Hindawi
Journal of Sensors
Volume 2019, Article ID 4581672, 14 pages
https://doi.org/10.1155/2019/4581672

Research Article

A Deep Learning Model for Concrete Dam Deformation Prediction Based on RS-LSTM

Xudong Qu[1], **Jie Yang**[1] **and Meng Chang**[2]

[1]*Institute of Water Resources and Hydro-Electric Engineering, Xi'an University of Technology, Xi'an 710048, China*
[2]*Department of Development, Sino Hydro Engineering Bureau 15 Co., Ltd, Xi'an 710016, China*

Correspondence should be addressed to Xudong Qu; qxd@stu.xaut.edu.cn

Received 23 June 2019; Accepted 14 September 2019; Published 31 October 2019

Guest Editor: Samir Mustapha

Deformation is a comprehensive reflection of the structural state of a concrete dam, and research on prediction models for concrete dam deformation provides the basis for safety monitoring and early warning strategies. This paper focuses on practical problems such as multicollinearity among factors; the subjectivity of factor selection; robustness, externality, generalization, and integrity deficiencies; and the unsoundness of evaluation systems for prediction models. Based on rough set (RS) theory and a long short-term memory (LSTM) network, single-point and multipoint concrete dam deformation prediction models for health monitoring based on RS-LSTM are studied. Moreover, a new prediction model evaluation system is proposed, and the model accuracy, robustness, externality, and generalization are defined as quantitative evaluation indexes. An engineering project shows that the concrete dam deformation prediction models based on RS-LSTM can quantitatively obtain the representative factors that affect dam deformation and the importance of each factor relative to the effect. The accuracy evaluation index (AVI), robustness evaluation index (RVI), externality evaluation index (EVI), and generalization evaluation index (GVI) of the model are superior to the evaluation indexes of existing shallow neural network models and statistical models according to the new evaluation system, which can estimate the comprehensive performance of prediction models. The prediction model for concrete dam deformation based on RS-LSTM optimizes the factors that influence the model, quantitatively determines the importance of each factor, and provides high-performance, synchronous, and dynamic predictions for concrete dam behaviours; therefore, the model has strong engineering practicality.

1. Introduction

Due to unique advantages in design, construction, and operational management, concrete dams account for a large proportion of all dams and have become the preferred dam type for the construction of high dams. However, most of the concrete dam projects are located in harsh alpine valleys. Thus, the dams are subjected to various dynamic, static, and special cyclic loads during service, and the design, construction, and operational management must be tailored to these conditions. Therefore, service safety behaviour involves a nonlinear dynamic process that includes material and structure interactions and multiple factors [1]. As a comprehensive variable that reflects the safety state of concrete dams, deformation can be used as an important index of structural behaviours and trends. Therefore, strengthening the prediction models for deformation, conducting safety monitoring, and establishing early warning systems are important ways to ensure long-term service safety of concrete dams [2].

In recent years, the successful application of dam engineering theory, finite element theory, and artificial intelligence (AI) technology has greatly promoted the development of concrete dam deformation prediction models. The most commonly used methods [3] for influential factor selection in concrete dam deformation prediction models include prior knowledge, linear correlation coefficient, stepwise regression, principal component analysis (PCA), and grey correlation analysis methods. However, in actual applications, the prior knowledge method relies too much on experience and has large errors. Notably, the water pressure, temperature, and dam age are generally selected as influential factors in hydrostatic seasonal temporal (HST) models considering

simplified physical models of dams and dam foundations, the burial conditions of monitoring equipment, prototype monitoring data, engineering mechanical analysis, and deductive investigation. The limitation of the PCA method is that only linear relations between variables are considered. If the dependence is nonlinear, the misinterpretation of results may occur. The grey correlation analysis method can only sort factors according to their relevance, and there is no clear criterion for selecting influential factors. Moreover, multiple collinearity can exist among the factors selected by conventional methods, which may reduce the accuracy of the model and adversely affect the prediction results [4]. Meanwhile, prediction models do not consider the influence of nonquantitative factors such as the seepage flow, crack opening degree, and lifting pressure; the dam construction materials; the construction quality; and the geological conditions. Additionally, model interpretation is important for evaluating the performance of prediction models, especially the model accuracy. The HST model has been traditionally used to identify the response of a dam to a considered action, such as a hydrostatic load, or to variations in factors such as temperature and time [5]. However, such analyses are only valid if the predictor variables are independent, which is not generally true [6]. In contrast, intelligent models (such as neural network, multilayer perceptron, and support vector machine models) have not been applied to interpret dam behaviour. Traditional models are frequently termed "black box" models, in reference to their lack of interpretability. Therefore, in the selection process of the factors that influence concrete dam deformation prediction models, imperfect selection criteria and neglecting important factors can seriously affect the prediction performance of the model. Single-point statistical models, deterministic models, and hybrid models [7–10] have evolved into multipoint intelligent models [11–16]. Based on the traditional statistical model, Gu et al. treated deformation at multiple measurement points and the spatial coordinates of these points as variables and established a spatiotemporal distributed prediction model of the deformation field of a concrete dam. Li et al. investigated the spatial and temporal expression of the factors that affected the deformation of an RCC dam and established a spatiotemporal deformation prediction model for RCC dams based on measured data. The prediction results agreed to the actual dam deformation data. Li et al. used the strong functional nonlinear mapping ability of a back propagation (BP) neural network to replace the complex factor subset in the traditional spatial deformation field model with water level, temperature, time, and measurement point variables as the input of the neural network. A BP network prediction model was established for dam deformation at multiple points. Chen et al. proposed a spectral decomposition method to decompose the monitoring data collected at multiple measurement points into several mutually independent latent variables for noise reduction and monitoring data processing. A least square support vector machine prediction model was established between the environmental data and latent variables, and the horizontal displacement of Mianhuatan Dam was successfully predicted. Many scholars have addressed these issues. The successful application of new methods has expanded the theoretical knowledge of dam deformation prediction and model establishment and provided important guiding significance for engineering practice. However, due to the complexity of concrete dam engineering, the structural volatility of dams, and the uncertainty of working conditions, there are still some shortcomings in existing prediction models. It is difficult for some models to process massive amounts of monitoring data in real time with extensive mining data mechanisms for high-performance prediction targets, such as those in practical applications. It is important to appropriately evaluate the prediction performance of a model from all angles because the practical value of the models can be guaranteed, different models can be compared, and different warning thresholds can be defined. There are various indexes [17] that can be used to assess how well a model matches the observed data, among which the most commonly used are the mean squared error (MSE), root mean squared error (RMSE), coefficient of determination (R^2), mean absolute error (MAE), mean absolute percentage error (MAPE), and average relative variance (ARV). The result of any of these indexes is frequently equivalent to a given prediction task. Specifically, an accurate model will have small MSE, RMSE, MAE, and MAPE values and high R^2 and ARV values. However, these accuracy indexes have differences that can be relevant but are often not considered [18]. Commonly, robustness and generalization ability are neglected in the model assessment, and quantitative evaluation indexes are not always used in practical applications. Therefore, it is necessary to explore methods for factor selection, establish high-performance, dynamic, synchronous prediction models, and design a scientific and comprehensive evaluation system which are urgent for concrete dam deformation prediction.

Attribute reduction is one of the core concepts of RS theory, which addresses incompleteness, redundancy, and ambiguity in data in the field of machine learning. This approach avoids the use of complex discernibility matrices and uses attribute importance as heuristic information to obtain inductive sets and importance analysis results; excellent results can be obtained in factor selection for prediction models based on RS theory [19–21]. Moreover, long short-term memory (LSTM) based on the memory architecture in deep learning (DL) can overcome the memory shortage and vanishing gradient issues of recurrent neural networks (RNNs). Besides, this method is characterized by controllable memory and rapid convergence. LSTM has achieved good practical application results in the dynamic and deep processing of massive, long-term, dependent data series [22–25]. To overcome the shortcomings of existing concrete dam deformation prediction models, RS theory and an LSTM network are applied to a concrete dam deformation prediction model in virtue of Tensor Flow. Finally, a concrete dam deformation prediction model based on RS-LSTM is established, and a new predictive model evaluation system is proposed.

2. Materials and Methods

2.1. Rough Set Theory. RS theory was proposed by Polish scholar Pawlak in the 1980s. The core objectives are the

mining and refining of essential information under the premise of maintaining equivalence relations. The main tasks in this approach are attribute reduction, correlation analysis, and importance evaluation for uncertain information systems.

2.1.1. Information System. To describe the samples that encompass the necessary information in RS theory, a quaternary information system S is established, and it can be expressed as follows:

$$S = \{U, R, V, f\}, \tag{1}$$

where U is a nonempty finite set of all samples; R is a set of attributes, including a set of conditional attributes C and a set of decision attributes D; V is the attribute value set; and f is the information function, also known as the decision table.

2.1.2. Attribute Reduction. For arbitrary $P \subseteq R$ and $P \neq \emptyset$, the indistinguishable relationship between P and U is defined as follows.

$$\mathrm{IND}(P) = \left\{(x, y) \in U^2 | \forall \alpha \in P, \alpha(x) = \alpha(y)\right\}. \tag{2}$$

For an arbitrary set of objects $X \subseteq U$ and attributes $B \subseteq C$ in a given information system S, the approximation of X is defined as $B\underline{X} = \{x|[x]_B \subseteq X\}$; the approximate definition of X is defined as $\bar{B}X = \{x|[x]_B \cup X \neq \emptyset\}$; and the boundary area of x is defined as $BN_B(X) = \bar{B}X - \underline{B}X$. In this case, $[x]_B$ represents the set of indistinguishable relations for the division of U by B.

If $BN_B(X)$ is not empty, then X is called a rough set of B. The positive region of B relative to D is as follows.

$$\mathrm{POS}_B(D) = \left\{\frac{\underline{B}X | X \in U}{\mathrm{IND}(D)}\right\}. \tag{3}$$

When $\mathrm{SIM} = \mathrm{POS}_C(D) - \mathrm{POS}_{C-\{a\}}(D) = 0$, where $a \in C$, a can be omitted. Additionally, when each element in C is not omissible from D, it can be concluded that C is independent of D. When $C' = C - C^*$, where C' is independent of D and all the elements in C^* can be omitted, then C' is called the relative reduction of D.

2.1.3. Importance Evaluation. In attribute reduction, the importance of the attribute can be defined by the degree of interdependence between the attribute sets B and D. The degree of interdependence between P and R is defined as follows:

$$\gamma_B(D) = \frac{|\mathrm{POS}_B(D)|}{|U|}, \tag{4}$$

where $|\cdot|$ represents the cardinality value of a set.

The importance of the conditional attribute a to the decision attribute D based on the attribute dependency degree is defined as follows.

$$\mathrm{Sig}(\alpha, B, D) = \gamma_B(D) - \gamma_{B-\{\alpha\}}(D). \tag{5}$$

2.2. LSTM Network Based on a Memory Architecture. LSTM is obtained by improving the hidden layer of the RNN structure. LSTM based on a memory architecture can overcome memory shortage and vanishing gradient problems. The LSTM model structure is shown in Figure 1. The key advantages of LSTM are twofold. Notably, the hidden layer includes a hidden state and a cell state, and a threshold mechanism is established in the RNN. These factors strengthen the ability of the model to learn current information, extract the information and rules associated with the data, and simultaneously transmit information to reduce memory use. The threshold mechanism uses input gates, forget gates, and output gates to selectively memorize the feedback parameters of the feedback error function as the gradient decreases, achieving rapid gradient convergence [26].

2.2.1. Input Gate Updates. The input gate controls the information $x^{(t)}$ transmitted from the input of the network at moment t and hidden state at the final moment $h^{(t-1)}$ to the cell state $C^{(t)}$. The function of the input gate is to filter new information. The structure of an input gate is shown in Figure 2.

Figure 2 shows that the input gate consists of two parts. The first part selects the sigmoid activation function, for which the output is $i^{(t)}$, and the second part selects the tanh activation function, for which the output is $a^{(t)}$. The two partial outputs are multiplied to update the cell state. The renewal process can be mathematically expressed as follows:

$$\begin{aligned} i^{(t)} &= \sigma\left(W_i h^{(t-1)} + U_i x^{(t)} + b_i\right), \\ a^{(t)} &= \tanh\left(W_a h^{(t-1)} + U_a x^{(t)} + b_a\right), \end{aligned} \tag{6}$$

where W_i, U_i, b_i, W_a, U_a, and b_a are the weights and biases of the input gate and σ is the sigmoid activation function.

2.2.2. Forget Gate Updates. The forget gate controls the information transmitted from the cell state $C^{(t-1)}$ at moment $t-1$ to the cell state $C^{(t)}$ at moment t, and the information that should be discarded is identified. The structure of the forget gate is shown in Figure 3.

Figure 3 shows that the hidden state $h^{(t-1)}$ at moment $t-1$ and the input $x^{(t)}$ at moment t activate the sigmoid function, and the output $f(t)$ is in the range of [0, 1]. This value represents the probability of forgetting the information associated with the cell state at a previous moment. The renewal process can be mathematically expressed as follows:

$$f^{(t)} = \sigma\left(W_f h^{(t-1)} + U_f x^{(t)} + b_f\right), \tag{7}$$

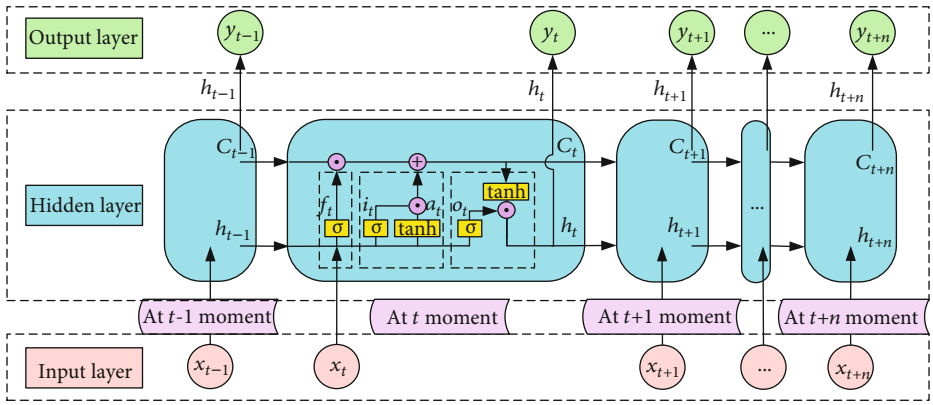

Figure 1: Long short-term memory network model structure.

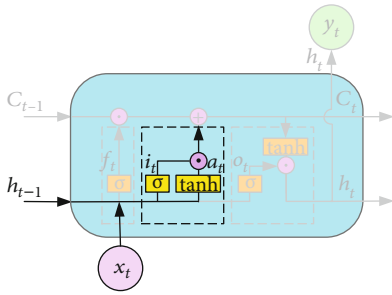

Figure 2: Input gate structure.

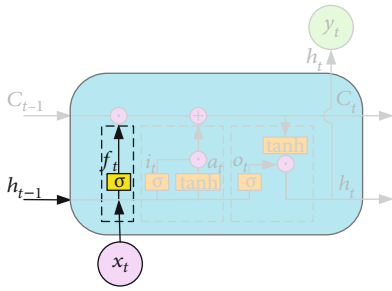

Figure 3: Forget gate structure.

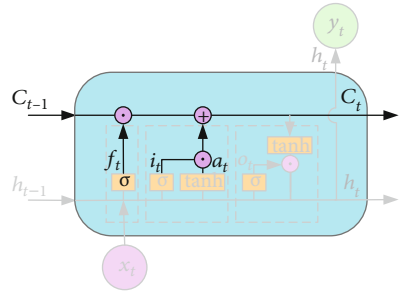

Figure 4: Cell state structure.

where W_f, U_f, and b_f are the weights and biases of the forget gate.

2.2.3. Cell State Updates. The cell state controls the information $a^{(t)}$ transmitted from the result of the input gate $f(t)$ and the result of the forget gate $i^{(t)}$ to the cell state $C^{(t)}$. The structure of a cell state is shown in Figure 4.

Figure 4 shows that the cell state updating result $C^{(t)}$ is mainly determined by the cell state $C^{(t-1)}$ at moment $t-1$ and the results of the input and forget gates ($f(t)$, $i^{(t)}$, and $a^{(t)}$) at moment t. The renewal process can be mathematically expressed as follows:

$$C^{(t)} = C^{(t-1)} \odot f^{(t)} + i^{(t)} \odot a^{(t)}, \tag{8}$$

where $\odot$ is the Hadamard product.

2.2.4. Output Gate Updates. The output gate controls the information transmitted from the hidden state $h^{(t-1)}$ at moment $t-1$, the cell state $C^{(t)}$ at moment t, and the input $x^{(t)}$ at moment t. The function of the output gate is to determine the final retained information. The structure of an output gate is shown in Figure 5.

Figure 5 shows that the hidden state $h^{(t)}$ at moment t contains two parts. The first part $o^{(t)}$ is determined by the hidden state $h^{(t-1)}$ at moment $t-1$, the input $x^{(t)}$ at moment t, and the sigmoid activation function. The other part is determined by the cell state $C^{(t)}$ at moment t and the tanh activation function. The renewal process can be mathematically expressed as follows:

$$\begin{aligned} o^{(t)} &= \sigma\left(W_o h^{(t-1)} + U_o x^{(t)} + b_o\right), \\ h^{(t)} &= o^{(t)} \odot \tanh\left(C^{(t)}\right), \end{aligned} \tag{9}$$

where W_f, U_f, and b_f are the weights and biases of the output gate.

2.2.5. Output Layer Updates. The output of the model is determined by the hidden state $h^{(t)}$ at moment t and the sigmoid activation function. The renewal process can be mathematically expressed as follows:

$$y\wedge^{(t)} = \sigma\left(Vh^{(t)} + c\right), \tag{10}$$

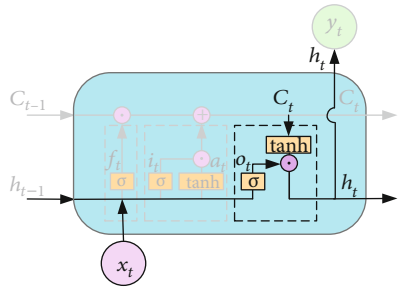

Figure 5: Output gate structure.

where V and c are the weight and bias of the output layer, respectively.

2.2.6. Model Parameter Updating. To obtain the optimal solution, this paper iteratively updates all the parameters in the LSTM model based on the gradient descent algorithm and error BP algorithm.

The objective loss function $L(t)$ is defined to minimize the sum of squared residuals between the predictions $y\wedge^{(t)}$ of the output layer and the target outputs $y^{(t)}$. $L(t)$ is divided into two parts: the loss $l(t)$ at moment t and the subsequent loss $l(t+1)$ moments later.

$$L(t) = \begin{cases} l(t) + l(t+1), & t < \tau, \\ l(t), & t = \tau. \end{cases} \tag{11}$$

The gradients of the hidden state $h^{(t)}$ and cell state $C^{(t)}$ are defined as $\delta_h^{(t)}$ and $\delta_C^{(t)}$, respectively, and the gradient at position τ can be expressed as follows.

$$\delta_h^{(\tau)} = \frac{\partial L(\tau)}{\partial h^{(\tau)}} = \frac{\partial L(\tau)}{\partial O(\tau)} \frac{\partial O(\tau)}{\partial h^{(\tau)}} = V^T \left(y\wedge^{(\tau)} - y^{(\tau)} \right), \tag{12}$$

$$\delta_C^{(\tau)} = \frac{\partial L(\tau)}{\partial C^{(\tau)}} = \frac{\partial L(\tau)}{\partial h(\tau)} \frac{\partial h(\tau)}{\partial C(\tau)} = \partial h(\tau) \odot o^{(\tau)} \odot \cdot \left(1 - \tanh^2 \left(C^{(\tau)} \right) \right). \tag{13}$$

The output gradient error at a given moment is determined in two parts, respectively, because $\delta_h^{(t)}$ and $\delta_C^{(t)}$ are obtained for $l(t)$ and $l(t+1)$. Thus, according to equations (12) and (13), the gradients of the hidden state $h^{(t)}$ and cell state $C^{(t)}$ can be expressed as follows.

$$\delta_h^{(t)} = \frac{\partial L}{\partial h^{(t)}} = V^T \left(y\wedge^{(t)} - y^{(t)} \right) + \delta_h^{(t+1)} \partial h^{(t+1)} / \partial h^{(t)}, \tag{14}$$

$$\delta_C^{(t)} = \frac{\partial L}{\partial C^{(t)}} = \delta_C^{(t+1)} f^{(t+1)} + \delta_h^{(t)} \odot o^{(t)} \odot \left(1 - \tanh^2 \left(C^{(t)} \right) \right). \tag{15}$$

According to equations (14) and (15), the following formula can be obtained.

$$\begin{aligned} \frac{\partial L}{\partial W_f} &= \sum_{t=1}^{\tau} \frac{\partial L}{\partial C^{(t)}} \frac{\partial C^{(t)}}{\partial f^{(t)}} \frac{\partial f^{(t)}}{\partial W_f} \\ &= \sum_{t=1}^{\tau} \left[\delta_C^{(t)} \odot C^{(t-1)} \odot f^{(t)} \odot \left(1 - f^{(t)} \right) \right] \left(h^{(t-1)} \right)^T. \end{aligned} \tag{16}$$

The other parameters in the model are derived similarly. The updating step size and learning rate of the model are defined as λ and α, respectively. The parameters in the LSTM model are iteratively updated using the gradient BP algorithm. The corresponding formula can be expressed as follows:

$$\beta_{t+1} = \beta_t - \alpha \frac{\partial L}{\partial \beta}, \tag{17}$$

where β represents the parameters in the LSTM model and $\partial L / \partial \beta$ represents the gradients of the parameters.

In summary, the updating process of the parameters in the LSTM network model based on a memory architecture can be expressed as follows. First, a parameter initialization process is implemented. Second, the iterative process is repeated by the gradient descent algorithm and the error BP algorithm until the target loss function converges. Finally, the parametric optimal solution of the LSTM model is obtained. Moreover, the Dropout algorithm is adopted in the training process of the LSTM model to avoid the overfitting phenomenon [27] and improve network performance by preventing feature detectors from working together.

2.3. Concrete Dam Deformation Prediction Model Based on RS-LSTM. With the advantages of RS, the mapping relationship between the factors that influence dam operating behaviour and the corresponding effects is established under the premise of retaining the key information. Additionally, the redundant information is eliminated, the expression space of the influential factors is simplified, and the importance of each factor is evaluated. Moreover, because the LSTM model overcomes the memory shortage and gradient dissipation issues of traditional RNNs and is characterized by controllable memory and fast gradient convergence, the model yields high-performance dynamic predictions based on long-term data series. Therefore, by combining the advantages of RS theory and the LSTM network, this paper establishes a concrete dam deformation prediction model based on RS-LSTM, and the prediction model is optimized considering the relevant influential factors and interactive mechanisms between these factors and concrete dam deformation in a quantitative manner. The process of establishing a concrete dam deformation prediction model based on RS-LSTM is shown in Figure 6. The specific modeling steps are as follows.

2.3.1. Data Acquisition. Statistical methods are used to perform gross error processing for concrete dam monitoring data. Such methods provide a reliable data foundation for the establishment of prediction models. Attribute reduction in

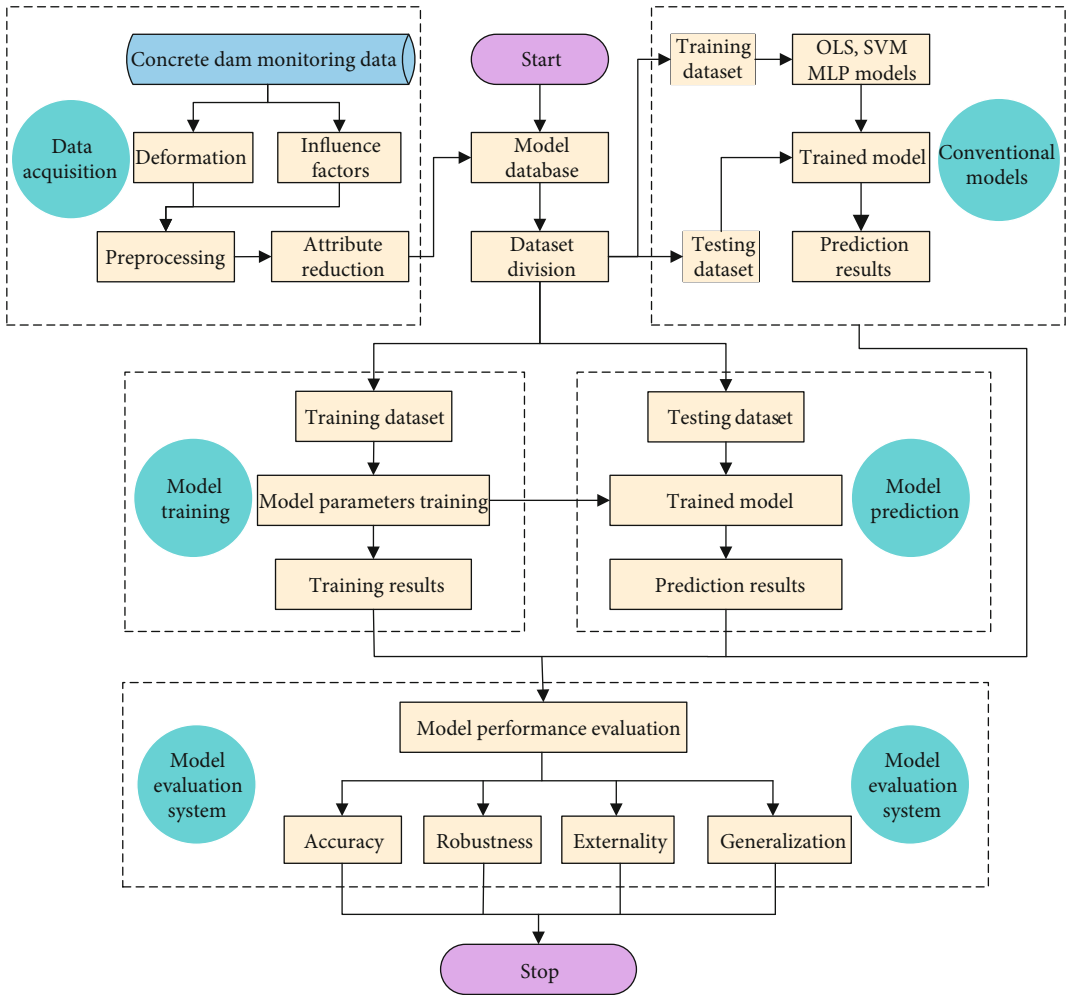

Figure 6: Process of establishing the concrete dam deformation prediction model based on RS-LSTM.

RS theory is conducted based on a complex multivariate dataset composed of water depth, temperature, seepage flow, fracture aperture, and uplift pressure information to accurately obtain the representative factors that affect the deformation behaviours of concrete dams. Deformation monitoring data and the representative influential factors corresponding to certain measurement points are selected as the model dataset. The representative factor dataset is standardized using an independent standardization formula, and the model dataset is divided into a training set and testing set by a cross-validation method.

2.3.2. Model Training. The preprocessed and standardized training set samples are used as model inputs. Error back propagation based on the gradient descent algorithm drives the model loss function to converge, and the optimal model parameters are obtained. The Dropout algorithm is used to overcome the problem of overfitting in training, and finally, a prediction model with optimal parameters is obtained.

2.3.3. Model Prediction. The testing set samples are input into the trained prediction model to obtain the corresponding deformation prediction results.

2.3.4. Model Performance Evaluation. According to the established evaluation system, the results of the concrete dam deformation prediction model based on LSTM, a classical least squares (OLS) model, a support vector machine (SVM) model, and a multilayer perceptron (MLP) model with 2 hidden layers are compared based on accuracy, robustness, externality, and generalization.

2.4. Evaluation System for the Concrete Dam Deformation Prediction Model. A concrete dam deformation prediction model plays an important role in operational behaviour monitoring, real-time abnormality detection, and decision-making, and its performance directly affects condition assessments and early warning strategies. In actual application processes, a single accuracy evaluation index may have certain limitations, and it is often impossible to evaluate the robustness, externality, and generalization of a model. Therefore, a complete evaluation system for concrete dam deformation prediction models must be established for practical applications. Therefore, this paper evaluates model performance from the aspects of accuracy, robustness, externality, and generalization, and quantitative evaluation indexes are used to comprehensively evaluate the performance of the

concrete dam deformation prediction model based on statistical theory.

2.4.1. Accuracy. The accuracy of the concrete dam deformation prediction model refers to the degree of agreement between the predicted and true values. This evaluation index is the most widely used in model assessment. In actual engineering, the MAPE, MSE, and RMSE are usually selected to evaluate the accuracy of a model. Considering the nonstationarity of deformation monitoring data and the overlap among evaluation indexes, the RMSE_P and MAPE_P are selected to establish the accuracy evaluation index (*AEI*) of the concrete dam deformation prediction model. The corresponding formulas are defined as follows.

$$\begin{aligned} \mathrm{RMSE}_P &= \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - y\wedge_t)^2}, \\ \mathrm{MAPE}_P &= \frac{100}{n}\sum_{t=1}^{n}\left|\frac{y_t - \widehat{y}_t}{y_t}\right|. \end{aligned} \tag{18}$$

2.4.2. Robustness. The robustness of a concrete dam deformation prediction model refers to its resistance to the inherent errors in training data. Model training and prediction are performed by establishing normal training samples and training samples with a certain degree of random error. The ability of a model to learn the true nonlinear mapping relationships when there is a small gross error in the training set is tested. The absolute difference between the RMSE_O of the training model prediction results with no gross error and the RMSE_E of the training model prediction results with gross error is selected as the robustness evaluation index (REI) for the concrete dam deformation prediction model. The corresponding formula is defined as follows.

$$\mathrm{REI} = \left|\mathrm{RMSE}_O - \mathrm{RMSE}_E\right|. \tag{19}$$

2.4.3. Externality. The externality of the concrete dam deformation prediction model refers to its adaptability to accurately process samples outside the training set with the same mapping relationship. A high-performance model based on its externality ability can learn the mapping relationships hidden in data through training set. Even if some samples are outside the training set, a model with a satisfactory externality can achieve accurate predictions. The samples outside the training set are fused with the testing samples, and the prediction performance of the model based on a training set with the same mapping relationship is tested. The accuracy index of the model under this condition, the RMSE_P, is selected as the externality evaluation index (EEI). The corresponding formula is defined as follows.

$$\mathrm{EEI} = \mathrm{RMSE}_P = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - y\wedge_t)^2}. \tag{20}$$

2.4.4. Generalization. The generalization of a concrete dam deformation prediction model refers to its adaptability to process samples with the same mapping relationship. A poor generalization ability can lead to overfitting. In such cases, the model error for the training set is very low, but the error is very large for the testing set. The model is optimized by adding training samples, performing regularization processing, and applying the Dropout algorithm to improve its generalization performance. Selecting the ratio of RMSE_T in the training process to RMSE_P in the prediction process is selected as the generalization evaluation index (GEI). The corresponding formula is defined as follows.

$$\mathrm{GEI} = \frac{\mathrm{RMSE}_P}{\mathrm{RMSE}_T}. \tag{21}$$

In each evaluation index formula above, $y(t)$ represents a measured value; $\widehat{y}(t)$ represents a predicted value; n represents the number of predicted samples; the subscript T represents the training process; the subscript P represents the prediction process; the subscript O represents samples with no gross error; and the subscript E represents samples with gross error.

2.5. Simulation Environment and Engineering Project. Concrete dam deformation prediction models based on OLS, SVM, MLP, and LSTM are established in accordance with the horizontal displacement of concrete gravity dams, and the evaluation system is used to evaluate the accuracy, robustness, externality, and generalization of each model. Additionally, a comparative analysis is performed. The simulation environment includes the Windows 10 operating system, an Intel Core i5 CPU, 8 GB of memory, the Python programming language version 3.7.2rcl, and the TensorFlow deep learning framework version 1.12.0.

2.5.1. Engineering Situation. Zhouning Hydropower Station is a diversion-type power station on the Muyang River in Fujian Province that performs step exploitation. The total installed capacity is 250 MW, the total storage capacity of the reservoir is 47 million m^3, and the designed flood level is 633.00 m. The power station consists of a barrage, a sluice building, a water conveyance system, an underground powerhouse, and a ground switch station. The barrage is an RCC gravity dam with a foundation plane elevation of 562.00 m, a maximum dam height of 72.40 m, and a dam crest length of 206.00 m. The body of Zhouning Dam is divided into nine dam sections, of which Nos. 1-4 and Nos. 7-9 are nonoverflow sections and Nos. 5-6 are overflow sections.

The deformation monitoring data collected by Zhouning Hydropower Station include horizontal and vertical dam displacement data. The horizontal displacement monitoring of the dam crest is performed by the extension wire alignment method. The fixed end of the extension wire with a total length of 200.75 m is arranged at Sta. R01+107.025 and the guide end is placed at Sta. L0+93.50. In total, 11 monitoring points are arranged along the dam, of which nine datum points are located at the top of each dam section and two checkpoints are set at the left and right ends of the extension wire to check the displacement of each end. The extension wire system was automated in April 2005 with an observation frequency of 1 time per day. The layout of the extension wire

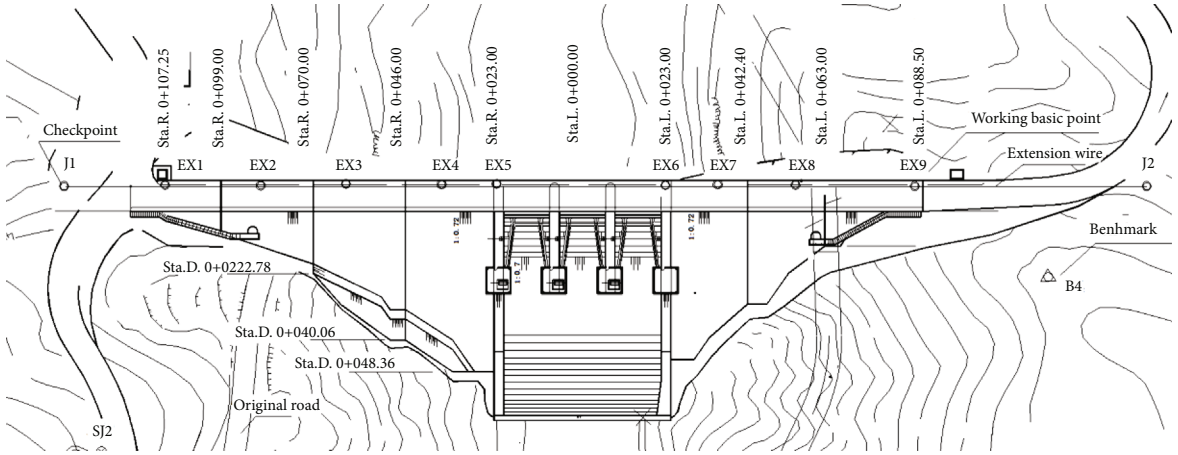

Figure 7: Layout of the extension wire measuring points for horizontal displacement.

measurement points for horizontal displacement is shown in Figure 7.

2.5.2. Selection and Optimization of Influential Factors in the Prediction Model. According to theoretical knowledge, monitoring data, expert experience, etc., the initial selection of empirical influence factors was as follows:

$$\{(H-H_0)^1, (H-H_0)^2, (H-H_0)^3, (T_5-T), (T_{20}-T), (T_{60}-T), (T_{90}-T), \theta, \ln(1+\theta), J, Q, U\}, \tag{22}$$

where H is the water depth on a day when observations are collected, H_0 is the water depth on the base day; T_i is the mean reservoir region temperature i days ago, and T is the annual mean temperature. Additionally, $\theta = (t - t_0)/100$, where t is the observation date and t_0 is the date of the base day. J is the average fracture aperture at measurement points, Q is the seepage flow, and U is the average uplift pressure at measurement points.

The initial empirical influential factors are selected as the conditional attributes X, and the horizontal displacements obtained by the dam crest extension wire (to the left bank is positive and to the right bank is negative) at point EX1 are set as the decision attributes Y in the single-point prediction model. Additionally, the horizontal displacements of EX1, EX2, EX4, EX5, EX6, and EX7 are selected as the decision attributes Y_M in the multipoint prediction model (because the extension wire at EX3 contacted a stainless-steel rod, the monitoring data at the point are not reliable). Overall, 864 monitoring samples of horizontal displacement and influential factors were selected as the sample set U. The attribute range V was determined based on the K-means clustering algorithm with adaptive discretization, and the number of clusters was experimentally determined to be 7. To eliminate irrelevant or weakly informative input variables and keep only the representative factors that affect concrete dam deformation, the RS theory is used to conduct an attribute reduction and importance evaluation and obtain an initial information table $S = \{U, X \cup Y, V, f\}$. The attribute reduction and importance evaluation results for the single-point and multipoint prediction models are shown in Table 1.

According to attribute reduction and importance evaluation results, the influential factors of the single-point prediction model are $\{H\text{-}H_0, (H\text{-}H_0)^2, (H\text{-}H_0)^3, (T_5\text{-}T), (T_{20}\text{-}T), \theta, \ln(1+\theta)\}$, and the importance evaluation values for each component of horizontal displacement at EX1 are 0.12, 0.08, 0.13, 0.42, 0.20, 0.02, and 0.03, respectively. The influential factors of the multipoint prediction model are $\{H\text{-}H_0, (H\text{-}H_0)^2, (H\text{-}H_0)^3, (T_5\text{-}T), (T_{20}\text{-}T), \theta, \ln(1+\theta), J, Q, U\}$, and the importance evaluation values of each component of the horizontal displacement at EX1, EX2, EX4, EX5, EX6, and EX7 are 0.10, 0.07, 0.06, 0.33, 0.19, 0.00, 0.00, 0.02, 0.04, 0.06, 0.08, and 0.05. Therefore, it can be concluded that the horizontal displacement of the extension wire is greatly affected by temperature changes and water level fluctuations. Specifically, the temperature component accounts for 60% of the horizontal displacement, and the lag period of the water level is approximately 20 days.

2.5.3. Sample Selection for Prediction Models. According to attribute reduction and importance evaluation results, the influential factor monitoring data of the single-point and multipoint prediction models are selected to obtain samples as independent variables, and the horizontal displacement at points EX1-EX7 (except EX3) is selected to obtain samples as dependent variables. The dataset is established between June 2, 2016, and October 22, 2018, and has a total of 864 samples of data. The dataset of 700 samples selected from June 2, 2016, to May 10, 2018, is used as the training set, and the dataset of 164 samples selected from May 11, 2018, to October 22, 2018, is adopted as the testing set. Investigations of the concrete dam deformation prediction model based on the OLS, SVM, MLP, and LSTM methods are performed using the dataset with 864 samples of data. Variations in the water depth and horizontal displacement are shown in Figures 8 and 9.

2.5.4. Model Parameter Setting. The performance of SVM, MLP, and LSTM models depends greatly on the setting of some parameters. According to experience and experiment

Table 1: Attribute reduction results of the single-point and multipoint prediction models.

Experience impact factors	Component name	Single-point model *SIM*	Reduction	Importance evaluation Sig (a, X, Y)	Multipoint model *SIM*	Reduction	Importance evaluation Sig (a, X, Y)
$H\text{-}H_0$	Water pressure	-5	No	0.12	-4	No	0.10
$(H\text{-}H_0)^2$		-2	No	0.08	-2	No	0.07
$(H\text{-}H_0)^3$		-2	No	0.13	-4	No	0.06
$(T_5\text{-}T)$	Temperature	-5	No	0.42	-7	No	0.33
$(T_{20}\text{-}T)$		-4	No	0.20	-2	No	0.19
$(T_{60}\text{-}T)$		0	Yes	0.00	0	Yes	0.00
$(T_{90}\text{-}T)$		0	Yes	0.00	0	Yes	0.00
θ	Aging	-2	No	0.02	-2	No	0.02
$\ln(1+\theta)$		-1	No	0.03	-3	No	0.04
J	Fracture	-1	Yes	0.00	-2	No	0.06
Q	Seepage	-2	Yes	0.00	-4	No	0.08
U	Uplift pressure	-3	Yes	0.00	-3	No	0.05

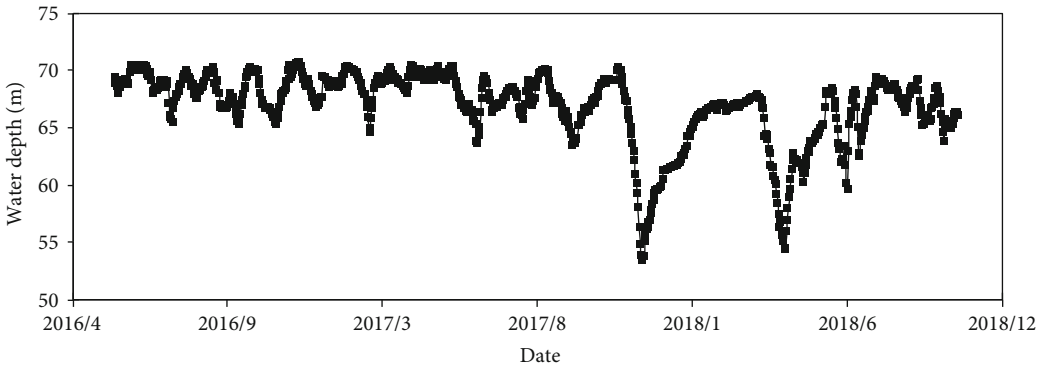

Figure 8: Variations in the water depth.

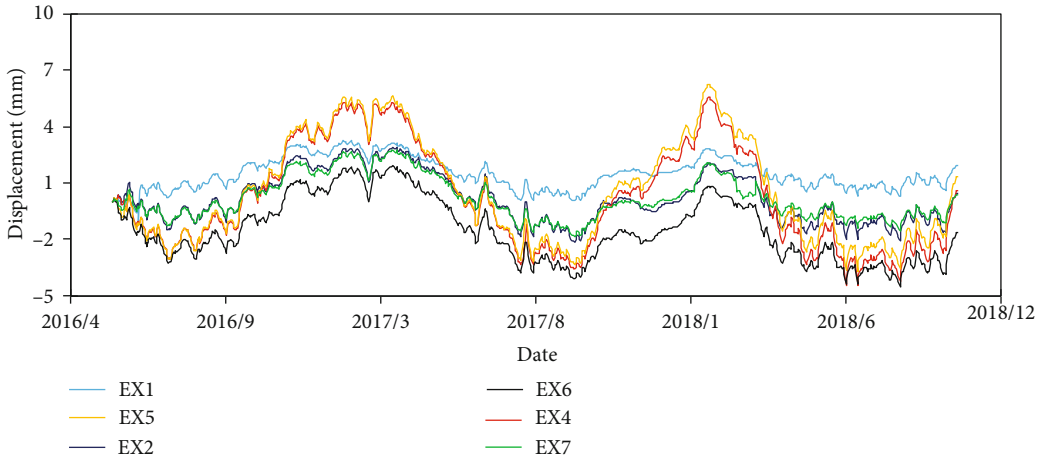

Figure 9: Variations in horizontal displacement.

results, parameters of the adapted algorithms, namely, regularization parameters, kernel function parameters, network parameters, learning rates, and so on, are given before the simulation.

Parameters in the SVM model: the kernel function is determined as a radial basis function (RBF) according to experience. Parameter range of the SVM model is determined based on experience, penalty parameter $C \in [-256,$

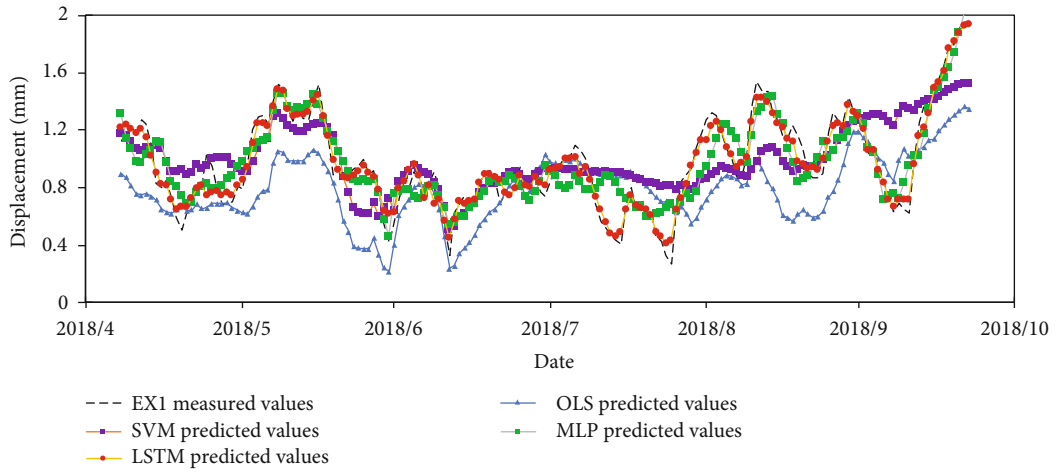

Figure 10: Measured and model-predicted values.

256], kernel parameter $\gamma \in [-256, 256]$. Parametric tuning is implemented with Grid Search, C is set to 8, and γ is set to 0.72 according to the experimental relationship between the objective functions and parameters.

Parameters in the MLP model: according to experiment results, the network is composed of input layer, hidden layers, and output layer with the three-layer topology of 7-15-15-1 for single-point prediction model and 10-15-15-1 for multipoint prediction model, and the learning rate is also set to 0.08. The network variable parameter weights and biases are initialized randomly and calculated by gradient descent algorithms, and the activation function is the ReLU function.

Parameters in the LSTM model: hyperparameter range of the LSTM model is determined based on experience, batch size $\in [0, 1000]$, timestep $\in [20, 300]$, hidden layers $\in [20,200]$, and the initial value of learning rate is set to 0.1. Taking the minimum RMSE value as the objective function, and according to the experimental relationship between the objective function and parameters, parametric tuning is implemented with Grid Search. Finally, batch size, timestep, hidden layers, and learning rate are set to 12, 46, 42, and 0.12, respectively. The network variable parameters including weights and biases are generated using the glorot_uniform initializer and calculated by gradient descent algorithms.

3. Results

To verify the superiority of the concrete dam deformation prediction model based on LSTM compared to other models in terms of accuracy, robustness, externality, and generalization, a comparative analysis of the OLS, SVM, MLP, and LSTM models is conducted based on the prediction results with preprocessed training and testing samples.

3.1. Single-Point Prediction Model. Single-point prediction models for concrete dam deformation based on the OLS, SVM, MLP, and LSTM algorithms are established to facilitate comparative analysis.

3.1.1. Model Prediction Analysis. Concrete dam deformation prediction models based on the OLS, SVM, MLP, and LSTM algorithms were established based on the preprocessed standardized environmental dataset and the unnormalized deformation dataset. Based on the objective functions, the training samples are used to train the models, and the optimal model parameters are obtained. Finally, a concrete dam deformation prediction and performance analysis are performed. The measured and predicted values of concrete dam deformation based on the OLS, SVM, MLP, and LSTM models are shown in Figure 10.

Figure 10 shows that the predicted values of the OLS model largely deviate from the measured values, but the overall trend is similar to that for the measured values. The deviation between the predicted values of the SVM, MLP, and LSTM models and the measured values is small, but the late-stage prediction trend of the SVM model deviates significantly from the measured values. The LSTM model not only exhibits the highest degree of agreement between the predicted and measured values but also yields the same trend as that for the measured values. Therefore, the prediction performance of the concrete dam deformation prediction model based on LSTM is significantly better than that based on the OLS, SVM, and MLP models.

3.1.2. Model Performance Evaluation. A prediction model for concrete dam deformation is an important tool for quantitatively evaluating the safety status of dams, revealing abnormalities in the service status and ensuring engineering safety. A high-performance deformation prediction model should meet the relevant accuracy, robustness, externality, and generalization requirements to implement effective early warning strategies and feedback control for engineering safety. To verify the reliability of the concrete dam deformation prediction model based on LSTM, concrete dam deformation prediction models based on OLS, SVM, MLP, and LSTM were compared and analyzed according to the evaluation system established in this paper. The horizontal displacement residuals of each prediction model are shown in Figure 11. Evaluation index values of all prediction models are shown in Table 2.

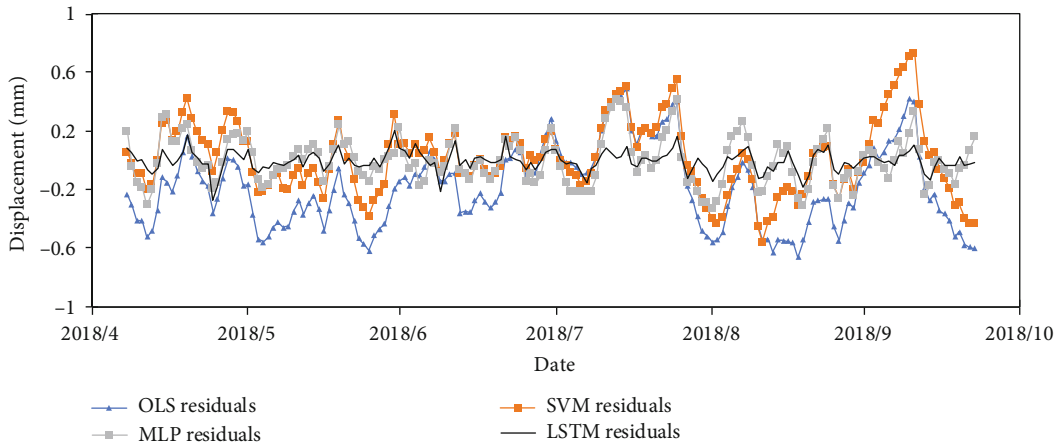

Figure 11: The horizontal displacement residuals of each prediction model.

Table 2: Evaluation index values of all prediction models.

	Prediction model			
	OLS	SVM	MLP	LSTM
AVI				
$RMSE_P$	0.32833	0.2466	0.1604	0.0690
$MAPE_P$	29.7077	24.7789	16.3695	6.3213
REI	0.3253	0.3672	0.2865	0.1760
EEI	0.4032	0.2739	0.1892	0.1023
GEI	0.8266	0.8788	0.9336	0.9610

(1) Accuracy. Figure 11 and AVI value of each model in Table 2 show that the horizontal displacement residuals of the concrete dam deformation prediction model based on LSTM are the smallest, $RMSE_P$ is lower than 0.1, $MAPE_P$ is lower than 10, and all these are in the low range compared with those of the models based on OLS, SVM, and MLP. Therefore, the concrete dam deformation prediction model based on LSTM displays better accuracy than the other models, and the prediction results better agree with the real data.

(2) Robustness. The REI value of each model in Table 2 shows that the concrete dam prediction models based on OLS, SVM, MLP, and LSTM are all affected by the gross error, resulting in different degrees of prediction accuracy. The REI value of the concrete dam deformation prediction model based on LSTM is the smallest among the REI values of all the models; thus, the gross error associated with the data sample has little impact on the prediction results of the proposed model, which displays the strongest robustness.

(3) Externality. The EEI value of each model in Table 2 shows that the accuracy of the models decreases after adding samples outside the training set to the model testing samples. Nevertheless, the concrete dam deformation prediction model based on LSTM exhibits the smallest EEI, representing the strongest externality and the most powerful learning ability.

(4) Generalization. The GEI value of each model in Table 2 shows that the generalization of the concrete dam deformation prediction models based on OLS and SVM is poor. These models likely experience overfitting during training, resulting in an increase in the error for the testing set and poor performance. The concrete dam deformation prediction models based on MLP and LSTM display good generalization performance, and the LSTM model yields the best performance.

In summary, the successful application of machine learning technology has greatly promoted the development of concrete dam deformation prediction model compared with using traditional statistical methods. The concrete dam deformation prediction models based on SVM, MLP, and LSTM all displayed high accuracy, but the performance of each model in terms of robustness, externality, and generalization varies. The concrete deformation prediction model based on LSTM displays the highest accuracy, robustness, externality, and generalization by comparison with the performance of the other models. Therefore, the application of LSTM to concrete deformation prediction models further promotes the development of concrete dam prediction model.

3.2. Multipoint Synchronized Prediction Model for Concrete Dam Deformation. According to the theoretical, mathematical, and mechanical principles of concrete dams, the concrete dam deformation is affected not only by loads such as water pressure and temperature loads but also by adjacent local factors. The sudden displacement of some dam parts will influence the surrounding areas, and a single-point prediction model for concrete dam deformation does not consider the relationships among points; therefore, it is difficult to grasp the displacement field under a given load. It is necessary to establish a multipoint synchronized prediction model for concrete dam deformation that can effectively improve the prediction performance compared to that of traditional models and the accuracy of mechanical parameter inversion and feedback analysis. Additionally, such a method could improve the

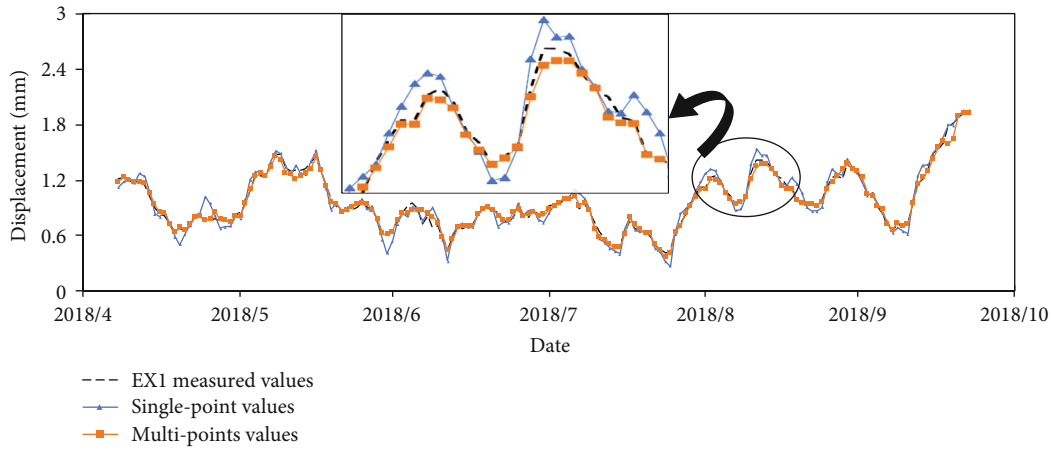

Figure 12: Actual values and predicted values with the single-point and multipoint models.

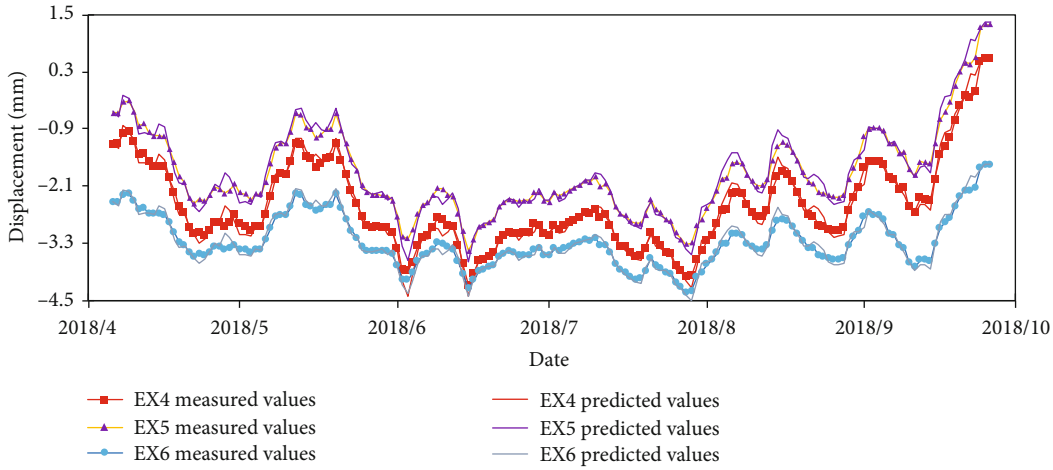

Figure 13: Measured values and values predicted with the multipoint deformation prediction model.

safety monitoring level of concrete dams. This paper establishes a multipoint synchronized prediction model for concrete dam deformation based on the data collected at multiple points and the advantages of LSTM for multiple inputs and outputs.

3.2.1. Model Prediction Analysis. The factors that influence the multipoint synchronized prediction model for concrete dam deformation are determined by attribute reduction. All the data are normalized and used as samples of the independent variables in the LSTM model, and the deformation monitoring data from points EX1-EX7 (except EX3) are selected as samples of the dependent variable (no normalization processing). The training data and testing set are divided in the same way as in the single-point model. The output layer is a multidimensional fully connected layer, the model learning rate is 0.18, and other parameters are the same as those in the single-point model. The six-point synchronized prediction model for concrete dam deformation with optimal parameters is obtained by training, and the deformation values are predicted based on the testing samples. The actual measured values and the predicted values of the single-point and multipoint models are shown in Figure 12 (taking the measured values at EX1 as an example). The measured values and values predicted with the multipoint synchronized deformation model are shown in Figure 13 (taking the measured values at EX4-EX6 as examples).

3.2.2. Model Performance Evaluation. Since the prediction model is based on the deformation values at multiple points, the error of the multipoint model includes the error at all points. According to error theory, RMSE of the multipoint model is the weighted average of RMSE at each point, and it can be expressed as follows:

$$S = \sqrt{\frac{ns_1^2 + ns_2^2 + \cdots + ns_k^2}{kn}} = \sqrt{\sum_{i=1}^{k} \frac{s_i^2}{k}}, \tag{23}$$

where S represents the RMSE of the multipoint model, s_i represents each point, and k represents the number of testing samples.

The RMSE values of the multipoint synchronized prediction model and single-point prediction model are compared and analyzed, as shown in Table 3.

Table 3: RMSE values of the multipoint prediction model and single-point prediction model.

Prediction model	Single-point model of EX1	Multipoint model of EX1	Average of all single-point models	Multipoint value
S	0.0690	0.0575	0.0739	0.0592

Figures 12 and 13 and Table 3 show that the performance of both models is good, and the error is within the acceptable precision range. The RMSE value of the multipoint prediction model is smaller than that of the single-point prediction model, and the predicted values of the multipoint model are closer to the measured values. Additionally, the weighted average RMSE of the single-point prediction model is larger than the RMSE of the multipoint model, which indicates that the prediction accuracy at each point in the multipoint model is high. Therefore, the multipoint synchronized prediction model for concrete dam deformation based on LSTM exhibits good performance, and the analysis results are locally meaningful and spatially representative at large scales.

4. Conclusions and Discussion

RS theory and an LSTM network are introduced for concrete dam safety monitoring in the TensorFlow framework, and single-point and multipoint concrete dam deformation prediction models based on LSTM are established. Moreover, a new evaluation system and quantitative evaluation indexes for the concrete dam deformation prediction model are proposed. The following conclusions were obtained from application examples.

(1) RS theory is applied to optimize the selection and evaluate the importance of the factors that influence concrete dam deformation based on the internal relationships among the monitoring dataset. This approach overcomes the deficiencies of intelligent prediction models related to the quantitative interpretation and ensures the objectivity of prediction model analysis

(2) According to statistical theory, an evaluation system is proposed, and accuracy, robustness, externality, and generalization evaluation indexes are given as performance inspection criteria to comprehensively evaluate the performance of concrete dam deformation prediction models in practical engineering

(3) The single-point prediction model for concrete dam deformation based on LSTM displays high prediction accuracy and strong robustness, externality, and generalization. Moreover, the multipoint synchronized prediction model for concrete dam deformation based on LSTM is locally pertinent and spatially representative at large scales. Thus, the multipoint approach can be effectively used in the deformation prediction of concrete dams at large scales

The continuous improvements in concrete dam technology have resulted in high requirements for prediction model performance, and establishing high-performance spatiotemporal prediction models will be important as concrete dam safety monitoring continues to progress. Therefore, the combination of AI, deep learning theory, online dynamic learning, and space-time deformation prediction models should be promoted to establish an ideal concrete dam monitoring system and achieve the goal of "intelligent monitoring."

Data Availability

"Dataset of Zhouning dam.xlsx" used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there is no conflict of interest.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (No. 51809212), the Key Projects of Natural Science Basic Research Program of Shaanxi Province (No. 2018JZ5010), and the Water Science Plan Project of Shaanxi Province (No. 2018SLKJ-5).

Supplementary Materials

The supplementary material submitted is the dataset used to build the prediction model based on RS-LSTM in the article. The dataset named "Dataset of Zhouning dam.xlsx" includes values of Zhouning dam's horizontal displacement at points EX1-EX7 (except EX3) and its impact factors from June 2, 2016, to October 22, 2018, which are selected to obtain samples. The dataset of 700 samples selected from June 2, 2016, to May 10, 2018, is used as the training set, and the dataset of 164 samples selected from May 11, 2018, to October 22, 2018, is adopted as the testing set. *(Supplementary Materials)*

References

[1] B. F. Zhu, "On the expected life span of concrete dams and the possibility of endlessly long life of solid concrete dams," *Journal of Hydraulic Engineering*, vol. 43, no. 1, pp. 1–9, 2012.

[2] C. S. Gu, H. Z. Su, and S. W. Wang, "Advances in calculation models and monitoring methods for long-term deformation behaviour of concrete dams," *Journal of Hydroelectric Engineering*, vol. 35, no. 5, pp. 1–14, 2016.

[3] X. Q. Li, D. J. Zheng, and Y. P. Ju, "Input factor optimization study of dam seepage statistical model based on copula entropy theory," *Journal of Hohai University (Natural Sciences)*, vol. 44, no. 4, pp. 370–376, 2016.

[4] J. Yang, D. X. Hu, and Z. R. Wu, "Multiple co-linearity and uncertainty of factors in dam safety monitoring model," *Journal of Hydraulic Engineering*, vol. 35, no. 12, pp. 99–105, 2004.

[5] F. Kang, L. S. Zhao, and Y. Wang, "Structural health monitoring of concrete dams using long-term air temperature for thermal effect simulation," *Engineering Structures*, vol. 180, pp. 642–653, 2019.

[6] F. Q. Li, *Research on Data Analysis Method of Dam Safety Monitoring*, Zhejiang University, 2012.

[7] Y. Zhao, X. H. Hua, and M. Li, "Application of stepwise regression model to dam radial displacement monitoring," *Journal of Geomatics*, vol. 21, no. 1, pp. 1–5, 2012.

[8] L. Pei, Z. Y. Wu, M. Cui, Q. Zhang, and J. K. Chen, "Research and application on the displacement hybrid-model of high earth dam," *Journal of Sichuan University*, vol. S1, pp. 7–12, 2012.

[9] B. Li, J. Li, K. Jiang, and L. H. Pi, "Spatial-temporal model of monitoring the displacement of roller compacted concrete dam," *Journal of Yangtze River Scientific Research Institute*, vol. 30, no. 1, pp. 90–92, 2013.

[10] D. Y. Li, Y. C. Zhou, and X. Q. Gan, "Research on deterministic displacement monitoring model for multiple survey points of concrete arch dam," *Journal of Hydraulic Engineering*, vol. 42, no. 8, pp. 981–985, 2011.

[11] C. S. Gu and Z. R. Wu, *Safety Monitoring of Dams and Dam Foundations -Theory, Methods and Application*, Hohai University Press, Nanjing, China, 1st edition, 2006.

[12] F. Kang, J. J. Li, and J. H. Dai, "Prediction of long-term temperature effect in structural health monitoring of concrete dams using support vector machines with Jaya optimizer and salp swarm algorithms," *Advances in Engineering Software*, vol. 131, pp. 60–76, 2019.

[13] D. Y. Li and Y. C. Zhou, "Application of BP network to multiple-spot model of dam deformation monitoring," *Journal of Yangtze River Scientific Research Institute*, vol. 25, no. 6, pp. 52–55, 2005.

[14] R. X. Chen and L. Cheng, "The monitoring model of multiple monitoring points with latent variables based on LS-SVM," *Water Resources and Power*, vol. 10, pp. 80–82, 2012.

[15] 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, article e1997, 2017.

[16] B. Dai, C. S. Gu, E. F. 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, article e2170, 2018.

[17] F. Salazar, R. Moran, 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*, vol. 24, no. 1, pp. 1–21, 2017.

[18] D.-A. Tibaduiza, M.-A. Torres-Arredondo, L. E. Mujica, J. Rodellar, and C.-P. Fritzen, "A study of two unsupervised data driven statistical methodologies for detecting and classifying damages in structural health monitoring," *Mechanical Systems & Signal Processing*, vol. 41, no. 1-2, pp. 467–484, 2013.

[19] Y. Yang, D. Chen, and H. Wang, "Active sample selection based incremental algorithm for attribute reduction with rough sets," *IEEE Transactions on Fuzzy Systems*, vol. 25, no. 4, pp. 825–838, 2016.

[20] Y. Jing, T. Li, H. Fujita, B. Wang, and N. Cheng, "An incremental attribute reduction method for dynamic data mining," *Information Sciences*, vol. 465, pp. 202–218, 2018.

[21] C. Yuan, X. Qin, L. Yang, G. Gao, and S. Deng, "A novel function mining algorithm based on attribute reduction and improved gene expression programming," vol. 7, IEEE Access, 2019.

[22] B. B. Yang, K. L. Yin, and J. Du, "Dynamic prediction model of landslide displacement based on time series and long and short time memory networks," *Chinese Journal of Rock Mechanics and Engineering*, vol. 37, no. 10, pp. 2334–2343, 2018.

[23] W. T. Tong, L. X. Li, X. L. Zhou, A. Hamilton, and K. Zhang, "Deep learning $PM_{2.5}$ concentrations with bidirectional LSTM RNN," *Air Quality, Atmosphere & Health*, vol. 12, no. 4, pp. 411–423, 2019.

[24] T. T. Huang and L. Yu, "Application of SDAE-LSTM model in financial time series prediction," *Computer Engineering and Applications*, vol. 55, no. 1, pp. 142–148, 2019.

[25] Y. X. Liu, Q. X. Fan, Y. Z. Shang, Q. M. Fan, and Z. W. Liu, "Short-term water level prediction method for hydropower stations based on LSTM neural network," *Advances in Science and Technology of Water Resources*, vol. 39, no. 2, pp. 56–60+78, 2019.

[26] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: continual prediction with LSTM," *Neural Computation*, vol. 12, no. 10, pp. 2451–2471, 2000.

[27] Y. Ying and M. Pontil, "Online gradient descent learning algorithms," *Foundations of Computational Mathematics*, vol. 8, no. 5, pp. 561–596, 2008.