IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 7, 2020, accepted January 3, 2021, date of publication January 6, 2021, date of current version January 15, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3049578

An Improved Random Forest Model for the Prediction of Dam Displacement

YAN SU, KAILIANG WENG, CHUAN LIN, AND ZHIMING ZHENG
School of Civil Engineering, Fuzhou University, Fuzhou 350108, China
Corresponding author: Chuan Lin (linchuan@fzu.edu.cn)

ABSTRACT Dam behavior prediction is a classic problem in the monitoring of dam structure. To obtain accurate results, different researchers have established various models. However, the models of predecessors rarely studied the nonlinear characteristics of dam displacement data and the abnormal values of monitoring data. It means that abnormal values will contaminate data set, and consequently reduce the accuracy of model predictions. In this article, an improved Random Forest (RF) model was proposed for analyzing dam displacement prediction and was coupled with a sliding time window strategy. The proposed model is developed by the following steps. First, for the purpose of alleviating the time-lag effect of impact factor phenomenon, a sliding time window strategy was introduced into the RF model to improve the time sensitivity. Second, aiming to determine the hyperparameters, Grid Search (GS) was introduced into RF model to improve the global optimization ability. This article takes masonry arch dam in China as an example, and adopts the horizontal displacement recorded by Global Navigation Satellite Systems (GNSS) as the research object. The accuracy and validity of the proposed model are verified and evaluated based on the evaluation criteria. The simulation results demonstrate that the proposed model could capture the long-term characteristics and provide better prediction based on short-term monitoring data. It also has strong robustness on the abnormal data series, has simpler structures and less parameters, and requires less time for model training, so it can be a potential tool for actual monitoring tasks.

INDEX TERMS Arch dam, health monitoring, random forest regression, grid search.

I. INTRODUCTION

According to statistics in September 2019, there are more than 58,000 dams distributed in 96 countries [1]. China has the largest number of dams and the largest irrigation area in the world. Since hydropower is the best choice for green energy, and China is rich in hydropower resources and has huge development potential, so hydropower will play a key role in China's energy structure [2]. However, with the development of hydropower, long-term maintenance of dams should be considered. Today, dam damage, or failures are caused by various reasons, increase in external load because of abnormal weather, and poor durability because of the aging of existing dams [3]. Structural failure of most dams does not happen suddenly, but it is a process that occurs gradually under the long-term action of various loads. Therefore, it is necessary to monitor dam safety through equipment. The dam safety monitoring aims to establish a long-term behavior monitoring and early warning model based on the observed data, and to identify abnormal behaviors as early as possible so as to eliminate or minimize adverse impact [4].

Compared with other engineering constructions, dams are extremely vulnerable to environmental, hydraulic and geomechanical factors (i.e. air and water temperature, water level, pore pressure, rock deformability and so on), each of which will affect the structural behavior [5]. As the most intuitive monitoring index, dam displacement is usually used to evaluate the overall structural performance and soundness. This indicates that dam displacement is an important element in ensuring the safe operation of the dam. The dam displacement series has a highly nonlinear and non-stationary characteristic due to the combined action of several random factors, and it is difficult to employ such an approach to directly quantify the displacement based on the current mathematical theory [6]. After all, now there is no unified and standardized method for modeling and forecasting the similar dynamic data at home and broad [7], so the establishment of a displacement prediction model is a key research topic in the field of dam

The associate editor coordinating the review of this manuscript and approving it for publication was Wai-Keung Fung.

safety monitoring. Since the 1950s, various statistical models were adopted for dam behavior modeling [8]. Some traditional statistical models, which are explicit physical interpretation, simple model structure, and fast execution, like the Hydrostatic-Seasonal-Time (HST) model, have been widely applied in the prediction of dam displacements. The model is explained by a simple function sum, for example, the thermal effect is treated as a periodic function. However, the realistic thermal effect cannot be accurately simulated by the periodic function. Therefore, the regression results are only the approximate fit of the actual relationship between variables. The accuracy and stability of dam displacement prediction has always been the difficult points in the dam displacement prediction, especially for the long-term time series prediction. Moreover, with the development of monitoring technology, the sampling frequency of dam monitoring devices has been increased from once a week to multiple times a day [9]. It is necessary to propose advanced intelligent methods to satisfy the requirements of big data processing.

Based on traditional statistical methods, improvements can be made from following two aspects. First, the influence of the temperature and water level on the displacements is nonlinear and delayed [10]. Second, the time-varying effect is complicated and could not be described by simple functions [11]. In recent years, with the rapid development of artificial intelligence, data-driven machine learning model, which have strong nonlinear processing capabilities, gradually replace linear regression methods in the dam displacement prediction [12]. Ranković V *et al.* [13] developed a model based on Support Vector Regression (SVR) to improve the prediction accuracy and training speed. Opyrchal [14] used adaptive neuro-fuzzy systems to qualitatively locate seepage paths in dams. Fernando Salazar *et al.* [15] used Boosted Regression Trees (BRT) as a tool for computing the relative influence and to identify the strength of each input-output relationship. Lin *et al.* [16] proposed a Gaussian Process Regression (GPR)-based model for dam displacement forecasting and also compared the model with SVR model and achieved comparatively better results. However, some recent studies compared the results of deep learning model, which is a new research direction of machine learning. In most cases the results of deep learning models performed well as compared to machine learning models. Mata J [17] proposed a prediction model for dam deformation based on Multilayer Perceptron (MLP) and utilized an example analysis to verify it feasibility. Furthermore, it was compared the results with Artificial Neural Network (ANN). Wen *et al.* [18] achieved good results on a Convolution Neural Network (CNN) and compared the results with SVR model. Qu *et al.* [19] used Long-Short Term Memory (LSTM) network for single-point and multipoint concrete dam deformation prediction which outperformed the traditional models and other machine learning models generated output results. These new technologies provide more accurate results and avoid the limitations of traditional statistical methods because they can test nonlinear relationships and correlations between predictive variables. However, there are still many shortcomings of these technologies, such as overfitting and parameter tuning [9].

Random Forest (RF) is a relative novel machine learning method, which has classification and regression functions [20]. The effectiveness of RF has been proven to be effective in many fields. For example, Christoph Behrens [21] used it when studying inflation forecasts. What's more, RF combines the predictions from a large number of decision trees and can score the importance of each predictor variable [22]. It has two attractive features: (a) with random sample selection and sample feature selection, it can improve the generalization ability of the model, and it does not produce the overfitting problem with the increase of decision trees; (b) The dam displacement can be predicted because it can handle the complex nonlinear relationship between high-dimensional data and input variables. For deep learning models, in order to suppress the model from overfitting, regularization or dropout layer needs to be introduced to improve the adaptabilities and robustness, but it may also lead to underfitting that occurs [23]. In order to solve the underfitting problem, it is necessary to increase the number of hidden layer units or the number of model layers, which will make the model architecture complicated and require a heavy price for model training [24]. Hence, compared with other models, RF has a clear advantage in solving high-dimensional, and nonlinear problems based on these features. The prediction accuracy of RF also depends on the reasonable setting of hyperparameters, but there is no consensus on selecting the optimal associated parameters, and the practical method is to determine the hyperparameters by the trial-and-error method. To solve the above problems, Grid Search (GS) optimization is utilized to fine-tune these parameters of RF model in this study. GS has been successfully applied to various fields such as power load forecasting [25], service satisfaction analysis [26], etc. These GS-related works demonstrated that GS is effective for hyperparameter optimization of RF model.

In short, the suppression of over-fitting or under-fitting and hyperparameter optimization are critical for the accuracy and validity of displacement prediction. However, although RF model does not produce the overfitting problem as the decision tree increases, it may cause the generalization error within a certain limit [20]. In view of the problem, the solution proposed in the article is to introduce the sliding time window strategy to couple the RF model. The sliding time window strategy can improve the time sensitivity of the statistical indicators in the frame and suppressing generalization errors [27]. Aiming at the problem of RF model hyperparameter adjustment, the GS method is used to select optimal parameters to speed up model convergence. Based on previous studies, predicting and estimating the displacement of a concrete dam must consider the seasonal thermal changes of the dam, otherwise it will lead to erroneous regression and prediction errors [28]. However, the seasonal prediction component of the thermal effect provided by the statistical model depends on the water temperature thermometer and the dam concrete

thermometer, and a dense and well-distributed sensor network needs to be embedded in the dam. However, many dams are not equipped with thermometers or thermometers have been damaged. Therefore, combined with the delayed and nonlinear characteristics of the influence of temperature on displacement, this study adds an average temperature variable around the dam, and implements a sliding time window strategy. Using the example of a real-world masonry arch dam in China, the results show that the optimized RF model has high accuracy and a certain degree of robustness to abnormal time series, and can be used for the prediction of long-term arch dam displacement. The novel contributions of this article can be summarized as follows. (1) Experimenting on a number of models over normal and abnormal time series in terms of prediction accuracy. After a deep analysis of these models, we advocate the usage of the optimized RF as optimal model, that is applicable for real-world dam displacement prediction problems. (2) The optimized RF model is evaluated on a shorter monitoring data set and achieved satisfied results as compared to other baseline models. Thus, it has the capability to capture the long-term characteristics of the dam displacement. (3) Verifying a nonlinear relationship between the average temperature and the dam displacement.

The remaining of this article is organized as follows. Section 2 presents a brief description of the HST, RF and the sliding time window strategy, and elaborates the proposed combined model in detail. Section 3 tells about the research design of the study case, input variable selection, evaluation indicators. In Section 4, the experimental results of the proposed model and benchmark methods for various displacement time series are illustrated and discussed. In Section 5, the conclusion is finally reached.

II. METHODOLOGY

The overall process of the proposed model is described in this section. Firstly, the theoretical basis of the traditional statistical model, RF models and the sliding time window strategy is described briefly. Then the proposed optimized RF model is formulated, and the specific steps of the proposed model are presented in detail.

A. TRADITIONAL STATISTICAL MODEL

The factors for the arch dam displacement are illustrated in FIGURE 1. Traditional statistical models mainly use the hydrostatic-seasonal-time (HST) model, $\delta(h, s, t)$ [29], which can be divided into three parts:

(a) the hydrostatic component (the influence of reservoir water), $\delta(h)$;

(b) the seasonal component (associated predominantly with temperature), $\delta(s)$;

(c) the irreversible component (the dissipation of hear of hydration, creep, alkali-aggregate reaction, and so on), $\delta(t)$;

The HST model can be written as follows:

$$\delta(h, s, t) = \delta(h) + \delta(s) + \delta(t) + \varepsilon \tag{1}$$

where ε is the residuals.

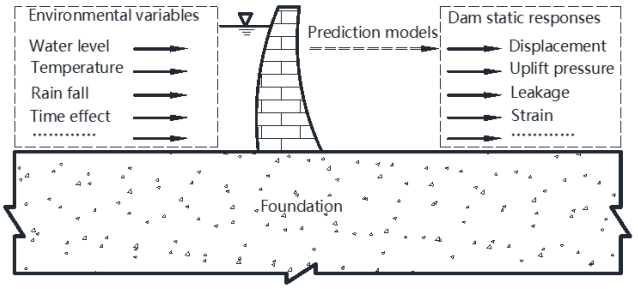

FIGURE 1. Modeling for dam safety monitoring.

In this study, three components are used as input variables. The three parts are calculated as follows:

(1) The hydrostatic component is always determined in accordance with a polynomial function of water height through the mechanical analysis:

$$\delta(h) = \sum_{i=0}^{K} \alpha_i h^i(t) \tag{2}$$

where $h(t)$ is the upstream water height, a_i is the coefficients, and i is the power exponent. K is always selected as 4 for arch dams and 3 for gravity dams.

(2) The seasonal component is caused by the internal temperature change in the dam, the foundation and the air temperature. For arch dams, the internal temperature after concrete hydration heat dissipates can be approximated as a steady state. Therefore, the seasonal component is calculated with a combination of harmonic sinusoidal functions:

$$\delta(s) = \sum_{i=1}^{m} (b_{1i} \sin \frac{2\pi it}{365} + b_{2i} \cos \frac{2\pi it}{365}) \tag{3}$$

where m is the temperature factor, selected as 2 for arch dams, t is the days from the initial survey date to the current survey date and b_{1i} and b_{2i} are the coefficients.

(3) The irreversible component is very complicated and is mainly related to the creep of concrete and rock. For arch dams, the time-dependent irreversible component increases rapidly at the beginning and slows down over time. A sum of the logarithmic function and the linear function are used:

$$\delta(t) = c_0 \ln(t) + c_1 t \tag{4}$$

where t is the days from the initial survey date to the current survey date and c_0 and c_1 are the coefficients.

B. CLASSIFICATION AND REGRESSION TREE MODEL

Classification and regression tree (CART) were proposed by Leo Breiman et al in 1984. It is a widely used decision tree algorithm [20]. Since the dam displacement monitoring data is essentially a regression problem, this article only describes the regression tree and the basic structure is shown in FIGURE 2. Assuming that X and Y are input and output variables, and Y is a continuous variable. Then, the given training data set is $D = \{(x_1, y_1), (x_2, y_2), \cdots, (x_N, y_N)\}$, where $x_i = (x_i^{(1)}, x_i^{(2)}, \cdots, x_i^{(n)})$ is the input instance

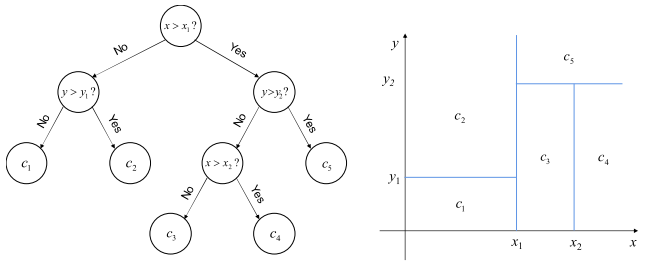

FIGURE 2. The schematic diagram of regression tree structure.

(feature vector), n is the number of sample feature. $i \in \{1, \ldots, N\}$ and N is the size of sample.

A heuristic method is adopted to divide the feature space, and the values of all the features in the current set are examined one by one, and the best one is selected as the cutting point according to the square error minimization criterion. The characteristic variable x^j and its own value s are used as the segmentation variable and segmentation point respectively. For the purpose of determining the optimal j and s, $R_1(j, s) = \{x|x^{(j)} \leq s\}$ and $R_2(j, s) = \{x|x^{(j)} > s\}$ are defined. Then, the (5) is used to solve the problem, that is, the j and s that minimize the sum of squared errors are determined after dividing the two regions.

$$\min_{j,s}\left[\min_{c_1}\sum_{x_i\in R_1(j,s)}(y_i-c_1)^2+\min_{c_2}\sum_{x_i\in R_2(j,s)}(y_i-c_2)^2\right] \quad (5)$$

where c_1 and c_2 are output values in the two regions after division.

In terms of the concept of square error, it can be known that the two optimal output values are the mean values of Y in their respective regions. Therefore, the (5) can be shown as:

$$\min_{j,s}\left[\min_{c_1}\sum_{x_i\in R_1(j,s)}(y_i-\hat{c}_1)^2+\min_{c_2}\sum_{x_i\in R_2(j,s)}(y_i-\hat{c}_2)^2\right] \quad (6)$$

where $\hat{c}_1 = \frac{1}{N_1}\sum_{x_i\in R_1(j,s)} y_i$ and $\hat{c}_2 = \frac{1}{N_2}\sum_{x_i\in R_2(j,s)} y_i$.

After all, its core is that the values of internal node features are yes and no, which is a binary tree structure. The regression is to determine the corresponding output value according to the feature vector. The regression tree divides the feature space into several units, and each division unit has a specific output. Because each node has a yes and no judgment, the boundaries of the partition are parallel to the axes. The relevant variable X_j corresponds to the predicted data. As long as it is classified into a certain unit according to the characteristics, the corresponding output value can be obtained. Decision trees have the advantages of handling both data type and conventional type attributes at the same time, and they can make feasible and effective results for large data sources in a relatively short time [30].

However, Chen *et al.* [31] discovered that if the regression tree takes all training data sets into consideration, the resulting regression tree will be too large. Although the fitting probability of the regression tree to the training data is 100%, theoretical and experimental evidence [32] shows that due to excessive consideration of all data, the regression tree learns some noise points and error points, even overfitting phenomenon.

C. RANDOM FOREST REGRESSION MODEL

Research [33] found that the depth of the regression tree is related to the suppression of the overfitting of the regression tree. Therefore, Leo Breiman [20] proposed the pruning principle of regression tree, which is achieved by minimizing the overall loss function or cost function of the regression tree to obtain the regression tree with the shortest height.

Nevertheless, pruning will sacrifice part of the training data learning, so Breiman [34] proposed RF model. RF model increases the randomness of guided clustering, and trees with higher resolution in the feature space will not be pruned. Therefore, the feature space will be divided into more and more smaller regions. Since the random forest is divided into Random Forest Classification (RFC) and the Random Forest Regression (RFR), the prediction of dam displacement is actually a regression problem, so this study only focuses on the RFR model.

RF is a supervised learning algorithm. Its base learner is a regression tree, but it can generate multiple trees without pruning. In the training process, it has two randomness [20]. One is the randomness of the sample. Each tree is built based on a random subset of the original data with replacement, that is, repeated sample data will be obtained when researchers make sampling. The other one is the randomness of features. When each regression tree is to be built, a certain number of candidate predictor subsets are randomly selected, and the most suitable value is selected as the split node.

It is assumed that the random regression forest prediction set is $\{h(x, \Theta_k), k = 1, 2, \ldots\}$, the appropriate output value of each tree is finally selected in the form of voting, and the predicted value is obtained based on averaging [20]. Moreover, all decision trees in the forest have the same distribution, and each regression tree has the right to vote. Therefore, the optimal output of each tree is up to the random vector $\{\Theta_k\}$, which be sampled independently, and the process is illustrated in the FIGURE 3.

Three parameters should be emphasized in RF modeling. The first one is *n_estimators*. It means the number of trees in the forest. The second one is *max_features*, which refers to maximum number of features in a single tree. The third one is *min_samples_leaf* and it is minimum size of sample for leaf nodes. The *max_features* are related to the prediction strength of each individual tree and the strength of the correlation between them. Increasing this parameter will also enhance the prediction ability of each tree and the strength of the correlation between them [35].

In terms of input vector X and output Y, the generalization error of regression tree $h(x)$ is $E_{X,Y}(Y - h(X))^2$. Leo Breiman [20] indicates that although RF model will not cause overfitting with the increase of the number of

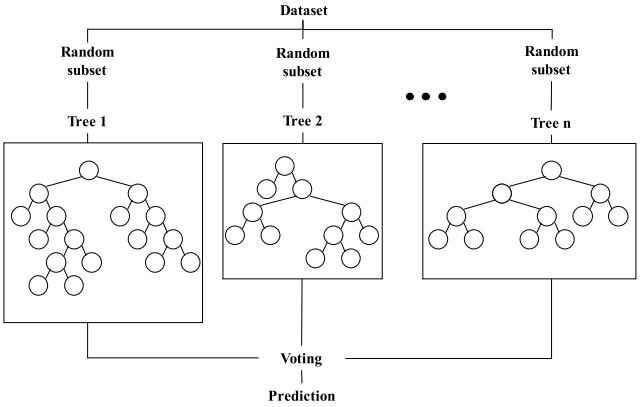

FIGURE 3. The schematic diagram of random forest structure.

regression trees, it may cause generalization errors within a certain range.

D. SLIDING TIME WINDOW STRATEGY

The original data of arch dam displacement monitoring is a one-dimensional time series, and the data collection frequency is 24 hours. In the real-time monitoring system, the data in the online module is continuously updated. In view of the characteristic of the time series, the model can be trained through capturing the data context. Therefore, the sliding time window strategy can be used to directly enter the newly generated data into the window, so as to alleviate the trouble of deleting the expired data. In addition, the model learning accuracy can be improved through the change trajectory in the captured time length, which can further improve the data processing efficiency.

When a data stream comes in, this time series is $t = \{x_1, x_2, \cdots, x_n\}$. At that time, the time window is moved to the starting point of the time sequence, and the time window intercepts a sub-sequence of length W, is the so-called the basic block of the sub-sequence. Then the time window continues to move back according to the step size. If the step size is one, the second point of the time sequence is taken as the starting point, and a subsequence of length W is continued to be extracted, and so on. In this way, a total of $n - W + 1$ subsequence basic block can be obtained.

In this study, the size of the sliding window is fixed, namely seven. The step size is one. The sliding time window strategy is shown in the FIGURE 4. It is assumed that the basic block of the sub-sequence numbered i is represented by B_i, with the length of 7. The optimized model can use the time series data of the first 5 days in B_i to predict that of the seventh day.

E. THE CONSTRUCTION PROCESS OF THE DAM DISPLACEMENT PREDICTION MODEL BY RF WITH A SLIDING TIME WINDOW

The generalization ability of the HST model is weak. In fact, it is based on the mathematical expression of dam safety theory. Each component can analyze the displacement data of the arch dam. The RF model requires fewer parameters to be

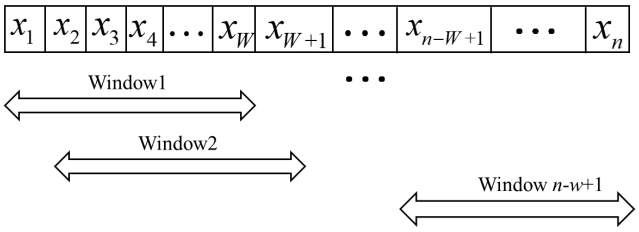

FIGURE 4. The schematic diagram of sliding time window strategy.

set by the user, and the implementation process is simple. For a good prediction result, it is necessary to improve the prediction strength of trees, lower the correlation between trees, and reduce the generalization error. Therefore, GS method is used to select the most appropriate value for each parameter.

The GS divides the training set data into K groups. Then, the value range of each parameter is set, each data subset is verified, and the remaining $K-1$ subset is used as the training set, so that each parameter combination will get K models. The final model regression result is the average regression result of K model validation sets. K is generally greater than or equal to two, and generally starts from three in reality. Only when the amount of data in the original data set is small, two will be tried. The GS schematic diagram is shown in the FIGURE 5.

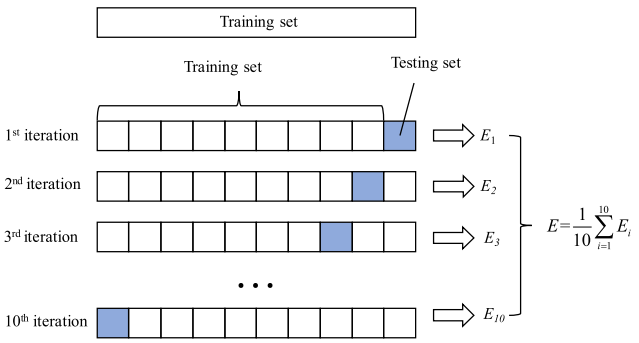

FIGURE 5. The schematic diagram of grid search.

In order to further improve the prediction of the RF model, this article uses segmented input vectors to capture the changing trend between variables and improve the ability to simulate dam response. Therefore, the RF model is coupled with a sliding time window strategy to adapt to the nonlinear interaction between high-dimensional input variables. The steps are shown in FIGURE 6.

Step 1: According to the length of the sliding time window, this study obtains the time sub-sequence basic block data set of the arch dam displacement monitoring data, and converts the two-dimensional tensor data into the three-dimensional tensor data. Therefore, the input variable is $\{h, h_2, h_3, h_4, T_7, T_{14}, \cdots, T_n, t\}$.

Step 2: Divide the basic block data set into a training set and a testing set, and the data are standardized to facilitate model training and speed up model convergence.

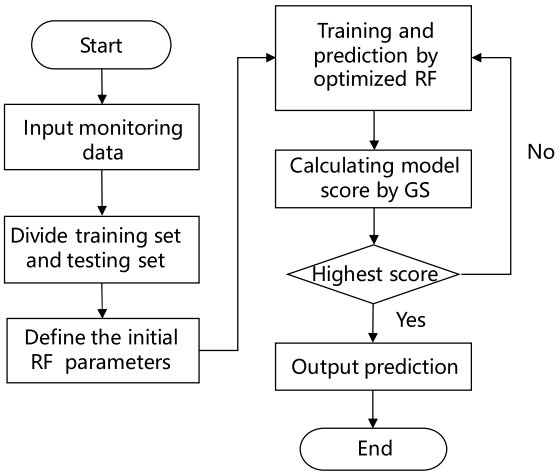

FIGURE 6. Flow chart of the optimized random forest model.

Step 3: Define the initial RF model parameters based on the input vector.

Step 4: Input the testing set into the optimized model and the predicted value of the arch dam displacement is obtained, which is the average value of each regression tree.

Step 5: Use the GS method to calculate the score of the model on the testing set, and the parameters are optimized according to the highest score. Repeat Step 4 to get the optimized forecast.

Step 6: Compare the model performance of the HST, RF, MLP, the optimized MLP, CNN, LSTM and the optimized RF.

III. CASE STUDY

A. PROJECT OVERVIEW

A parabolic hyperbolic stone masonry arch dam located in China. The main water retaining structure is a rubble masonry from fine aggregate concrete, as shown in FIGURE 7. It is composed of eleven sections in total, with sections 1# to 11# from the left bank to the right bank. It has a maximum height of 68.8 m and a crest length of 237.42 m. The total storage capacity of the reservoir is estimated to be 24.28 million m^3 and the dam crest elevation is 97.0 m. It is an annual adjustment and valley-type reservoir. The water level of the reservoir varies slightly as the season changes.

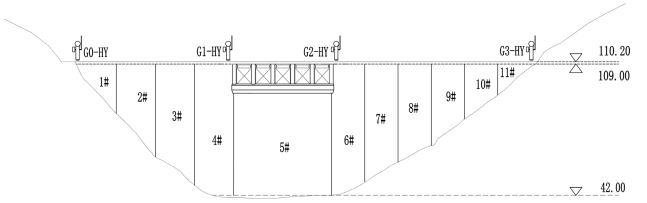

FIGURE 7. The spatial location of global navigation satellite systems.

Various sensors are installed on the dam to detect the performance of the dam. Environmental variables include temperature, hydrology, water level, and rainfall. For structural variables, the arch dam has displacement and seepage around the dam. The sensor has recorded data for more than 10 years. Therefore, through the analysis of different monitoring data, not only the current structural characteristics of the arch dam can be obtained, but also the long-term change law of the arch dam can be obtained.

The displacement of the arch dam is measured by the global navigation satellite systems (GNSS), and the measuring points are arranged respectively at 1#, 4#, 6# and 11# on the dam sections, as shown in FIGURE 7 for measuring the displacement of the dam crest and body. This study uses radial displacement monitoring data as the research object. For example, G0-HY represents the absolute displacement monitored by the GNSS at the 1# dam section.

B. DATA SET

The input environmental variables are the daily average temperature and daily water level measurement values from the beginning of the monitoring to the survey date. The water level and daily average temperature changes are shown in FIGURE 8. It can be seen from the figure that the water level varies from 66.3 m to 100.2 m, and an overall slight increasing trend can be observed. The temperature changes with the changing seasons, ranging from 4 °C to 36 °C, and the mean temperature is 24°C. Output variable is the observed displacement value. The radial displacement changes of each measuring point are shown in FIGURE 9. G0-HY varies from −0.8 mm to 5 mm, G1-HY varies from −6 mm to 23 mm, G2-HY ranges from −5 mm to 33 mm, and G3-HY ranges from −3 mm to 4 mm. The displacement of each measuring point has seasonal characteristics. This study takes June 2009 to December 2019 as the analysis period.

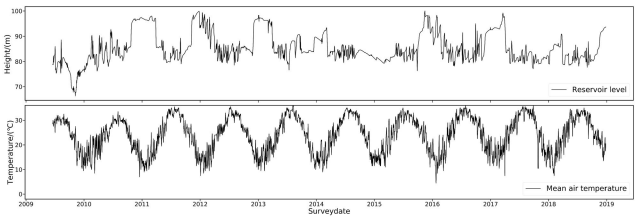

FIGURE 8. The environmental variables of water level and daily average temperature.

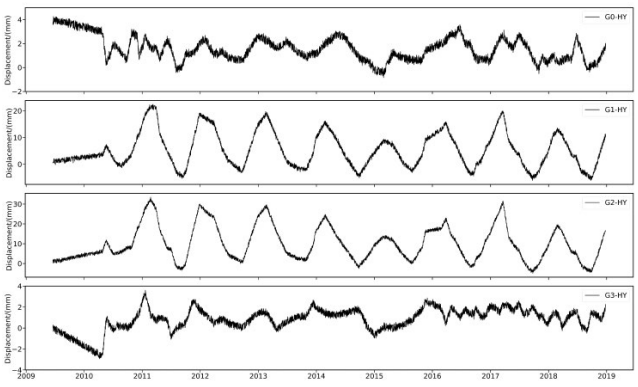

FIGURE 9. Radial displacement change of each measuring point.

G2-HY suddenly generated a small peak during the April 27, 2010 to July 11, 2010, which is the abnormal data.

Therefore, it is necessary to select the normal dam deformation time series and abnormal dam deformation time series to explore the universality of the proposed prediction model. We take G2-HY and G3-HY as representative examples to explore the effectiveness and robustness of the proposed prediction model in this study. Moreover, for the purpose of verifying the long-term prediction accuracy of the optimized RF model, 30% data and 60% data are selected to train the model, and the remaining data are used as the testing set for comparison. All models are trained in the same training set and verified in the same tesing set.

C. EVALUATION INDICATORS

The accuracy of the prediction models is evaluated based on the mean absolute error (MAE) on the testing set. MAE can effectively measure the performance of a forecasting model when tracking dam displacement [36]. The definition of MAE is provided below:

$$MAE = \frac{1}{M}\sum_{i=1}^{M}|y_i - \hat{y}_i| \tag{7}$$

where M is the size of the training or testing sets, y_i is the actual observed value, and $\hat{y}_i$ is the predicted value.

D. MODEL IMPLEMENTATION

In this study, for the purpose of alleviating the time-lag effect of impact factor phenomenon, a sliding time window strategy was introduced into the RF model to improve the time sensitivity. To execute the sliding time window strategy, the size of sliding window is needed to be predetermined, namely seven. Then, seven models, including HST model, MLP model, the optimized MLP model, CNN model, LSTM model, RF model and optimized RF model, are taken into consideration to explore the effectiveness of the proposed prediction model. These models for comparison have the same input variable set, $\{h, h_2, h_3, h_4, T_7, T_{14}, \cdots, T_n, t\}$, namely, the three-dimensional tensor data set. They are implemented using the Python module.

Moreover, GS is used for cross-validation and parameter selection, and the determination coefficient, R^2, is used to measure the final goodness of fit and score the final model. When R^2 approaches 1, it means that the model is much more effective for predicting the divided testing set. Conversely, the smaller the value of R^2 is, the worse the model's prediction on the divided testing set. Equation of R^2 can be rewritten as:

$$R^2 = \sum_{i=1}^{n}(\hat{y}_i - \frac{1}{n}\sum_{i=1}^{n}y_i)^2 / \sum_{i=1}^{n}(y_i - \frac{1}{n}\sum_{i=1}^{n}y_i)^2 \tag{8}$$

where n is the size of the training or validation sets, y_i is the actual observed value, and $\hat{y}_i$ is the predicted values.

Determine the values of the parameters of the best model score by dividing a wide range, and then narrow the range to be close to this value for further parameter optimization, and then loop until the model score decreases or no longer changes. Use GS to seek the optimal parameters of each model so as to ensure that the model converges to the optimal status and improve the prediction accuracy of each model.

In addition, in order to prevent overfitting, this study introduces the dropout technique to regularize some networks. The hyperparameters of the MLP models including the number of neurons in the hidden layers, and dropout rate. Four hyperparameters of the CNN model (the number of neurons in the convolution layers, the size of kernel, the number of layers, and dropout rate) make up a four-dimension space. The LSTM model including the number of layers, the number of neurons in each layer, and the dropout rates make up a three-dimension space. The RF model needs to define three parameters: the total number of trees, *n_estimators* (the initial value is 100), the maximum number of features of a single tree, *max_ features* with an initial value of 1, and the minimum number of leaf nodes sample, *min_samples_leaf* with an initial value is 1. TABLE 1 shows the final determined hyperparameters of these models.

TABLE 1. The Determined Hyperparameters of Different Models

Model	Hyperparameters	Domain	G2-HY	G3-HY
MLP	Number of neurons	[50,80]	60	64
	Dropout rate	{0.1,0.2,0.3}	0.2	0.2
RF	n_estimators	[50,100]	88	55
	max_features	[0,1]	0.7	0.7
	min_samples_leaf	[1,6]	4	4
CNN	Number of layers	{1,2,3}	2	2
	Number of neurons	[50,150]	128 70	62 55
	Convolution kernel size	[5,9]	7	7
	Dropout rate	{0.3,0.4,0.5}	0.4	0.5
LSTM	Number of layers	{1,2,3}	3	2
	Number of neurons	[50,150]	120 70 76	110 56
	Dropout rate	{0.3,0.4,0.5}	0.5	0.4

IV. RESULTS AND DISCUSSION

A. THE ADVANTAGE OF THE OPTIMIZED RF MODEL FOR PREDICTION ACCURACY

MAE of the training set and testing set are shown in TABLE 2, and are illustrated in FIGURE 10 for the sake of comparison. Due to the abnormality of the data series and the large variation range, the value of MAE for G2-HY is out of the normal range. The analysis is given in detail in section IV-C. The value of MAE for G3-HY varies from 0.12 to 3.12, with high accuracy.

Although HST model captures the trend of the time series in training set, it can be defined as a deviation from the normal displacement and the prediction accuracy is reduced for some cases where the amplitude is significantly larger. Compared with the HST model, the fitting of RF model and MLP model are stable on the training set, and even the abnormal peak and trough data series can be better fitted. Although the MAE value is significantly reduced, but MLP model could not accurately capture the trend of original dam displacement time series the testing set, meanwhile, RF has the generalization

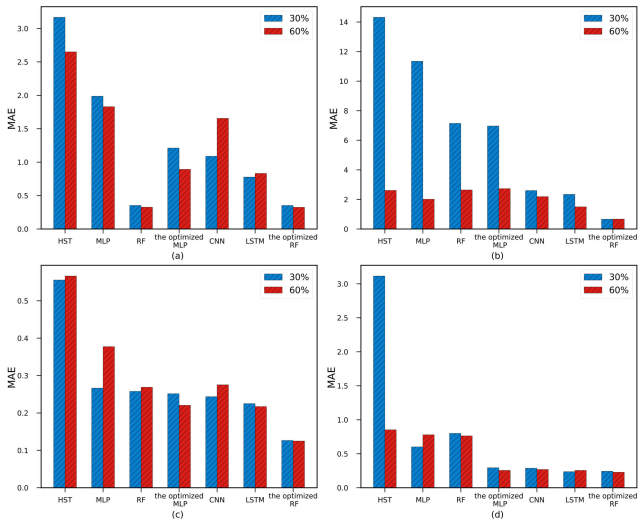

FIGURE 10. MAE of the prediction models for the selected monitoring variables: (a) Training set for G2-HY; (b) Testing set for G2-HY; (c) Training set for G3-HY; (d) Testing set for G3-HY.

TABLE 2. MAE of Different Models

	Models	G2-HY		G3-HY	
		Train	Test	Train	Test
60% data	**The optimized RF**	**0.3251**	**0.6699**	**0.1249**	**0.2276**
	HST	2.6499	2.6102	0.5661	0.8529
	MLP	1.8299	2.0126	0.3773	0.7791
	RF	0.3262	2.6436	0.2689	0.7631
	The optimized MLP	0.8927	2.7304	0.2205	0.2557
	CNN	1.6551	2.1887	0.2753	0.2685
	LSTM	0.8314	1.4963	0.2171	0.2556
30% data	**The optimized RF**	**0.3523**	**0.6626**	**0.1265**	**0.2328**
	HST	3.1672	14.3218	0.5554	3.1140
	MLP	1.9862	11.3459	0.2664	0.6015
	RF	0.3525	7.1462	0.2578	0.8003
	The optimized MLP	1.2113	6.9675	0.2515	0.2944
	CNN	1.0885	2.5986	0.2434	0.2872
	LSTM	0.7774	2.3448	0.2251	0.2367

error problem described in section II-C. And the prediction accuracy of both has not been effectively improved. CNN model, LSTM model, the optimized MLP model and the optimized RF model have similar MAE on the training set and are smaller than that from the HST model. It can be seen that by introducing time-sensitive factors, the optimized RF model captures the trend of displacement and improves the generalization ability of the model. Moreover, the prediction accuracy on the testing set is significantly higher than other models.

FIGURE 11 shows the predicted radial displacement for the G3-HY obtained from the seven models. These are also true for other prediction results, which are not illustrated herein. The training set selects 60% data from June 2009 to March 2015. The testing set selects data from April 2015 to December 2018. It can be seen that, due to the use of non-linear relationship to describe the evolution law, MLP model and RF model perform better at the training set than the HST model.

MLP model and RF model can capture the changing trend at the testing set as same as the HST model, meanwhile,

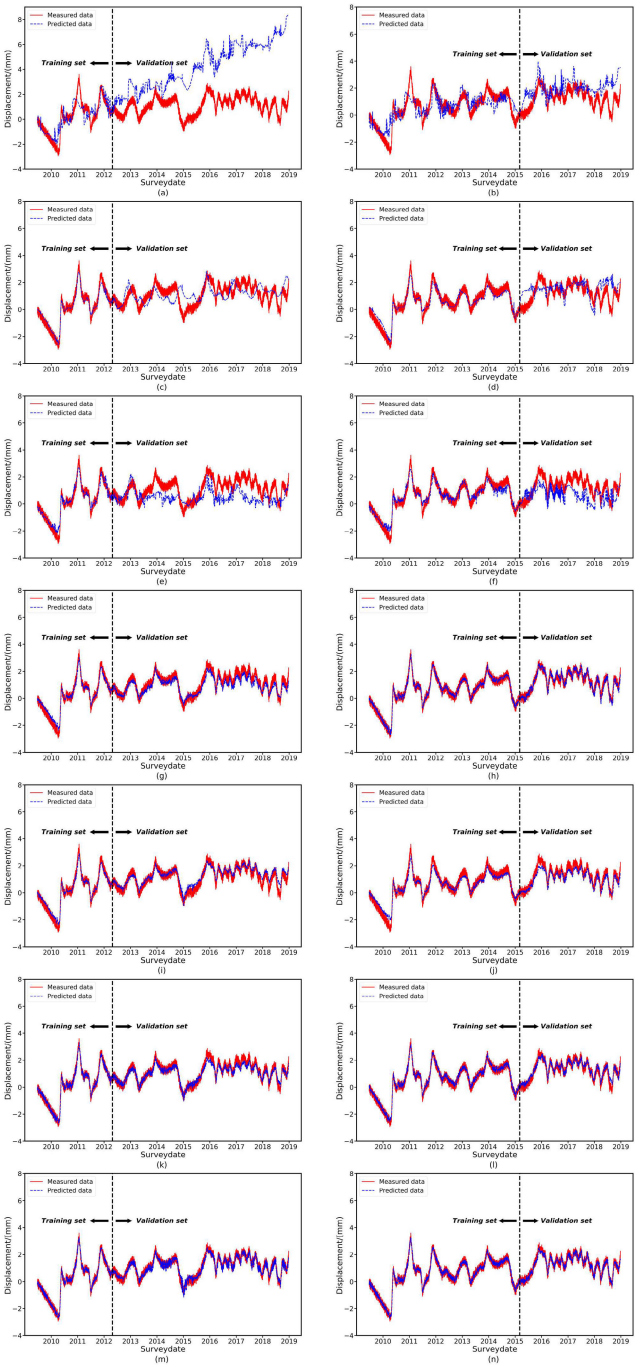

FIGURE 11. The predicted displacements of G3-HY from different models using training data with different lengths: (a) HST with 30% training data; (b) HST with 60% training data; (c) MLP with 30% training data; (d) MLP with 60% training data; (e) RF with 30% training data; (f) RF with 60% training data; (g) the optimized MLP with 30% training data; (h) the optimized MLP with 60% training data; (i) CNN with 30% training data; (j) CNN with 60% training data; (k) LSTM with 30% training data; (l) LSTM with 60% training data; (m) the optimized RF with 30% training data; (n) the optimized RF with 60% training data.

they improve predictions at the peaks and troughs of the curve. Compared with the previous models, CNN and LSTM perform better at the peak of the curve, confirming the conclusion that the deep learning model generally has better prediction than the machine learning model. However, thanks to the

introduction of a sliding time window strategy, the optimized MLP model and the optimized RF model not only fit the training set data well, but also achieve similar or better results with the deep learning model on the testing set.

But the optimized RF model has higher accuracy for the testing set compared with the optimized MLP model. Compared with RF model, MLP model is short of the characteristics of random feature selection. For large data sets, overfitting problems are prone to occur. As a result, the overfitting of the training set data reduces the prediction accuracy of the testing set. In contrast, the optimized RF model can capture the trend of changes between variables, thereby providing sufficient accuracy. Therefore, compared with the MLP model, the RF model is more suitable for coupling with the sliding time window strategy in terms of the prediction of the dam behavior. The optimized RF model has similar or better accuracy than the deep learning model, but has a simpler model architecture than them. Moreover, it can reflect the nonlinear effects of mean air temperature on the dam displacements.

B. THE ADVANTAGE OF THE OPTIMIZED RF MODEL FOR LONG-TERM PREDICTION

The testing set divided by the time series can be regarded as future data, which is a method to verify the generalization of the model. Therefore, this study selects the original data set within 30% as the training set to verify the training effect of the training set from June 2009 to April 2012. The remaining 70% data, which are from May 2012 to December 2018, are used as the testing set. The seven models are used to test the ability of long-term prediction in the future. Previously, 60% of the data is selected as the training set and 40% of the data is used as a contrast, and these seven models are used to test the future long-term prediction ability.

FIGURE 11 shows the predicted radial displacement for G3-HY obtained from the seven models. Other predicted results have similar phenomena, which are not illustrated in this article. It can be seen that the HST model fits well on the 30% training data set, but has a large deviation on the testing set, which is significantly higher than the prediction using 60% data for training. Because the reason behind is that the HST model only calculates the seasonal effect based on a simple linear function, so prediction of long-term displacement requires more data to fit.

MLP model has a great capability of nonlinear function approximation, so it is suitable for nonlinear time series prediction. Compared with the HST model, it has a significant improvement on the training set, but there are still errors on the testing set.

RF model performs better on the 30% training data set. Although RF model can fit the variation trend, but the prediction deviation of the displacement value is large, and it will eventually oscillate within an average value because RF model feeds back the average value of the prediction of each regression tree.

On the contrast, the prediction of CNN and LSTM reveals that using of a shorter training set could get the results as good as those from a longer training data set. This phenomenon could be explained with the feature's representation ability of CNN and the long-term dependency of LSTM. They store the useful history information and captures the long-term characteristics of time series, and thus, they could be trained with shorter monitoring data set and simulate the time effect of the displacements of the dam.

Meanwhile, the prediction of the optimized MLP and the optimized RF model reveals that using a shorter training set can achieve the same results as a longer training data set. The phenomenon can be explained by using the sliding time window strategy to capture the trend of monitoring data, making the model with the time sensitivity. The randomness of RF model suppresses the overfitting phenomenon and improves the generalization ability of the optimized model. Thus, a shorter monitoring data set can be used to train the optimized RF model to obtain a better dam displacement prediction model.

C. THE ADVANTAGE OF THE OPTIMIZED RF MODEL FOR ABNORMAL DATA SERIES

G2-HY suddenly generated a small peak during the April 27, 2010 to July 11, 2010. The abnormal phenomenon of displacement is caused by some complicated factors. It will pollute the data set, and affect the model identification, parameter estimation, diagnostic test and prediction of time series [37]. In this section, this study explores the reaction of the seven models to abnormal data series and compare how they are affected by the abnormal data.

FIGURE 12 shows the prediction results from the seven models, of which 30% of the monitoring data and 60% of the monitoring data are used for training. In terms of 30% monitoring data for training, only LSTM and the optimized RF model can provide satisfactory predictions. The other models have great deviations. Moreover, the HST model and MLP model cannot fit the abnormal data. In contrast, due to the randomness of the RF model, and LSTM forgets the old error information and storing the new useful information, which are robust to abnormal data, the fitting of abnormal data is good.

When 60% of data are used for training, great deviations are still found in the HST model. This is because the functions of the time effect in the HST model are simple and can lead to spurious regression. With the increase of training data, the prediction of CNN and MLP model is improved, but their architectures are susceptible to abnormal data and there are still great deviations on the testing set. On the contrary, RF model, the optimized MLP model and LSTM can fit the training data well, and only the prediction results of the testing set data at peaks and troughs have a small deviation. The optimized RF model inherits the fitting ability of RF model at training data set, and it also has sound predictive ability than other models. This is due to time sensitivity, which makes the model more robust to abnormal data sequences.

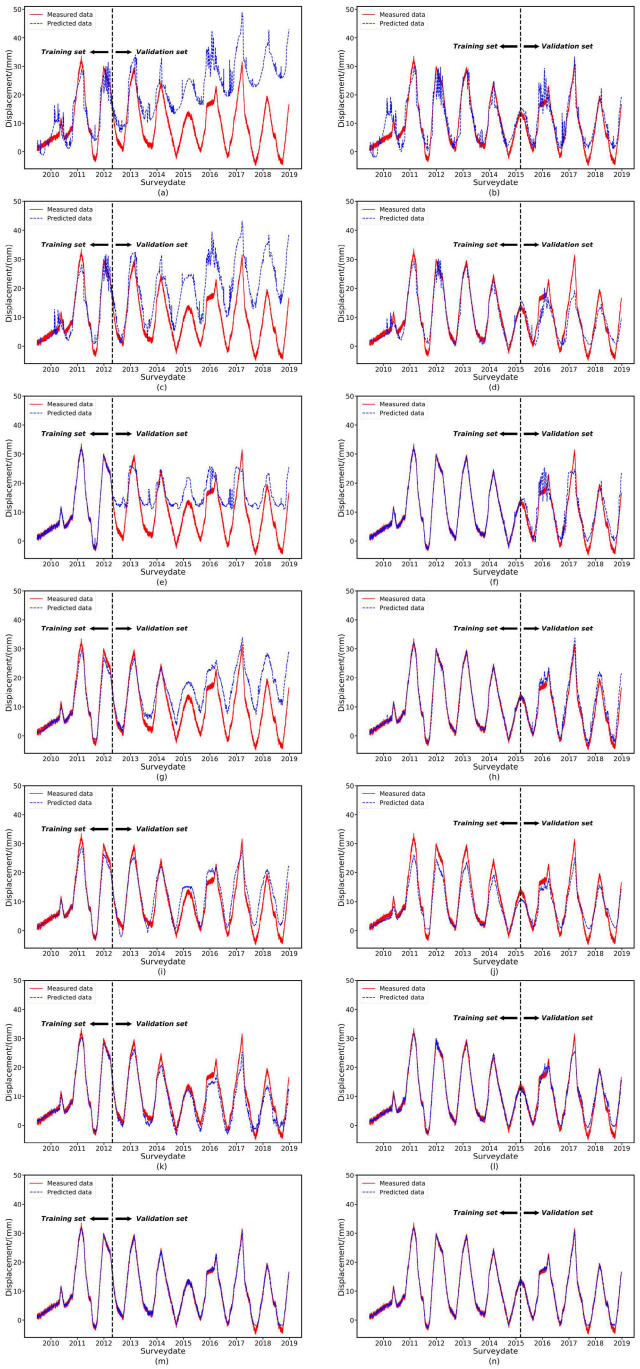

FIGURE 12. Comparison of predicted displacements from different models for abnormal time series: (a) HST with 30% training data; (b) HST with 60% training data; (c) MLP with 30% training data; (d) MLP with 60% training data; (e) RF with 30% training data; (f) RF with 60% training data; (g) the optimized MLP with 30% training data; (h) the optimized MLP with 60% training data; (i) CNN with 30% training data; (j) CNN with 60% training data; (k) LSTM with 30% training data; (l) LSTM with 60% training data; (m) the optimized RF with 30% training data; (n) the optimized RF with 60% training data.

V. CONCLUSION

The accuracy of the prediction model is of great importance for arch dam monitoring and safety assessment. This research proposes a RF model combined with sliding time window strategy to predict dam deformation. Water level, time and daily average temperature are input into the model. Compared to the HST model, the response of the dam can be predicted more accurately by simply adding the average temperature data. The performance of the proposed model was verified on the measured data of a masonry arch dam. In terms of thermal effects, the RF model based on long-term average temperature coupled with the sliding time window strategy can achieve better performance than the model using harmonic sine functions. As MAE is reduced observably, the accuracy is greatly improved. Therefore, the model can capture the hysteresis of temperature and reflect the influence of air temperature changes on structural behavior.

A masonry arch dam with the max height of 68.8 m was taken as the example. It can be seen that after adding mean air temperature, the prediction accuracy of the model can be significantly improved, because temperature in dam is affected by external temperature. Besides, as there is a hysteresis between internal temperature and external temperature changes, a sliding time window strategy is used to capture the time relationship of the original data. Numerical experiments show that the introduction of this strategy can obviously improve accuracy. For masonry arch dams, the appropriate time window size requires further tests on the actual monitoring data of masonry arch dams.

Compared with the HST model, MLP model, RF model, and the optimized MLP model, the proposed model provides higher accuracy for the displacement prediction. In contrast, the accuracy of the proposed model to predict dam displacement is almost the same as or even better than that of CNN and LSTM. Given that the proposed model has simpler structures and less parameters, and requires less time for model training, it may be the preferred method for short term dam displacement prediction. In addition, in terms of shorter monitoring data sets, the performance of this model is better than other models. The proposed model also has high robustness to abnormal time series to abnormal time series, because it has the randomness of sample extraction and random data feature selection and can increase the generalization ability of the model. To build predictive models, it relies entirely on data rather than previous assumptions about physical properties of the phenomenon. Moreover, the dam safety monitoring system is usually arranged with multiple measuring points, and there may be spatial correlation between the measuring points, but the proposed model cannot use multiple measuring points to explain these correlations. In the further, patio-temporal diagnosis method such as clustering analyses will be introduced into the field of arch dam safety monitoring to mine the dam deformation to avoid accident.

REFERENCES

[1] *International Commission on Large Dams*. Accessed: Sep. 9, 2019. [Online]. Available: https://www.icoldcigb.org/article/GB/world_register/general_synthesis/number-of-dams-by-country-members

[2] C. Penghao, L. Pingkuo, and P. Hua, "Prospects of hydropower industry in the yangtze river basin: China's green energy choice," *Renew. Energy*, vol. 131, pp. 1168–1185, Feb. 2019, doi: 10.1016/j.renene.2018.08.072.

[3] J. Jeon, J. Lee, D. Shin, and H. Park, "Development of dam safety management system," *Adv. Eng. Softw.*, vol. 40, no. 8, pp. 554–563, Aug. 2009, doi: 10.1016/j.advengsoft.2008.10.009.
[4] J. M. W. Brownjohn, "Structural health monitoring of civil infrastructure," *Phil. Trans. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 365, no. 1851, pp. 589–622, Dec. 2006, doi: 10.1098/rsta.2006.1925.
[5] A. De Sortis and P. Paoliani, "Statistical analysis and structural identification in concrete dam monitoring," *Eng. Struct.*, vol. 29, no. 1, pp. 110–120, Jan. 2007, doi: 10.1016/j.engstruct.2006.04.022.
[6] M. S. Al-Musaylh, R. C. Deo, J. F. Adamowski, and Y. Li, "Short-term electricity demand forecasting with MARS, SVR and ARIMA models using aggregated demand data in queensland, Australia," *Adv. Eng. Informat.*, vol. 35, pp. 1–16, Jan. 2018, doi: 10.1016/j.aei.2017.11.002.
[7] M. Li, Y. Shen, Q. Ren, and H. Li, "A new distributed time series evolution prediction model for dam deformation based on constituent elements," *Adv. Eng. Informat.*, vol. 39, pp. 41–52, Jan. 2019, doi: 10.1016/j.aei.2018.11.006.
[8] F. Li, Z. Wang, and G. Liu, "Towards an error correction model for dam monitoring data analysis based on cointegration theory," *Structural Saf.*, vol. 43, pp. 12–20, Jul. 2013, doi: 10.1016/j.strusafe.2013.02.005.
[9] Y. Li, T. Bao, J. Gong, X. Shu, and K. Zhang, "The prediction of dam displacement time series using STL, extra-trees, and stacked LSTM neural network," *IEEE Access*, vol. 8, pp. 94440–94452, May 2020, doi: 10.1109/ACCESS.2020.2995592.
[10] I. Penot, B. Daumas, and J. Fabre, "Monitoring Behavior," *Int. Water Power Dam Construct.*, vol. 57, no. 12, pp. 24–27, May 2005.
[11] S. Wang, C. Gu, and T. Bao, "Observed displacement data-based identification method of deformation time-varying effect of high concrete dams," *Sci. China Technol. Sci.*, vol. 61, no. 6, pp. 906–915, Jun. 2018, doi: 10.1007/s11431-016-9088-9.
[12] H. Su, Z. Wen, X. Sun, and M. Yang, "Time-varying identification model for dam behavior considering structural reinforcement," *Struct. Saf.*, vol. 57, pp. 1–7, Nov. 2015, doi: 10.1016/j.strusafe.2015.07.002.
[13] V. Ranković, N. Grujović, D. Divac, and N. Milivojević, "Development of support vector regression identification model for prediction of dam structural Behaviour," *Struct. Saf.*, vol. 48, pp. 33–39, May 2014, doi: 10.1016/j.strusafe.2014.02.004.
[14] 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," *Arch. Comput. Method Eng.*, vol. 24, no. 1, pp. 1–21, Jul. 2015, doi: 10.1007/s11831-015-9157-9.
[15] F. Salazar, M. Á. Toledo, E. Oñate, and B. Suárez, "Interpretation of dam deformation and leakage with boosted regression trees," *Eng. Struct.*, vol. 119, pp. 230–251, Jul. 2016, doi: 10.1016/j.engstruct.2016.04.012.
[16] C. Lin, T. Li, S. Chen, X. Liu, C. Lin, and S. Liang, "Gaussian process regression-based forecasting model of dam deformation," *Neural Comput. Appl.*, vol. 31, no. 12, pp. 8503–8518, Dec. 2019, doi: 10.1007/s00521-019-04375-7.
[17] J. Mata, "Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models," *Eng. Struct.*, vol. 33, no. 3, pp. 903–910, Mar. 2011, doi: 10.1016/j.engstruct.2010.12.011.
[18] W. Xi, J. Yang, J. Song, and X. Qu, "Deep learning model of concrete dam deformation prediction based on CNN," presented at the IOP Conf. Ser., Earth and Environ. Sci., Xi'an, China, Aug. 2020.
[19] X. Qu, J. Yang, and M. Chang, "A deep learning model for concrete dam deformation prediction based on RS-LSTM," *J. Sensors*, vol. 2019, pp. 1–14, Oct. 2019, doi: 10.1155/2019/4581672.
[20] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.
[21] C. Behrens, C. Pierdzioch, and M. Risse, "Testing the optimality of inflation forecasts under flexible loss with random forests," *Econ. Model.*, vol. 72, pp. 270–277, Jun. 2018, doi: 10.1016/j.econmod.2018.02.004.
[22] R. Genuer, J.-M. Poggi, and C. Tuleau-Malot, "Variable selection using random forests," *Pattern Recognit. Lett.*, vol. 31, no. 14, pp. 2225–2236, Oct. 2010, doi: 10.1016/j.patrec.2010.03.014.
[23] S. Lawrence and C. L. Giles, "Overfitting and neural networks: Conjugate gradient and backpropagation," in *Proc. IEEE-INNS-ENNS Int. Joint Conf. Neural Netw., IJCNN Neural Comput., New Challenges Perspect. New Millennium*, Jul. 2000, pp. 114–119.
[24] J. Kolluri, V. K. Kotte, M. S. B. Phridviraj, and S. Razia, "Reducing overfitting problem in machine learning using novel L1/4 regularization method," in *Proc. 4th Int. Conf. Trends Electron. Informat. (ICOEI)*, Jun. 2020, pp. 934–938.
[25] Y. Yang, J. Che, C. Deng, and L. Li, "Sequential grid approach based support vector regression for short-term electric load forecasting," *Appl. Energy*, vol. 238, pp. 1010–1021, Mar. 2019, doi: 10.1016/j.apenergy.2019.01.127.
[26] S. Smith and C. Costello, "Culinary tourism: Satisfaction with a culinary event utilizing importance-performance grid analysis," *J. Vacation Marketing*, vol. 15, no. 2, pp. 99–110, Apr. 2009, doi: 10.1177/1356766708100818.
[27] M. Herrera, L. Torgo, J. Izquierdo, and R. Pérez-García, "Predictive models for forecasting hourly urban water demand," *J. Hydrol.*, vol. 387, nos. 1–2, pp. 141–150, Jun. 2010, doi: 10.1016/j.jhydrol.2010.04.005.
[28] B. Dai, C. Gu, E. Zhao, and X. Qin, "Statistical model optimized random forest regression model for concrete dam deformation monitoring," *Struct. Control Health Monitor.*, vol. 25, no. 6, Mar. 2018, Art. no. e2170, doi: 10.1002/stc.2170.
[29] Y. Hu, C. Gu, Z. Meng, and C. Shao, "Improve the model stability of Dam's displacement prediction using a numerical-statistical combined model," *IEEE Access*, vol. 8, pp. 147482–147493, 2020, doi: 10.1109/ACCESS.2020.3014919.
[30] J. N. Morgan and J. A. Sonquist, "Problems in the analysis of survey data, and a proposal," *J. Amer. Stat. Assoc.*, vol. 58, no. 302, pp. 415–434, Jun. 1963.
[31] W. Chen, X. Xie, J. Wang, B. Pradhan, H. Hong, D. T. Bui, Z. Duan, and J. Ma, "A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility," *Catena*, vol. 151, pp. 147–160, Apr. 2017, doi: 10.1016/j.catena.2016.11.032.
[32] V. R. Basili, L. C. Briand, and W. L. Melo, "A validation of object-oriented design metrics as quality indicators," *IEEE Trans. Softw. Eng.*, vol. 22, no. 10, pp. 751–761, Oct. 1996, doi: 10.1109/32.544352.
[33] T. M. Khoshgoftaar, E. B. Allen, and J. Deng, "Controlling overfitting in software quality models: Experiments with regression trees and classification," in *Proc. 7th Int. Softw. Metrics Symp.*, Apr. 2001, pp. 190–198, doi: 10.1109/METRIC.2001.915528.
[34] L. Breiman, "Bagging predictors," *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, Aug. 1996.
[35] J. Peters, N. E. C. Verhoest, R. Samson, P. Boeckx, and B. De Baets, "Wetland vegetation distribution modelling for the identification of constraining environmental variables," *Landscape Ecol.*, vol. 23, no. 9, pp. 1049–1065, Sep. 2008, doi: 10.1007/s10980-008-9261-4.
[36] W.-C. Wang, K.-W. Chau, L. Qiu, and Y.-B. Chen, "Improving forecasting accuracy of medium and long-term runoff using artificial neural network based on EEMD decomposition," *Environ. Res.*, vol. 139, pp. 46–54, May 2015, doi: 10.1016/j.envres.2015.02.002.
[37] M. Munir, S. A. Siddiqui, A. Dengel, and S. Ahmed, "DeepAnT: A deep learning approach for unsupervised anomaly detection in time series," *IEEE Access*, vol. 7, pp. 1991–2005, 2019, doi: 10.1109/ACCESS.2018.2886457.

YAN SU received the B.S. and M.S. degrees from Fuzhou University, Fuzhou, China, in 1996 and 1999, respectively, and the Ph.D. degree from Tongji University, Shanghai, China, in 2005. She is currently an Associate Professor and a M.S. Tutor with Fuzhou University. Her research interests include safety monitoring, the evaluation and feedback analysis of hydraulic buildings, and machine learning.

KAILIANG WENG received the B.S. degree from Fuzhou University, Fuzhou, China, in 2019, where he is currently pursuing the M.S. degree in hydraulic structural engineering. His research interests include data mining and the application of machine learning technology in the field of dam safety monitoring.

CHUAN LIN received the B.S. and Ph.D. degrees in hydraulic structural engineering from Hohai University, Nanjing, China, in 2013 and 2019, respectively. He is currently a M.S. Tutor with Fuzhou University. His research interests include safety assessment and the stability analysis of dam structure.

ZHIMING ZHENG received the B.S. degree from Fuzhou University, Fuzhou, China, in 2020, where he is currently pursuing the M.S. degree in hydraulic structural engineering. His research interests include machine learning and the stability analysis of dam structure.