect to the weight and bias variables. Each variable is adjusted according to gradient descent with momentum:

$$\Delta w_t = \alpha \times \Delta w_{t-1} + \mu \times \alpha \times g_t \tag{4}$$

where Δw_{t-1} is the previous change to the weight or bias, α is a constant parameter called momentum coefficient and its value may vary between 0 and 1, and μ is a scalar called the learning rate. Similarly, the weight and bias vectors are adjusted as follows in GDM algorithm:

$$\Delta w_t = \alpha \times \Delta w_{t-1} + \mu \times (1 - \alpha) \times g_t \tag{5}$$

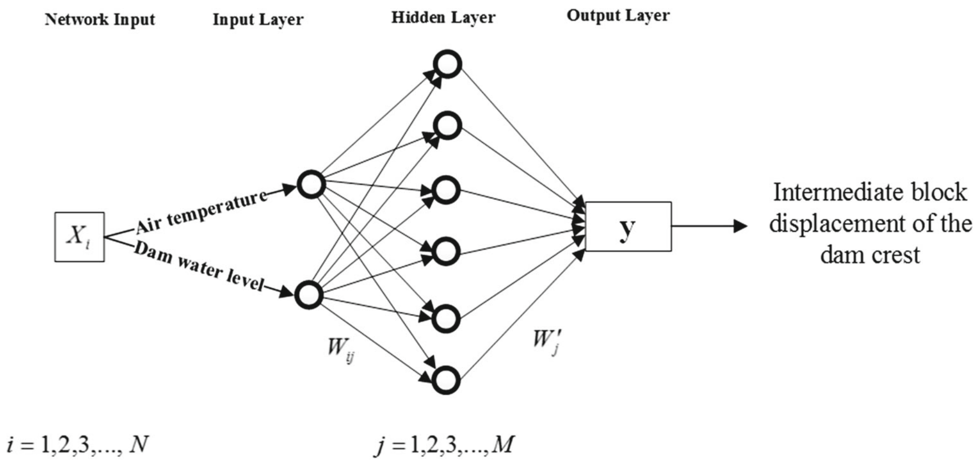


Fig. 3 Schematic diagram of the ANN model used in this study

Unlike the GDX and GDM methods, the network weight updating equation for each variable in GD training algorithm is adjusted according to gradient descent without momentum as given below:

$$\Delta w_t = \mu \times g_t \tag{6}$$

3.2 Support Vector Regression

The SVR method is based on the principle of minimizing the structural risk. In this method, the input vector x is transferred into a space with higher dimensions by nonlinear mapping. In this space, the linear regression applies to the input vector. A set of data (x_n, y_n) is considered, where x_n and y_n are independent and dependent variables, respectively, and $n = 1, 2, \ldots, N$ where N is the total number of input–output data pairs. In fact, the linear regression function can be written as Vapnik (1998):

$$f(x) = \sum_{n=1}^{N} w_n \varphi_n(x) + b \tag{7}$$

where w_n is the weight vector of nth data pairs, b is bias, and $\varphi(x)$ represents a nonlinear transfer function that maps the input vectors into a higher-dimensional space. The coefficients w and b are determined by minimizing the following regularized risk function:

$$P(f(x)) = \frac{C}{N} \sum_{n=1}^{N} E_\varepsilon(y_n, f(x_n)) + \frac{\|w\|^2}{2} \tag{8}$$

where

$$E_\varepsilon(y_n, f(x_n)) = \begin{cases} 0 & \text{if } \ |y - f(x_n)| \le \varepsilon \\ |y - f(x_n)| - \varepsilon & \text{otherwise} \end{cases} \tag{9}$$

where $E_\varepsilon(y_n, f(x_n))$ is called ε-insensitive loss function, and C is a positive regularization constant that determines the trade-off between an approximation error and the weight vector $\|w\|$, and ε is a precision parameter representing the radius of tube size located around the regression function. Both parameters ε and C must be selected in advance by the user. The $0.5\|w\|^2$ is the flatness term (Shirzad et al. 2014).

By introducing two positive slack variables ξ and ξ^* into Eq. (8) representing the distance from actual values to the corresponding boundary values of ε-tube, the overall optimization is formulated as follows:

$$\text{Minimize} \left(\varphi(w, \xi, \xi^*) = C \sum_{n=1}^{N} (\xi + \xi^*) + \frac{\|w\|^2}{2} \right) \tag{10}$$

Subjected to:

$$\begin{cases} y_i - w.\varphi(x) - b \le \varepsilon + \xi_i, & \xi_i \ge 0 \\ w.\varphi(x) - y_i + b \le \varepsilon + \xi_i^*, & \xi_i^* \ge 0 \end{cases} \tag{11}$$

This constrained optimization problem is usually solved by sequential minimal optimization algorithm in a dual form using Lagrangian multipliers and imposing the Karush–Kuhn–Tucker (KKT) optimality condition (Yoon et al. 2011; Elbisy 2015). The Lagrangian form of optimization problem is as follows:

$$\text{Maximize} \left(\begin{array}{l} H(\alpha, \alpha^*) = -0.5 \sum_{i=1}^{N} \sum_{j=1}^{N} (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) K(x_i, x_j) + \\ \sum_{i=1}^{N} y_i (\alpha_i - \alpha_i^*) - \varepsilon \sum_{i=1}^{N} y_i (\alpha_i + \alpha_i^*) \end{array} \right) \tag{12}$$

Subjected to:

$$\sum_{i=1}^{N} (\alpha_i - \alpha_i^*) = 0, \quad 0 \le \alpha_i, \ \alpha_i^* \le C \tag{13}$$

where α_i and α_i^* are Lagrangian multipliers such that $\alpha_i \alpha_i^* = 0$ and $\alpha_i, \alpha_i^* \geq 0$, and $K(x_i, x_j) = \varphi(x_i)\varphi(x_j)$ is the kernel function that yields the inner product in the N-dimensional space. By using kernel functions, one can perform directly all required computations in the input space without calculating the explicit map $\varphi(x)$. In this research, five kernel functions are used as follows:

- The radial basis function (RBF): $K(x_i, x_j) = \exp\left(-\sigma \left\| x_i - x_j \right\|^2\right)$
- The polynomial function (PoL): $K(x_i, x_j) = (x_i x_j + 1)^\sigma$
- The Gaussian RBF function (Gau): $K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)$
- The exponential RBF function (Exp-RBF): $K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|}{2\sigma^2}\right)$
- The linear function (Lin): $K(x_i, x_j) = x_i x_j$

where σ represents the kernel parameter.

Having considered the Lagrangian and optimal conditions, the nonlinear regression function can be given as:

$$f(x) = \sum_{i=1}^{M} \left(\alpha_i - \alpha_i^*\right) K(x, x_i) + b \tag{14}$$

Schematic diagram of the SVR used in this study is shown in Fig. 4.

3.3 Data acquisition and preprocessing

In this study the average monthly temperature and the monthly mean water level at Karaj dam over a period of 8 years (from 2002 to 2009) are used as input parameters. Also, according to the purpose of this study which is to predict the monthly rate of the displacement of the intermediate block of the dam crest, the amount of displacement is simulated by ABAQUS software. These simulated data are considered as output parameter.

Since each input and output parameters have different dimensions and their values do not represent the same quantities, in order to obtain the consistency of the soft model, all data sets are firstly normalized, and then the model outputs are returned to the their original scale after the simulation by using Eq. (15). This scheme can prevent the model from being dominated by the variables with large values and is commonly used in data-driven models such as SVR and ANNs (Yu et al. 2006). Moreover, the results of Bray and Han (2004) showed that the SVR model with normalized input data from zero to one outperforms the models with unscaled input data.

To apply the soft models on the data sets and accelerate the learning process models, it is necessary that all data sets are normalized using the following equation:

$$X_{\text{nor}} = (y_{\max} - y_{\min})\frac{X - X_{\min}}{X_{\max} - X_{\min}} + y_{\min} \tag{15}$$

where X_{nor} is the normalized value of X (as input vector), X is the actual value of each data, and $X_{\min}$ and $X_{\max}$ are the minimum and maximum values of the input vectors, respectively. Also, $y_{\min}$ and $y_{\max}$ are scaling factors. In this paper, zero and one are selected for $y_{\min}$ and $y_{\max}$, respectively.

3.4 Performance evaluation of various models

To evaluate the soft models used in this study, all data sets are divided into three subsets: training data set (including 85% of data), validation data set (including 10% of data), and testing data set (including 5% of data). Based on this classification, the performance of each soft model in the model-building process is evaluated using the correlation coefficient (CORR) and mean squared error (MSE) criteria. In fact, the overall performance is assessed using CORR index and if two sets of simulated and real data have the same trend, the value of this CORR is very close to one. Therefore, based on these criteria one cannot determine the difference between simulated and the real data. For this purpose, the MSE indicator is used to determine the difference between the two mentioned values. The predictive capabilities of the soft model would be high if the value of this indicator is close to zero. The MSE shows the average magnitude of the error between simulated and measured values. Because the errors are squared before taking average, great errors take a relatively high weight. Therefore, MSE is a useful error index when great errors are particularly undesirable (Yoon et al. 2011).

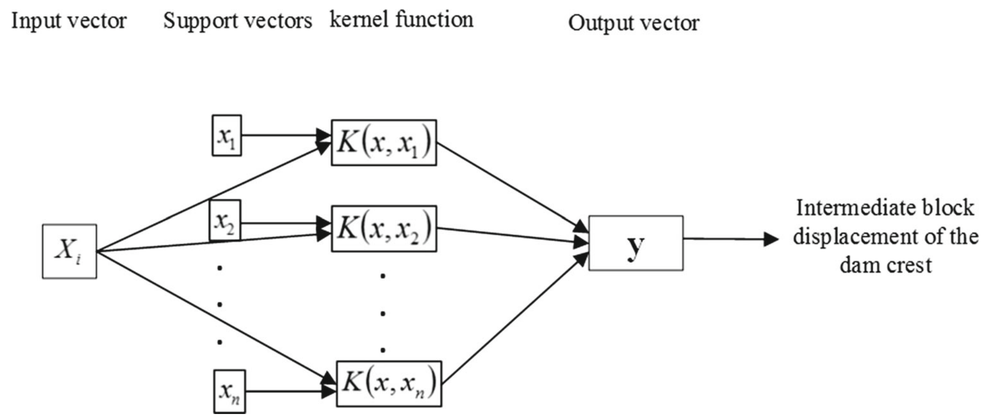
For evaluating the efficiency and superiority of the soft models, the mean error (ME), the mean absolute percentage error (MAPE), the Nash–Sutcliffe efficiency (NS) are used as model performance criteria. The ME measures the bias of overall errors; MAPE shows the relative magnitude of errors, and NS represents the percentage of the initial uncertainty explained by the model. Its value varies between −1 and 1. The values closer to one show that the model performance is more appropriate. The mathematical forms of these indicators are as follows:

$$\text{CORR} = \frac{\sum_{i=1}^{n}(x_{mi} - \bar{x}_m)(x_{ci} - \bar{x}_c)}{\sqrt{\sum_{i=1}^{n}(x_{ci} - \bar{x}_c)^2 \sum_{i=1}^{n}(x_{mi} - \bar{x}_m)^2}} \tag{16}$$

$$\text{MSE} = \frac{\sum_{i=1}^{n}(x_{mi} - x_{ci})^2}{n} \tag{17}$$

$$\text{ME} = \frac{\sum_{i=1}^{n}(x_{mi} - x_{ci})}{n} \tag{18}$$

Fig. 4 Schematic diagram of the SVR model used in this study



$$\text{MAPE} = 100 \times \frac{1}{n} \times \left(\sum_{i=1}^{n} \frac{|x_{mi} - x_{ci}|}{x_{mi}} \right) \tag{19}$$

$$\text{NS} = 1 - \frac{\sum_{i=1}^{n} (x_{mi} - x_{ci})^2}{\sum_{i=1}^{n} (x_{mi} - \bar{x}_m)^2} \tag{20}$$

where x_{mi} and x_{ci} are the measured and simulated values, respectively. Also, $\bar{x}_m$ and $\bar{x}_c$ are the average of measured and simulated values.

4 Results and discussion

In order to predict the displacement of the intermediate block of the dam crest with respect to variations of water level and temperature based on data generated by ABAQUS simulation model, the miscellaneous structures of soft models described in this study are investigated. Then, the selected structures of each soft model are compared in terms of prediction accuracy. It should be noted that in this study, the ANN and SVR modeling was performed using MATLAB R2017b software.

4.1 Development of ANN model

In this study, in order to build a neural network simulation model the following steps need to be taken:

- selection of the input and the output data and statistical analyses of them
- data normalization
- selection of the transfer function and ANN architecture
- selection of the network learning algorithm
- selection of the appropriate indicators to assess the network performance
- selection of stopping criteria

4.1.1 Selection of ANN effective parameter

The basic structure of neural network usually consists of the architecture, the learning algorithm, and the transfer functions. A multilayer perceptron feed-forward neural network (MLP–FNN) is the most famous of neural network models that were used in many water resources systems [such as Kim et al. (2013) and Seckin et al. (2013)] and is selected to be used in this study. According to the nonlinear nature of the data set used in soft simulation models, it is necessary to establish a clear relationship between these data by selecting the parameters that provide the most suitable conditions for simulation with minimum error. In ANN, the important effective parameters are: number of layers, type of transfer function in each layer, the number of neurons in each layer, and criteria to stop the training network. To provide supplementary flexibility to the ANN model, linear function is used in the output layer where extrapolation is needed beyond the range of training data (Mustafa et al. 2012). Also, the log-sigmoid transfer functions are considered for the hidden layer to accelerate the training network. The mathematical transfer functions are presented in the following equations:

$$\text{Logsig}(n) = \frac{1}{1 + e^{-n}} \tag{21}$$

$$\text{Purelin}(n) = n \tag{22}$$

where n in Eqs. (16) and (17) represents the weighted sum for a neuron in the hidden layer and is the weighted sum of inputs at the output layer, respectively.

The number of layers and neurons in each layer constitutes the architecture of an ANN model. Appropriate selection of neurons number and layers is necessary for good predicting results. Investigations done in this study show that network complexity increases and accuracy of the output decreases by the increase in the number of layers. So, in the present study, three layers (input, hidden, and output) are considered for the network.

Table 3 The correlation between the delayed input data and the displacement of the intermediate block of the dam crest

Lag time (month)	Temperature	Reservoir water level
1	0.48	0.66
2	0.44	0.6
3	0.28	0.42
4	0.02	0.11

In order to determine the number of neurons in the input layer, the correlation between the delayed input data and the displacement of the intermediate block of the dam crest (as output of the network) is determined. It should be noted that the temperature and the reservoir water level in the antecedent times may affect the amount of intermediate block displacement of the dam crest. This issue is studied in this research. If good correlation between input data at present and antecedent times exists, it is necessary that the delayed values for each of the input parameters of soft models be considered. Based on Table 3, it can be found that none of the inputs show good correlation at different delays with output. Thus, application of the values at antecedent times (months) on input data in soft models is not appropriate.

Therefore, the number of neurons in the input layer was considered to be 2. In order to determine the appropriate number of neurons for the hidden layer, the network is trained for different values of neurons in this layer. This process is performed by trial-and-error method suggested by Maier and Dandy (2000). In the developed neural network structure, the mean square error (MSE) between observed and simulated data is considered as stopping criteria of training. Training is stopped when the MSE error index in the validation stage starts to increase. Under these conditions, the number of training epochs is determined. In this section, the results related to the training of the ANN model are presented using various training algorithms.

In GD and GDM algorithms, the network is trained with the different learning rate. By using trial-and-error method and network training with different values of the learning rate, it is found that the learning rate of 0.05 could improve network performance using GD (the model ANN–GD2) and GDM (ANN–GDM model) algorithms. However, in these conditions, a small number of output values are negative. Accordingly, we conclude that both GD and GDM algorithms do not have good capability to predict the displacement of the intermediate block of the dam crest.

The results of the GDX algorithm indicate that this algorithm has better performance than the ANN–GD2 and ANN–GDM models in training and validation stage. This algorithm in testing stage has a relatively high error rate. According to the results obtained from the testing data, which have not previously been used in the training process, it can be found that the ANN–GDM model, despite having a higher error rate than the ANN–GDX model in the training stage, is able to provide better in prediction of the displacement of the intermediate block of the dam crest under unexperienced conditions. The comparison of the MSE error index associated with the five neural network models in the testing stage suggests the superiority of the ANN–LM training algorithm in estimating the displacement of the intermediate block of the dam crest with a decrease of 68.6% in the prediction error compared to the ANN–GDM model.

However, investigation of error index values of models shows that the LM algorithm attains the acceptable accuracy with less number of epochs and MSE error index. The LM algorithm converges faster than other training algorithms. (It achieves the goal before 8 iterations.) To determine the number of suitable neurons for the hidden layer and for each training algorithm, the structures of the developed neural network were trained 30 times for neurons 2–25 in the hidden layer. For example, for neuron 5 in the hidden layer, the structure of neural network is trained 30 times and the best structure with minimum MSE error index is extracted. This process was performed for all the number of examined neurons. The results of the comparison between the amounts of MSE error index in the number of different neurons can be observed for the LM training algorithm in Fig. 5. The results show that in LM algorithm the structure including 20 neurons in the hidden layer leads to the lowest error. The number of output neurons is restricted to one because only one output (displacement of the intermediate block of the dam crest) is considered. Hence, the most appropriate architecture of the network is selected as 2–20–1 (where 2, 20, and 1 are the number of neurons in the input, hidden, and output layers, respectively) with three layers and logistic sigmoid and linear transfer functions for the hidden and output layers, respectively.

Finally, review of the MSE and CORR error index during training, validation, and testing of different models shows that the ANN–LM model has relative advantage in prediction of the displacement of the intermediate block of the dam crest (Table 4). This finding is approved by other researchers such as Mohanty et al. (2010) and He et al. (2014).

In order to evaluate the performance of the ANN–LM model, the scatter plots related to predicted and observed values for the training and testing are presented in Figs. 6 and 7. According to these figures, it is clear that the correlation coefficient of testing stage is higher than the training stage. This clearly shows the suitability of LM algorithm to predict the displacement of the intermediate block of the dam crest even for the data which are not used in the training process.

Comparison between the observed and the predicted intermediate block displacement during training (Fig. 8) shows that the ANN–LM model is not able to predict accurately the intermediate block displacement. Significant differences

Fig. 5 Comparison between MSE error index in testing stage for different neurons located on the hidden layer

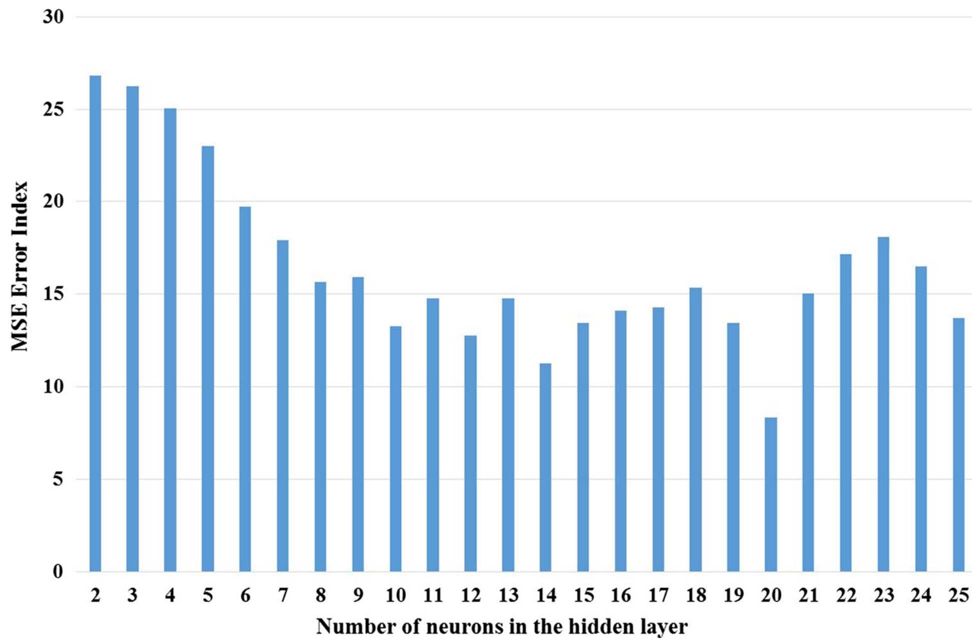


Table 4 Comparison of performance of ANN models during training and testing with different algorithms

Model name	Algorithm	Number of epochs	Time (s)	Train		Validation		Test	
				CORR	MSE	CORR	MSE	CORR	MSE
ANN–LM	LM	8	0.37	0.73	57.36	0.77	25.44	0.93	8.34
ANN–GD1	GD ($\mu = 0.01$)	474	0.59	0.47	99.73	0.54	46.56	0.74	30.54
ANN–GD2	GD ($\mu = 0.05$)	1000	1.11	0.58	80	0.48	46	0.72	34.79
ANN–GDM	GDM ($\mu = 0.05,\ \alpha = 0.9$)	512	0.62	0.54	85.16	0.5	42.8	0.78	26.56
ANN–GDX	GDX ($\mu = 0.05,\ \alpha = 0.9$)	16	0.54	0.61	81.57	0.62	41.43	0.64	55.92

Fig. 6 Scatter plot of predicted values versus observed values of displacement of the intermediate block for the training stage of ANN–LM model

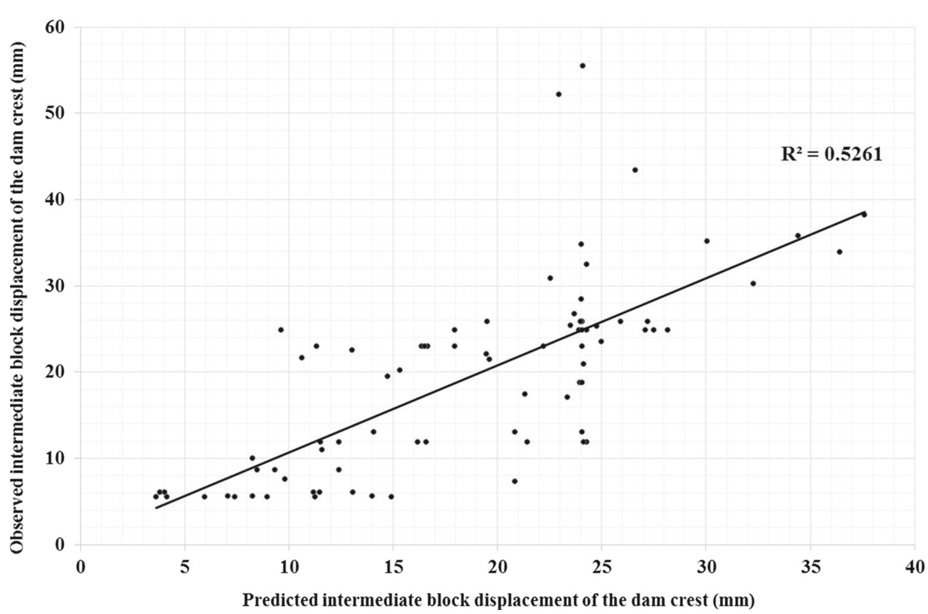


Fig. 7 Scatter plot of predicted values versus observed values of displacement of the intermediate block for the testing stage of ANN–LM model

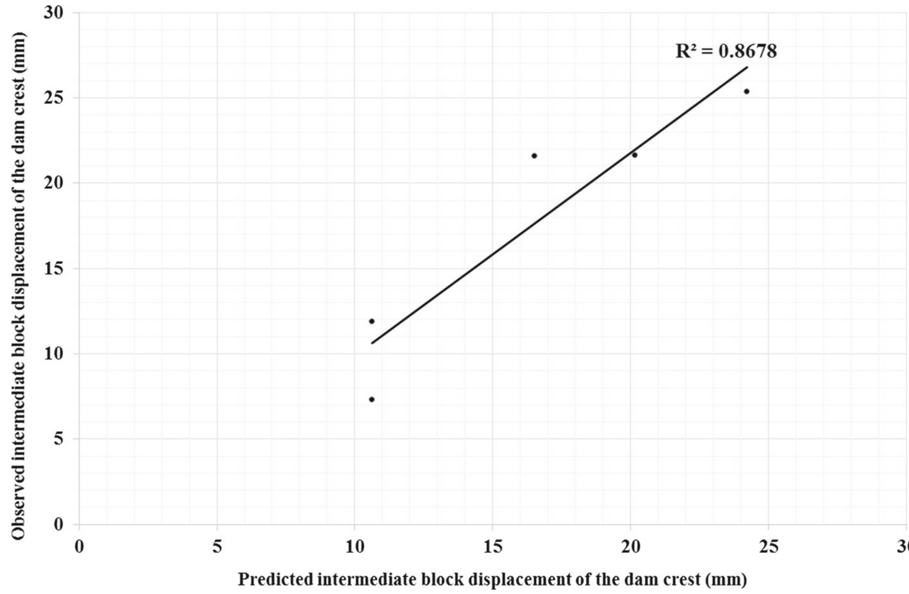
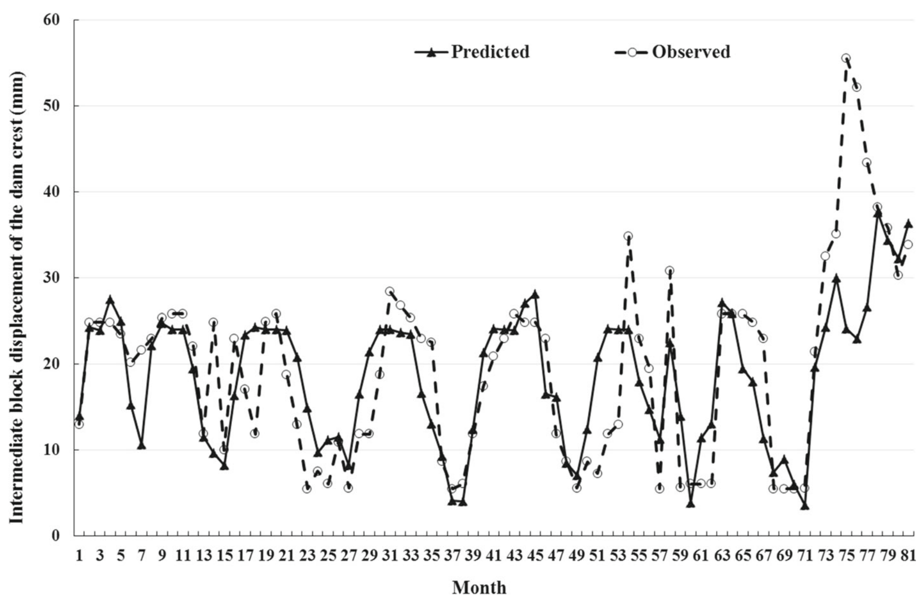


Fig. 8 Comparison between time series of observed and predicted intermediate block displacement using ANN–LM model (training stage)



between predicted and observed values exist in months when the values of measured input parameters (mean of temperature and reservoir water level) are high. This is also true about the obtained results of testing stage (Fig. 9). In fact, this model (ANN model) has serious weakness at maximum values of inputs and it is necessary to be implemented with other soft models.

4.2 Development of SVR model

One of the most important parameters for prediction of the displacement of the intermediate block of the dam crest using support vector regression model is to determine the appropriate kernel functions. In this study, the polynomial (POLYNOMIAL), radial basis function (RBF), Gaussian (GAUSSIANRBF), power (EXPONENTIALRBF), and linear kernels are evaluated. With the implementation of SVR models for different kernel functions, it can be inferred that the RBF kernel function used in the research of Yu et al. (2006), Kalteh (2015), Mirzavand et al. (2015), and Elbisy (2015) is more accurate, and therefore this kernel is used to predict the displacement of the intermediate block of the dam crest (Table 5).

In order to increase the accuracy of SVR–RBF model, it is necessary to determine the governing parameters of this model (such as C, ε, and γ) by trial-and-error approach. It is worth to mention that parameters C and ε influence the quality and time of training. Also, the value of γ can also be effective in creating overfitting and underfitting. Based on repeatedly running of SVR–RBF model for different men-

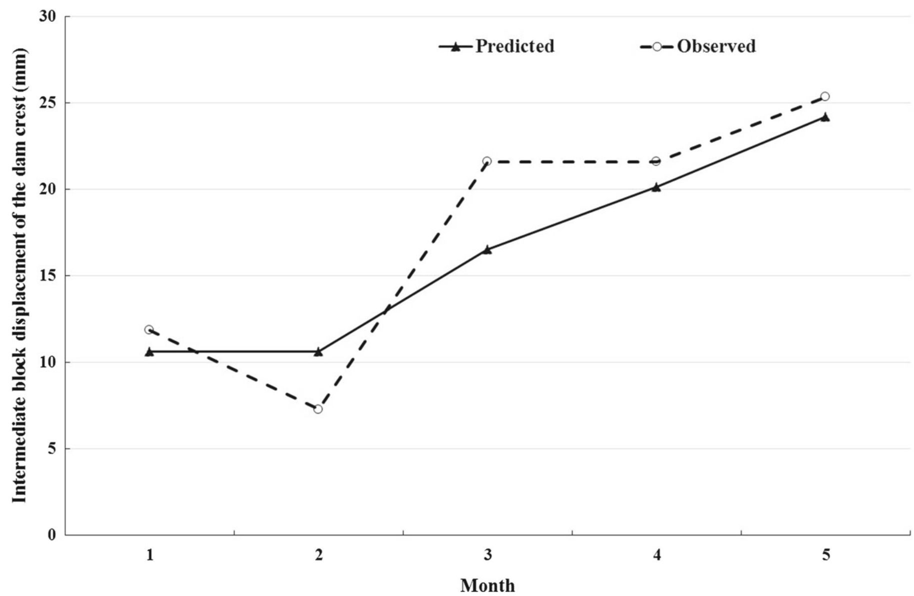


Fig. 9 Comparison between time series of observed and predicted intermediate block displacement using ANN–LM model (testing stage)

Table 5 Comparison of different SVR models to predict the displacement of the intermediate block of the dam crest

Model	Kernel function type	Train		Validation		Test	
		CORR	MSE	CORR	MSE	CORR	MSE
SVR–RBF	RBF	0.96	8.47	0.99	0.278	0.98	4.7
SVR–PoL	POLYNOMIAL	0.67	61.78	0.7	52.81	0.73	31.42
SVR–Gau	GAUSSIANRBF	0.55	95.15	0.4	94	0.49	73.7
SVR–Exp	EXPONENTIALRBF	0.58	76.52	0.511	80.03	0.53	46.41
SVR–Lin	LINEAR	0.59	73.97	0.67	59	0.56	47.75

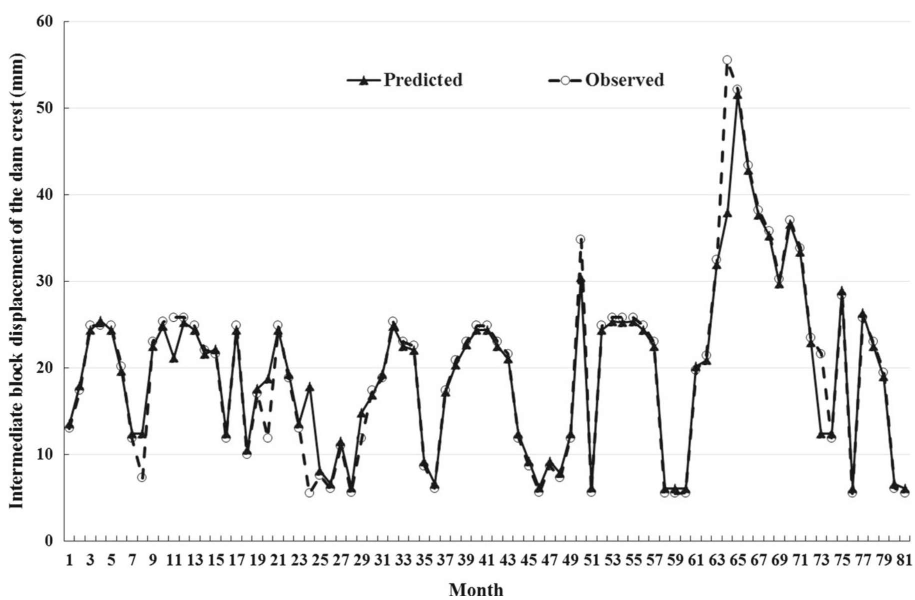


Fig. 10 Comparison between time series of observed and predicted intermediate block displacements using SVR–RBF model (training stage)

tioned parameters, the optimal values of the user-defined SVR–RBF parameters for C, ε, and γ are 32.35, 0.53, and 0.08, respectively. Figures 10 and 11 show the comparison of the results of the simulation by SVR–RBF model and the observed data in training and testing stages. Evaluation of the predicted values of intermediate block displacement by the SVR–RBF model shows that the MSE values in training and testing stages are reduced to the 85.23 and 43.64%, respectively, compared to the ANN model (Table 6). These results indicate the relatively better performance

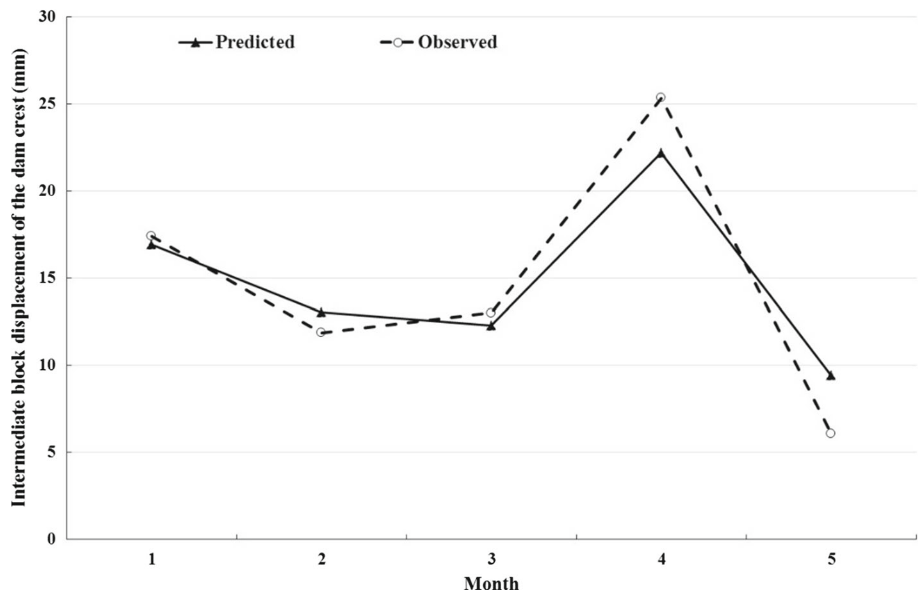


Fig. 11 Comparison between time series of observed and predicted intermediate block displacements using SVR–RBF model (testing stage)

of SVR–RBF model compared to the ANN model for nonlinear systems modeling.

4.3 Assessment of the ability of SVR and ANN soft models

In this section for overall assessment of the ability of SVR and ANN models based on five defined error indicators, the performance of two soft models in prediction of the displacement of the intermediate block of the dam crest is compared according to input parameters (i.e., the water level in the reservoir and monthly temperature average) as shown in Table 6.

Based on the ME index, it can be inferred that ANN–LM model is overestimated in testing stage. The magnitude of the ME index in ANN–LM model is higher than SVR–RBF model which reflects the fact that this model gives lower values for intermediate block displacement than the observed values.

MAPE and MSE values also show that the SVR–RBF model is better than the ANN–LM model. Moreover, NS and CORR values of the SVR–RBF model are greater than the ANN–LM model. Investigation of various error indices presented in Table 6 shows that the SVR–RBF model predicts the displacement of the intermediate block of the dam crest better than the ANN–LM model.

In fact, the presented error indicators represent the average value of the errors of the soft models and accordingly cannot provide an accurate assessment of the distribution of errors. So, in this study, normal cumulative distribution functions (CDFs) of observed and predicted intermediate block displacement values in both training and testing stages are presented for better assessment of two soft models (Figs. 12, 13). This approach was previously followed by Yoon et al. (2011) for selection of the best model. Comparison of CDFs shows that the amount of deviation of the CDF for ANN–LM model is higher than SVR–RBF model particularly in the training stage. The reason is due to the fact that the SVR–RBF model is generally based on the structural risk minimization (SRM) and the SRM minimizes the empirical risk and complexity of model simultaneously which can improve the ability of the SVR for regression problems. Therefore, the SVR–RBF model can predict the dam stability with acceptable accuracy only by measuring two parameters including water level of the dam and the air temperature.

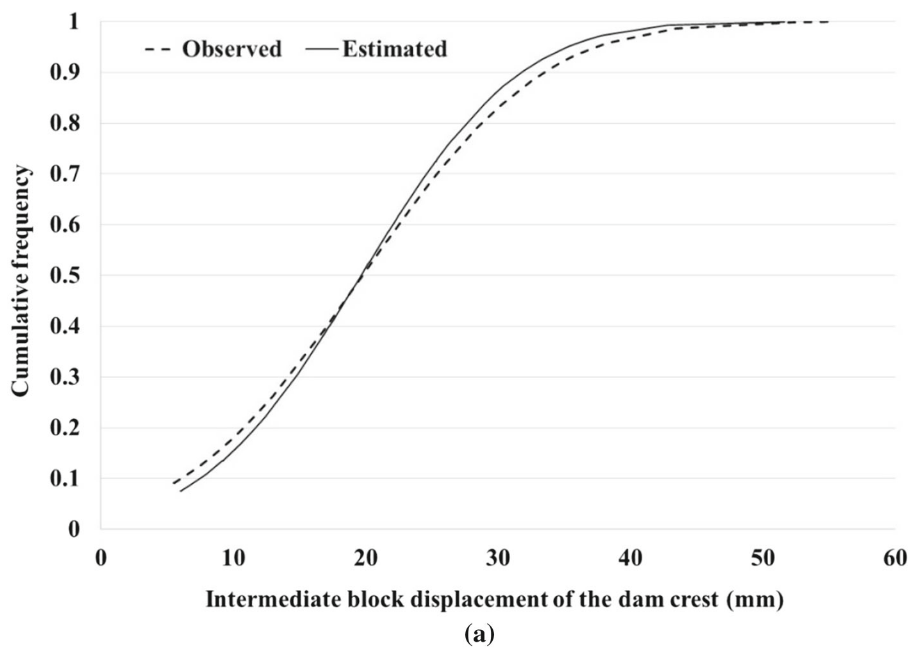
5 Conclusions

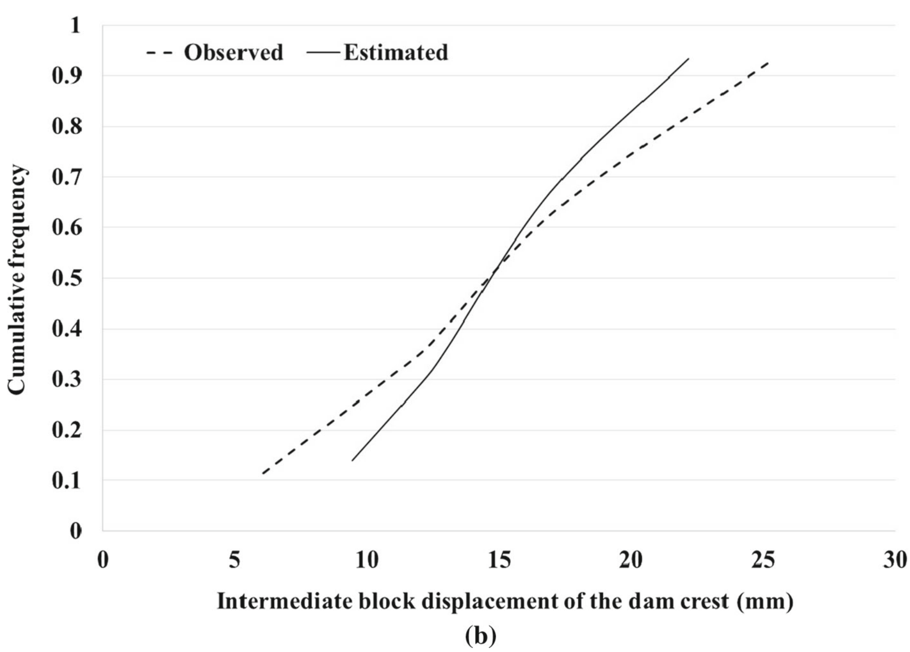
In this research, time series models were developed for prediction of the dam crest displacement using support vector regression and artificial neural network. In order to apply ABAQUS simulation model, the Karaj dam thermal behavior over a period of 8 years (2002–2009) was simulated under different water levels in the reservoir and monthly average temperature. Based on simulated data and in order to obtain the nonlinear relationship between the crest displacement and the average temperature of the reservoir dam, the capability of SVR and ANN models to predict the displacement of the dam crest was investigated. The results of different structures of the ANN to create described nonlinear relationship show that the ANN–LM model has better performance and fastest training time among five training algorithms. Based on defined error indicators, it can be found that the accuracy of predicted intermediate block displacements of ANN–LM model is relatively low. In order to increase the accuracy of

Table 6 Comparison of different error indices for prediction of the displacement of the intermediate block of the dam crest using ANN–LM and SVR–RBF models

Error index	Model					
	ANN–LM			SVR–RBF		
	Train	Validation	Test	Train	Validation	Test
NS	0.52	0.54	0.82	0.93	0.99	0.88
MAPE (percentage)	34.54	22.83	18.15	9.35	3.99	17.2
ME (mm)	0.8	– 0.29	1.11	0.13	0.001	– 0.04
CORR	0.73	0.77	0.93	0.96	0.99	0.98
MSE (mm)	57.36	25.44	8.34	8.47	0.278	4.7

Fig. 12 Cumulative distribution function (CDFs) of observed and estimated intermediate block displacement using SVR model. **a** Training stage and **b** testing stage





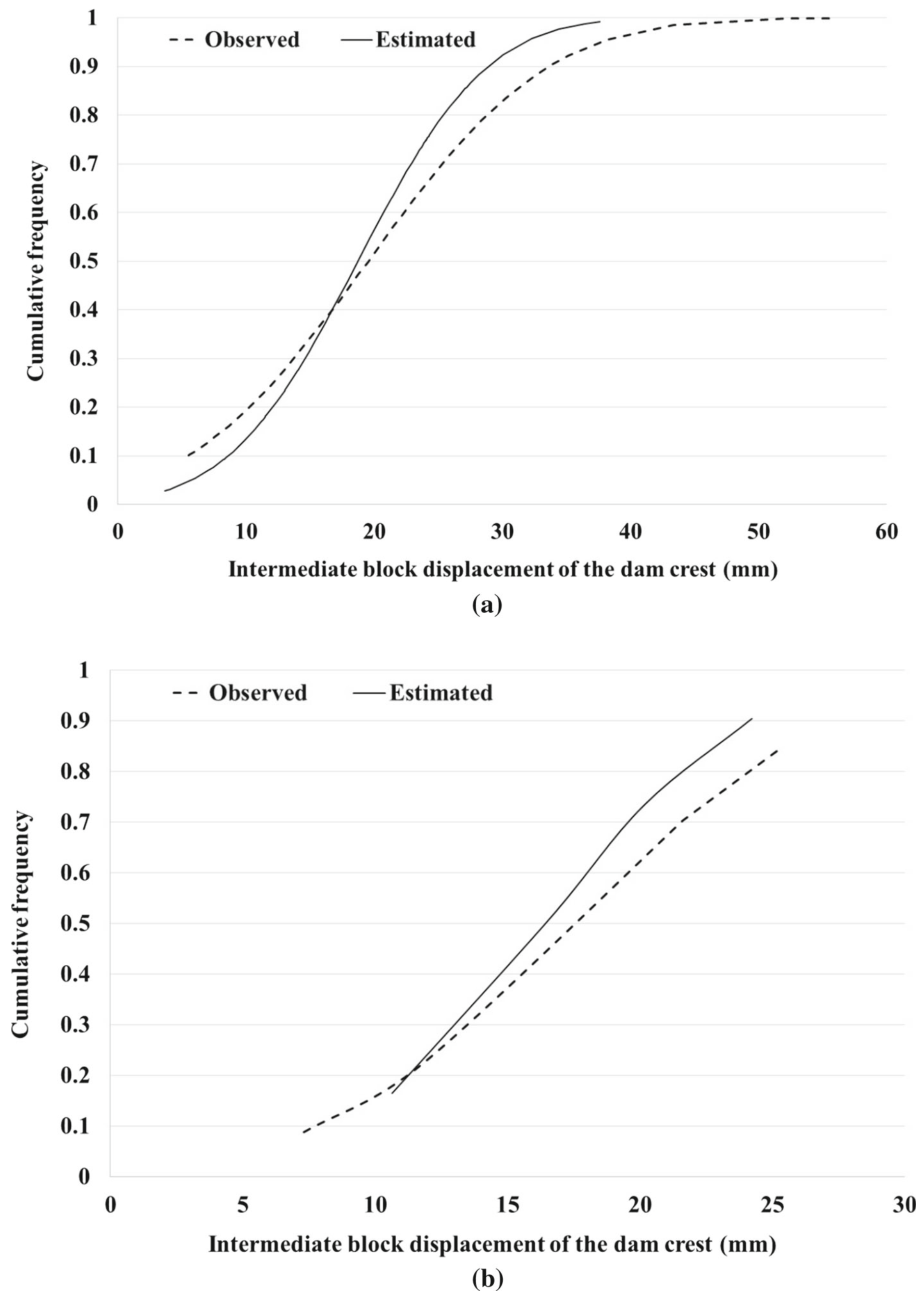


Fig. 13 Cumulative distribution function (CDFs) of observed and estimated intermediate block displacement using ANN model. **a** Training stage and **b** testing stage

the prediction, it was proposed to apply the support vector regression as a model that has high accuracy for nonlinear modeling. In this model, the appropriate parameters were determined by trial-and-error method and the results were compared with the ANN model. The evaluation of results of ANN and SVR models with five error indicator shows that the SVR model predicts better the intermediate block displacement of the dam crest. Furthermore, comparison of cumulative distribution function of observed and estimated (using ABAQUS model) displacement values showed significant differences between CDFs in ANN model compared with the SVR model. This shows the inability of ANN model to predict the intermediate block displacement of the dam crest. The results of this paper can be used to develop the optimal operation policy leading to increased stability of the reservoir with minimum amount of displacement of the dam crest. Also, the displacement of the dam crest for any given period can be predicted by only measuring the air temperature and the water level of the dam reservoir without any need to re-simulate with ABAQUS model.

Acknowledgements This work has been financially supported by the Research Deputy of Shahrekord University. The Grant Number was 95GRN1M1156.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval His article does not contain any studies with human participants or animals performed by any of the authors. This work was carried out in collaboration between all authors. All authors read and approved the final manuscript.

References

Agullo L, Aguado A, Mirambell E (1991) A model for the analysis of concrete dams due environmental effects. Int J Numer Methods Heat Fluid Flow 6(4):25–36

Barzegar R, Fijani E, Asghari Moghaddam A, Tziritis E (2017) Forecasting of groundwater level fluctuations using ensemble hybrid multi-wavelet neural network-based models. Sci Total Environ 599–600:20–31

Behnia D, Ahangari K, Noorzad A, Moeinossadat SR (2013) Predicting crest settlement in concrete face rockfill dams using adaptive neuro-fuzzy inference system and gene expression programming intelligent methods. J Zhejiang Univ Sci A (Appl Phys Eng) 14(8):589–602

Bofang Z (2014) Thermal stresses and temperature control of mass concrete. China Institute of Water Resources and Hydropower Research and Chinese Academy of Engineering, Elsevier, Amsterdam

Bray M, Han D (2004) Identification of support vector machines for runoff modeling. J Hydroinform 6(4):265–280

Cao M, Qiao P, Ren Q (2009) Improved hybrid wavelet neural network methodology for time-varying behavior prediction of engineering structures. Neural Comput Appl 18(7):821–832

Csábrági A, Molnár S, Tanos P, Kovács J (2017) Application of artificial neural networks to the forecasting of dissolved oxygen content in the Hungarian section of the river Danube. Ecol Eng 100:63–72

Ebrahimi H, Rajaee T (2017) Simulation of groundwater level variations using wavelet combined with neural network, linear regression and support vector machine. Global Planet Change 148:181–191

Elbisy MS (2015) Support vector machine and regression analysis to predict the field hydraulic conductivity of sandy soil. KSCE J Civ Eng 19(7):2307–2316

Emamgholizadeh S, Moslemi K, Karami G (2015) Prediction the groundwater level of Bastam Plain (Iran) by artificial neural network (ANN) and adaptive neuro-fuzzy inference system (ANFIS). Water Resour Manag 28(15):5433–5446

Fedele R, Maier G, Miller B (2006) Health assessment of concrete dams by overall inverse analyses and neural networks. Int J Fract 137(1–4):151–172

Fisher WD, Camp TK, Krzhizhanovskaya VV (2016) Crack detection in earth dam and levee passive seismic data using support vector machines. Proc Comput Sci 80:577–586

French MN, Krajewshi WF, Cuendall PP (1992) Rainfall forecasting in space and time using artificial neural network. J Hydrol 13(7):1–13

Gaziev EG (2000) Safety provision and an expert system for diagnosing and predicting dam behavior. Hydrotech Constr 33(6):285–289

Guang-yong X, Jian-ping Y, Bao-xing Z, Pu T (2011) Application of an artificial immune algorithm on a statistical model of dam displacement. Comput Math Appl 62(10):3980–3986

He Z, Zhang Y, Guo Q, Zhao X (2014) Comparative study of artificial neural networks and wavelet artificial neural networks for groundwater depth data forecasting with various curve fractal dimensions. Water Resour Manag 28(15):5297–5317

Hipni A, El-shafie A, Najah A, Othman A, Aini H, Muhammad M (2013) Daily forecasting of dam water levels: comparing a support vector machine (SVM) model with adaptive neuro fuzzy inference system (ANFIS). Water Resour Manag 27(10):3803–3823

Hosseini SM, Mahjouri N (2016) Integrating support vector regression and a geomorphologic artificial neural network for daily rainfall-runoff modeling. Appl Soft Comput 38:329–345

Humphrey GB, Gibbs MS, Dandy GC, Maier HR (2016) A hybrid approach to monthly stream flow forecasting: integrating hydrological model outputs into a Bayesian artificial neural network. J Hydrol 540:623–640

Kalteh AM (2015) Wavelet genetic algorithm-support vector regression (wavelet GA–SVR) for monthly flow forecasting. Water Resour Manag 29(4):1283–1293

Kao CY, Loh CH (2013) Monitoring of long-term static deformation data of Fei-Tsui arch dam using artificial neural network-based approaches. Struct Control Health Monit 20(3):282–303

Karimi I, Khaji N, Ahmadi MT, Mirzayee M (2010) System identification of concrete gravity dams using artificial neural networks based on a hybrid finite element-boundary element approach. Eng Struct 32(11):3583–3591

Khashei-Siuki A, Sarbazi M (2013) Evaluation of ANFIS, ANN, and geo-statistical models to spatial distribution of groundwater quality (case study: Mashhad plain in Iran). Arab J Geosci 8(2):903–912

Kim S, Shiri J, Kisi O, Singh VP (2013) Estimating daily pan evaporation using different data-driven methods and lag-time patterns. Water Resour Manag 27(7):2267–2286

Labibzadeh M, Khajehdezfuly A (2010) Hydro-thermal safety control of Karun-1 Dam under unusual reservoir level reduction. J Am Sci 6(11):179–184

Maier HR, Dandy GC (2000) Neural networks for the prediction and forecasting of water resources variables: a review of modeling issues and applications. Environ Model Softw 15:101–124

Mata J (2011) Interpretation of concrete dam behavior with artificial neural network and multiple linear regression models. Eng Struct 33:903–910

Mirzavand M, Khoshnevisan B, Shamshirband S, Kisi O, Ahmad R, Akib S (2015) Evaluating groundwater level fluctuation by support vector regression and neuro-fuzzy methods: a comparative study. Nat Hazards. https://doi.org/10.1007/s11069-015-1602-4

Mohanty S, Jha MK, Kumar A, Sudheer KP (2010) Artificial neural network modeling for groundwater level forecasting in a river island of eastern India. Water Resour Manag 24(9):1845–1865

Mustafa MR, Rezaur RB, Saiedi S, Isa MH (2012) River suspended sediment prediction using various multilayer perceptron neural network training algorithms: a case study in Malaysia. Water Resour Manag 26(7):1879–1897

Popescu TD (2012) Neural network learning for blind source separation with application in dam safety monitoring. Neural Inf Process Lect Notes Comput Sci 7666:1–8

Rankovic V, Nenad A, Dejan DB, Nikola M (2014) Development of support vector regression identification model for prediction of dam structural behavior. Struct Saf 48:33–39

Rumelhart DE, McClelland JL (1986) Parallel distributed processing: explorations in the microstructure of cognition. The PDP Research Group, MIT Press, Cambridge, p 516

Sahoo S, Jha MK (2013) Groundwater-level prediction using multiple linear regression and artificial neural network techniques: a comparative assessment. Hydrogeol J 21(8):1865–1887

Salazar F, Toledo MA, Oñate E, Morán R (2015) An empirical comparison of machine learning techniques for dam behavior modelling. Struct Saf 56:9–17

Saqib M, Ansari MI (2017) Computation of stresses in concrete gravity dam under seismic loading through ANN and FEM. Proc Eng 173:1779–1783

Seckin N, Cobaner M, Yurtal R, Haktanir T (2013) Comparison of artificial neural network methods with L moments for estimating flood flow at ungauged sites: the case of East Mediterranean River Basin, Turkey. Water Resour Manag 27(7):2103–2124

Senthil Kumar AR, Goyal MK, Ojha CSP, Singh RD, Swamee PK, Nema RK (2013) Application of ANN, fuzzy logic and decision tree algorithms for the development of reservoir operating rules. Water Resour Manag 27(3):911–925

Seo I, Yun SH, Choi SY (2016) Forecasting water quality parameters by ANN model using pre-processing technique at the downstream of Cheongpyeong Dam. Proc Eng 154:1110–1115

Shirzad A, Tabesh M, Farmani R (2014) A comparison between performance of support vector regression and artificial neural network in prediction of pipe burst rate in water distribution networks. KSCE J Civ Eng 18(4):941–948

Stojanovic B, Milivojevic M, Milivojevic N, Antonijevic D (2016) A self-tuning system for dam behavior modeling based on evolving artificial neural networks. Adv Eng Softw 97:85–95

Tayfur G, Nadiri AA, Moghaddam AA (2014) Supervised intelligent committee machine method for hydraulic conductivity estimation. Water Resour Manag 28(4):1173–1184

Us Army Corps of Engineering (1994) Arch dam design. Engineer Manual 110-2-2201, Chapter 8, pp 1–15

Vapnik VN (1995) The nature of statistical learning theory. Springer, New York, p 187

Vapnik VN (1998) Statistical learning theory. Wiley, New York, p 740

Yoon H, Jun SC, Hyun Y, Bae GO, Lee KK (2011) A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. J Hydrol 396:128–138

Yu P, Chen ST, Chang IF (2006) Support vector regression for real-time flood stage forecasting. J Hydrol 328:704–716

Zhou CB, Liu W, Chen YF, Hu R, Wei K (2015) Inverse modeling of leakage through a rockfill dam foundation during its construction stage using transient flow model, neural network and genetic algorithm. Eng Geol 187:183–195