Research Article

An Empirical Comparison of Multiple Linear Regression and Artificial Neural Network for Concrete Dam Deformation Modelling

Mingjun Li and Junxing Wang

State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China

Correspondence should be addressed to Junxing Wang; jxwang@whu.edu.cn

Received 12 January 2019; Revised 6 March 2019; Accepted 24 March 2019; Published 17 April 2019

Academic Editor: Łukasz Jankowski

Copyright © 2019 Mingjun Li and Junxing Wang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Deformation predicting models are essential for evaluating the health status of concrete dams. Nevertheless, the application of the conventional multiple linear regression model has been limited due to the particular structure, random loading, and strong nonlinear deformation of concrete dams. Conversely, the artificial neural network (ANN) model shows good adaptability to complex and highly nonlinear behaviors. This paper aims to evaluate the specific performance of the multiple linear regression (MLR) and artificial neural network (ANN) model in characterizing concrete dam deformation under environmental loads. In this study, four models, namely, the multiple linear regression (MLR), stepwise regression (SR), backpropagation (BP) neural network, and extreme learning machine (ELM) model, are employed to simulate dam deformation from two aspects: single measurement point and multiple measurement points, approximately 11 years of historical dam operation records. Results showed that the prediction accuracy of the multipoint model was higher than that of the single point model except the MLR model. Moreover, the prediction accuracy of the ELM model was always higher than the other three models. All discussions would be conducted in conjunction with a gravity dam study.

1. Introduction

Deformation modelling is an important component of dam safety systems, both for the daily operation and for long-term behavior evaluation [1]. They are built to calculate the dam response under safe conditions for a given load combination, which is compared to actual measurements of dam performance with the aim of detecting anomalies and preventing failures. Current predictive models for simulating dam deformation can be classified as three types: deterministic models, statistical models, and hybrid models [2], i.e., a mixture of the first two.

Deterministic models based on physical laws such as load, material properties, and stress-strain relationships are often used to design dams and function throughout the life of concrete dams [3]. Anomalies in the operation of concrete dams can be fundamentally explained by deterministic

models whose parameters have specific physical meanings; uncertainties existing in geological conditions and in material properties of rock base and concrete hinder the implementation of deterministic models.

Statistical models are mathematical equation that quantitatively describe the variation law of dam monitoring values and are the abstraction and simplification of the actual working state of the dam [4]. Without regard to the specific physical mechanism of dam operation, the statistical model is essentially an empirical model based on dam measurement data over years. As the most widely used model for characterizing concrete dam deformation, the statistical model usually consists of three parts temperature component, hydrostatic pressure component, and aging component. The statistical model assumes that the components are completely independent and may not match the actual situation [5]. At present, the most common statistical

models are based on regression methods such as the multiple linear regression (MLR) [6], stepwise regression (SR) [7], principal component regression (PCR) [8], and partial least squares regression (PLSR) [9]. By gradually screening regression factors of MLR models, SR models obtain regression coefficients at a certain significance level, results more accurate than those obtained by the general least square method.

In recent years, more and more scholars have begun to apply machine learning algorithms or intelligent algorithms to dam safety effect prediction or dam safety diagnosis analysis [2, 10, 11]. Mata [12] found that ANN models can be a very powerful tool in evaluating dam behavior by comparing the multiple linear regression model with the multilayer perceptron model for the horizontal displacement of a concrete arch dam. F. Salazar [2] assessed the potential of some state-of-the-art machine learning techniques including random forests, boosted regression trees, neural networks, and multivariate adaptive regression splines to build models for the prediction of dam behavior in the field of displacement and leakage. Kao [13] studied the feasibility of ANN-based approaches for dam health monitoring and set an early warning threshold level of the Fei-Tsui dam based on the analysis results. But the artificial neural network based on gradient descent method is still relatively slow and easily gets stuck. Su [14] proposed a dam safety monitoring model based on support vector machine, which could overcome the disadvantages of the above artificial neural network, but the selection of kernel function parameters is a difficult point. The extreme learning machine (ELM) is a new method training single hidden layer feedforward neural networks proposed by Huang et al. [15, 16]. The extreme learning machine first randomly generates the hidden layer deviation and weight of the connected input and the hidden layer and then directly determines the weight value between the hidden layer and the output layer by Moore-Penrose generalized inverse method. While overcoming the shortcomings of the gradient descent learning method, the ELM greatly improves the learning speed of artificial neural networks and ensures good generalization ability. Kang [17] proposed an ELM-based gravity dam deformation prediction model and explored the application of ELM algorithm from the perspective of prediction accuracy. However, the adaptability of the ELM model to the interpretation of dam deformation has not yet been elucidated.

The dam deformation monitoring model can be divided into two types: the single point deformation monitoring model and the multipoint deformation monitoring model. The current research mainly focuses on single point monitoring deformation model, which cannot reflect the spatial distribution of deformation. And the multipoint deformation monitoring model can better reflect the mutual relationship between the deformation points of the dam body, which is more reasonable than the single point model. As a kind of statically indeterminate shell structure, the concrete arch dam is obviously affected by the spatial integrity of the concrete arch dam.

This paper studies the application characteristics and effects of the multiple linear regression (MLR), stepwise regression (SR), backpropagation (BP) neural network, and extreme learning machine (ELM) on concrete dam deformation modelling based on the monitoring data of the Dongjiang arch dam. The similarities and differences between the single point model and the one-dimensional multipoint model are discussed. The work focuses on prediction accuracy and the suitability for interpreting dam behavior. All discussions will be carried out in conjunction with the results of a gravity dam [17] study.

2. Statistical Model

Statistical models are established by the correlation between observed effect quantities and environmental variables. With the environment treated as an independent variable, the structural response of the dam is affected by three effects the reversible effect of the hydrostatic load, the reversible thermal influence of the temperature, and the irreversible term due to the evolution of the dam response over time [4, 18]. According to the influencing factors, the displacement δ_z of the arbitrary point $k(x, y, z)$ in the z direction can be expressed as

$$\begin{aligned}\delta &= f\{g(H, T, \theta), l(z)\} \\ &= f_1\{g_1(H), l_1(z)\} + f_2\{g_2(T), l_2(z)\} \\ &\quad + f_3\{g_3(\theta), l_3(z)\}\end{aligned}\quad (1)$$

where $g_1(H)$, $g_2(T)$, $g_3(\theta)$ represent the hydraulic displacement component, temperature displacement component, and aging displacement component, respectively. $l_i(x, y, z)$ indicates the deformed surface of the dam fixed point $k(x, y, z)$ under the action of water pressure ($i = 1$), temperature ($i = 2$) and aging ($i = 3$) alone, where $l(z)$ is approximated by multiple power series, and

$$l(z) = \sum_n^3 a_n z^n \quad (2)$$

And the hydraulic displacement component $g_1(H)$, temperature displacement component $g_2(T)$, and aging displacement component $g_3(\theta)$ can expressed as follows [19]:

$$g_1(H) = \sum_{i=1}^4 a_i H^i \quad (3)$$

$$g_2(T) = \sum_{j=1}^{m_1} b_j T_{p_j - q_j} \quad (4)$$

$$\text{or } g_2(T) = \sum_{i=1}^2 \left(b_{1i} \sin \frac{2\pi i t}{365} + b_{2i} \cos \frac{2\pi i t}{365} \right) \quad (5)$$

$$g_3(\theta) = c_1 \theta + c_2 \ln \theta \quad (6)$$

where H is the upstream and downstream water level difference; i represents the period, $i = 1$ represents the annual period, and $i = 2$ represents the half-year period in (4); m_1 is the number of days since the initial date. T_{p-q} is the average temperature from p to q days before the observation day; $\theta = (t - t_0)/100$, t is the measured date and t_0 is the initial date; $a_i, b_{1i}, b_{2i}, b_j, c_1, c_2$ are the regression coefficients.

More attention should be placed on the choice of two calculation methods for temperature displacement (see (4) and (5)). When the temperature data is complete and continuous, (4) is adopted to consider the influence of the actual temperature. When the temperature data is incomplete or discontinuous, (5) is used.

Substituting (2), (3), (4), (6) or (2), (3), (5), and (6) into (1), using Taylor series expansion, omitting high-order terms, and combining similar items, we can obtain the space-time distribution model of the fixed point $k(x, y, z)$ in the z direction, that is, the one-dimensional multipoint statistical deformation model.

$$\delta = \sum_{k=0}^4 \sum_{p=0}^3 A_{kp} H^k z^p + \sum_{k,j=1}^{m_1} \sum_{p=0}^3 B_{kjp} T_{p_j-q_j} z^p + \sum_{k,j=0}^1 \sum_{p=0}^3 C_{kjp} \theta_k \ln \theta_j z^p \quad (7)$$

$$\begin{aligned} \text{or } \delta = & \sum_{k=0}^4 \sum_{p=0}^3 A_{kp} H^k z^p \\ & + \sum_{k,j=0}^1 \sum_{p=0}^3 B_{kjp} \sin \frac{2\pi kt}{365} \cos \frac{2\pi jt}{365} z^p \\ & + \sum_{k,j=0}^1 \sum_{p=0}^3 C_{kjp} \theta_k \ln \theta_j z^p \end{aligned} \quad (8)$$

When the z coordinate of measuring point remains unchanged, a displacement statistical model of the single measuring point is obtained:

$$\delta = a_0 + \sum_{i=1}^4 a_i H^i + \sum_{k,j=1}^{m_1} B_{kjp} T_{p_j-q_j} + c_1 \theta + c_2 \ln \theta \quad (9)$$

$$\begin{aligned} \text{or } \delta = & a_0 + \sum_{i=1}^4 a_i H^i + \sum_{i=1}^2 \left(b_{1i} \sin \frac{2\pi it}{365} + b_{2i} \cos \frac{2\pi it}{365} \right) \\ & + c_1 \theta + c_2 \ln \theta \end{aligned} \quad (10)$$

According to the reasons mentioned above, this paper chooses (7) and (9) to study the deformation monitoring model of concrete dams. Therefore, the input variable of the single point deformation prediction model is $X = (H, H^2, H^3, H^4, T_{1-10}, T_{11-30}, T_{31-60}, T_{61-100}, \theta, \ln \theta)$, the input variable of the one-dimensional multipoint deformation prediction model is

$$X = \begin{pmatrix} H, H^2, H^3, H^4, z, z^2, z^3, T_{1-10}, T_{11-30}, T_{31-60}, T_{61-100}, \theta, \ln \theta \\ zH, zH^2, zH^3, zH^4, z^2H, z^2H^2, z^2H^3, z^2H^4, z^3H, z^3H^2, z^3H^3, z^3H^4, \\ zT_{1-10}, zT_{11-30}, zT_{31-60}, zT_{61-100}, z^2T_{1-10}, z^2T_{11-30}, z^2T_{31-60}, z^2T_{61-100}, z^3T_{1-10}, z^3T_{11-30}, z^3T_{31-60}, z^3T_{61-100}, \\ z\theta, z\ln \theta, z^2\theta, z^2\ln \theta, z^3\theta, z^3\ln \theta \end{pmatrix}, \quad (11)$$

and the output variable is the radial displacement of the measuring point.

3. Methodology

3.1. Multiple Linear Regression. Multiple linear regression (MLR) models are based on the linear correlation between dam effect quantities and environmental variables. When considering the relationship between the k independent variables ($X_1, X_2, X_3 \dots X_k$) and the dependent variable Y , a regression equation is established: $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + u_i$, where $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ are the regression coefficients to be estimated; $i = 1, 2, \dots, n$ (n is the sample size); u_i is the random error [20].

Assuming that the random errors are generally normal distributed $N(0, \sigma^2)$ and independent of each other, the multiple linear regression equation is represented by a matrix: $\mathbf{Y} = \mathbf{X}\beta + \mathbf{u}$, where $\mathbf{Y}$ is the vector of observations; β is the parameter vector; $\mathbf{X}$ is the constant vector; $\mathbf{u}$ is the random

error vector. There is a set of parameter estimates $\hat{\beta}$ such that the residual sum of squares $Q = \sum_{i=1}^n u_i^2 = \mathbf{u}^T \mathbf{u} = (\mathbf{Y} - \mathbf{X}\hat{\beta})^T (\mathbf{Y} - \mathbf{X}\hat{\beta})$ is the smallest; that is, the system of equations $\partial Q / \partial \hat{\beta} = 0$ is solved. Therefore, the overall parameter of the least squares estimation is $\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}$, the fitted model is $\mathbf{Y} = \mathbf{X}\hat{\beta}$, and the vector of the residuals is denoted by $\hat{\mathbf{u}} = \mathbf{Y} - \hat{\mathbf{Y}}$. The ultimate goal of the overall model is to minimize the sum of the squared deviations between the model predictions and the observations.

3.2. Stepwise Regression. For the MLR method, the more independent variables, the smaller the residual square sum Q , the better the regression equation effect, and the higher the prediction accuracy. In the optimal regression equation, it is always desirable to include as many independent variables as possible, especially the independent variables that have a significant influence on the dependent variable. Nonetheless, too many independent variables may also result in some

FIGURE 1: Structure of a typical single hidden layer feedforward neural network.

disadvantages of the regression equation. Firstly, if more independent variables are required, many quantities must be measured and calculations are inconvenient. Secondly, if the regression equation includes an independent variable that has no effect on the dependent variable or has a very small effect, then the residual square sum Q will not decrease, thus affecting the accuracy of the regression equation. Thirdly, the existence of independent variables that have no significant influence on the dependent variable affects the stability of the regression equation and reduces the prediction accuracy. Thus, in the optimal regression equation, it is desirable to exclude independent variables that have no significant effect on the dependent variable.

Stepwise regression (SR) is a method for a linear regression model to select independent variables [21]. The basic idea is to introduce variables one by one, with the condition that its partial regression squared and experience are significant. According to the above principle, stepwise regression can be used to screen and eliminate the variables causing multicollinearity. The specific steps are as follows: first, use Y to make a simple regression for each X_i considered and then gradually introduce the remaining X_{i-1} based on the regression equation corresponding to the X_i that contributes the most to Y . After a stepwise regression, X_i that is finally retained in the model is both important and not heavily multicollinear. The effect of stepwise regression on the improvement of multiple linear regression is still controversial, which is also a focus of this paper.

3.3. Backpropagation Neural Network. Artificial neural networks are often divided into two categories: one is a recursive network that generates loops through feedback connections, and the other is a feedforward neural network [22] in which the network structure has no loops. The typical single hidden layer feedforward neural network structure is shown in Figure 1. Both the ELM and BP neural network belong to the feedforward neural network, except that the learning methods of the two are different. The BP neural network is a learning method that uses backpropagation by the gradient descent method, which requires constant iteration to update the weights and thresholds, while the ELM randomly determines the initial weights and thresholds without adjustment.

The traditional BP neural network adopts the error backpropagation algorithm, whose guiding idea is that the weight and threshold should be adjusted along the direction

of negative gradient, which is the fastest descending error function. Supervised BP neural network learning algorithm usually consists of three stages [10].

The first stage is to feed the data forward, and the computed output of the i th node in the output layer is as follows:

$$y_i = g \left(\sum_{j=1}^{N_h} \left(w_{ij} g \left(\sum_{k=1}^{N_i} v_{jk} x_k + \theta_{vj} \right) + \theta_{wi} \right) \right) \quad (12)$$

$i = 1, 2, \dots, N_o$

where w_{ij} is the connective weight between nodes in the hidden layer and those in the output layer; v_{jk} represents the connective weight between the nodes in the input layer and those in the hidden layer; θ_{wi} (or θ_{vj}) are bias terms that represent the threshold of the transfer function g ; x_k is the input to the k th node in the input layer; N_i , N_h , and N_o are the number of nodes in the input, hidden layer, and output layer, respectively.

The second stage is the backpropagation of the error. The learning process of error backpropagation is the process of propagating errors from the output layer to the input layer and correcting the corresponding network parameters. The goal of learning is to minimize or reduce the total error of the network.

The third and final stage is to adjust the weights and thresholds. The training is performed using a gradient descent method with a learning ratio in the standard BP algorithm and is defined as follows:

$$\mathbf{W}^{(r+1)} = \mathbf{W}^{(r)} + \Delta \mathbf{W}^{(r)} \quad (13)$$

where $\mathbf{W} = (v_{11}v_{12} \cdots v_{1N_i} \cdots v_{N_hN_i} \theta_{v1} \theta_{v2} \cdots \theta_{vN_h} \times w_{11}w_{12} \cdots w_{ij} \cdots w_{N_hN_i} \theta_{w1} \theta_{w2} \cdots \theta_{wN_o})$; $\mathbf{W}$ is the parameter vector to be determined in the BP neural network. The weight is adjusted as follows:

$$\Delta \mathbf{W}^{(r)} = -\eta \frac{\partial E(\mathbf{W})}{\partial \mathbf{W}^{(r)}} \quad (14)$$

where η is the learning ratio; the superscript (r) refers to the r th learning iteration; and $E(\mathbf{W})$ is the system error function. The convergence speed of BP neural network depends on the learning speed. For computational efficiency, Levenberg-Marquardt [23] algorithm is applied to obtain w_{ij} , v_{jk} , θ_{wi} , and θ_{vj} by minimizing the system error function.

3.4. Extreme Learning Machine. Extreme learning machine is an algorithm for single hidden layer feedforward neural network. Suppose there are N arbitrary samples, where $\mathbf{X}_i = [x_{i1}, x_{i2}, \dots, x_{iN}]$ and $\mathbf{t}_i = [t_{i1}, t_{i2}, \dots, t_{im}]^T \in \mathbb{R}^m$. The output of a standard single hidden layer neural network with L hidden nodes can be mathematically described as follows:

$$\mathbf{o}_j = \sum_{i=1}^L \beta_i g(\mathbf{W}_i \cdot \mathbf{X}_j + b_i), \quad j = 1, 2, \dots, N \quad (15)$$

where $\mathbf{o}_j$ is the output vector relative to the input $\mathbf{X}_j$; $g(x)$ is the activation function; $\mathbf{W}_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T$ is the input

weight vector; $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T \in R^m$ is the output weight vector; b_i is the offset of the i th hidden layer; $W_i \cdot X_j$ is the inner product of W_i and X_j .

The learning goal of single hidden layer neural network is to minimize the output errors, which can be expressed as follows:

$$\sum_{j=1}^L \|\mathbf{o}_j - \mathbf{t}\| = 0 \quad (16)$$

That is, there exist specific β_i , W_i , and b_i such that

$$\sum_{i=1}^L \beta_i g(W_i \cdot X_j + b_i) = \mathbf{t}_j, \quad j = 1, 2, \dots, N \quad (17)$$

Equation (17) can be simplified as

$$H\beta = T \quad (18)$$

where T is the target matrix of training samples.

$$H(W_1, \dots, W_L, b_1, \dots, b_L, X_1, \dots, X_L) = \begin{bmatrix} g(W_1 \cdot X_1 + b_1) & \dots & g(W_L \cdot X_1 + b_L) \\ \vdots & \dots & \vdots \\ g(W_1 \cdot X_N + b_1) & \dots & g(W_L \cdot X_N + b_L) \end{bmatrix}_{N \times L} \quad (19)$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L \times m}, \quad (20)$$

$$T = \begin{bmatrix} T_1^T \\ \vdots \\ T_N^T \end{bmatrix}_{N \times m}$$

There are $\hat{\beta}_i$, $\hat{W}_i$, and $\hat{b}_i$ such that

$$\|H(\hat{W}_i, \hat{b}_i) \hat{\beta}_i - T\| = \min_{W, b, \beta} \|H(\hat{W}_i, \hat{b}_i) \hat{\beta}_i - T\| \quad (21)$$

where $i = 1, 2, \dots, L$; (21) is equivalent to the following minimization loss function:

$$E = \sum_{j=1}^N \left(\sum_{i=1}^L \beta_i g(W_i \cdot X_j + b_i) - \mathbf{t}_j \right)^2, \quad j = 1, 2, \dots, N \quad (22)$$

Conventional gradient-based learning algorithms require adjustment of all parameters over multiple iterations. In the ELM algorithm, once the input weights W_i and the hidden layer offsets b_i are randomly determined, the output matrix of the hidden layer is uniquely determined [15]. The training single hidden layer neural network can be transformed into solving a linear system $H\hat{\beta} = T$. And the output weight $\hat{\beta}$ can be determined

$$\hat{\beta} = H^T T \quad (23)$$

where H^T is the Moore-Penrose generalized inverse of H .

4. Data and Processing

4.1. Case Dam and Operation Data. The Dongjiang (DJ) dam is located in Zixing City, Hunan Province, China. It is a variable center double-curved concrete arch dam with a maximum height of 157m and a center arc length of 438m. Its left and right shores are basically symmetrical. The designed normal storage level is 285m above the sea level corresponding to a storage capacity of 8.12 billion m^3 . The first water impounding began in 1986 and has been going on for more than 20 years now. The DJ dam is equipped with monitoring items such as deformation and seepage. The deformation monitoring items include forward intersection, inverted perpendicular, vertical displacement, and cross-river length monitoring system. The layout of dam vertical monitoring system is shown in Figure 3. There are 5 sets of vertical lines, namely, L1, L3, L5, L7, and L9, and each vertical line has a vertical reversal line. The vertical line monitoring is available in both manual and automated monitoring.

The observation data of the measuring point L5-205 from February 2003 to December 2013 is the basis of this study. The data includes temperature, upstream and downstream water levels, and the radial deformation of measuring points. The position of the L5-205Z measuring point is shown in Figure 2.

Kang [17] has initially discussed the application of ELM in deformation prediction of Fengman (FM) concrete gravity dam, which will be presented here to be a reference for the research of this paper. Table 1 shows the basic parameters of the two dams, and Figure 3 shows a cross-sectional view of the two dams.

Figure 4 shows the time history curve of environmental variables and corresponding radial horizontal displacement of the L5-205Z measuring point, and the vertical dotted line marks the division between training and predicted sets. In Figure 4, sign (+) denotes the displacement towards the upstream, and sign (-) indicates the displacement towards the downstream. The radial displacement of the measuring point changes periodically with temperature, that is, the temperature rises, the measuring point deforms upstream, the temperature drops, and the measuring point deforms downstream. The change in displacement lags significantly behind the change in temperature. The highest temperature generally occurs between July and October, and the lowest temperature occurs between January and March. However, the displacement generally reaches a maximum value (maximum to the downstream) from April to June and reaches a minimum value (maximum to the upstream) from October to December. The radial displacement of the L5-205Z measuring point is in accordance with the general law of the arch dam deformation, which also proves the validity of the data samples from the side.

As shown in Figure 5 [17], the water level of the FM dam also changes periodically with the season. Unlike the DJ dam, the FM dam has hardly observed the hysteresis of the displacement change with respect to the temperature change. The existence of correlation among water level changes, temperature changes, and time changes may affect the prediction accuracy of conventional linear equations.

FIGURE 2: Distribution of vertical lines of the Dongjiang dam.

FIGURE 3: Cross section of dam (unit: m): (a) the DJ dam; (b) the FM dam.

4.2. Parameter Settings. For the BP neural network, the initial weights and thresholds were obtained by the most widely used Levenberg-Marquardt. The optimal number of hidden neurons and the optimal learning rate were determined by trial and error, while the transfer functions of the hidden layer and the output layer, respectively, selected the sigmoid function and the linear function. The number of training epochs was set as 10^3 and the training goal for the MSE was set as 10^{-3} . The activation function of the ELM model was also a sigmoid function. Compared with the BP model, the ELM model only needs to determine the number of hidden nodes to obtain satisfactory results.

Because the weights are randomly initialized, different results can be achieved by separate runs. In order to enhance

the reliability of calculation results, both the ELM and BP neural network are continuously trained 20 times, the result of a calculation with a small difference between the MSE of the training set and the predicting set will be the final result. Kang [17] used the average of 5 calculations as the final result, and this paper believes that 5 times may not be sufficient to prove the reliability of the results.

Figures 6 and 7 show the effect of the number of hidden layer neurons to the predictive performance of ANN models for the DJ dam. For the single point deformation monitoring model, the training and predicting errors of the BP model are relatively small when the number of hidden nodes is 16. And, for the ELM model, when the number of hidden nodes is 17, the error of training and predicting is relatively small. For

TABLE 1: Dimensional parameters of the dam considered.

Dam	Typology	Height (m)	Top elevation (m ab.sea level)	Impact factors	Output
Dongjiang(DJ)	ARC	157	294	H, T, t/ H, T, t, z	RAD
Fengman [17] (FM)	GRA	91	268.9	H, T, t	RAD

Typology: ARC= arch; GRA= gravity; outputs: RAD= radial displacement; H= water level; T= air temperature; t= time; z= Z coordinate;

FIGURE 4: Time series of environmental variables and response displacements at the DJ dam site. From top to bottom: mean daily air temperature, upper and lower water level difference, and measured horizontal displacements. The vertical dotted line marks the division between training and predicting periods.

the FM dam, Kang [17] determined the optimal number of hidden nodes of the BP neural network model is 15, and the ELM model is 22. For the multipoint deformation monitoring model, in the BP network and ELM models, the numbers of hidden nodes are set as 15 and 14, respectively.

4.3. Performance Evaluation. It is important to appropriately estimate the prediction error of a model, since (a) it provides insight into its accuracy, (b) it allows comparison of different models, and (c) it is used to define warning thresholds [24, 25]. In order to facilitate the analysis of final calculation results, different performance evaluation functions are adopted in this paper, that is, mean absolute error (MAE), mean square error (MSE), maximum absolute error (S), and correlation coefficient (R), as shown below [17].

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_D(i) - y(i)| \quad (24)$$

FIGURE 5: Time series of environmental variables and response displacements at the FM dam site (Kang F. et al., 2017). From top to bottom: mean daily air temperature, upper and lower water level difference, and measured horizontal displacements. The vertical dotted line marks the division between training and predicting periods.

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_D(i) - y(i))^2 \quad (25)$$

$$S = \max |y_D(i) - y(i)|, \quad i = 1, 2, \dots, N \quad (26)$$

$$R = \frac{\sum_{i=1}^N (y_D(i) - \bar{y}_D)(y(i) - \bar{y})}{\sqrt{\sum_{i=1}^N (y_D(i) - \bar{y}_D)^2 \sum_{i=1}^N (y(i) - \bar{y})^2}} \quad (27)$$

where y_D and $\bar{y}_D$ are simulation values and simulation averages; y and $\bar{y}$ are observed values and observed average values, respectively; N is the number of measured samples.

To estimate the uncertainty associated with model simulations, the residuals of predicting sets are computed and analyzed [26]. The independence analysis, heteroscedastic analysis, and normality analysis of residuals are performed by plotting graphs of residual autocorrelation, residual variation relative to observed values, and residual probability distributions. If the residual sequence is autocorrelated, then the

FIGURE 6: Effects of hidden nodes on model performance for the single point model.

FIGURE 7: Effects of hidden nodes on model performance for the one-dimensional multipoint model.

model fails to fully explain the variation rule of the variable. On the other hand, low residual heteroscedasticity and a close approximation to the normal distribution indicate the model is closer to unbiased estimation and has low uncertainty. In this paper, the standardized residual of the model is shown in

$$r_s = \frac{e_i}{\sigma} = \frac{y - y_D}{\sigma} \quad (28)$$

where y represents observed values, y_D represents predicted values, r_s represents standardized residuals, and σ represents the standard deviation. $e_i = y - y_D$, and e_i represents residuals.

5. Results and Discussion

5.1. Comparison of Simulation Results. In this section, the observed DJ dam deformations are compared with simulated results based on the four different models, i.e., the MLR, SR,

BP neural network, and ELM model. The specific calculation results for both the DJ arch dam and the FM gravity dam of the four models are shown in Tables 2-3 and Figures 8-10. The calculation results show that, in the single point deformation monitoring model, the best MAE, MSE, S, and R values are obtained by the ELM models for both the DJ arch dam and the FM gravity dam. And the best results are highlighted in black bold. According to the comparison among different models, the best accuracy ranking for the DJ dam is the ELM model > BP neural network > SR > MLR, while for the FM dam, the ranking is the ELM model > BP neural network > MLR > SR. The stepwise regression does not always play a positive role in the improvement of multiple linear regression and should be analyzed in specific situations.

It can be seen from Figure 8 that, except for the multimeasurement MLR model, the fitting and predicted displacements of all the models are consistent with the trend

TABLE 2: Statistical performance of the MLR, SR, BP, and ELM single-point models.

DAM	model	MAE		MSE		S		R		Time (s)
		Training	Predicting	Training	Predicting	Training	Predicting	Training	Predicting	
DJ	MLR	0.5194	0.4962	0.4710	0.5599	3.3345	3.2524	0.9519	0.9054	1.3800
	SR	0.5210	0.6517	0.4999	0.6951	3.4872	3.1291	0.9489	0.9143	1.3910
	BP	0.4325	0.5889	0.3595	0.7087	3.7087	3.4262	0.9649	0.9360	2.4140
	ELM	0.4240	0.4555	0.3504	0.5005	3.3246	2.4620	0.9652	0.9420	1.4360
FM [17]	MLR	0.8355	1.0673	0.9947	1.7170	2.0376	2.8455	0.9388	0.9190	2.1619
	SR	0.9097	1.0959	1.1750	1.6609	2.5263	2.3165	0.9361	0.9099	2.7656
	BP	0.4270	0.8491	0.3505	1.1311	2.2703	2.3678	0.9824	0.9466	5.3310
	ELM	0.3978	0.7020	0.2653	0.8929	1.6336	2.2076	0.9840	0.9564	2.9265

MAE = mean absolute error; MSE = mean square error; S = maximum absolute error; R = the correlation coefficient; MLR = multiple linear regression; SR = stepwise regression; BP = backpropagation; ELM = extreme learning machine. DJ = the Dongjiang arch dam; FM = the Fengman gravity dam.

TABLE 3: Statistical performance of the MLR, SR, BP, and ELM multipoint models.

model	MAE		MSE		S		R		Time (s)
	Training	Predicting	Training	Predicting	Training	Predicting	Training	Predicting	
MLR	0.3586	1.1977	0.3510	2.2516	3.5438	4.0139	0.7621	0.7541	1.4200
SR	0.5224	0.4225	0.4729	0.4777	3.2501	3.2718	0.9517	0.9188	1.6910
BP	0.5609	0.5824	0.3609	0.4542	2.7205	2.8618	0.9759	0.9543	3.0750
ELM	0.3571	0.4117	0.3073	0.4016	1.7774	2.5375	0.9792	0.9558	1.7010

MAE = mean absolute error; MSE = mean square error; S = maximum absolute error; R = the correlation coefficient; MLR = multiple linear regression; SR = stepwise regression; BP = backpropagation; ELM = extreme learning machine.

of the measured displacement. In June 2004, all models showed similarly large errors. The reason is that the high temperature generated by the high voltage line affects the measurement accuracy.

With respect to the multipoint deformation monitoring model, the best MAE, MSE, S, and R values are also obtained by the ELM models for the DJ arch dam. And the best accuracy ranking is the same as the point deformation monitoring model, which is the ELM model > BP neural network > SR > MLR. Nonetheless, in the multipoint deformation monitoring model, the prediction accuracy of the MLR model drops sharply due to the inclusion of too many redundant independent variables. And the SR, BP, and ELM perform better in the multipoint deformation modelling, and the prediction accuracy is higher, which indicates that the multipoint model is more reasonable than the single point model.

In addition to the simulation accuracy, the calculation speed is also an important index to measure the performance of a model. In this paper, the time consumption is used as an evaluation index to compare the calculation speed of the four models. In general, the time consumption is ranked as the BP neural network > ELM > SR > MLR model among the different models and as the multipoint deformation model > the single point deformation model among the different measuring points. It should be noted that the BP neural network is the longest-running model and the time consumption of the ELM model is significantly lower than that of the BP neural network. Table 2 (FM dam) is Table 5 in [Kang et al., 2017].

5.2. Residuals Analysis. As shown in Figures 11–12, to evaluate the uncertainty of the models, residual analysis is performed on the statistical results of the four models. In the single point deformation monitoring models, the experimental results show that the autocorrelation of r_s is almost not found in all four models, and the ACF lies mainly in the 95% confidence interval (Figures 11(a)–11(d)). Figures 11(e)–11(h) show the scatter points of r_s as a function of observed deformations. It is clear that the r_s values do not appear to be randomly distributed over the deformation interval except for the ELM model. And the r_s of the other three models show a decreasing trend with an increase in deformation. Figures 11(i)–11(l) display the probability density distribution of r_s for all the four models. The results show that the probability density distribution curve of r_s for all four models is unimodal without considering the influence of two abnormal points, and the values of r_s are mainly distributed between -2 and 2 (Figures 11(e)–11(h)). The existence of two abnormal points was caused by the measurement anomaly in December 2013, when the vertical line was being overhauled.

In the multipoint deformation monitoring models, the experimental results show that the autocorrelation of r_s is almost not found in the SR, BP and ELM models, whereas the r_s values of the MLR model show remarkable autocorrelation (Figures 12(a)–12(d)). The r_s values of the MLR model exhibit heteroscedasticity as the observed outflow changes. Compared with the MLR, SR, and BP neural network, the spatial distribution of r_s with observed deformation for the ELM model is relatively uniform (Figures 12(e)–12(h)). The probability density of r_s for the MLR displays a multimodal

FIGURE 8: Performance of the MLR (a), SR (b), BP (c), and ELM (d) single point models for the DJ dam: fitting of measured values and simulated values. The vertical dotted line marks the division between training and predicting periods.

FIGURE 9: Performance of the MLR (a), SR (b), BP (c), and ELM (d) single point models for the FM dam [17]: fitting of measured values and simulated values (above) and residual error (below). The vertical dotted line marks the division between training and predicting periods.

FIGURE 10: Performance of the MLR (a), SR (b), BP (c), and ELM (d) multipoint models for the DJ dam: fitting of measured values and simulated values. The vertical dotted line marks the division between training and predicting periods.

FIGURE 11: Investigation of residuals of MLR (a, e, i), SR (b, f, j), BP neural network (c, g, k), and ELM (d, h, l) for the single point model. (a-d) Autocorrelation function (ACF) plots of r_s at with 95% significance levels. (e-h) r_s as a function of the observed deformation. (i-l) Fitted (solid line) and actual (bars) probability density function (PDF) of r_s .

FIGURE 12: Investigation of residuals of MLR (a, e, i), SR (b, f, j), BP neural network (c, g, k), and ELM (d, h, l) for the one-dimensional multipoint model. (a–d) Autocorrelation function (ACF) plots of r_s at with 95% significance levels. (e–h) r_s as a function of the observed deformation. (i–l) Fitted (solid line) and actual (bars) probability density function (PDF) of the r_s .

distribution, with four peaks (one high and three low) distributed at -1.4 , 0.4 , 1 , and 2.2 , respectively (Figure 12(i)). The r_s of the ELM model presents a unimodal distribution with a sharp peak without considering the influence of two abnormal points, and r_s is mainly distributed between -1.5 and 1.2 (Figure 12(l)).

6. Conclusion

This paper investigated the usefulness of two traditional multiple regression models (MLR and SR) and two artificial neural network models (ELM and BP neural network) in predicting dam deformation. All the four models presented here have the advantages of simple operation and fast application, which increases the confidence in using these models.

The artificial neural networks (ELM and BP) can significantly improve the accuracy of conventional statistical methods (MLR and SR) for predicting the behavior of concrete dams and have good adaptability and generalization ability for deformation predicting of concrete dams. Compared with the BP model, the ELM model has fewer adjustment parameters, faster learning, and higher efficiency. If there is a high accuracy requirement for concrete dam deformation prediction, the ELM model would be optimal.

The one-dimensional deformation multipoint monitoring model can reflect the deformation distribution in the one-dimensional direction of the arch dam, with clear physical

concepts and spatial characteristics. Compared with the single point model, it has better anti-interference ability and higher prediction accuracy. In general, for the single point deformation monitoring model, the four models mentioned in this paper can meet the engineering needs. Nonetheless, artificial neural networks are a better choice when considering the interaction of measuring points. Among them, the ELM model can effectively solve the time consumption problem associated with the BP neural network, and it has superior performances over other three models in simulating dam deformation.

Obviously, artificial neural network-based models are more suitable for reproducing nonlinear effects and complex interactions between input variables and dam responses. Nonetheless, the determination of the number of hidden nodes is the key and difficult point that artificial neural networks are difficult to avoid. In order to overcome the error caused by randomness and improve the generalization ability model based on ELM, evolutionary algorithms, such as artificial bee colony (ABC) algorithm [27] or particle swarm algorithm, can be used to optimize the ELM model, which is the next research goal.

Data Availability

(1) The initial observation data of Dongjiang dam used to support the findings of this study were supplied by Hunan

Electric Power Company Science Research Institute under license and so cannot be made freely available. Requests for access to these data should be made to Tianhaiping, 329971674@qq.com. (2) The calculated data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] F. Salazar, R. Morán, M. Á. Toledo, and E. Oñate, "Data-based models for the prediction of dam behaviour: a review and some methodological considerations," *Archives of Computational Methods in Engineering: State-of-the-Art Reviews*, vol. 24, no. 1, 2017.
- [2] F. Salazar, M. A. Toledo, E. Oñate, and R. Morán, "An empirical comparison of machine learning techniques for dam behaviour modelling," *Structural Safety*, vol. 56, pp. 9–17, 2015.
- [3] G. Liang, Y. Hu, and Q. Li, "Safety monitoring of high arch dams in initial operation period using vector error correction model," *Rock Mechanics and Rock Engineering*, vol. 51, no. 8, pp. 2469–2481, 2018.
- [4] Q. B. Li, G. H. Liang, Y. Hu, and Z. Zuo, "Numerical analysis on temperature rise of a concrete arch dam after sealing based on measured data," *Mathematical Problems in Engineering*, vol. 2014, Article ID 602818, 12 pages, 2014.
- [5] J. Mata, A. Tavares de Castro, and J. Sá da Costa, "Constructing statistical models for arch dam deformation," *Structural Control and Health Monitoring*, vol. 21, no. 3, pp. 423–427, 2014.
- [6] A. Bayer, M. Bachmann, A. Müller, and H. Kaufmann, "A Comparison of feature-based MLR and PLS regression techniques for the prediction of three soil constituents in a degraded South African ecosystem," *Applied and Environmental Soil Science*, vol. 2012, Article ID 971252, 20 pages, 2012.
- [7] M. O. Akinwande, H. G. Dikko, and A. Samson, "Variance inflation factor: as a condition for the inclusion of suppressor variable(s) in regression analysis," *Open Journal of Statistics*, vol. 05, no. 07, pp. 754–767, 2015.
- [8] A. G. Mulat and S. A. Moges, "Assessment of the impact of the grand ethiopian renaissance dam on the performance of the high aswan dam," *Journal of Water Resource and Protection*, vol. 06, no. 06, pp. 583–598, 2014.
- [9] M. H. Ali, M. R. Alam, M. N. Haque, and M. J. Alam, "Comparison of design and analysis of concrete gravity dam," *Natural Resources*, vol. 03, no. 01, pp. 18–28, 2012.
- [10] C. H. Chen, "Structural identification from field measurement data using a neural network," *Smart Materials and Structures*, vol. 14, no. 3, pp. S104–S115, 2005.
- [11] B. Stojanovic, M. Milivojevic, M. Ivanovic, N. Milivojevic, and D. Divac, "Adaptive system for dam behavior modeling based on linear regression and genetic algorithms," *Advances in Engineering Software*, vol. 65, pp. 182–190, 2013.
- [12] J. Mata, "Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models," *Engineering Structures*, vol. 33, no. 3, pp. 903–910, 2011.
- [13] C. Y. Kao and C. H. Loh, "Monitoring of long-term static deformation data of Fei-Tsui arch dam using artificial neural network-based approaches," *Structural Control and Health Monitoring*, vol. 20, no. 3, pp. 282–303, 2013.
- [14] H. Su, Z. Chen, and Z. Wen, "Performance improvement method of support vector machine-based model monitoring dam safety," *Structural Control and Health Monitoring*, vol. 23, no. 2, pp. 252–266, 2016.
- [15] G. Huang, S. Song, and K. You, "Trends in extreme learning machines," *Neural Networks*, vol. 61, pp. 32–48, 2015.
- [16] G. B. Huang, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, 2006.
- [17] F. Kang, J. Liu, J. Li, and S. Li, "Concrete dam deformation prediction model for health monitoring based on extreme learning machine," *Structural Control and Health Monitoring*, vol. 24, no. 10, 2017.
- [18] M. A. Hariri-Ardebili and H. Mirzabozorg, "Feasibility study of dez arch dam heightening based on nonlinear numerical analysis of existing dam," *Archives of Civil Engineering*, vol. 59, no. 1, pp. 21–49, 2013.
- [19] B. Dai, C. Gu, E. Zhao, and X. Qin, "Statistical model optimized random forest regression model for concrete dam deformation monitoring," *Structural Control and Health Monitoring*, vol. 25, no. 6, 2018.
- [20] V. Ranković, A. Novaković, N. Grujović, D. Divac, and N. Milivojević, "Predicting piezometric water level in dams via artificial neural networks," *Neural Computing and Applications*, vol. 24, no. 5, pp. 1115–1121, 2014.
- [21] Z. Pengfei and L. Zeyu, "Prediction of urban water consumption based on SPSS multiple linear regression model," *Water Science and Technology and Economy*, vol. 05, pp. 6–10, 2018.
- [22] A. Zeroual, A. Fourar, and M. Djeddou, "Predictive modeling of static and seismic stability of small homogeneous earth dams using artificial neural network," *Arabian Journal of Geosciences*, vol. 12, no. 2, 2019.
- [23] A. K. Singh, B. Tyagi, and V. Kumar, "ANN controller for binary distillation column - A Marquardt-Levenberg approach," in *Proceedings of the 2011 Nirma University International Conference on Engineering: Current Trends in Technology, NUiCONE 2011*, India, December 2011.
- [24] F. Salazar, M. Á. Toledo, J. M. González, and E. Oñate, "Early detection of anomalies in dam performance: A methodology based on boosted regression trees," *Structural Control and Health Monitoring*, vol. 24, no. 11, 2017.
- [25] A. Swanepoel, S. Barnard, F. Recknagel, and H. Cao, "Evaluation of models generated via hybrid evolutionary algorithms for the prediction of Microcystis concentrations in the Vaal Dam, South Africa," *Water SA*, vol. 42, no. 2, pp. 243–252, 2016.
- [26] D. Zhang, J. Lin, Q. Peng et al., "Modeling and simulating of reservoir operation using the artificial neural network, support vector regression, deep learning algorithm," *Journal of Hydrology*, vol. 565, pp. 720–736, 2018.
- [27] B. Yilmaz, E. Aras, S. Nacar, and M. Kankal, "Estimating suspended sediment load with multivariate adaptive regression spline, teaching-learning based optimization, and artificial bee colony models," *Science of the Total Environment*, vol. 639, pp. 826–840, 2018.

Submit your manuscripts at
www.hindawi.com