Article

Exploration of Multi-Scale Reconstruction Framework in Dam Deformation Prediction

Rongyao Yuan , Chao Su *, Enhua Cao *, Shaopei Hu and Heng Zhang

College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing 210098, China; 190802020006@hhu.edu.cn (R.Y.); hushaopei@hhu.edu.cn (S.H.); zhangheng@hhu.edu.cn (H.Z.)

* Correspondence: csu_hhu@126.com (C.S.); caoenhua@hhu.edu.cn (E.C.)

Abstract: Affected by various complex factors, dam deformation monitoring data usually reflect volatility and non-linear characteristics, and traditional prediction models are difficult to accurately capture the complex laws of dam deformation. A multi-scale deformation prediction model based on Variational Modal Decomposition (VMD) signal decomposition technology is proposed in this study. The method first decomposes the original deformation sequence into a series of sub-sequences with different frequencies, then the decomposed sub-sequences are modeled and predicted by Long Short-Term Memory neural network (LSTM) and Random Forest (RF) according to different frequencies. Finally, the prediction results of all sub-sequences are reconstructed to obtain the final deformation prediction results. In this process, it is proposed to use the instantaneous frequency mean method to determine the decomposition modulus of VMD. The innovation of this paper is to decompose the monitoring data with high volatility, and use LSTM and RF prediction, respectively, according to the frequency of the monitoring data, so as to realize the more accurate capture of volatility data during the prediction process. The case analysis results show that the proposed model can effectively solve the negative impact of the original data volatility on the prediction results, and is superior to the traditional prediction models in terms of stability and generalization ability, which has an important reference value for accurately predicting dam deformation and has far-reaching engineering significance.

Keywords: dam deformation; multi-scale prediction; VMD; LSTM; random forest

Citation: Yuan, R.; Su, C.; Cao, E.; Hu, S.; Zhang, H. Exploration of Multi-Scale Reconstruction Framework in Dam Deformation Prediction. *Appl. Sci.* **2021**, *11*, 7334. <https://doi.org/10.3390/app11167334>

Academic Editor: Daniel Dias

Received: 12 July 2021

Accepted: 6 August 2021

Published: 9 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

So far, there are approximately 98,000 dams in China, which play a significant role in flood control and power generation. Once the dam bursts, it will cause irreversible damage to the environment, society and economy [1,2]. Generally, the collapse of the dam is caused by the long-term deformation. Therefore, it is the primary task to accurately predict the deformation of the dam at present. Furthermore, how to improve the prediction accuracy is an important issue in the field of the dam's safety monitoring [3]. From the perspective of the current relevant research, the common dam deformation analysis and prediction models are mainly divided into the following three categories: the statistical model [4–6], the deterministic model [7,8], and the mixed model [9–11].

Based on the regression theory, the statistical model is usually applied to output the linear relationship between the influencing factor and the target variable. Wu Zhongru et al. [12,13] have conducted more comprehensive research in this area, and their findings have been widely utilized. However, the dam deformation is usually affected by factors such as temperature and water level. These influencing factors present a strong nonlinear relationship, which can be quantified by the long-term persistence (LTP). The global scale analysis reveals that there is long-term persistence in all the above factors [14]. Hence, the existence of these LTP factors makes the deformation of the dam exhibit strong volatility, which increases the difficulty to accurately predict the deformation of the dam. Although the regression model can establish a nonlinear relationship between variables, its

generalization ability is weak. Based on mechanical assumptions, the deterministic model simulates the working process of the dam through the finite element method and calculates its displacement and stress [15]. However, this method requires a heavy workload and there is a strong nonlinear relationship between dam deformation and influencing factors, so that it can only implement qualitative analysis. In recent years, with the development of artificial intelligence, machine learning has made breakthroughs in data mining, especially the neural network models represented by BP, ARIMA, SVM, ELM, etc. Many scholars applying neural network to dam deformation prediction have achieved good results [16–18], which has further improved the processing of nonlinear problems in comparison with the regression model. The ARIMA model can capture processes with strong volatility, while it requires a large number of parameters when the process exhibits long-range dependence [19]. It makes the model complicated and difficult to handle. However, multivariate forecasting models usually only consider environmental factors and ignore the time dependence of the deformation sequence itself. When the deformation sequence has strong volatility, the above methods have poor robustness and generalization ability.

In order to solve the shortcomings of the above models, many scholars have proposed to adopt the LSTM [20] model to solve the problem of the dam's deformation prediction. Due to a long-term memory function, LSTM possesses a unique advantage in time series prediction. Mei-Li Shen et al. [21] proposed employing LSTM to predict import and export. They found that LSTM can effectively simulate the uncertainty trend in the data. Zhangtao et al. [22] applied LSTM to predict ship motion and proved that this approach has better performance. Yan.s et al. combined LSTM with the attention mechanism, so that the model could attach great importance to other more important factors in the time dimension [23]. Through the analysis of the results, it is found that although the model has made great improvements in the overall prediction accuracy, the extreme value prediction is still insufficiently accurate when the data's volatility is large.

In this case, this study proposes a multi-scale dam deformation prediction model based on signal decomposition. The model decomposes the original deformation sequence into a finite number of sub-sequences with a variety of frequencies, which can greatly reduce the model's nonlinearity with the potential to improve the model's prediction accuracy and robustness.

The traditional signal decomposition techniques mainly include Fourier transform, Wavelet Decomposition (WD) and Empirical Mode Decomposition (EMD) which are frequently applied to decompose the dam deformation sequence [24]. Fourier transform can realize the mutual conversion from time domain to frequency domain, while its conditions are relatively harsh, and there is no Fourier transform for considerable useful signals; WD has a certain priority, while it is difficult to choose the basis function and decomposition scale. Furthermore, its practical application is more complicated [25]; EMD does not need to set a basis function, which is capable to be adaptively decomposed into a finite number of modal functions. However, it is prone to modal confusion and lacks the support of mathematical theory [26,27]. Aiming at solving the above problems, this paper introduces a new adaptive signal decomposition method—Variational Modal Decomposition (VMD) [28]. In contrast to the above decomposition method, it is a more accurate mathematical model, which can decompose the original data into a set of eigenmode functions (IMFs) fluctuating around the center frequency [29–31], with a better decomposition effect as well as higher robustness [9,32]. However, the decomposition effect of VMD is affected by the decomposition modulus, and too many or too few subsequences will affect the decomposition effect [33,34]. In order to solve this problem, this paper proposes the instantaneous frequency average method to analyze the decomposition modulus K so as to avoid the modal confusion caused by excessive modulus or insufficient decomposition due to too small modulus.

In summary, this paper proposes a multi-scale dam-deformation prediction model based on VMD-LSTM-RF, which greatly reduces the complexity of the original sequence. In this model, the original sequence is first decomposed into sub-sequences with various

frequencies and residuals. The LSTM is suitable for dealing with complex nonlinear problems with good long-term memory capabilities; it has the following characteristics: its results are reliable in predicting weak volatility series; the model is stable; it is not prone to overfitting, and the model performance is robust. Therefore, LSTM is applied to predict high-frequency components and residuals, and the low-frequency components are input to RF for prediction. Finally, the prediction results of the two parts are superimposed and reconstructed to obtain the final prediction results, which can not only avoid the inaccurate prediction problem caused by too large volatility of LSTM, but can also reduce the workload to a large extent, and effectively improve the accuracy of dam prediction. In order to verify the superiority of the proposed model, the most commonly applied methods are selected as benchmark methods, and the three quantitative evaluation indicators are utilized to evaluate the prediction performance of the proposed model. The main contributions of this paper are as follows:

- (1) The VMD decomposition technology is introduced, and the instantaneous frequency average method is adopted to determine the decomposition modulus K , and the dam displacement and deformation data with strong volatility are decomposed into periodic and stable sub-sequences.
- (2) On the basis of LSTM and RF, the high and low frequency sub-sequences are modeled and predicted, respectively. The analysis indicates that this approach can accurately capture the nonlinear characteristics of the dam deformation.
- (3) This paper verifies the feasibility of the signal decomposition technology as well as the machine learning in dam deformation prediction, and predicts high frequency and low frequency separately, which reduces the data's complexity of data and greatly improves the accuracy of the dam's deformation prediction and has the obvious significance for actual project.

The rest contents of this paper are as follows: the second part briefly introduces VMD, LSTM and RF-related theories, as well as the method of determining the K value of VMD; the third part introduces the two cases' research design, evaluation indicators, model realization, and the determination of relevant model parameters; the fourth part compares and analyzes the prediction performance of the proposed model and other models; the fifth part draws conclusions for the whole paper.

2. Materials and Methods

This section describes the proposed model, and briefly introduces the basic principles of VMD, LSTM, and RF, respectively, and then establishes the VMD-LSTM-RF model, and gives the modeling process and the specific steps of the model in detail.

2.1. VMD-Based Decomposition Technology

VMD is a new type of self-adaptable signal decomposition technology with preset scale. It can decompose a real-valued signal into K modal components around the center frequency. The process of signal decomposition is the process of solving the variational problem, as shown in Formula (1):

$$\left\{ \begin{array}{l} \min \\ \{u_k\}, \{\omega_k\} \end{array} \right. = \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) u_k(t) \right] * e^{-j\omega_k t} \right\|_2^2 \right\} \quad (1)$$

$$\sum_k u_k = f$$

where $\{u_k\} = \{u_1, u_2, \dots, u_k\}$ are the K modal components, $\{\omega_k\} = \{\omega_1, \omega_2, \dots, \omega_k\}$ are the center frequency of each modal component, f is the original signal, and δ_t is the pulse function. In order to obtain the optimal solution of the constrained variational problem, the Lagrange multiplication operator $\lambda(t)$ and the quadratic penalty factor a are introduced to transform the constrained variational problem into an unconstrained variational problem. The extended Lagrange function is expressed as:

$$L(\{u_k\}, \{\omega_k\}, \lambda) = a \sum_k \|\partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) u_k(t) \right] * e^{-j\omega_k t}\|_2^2 + \|f(t) - \sum_k u_k(t)\|_2^2 + \langle \lambda(t), f(t) - \sum_k u_k(t) \rangle \quad (2)$$

Alternate direction operator multiplication is used to obtain the saddle point of the above Lagrangian function, then the optimal solution can be obtained. Specific steps are as follows:

- (1) Initializing $\{u_k^1\}$, $\{\omega_k^1\}$, λ^1 , $n \leftarrow 0$ and converting each parameter from time domain to frequency domain.
- (2) $n \leftarrow n + 1$. On the non-negative frequency domain interval, $\omega \geq 0$, for $k = 1 : K$, updating $\hat{u}_k$, ω_k :

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i < k} \hat{u}_i^{n+1}(\omega) + \hat{\lambda}^n(\omega)/2}{1 + 2a(\omega - \omega_k^n)^2} \quad (3)$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k^{n+1}(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k^{n+1}(\omega)|^2 d\omega} \quad (4)$$

- (3) For all $\omega \geq 0$, updating $\hat{\lambda}^n$:

$$\hat{\lambda}^{n+1}(\omega) = \hat{\lambda}^n + \tau(\hat{f}(\omega) - \sum_k \hat{u}_k^{n+1}(\omega)) \quad (5)$$

In the above Formulas, $\hat{u}_k^{n+1}(\omega)$, $\hat{f}(\omega)$ and $\hat{\lambda}^{n+1}(\omega)$ are the Fourier transforms of $u_k^{n+1}(\omega)$, $f(\omega)$ and $\lambda^{n+1}(\omega)$, respectively.

- (4) Judging whether the convergence condition (judgment accuracy) is satisfied, if the following Formula is satisfied, stop the iteration, otherwise return to step (2). When the entire period is over, the result is output, and K eigenmode components (IMF) are obtained.

$$\frac{\sum_k \|\hat{u}_k^{n+1} - \hat{u}_k^n\|_2^2}{\|\hat{u}_k^n\|_2^2} < \varepsilon \quad (6)$$

2.2. Method of Determining K Value Based on the Mean Value of Instantaneous Frequency

When using VMD to decompose the original signal, the value of the decomposition modulus K will greatly affect the decomposition effect. If the K value is too small, the decomposed subsequence will lose information or cause modal heterogeneity; on the contrary, if the K value is too large, it will lead to excessive decomposition, increase the amount of calculation of the subsequent neural network, and then affect the prediction effect [35]. Therefore, it is very important to select an appropriate K value and fully extract the characteristics of the original data. This article introduces the instantaneous frequency mean value method of the component to select the K value. The main idea is that the IMFs obtained based on VMD decomposition are with obviously different frequencies. By dividing the original sequence into sub-sequences with different modulus, the mean value of the component instantaneous frequency under various K values is calculated. If the average change corresponding to two adjacent K values is significantly reduced, it is considered that the over-decomposition has caused the modal confusion phenomenon at this time, and this critical K value is the optimal decomposition modulus. In order to verify the effectiveness of this method, an analog signal is established, which is shown in Formula (7).

$$x(t) = \cos(4\pi t) + \frac{1}{8} \cos(48\pi t) + \frac{1}{32} \cos(576\pi t) \quad (7)$$

After the signal is decomposed by VMD, three IMFs should be obtained, as shown in Figure 1:

Figure 1. Decomposition results of analog signals based on VMD.

Using the instantaneous frequency mean value method, the decomposition is divided into deformed sub-sequences with different modulus, and the instantaneous frequency mean values of the components under different K values are calculated. The specific calculation results are shown in Figure 2. It can be seen that when $K \geq 4$, each component has a certain degree of modal confusion. Therefore, for analog signals, $K = 3$ is the best decomposition modulus, which is consistent with the preset K value. It is proved that the method is effective in selecting the modulus of VMD.

Figure 2. Variation of the mean values of the instantaneous frequencies of the components corresponding to different moduli.

2.3. Long Short-Term Memory Network (LSTM)

LSTM is transformed from Recurrent Neural Network (RNN), which solves the problem of gradient disappearance and gradient explosion caused by the long-term dependence of RNN on the time series, and can effectively deal with the prediction problem of the time

series [18]. The essence of RNN is to add a feedback mechanism to the fully connected neural network, which determines that the output of the network is not only related to the current input, but also related to the previous output while LSTM adds a memory unit on the basis of RNN. Compared with RNN, LSTM has a longer memory time, but its control process is similar to RNN. They both process the data flowing through the cell in the process of forward propagation. The difference is that the structure and operation of the cells in LSTM have changed.

The structure of LSTM is shown in Figure 3. It is mainly composed of a memory unit that stores the information state and three gate control units to regulate the flow of information in and out of the memory unit. The memory unit can retain time sequence hidden information to achieve the purpose of using longer time sequence information. The three gating units mainly use the activation function to change the information state in the memory unit, while the input gate is mainly to update and store the information in the cell state, and the output gate selects the useful part for prediction from the final memory unit. The cell unit update status of LSTM is as follows below.

Figure 3. LSTM schematic.

Forgetting gate: It outputs a 0–1 vector by checking the information of the sum to determine how many memories of the cell state in the previous period need to be forgotten. At this time, the output of the cell state is:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (8)$$

Input gate: Using h_{t-1} and x_t to decide what new information to add to the cell state, the update Formula is as follows

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (9)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (10)$$

Cell state update: Using the state f_t output by the forget gate and the i_t and $\tilde{C}_t$ obtained from the input gate to update the cell state so that the next state can be used. The update Formula is as follows:

$$C_t = C_{t-1} \odot f_t + i_t \odot \tilde{C}_t \quad (11)$$

Output gate: After updating the cell state, it is necessary to determine which state features of the output cell are based on the sum of the inputs. The calculation Formula is:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (12)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (13)$$

In the Formulas (8)–(13), W_f , W_i , W_o and b_f , b_i , b_o are the weight matrix and bias of the forgetting gate, input gate, and output gate, respectively, and σ is the activation function.

As shown in Figure 3, the output of one cell state on the hidden layer will be used as the input of the next state. This transfer mechanism makes LSTM have better learning and memory capabilities. This allows the model to obtain the optimal parameters through iteration. In essence, iteration is a process of seeking optimization, and its loss function is the objective function. In this article, the Adam optimizer is used for stochastic gradient descent. Compared with other optimizers, Adam is more efficient and has fewer parameters.

2.4. Random Forest

RF is a special bagging method, which is an organic combination of bagging algorithm and decision tree. In each decision tree in the random forest algorithm, there is no connection between each other. After the model is established, whenever a new deformed input sample appears, each decision tree in the model will make a judgment and judge which type the sample belongs to, then the more frequent ones will be the selection result of the final data analysis [36]. The RF algorithm flow is as follows.

- (1) Among N samples, T training sets $S_1, S_2, \dots, S_T$ are formed by Bootstrap sampling.
- (2) Each training set S generates a corresponding decision tree $C_1, C_2, \dots, C_T$.
- (3) The node of each decision tree is split, and an attribute is randomly selected from the attributes as the attribute set of the current node split, and a feature is selected for splitting based on the principle of minimum absolute average error or square average error.
- (4) After the decision tree is constructed, there is no need to perform pruning processing.
- (5) Testing set sample X , after testing with each decision tree C , get the predicted value.
- (6) The average value of all decision trees is used as the final predicted value.

2.5. Proposed Model (VMD-LSTM-RF)

Combining the above-mentioned theories, this section proposes a hybrid model based on multi-scale prediction, and its process is shown in Figure 4. The implementation of the hybrid model is mainly divided into the following steps:

Step 1: Decomposition. Decomposing the original displacement sequence of the dam into several components such as low frequency, high frequency and residual by VMD method.

Step 2: Components Prediction. Using LSTM to predict high-frequency components and residuals (ER). The low-frequency components are predicted by RF. The optimal parameters of the model are determined by grid search and five-fold cross-validation.

Step 3: Refactor. Adding and reconstructing the components obtained by the above prediction model to obtain the final dam deformation prediction value.

Step 4: Evaluation. In order to verify the accuracy of the model prediction, the average absolute error (MAE), root mean square error (RMSE) and coefficient of determination (R^2) are used for evaluation and compared with the basic model.

2.6. Comparison Model Extreme Learning Machine (ELM)

ELM is a new type of feedforward neural network proposed by Huang et al. in 2006, which consists of an input layer, a hidden layer and an output layer. In this article, the neurons in the input layer correspond to the n input variables ($x_1, x_2, x_3, \dots, x_n$) in the model, and the neurons in the output layer correspond to the horizontal displacement of the concrete dam crest. The specific principles of ELM are as follows:

Given a Q group sample (x_n, t_m) , where $x_n = [x_{n1}, x_{n2}, \dots, x_{nQ}]^T$, and $t_m = [t_{m1}, t_{m2}, \dots, t_{mQ}]^T$. Assuming that the hidden layer activation function is $g(x)$, the output vector is:

$$T = [t_1, t_2, \dots, t_Q]_{m \times Q} \quad (14)$$

$$t_j = [t_{1j}, t_{2j}, \dots, t_{mj}]^T = \sum_{i=1}^l \beta_{im} g(\omega_i x_i + b_i) (j = 1, 2, \dots, Q) \quad (15)$$

Figure 4. Flow chart of the proposed model.

Among them, b_i represents the threshold of the i -th neural node in the hidden layer; $\omega_i = [\omega_{1i}, \omega_{2i}, \dots, \omega_{mi}]^T$ is the weight from the input layer of the i -th neural node to the hidden layer; $\beta_i = [\beta_{1i}, \beta_{2i}, \dots, \beta_{mi}]^T$ represents the weight from the hidden layer to the i -th neural node of the output layer. Formula (15) can be expressed as:

$$H\beta = T' \quad (16)$$

According to the ELM theorem, when the number of training set samples is large, the number of hidden neurons K is usually less than Q . At this time, the training error of ELM can be infinitely approximated to any real number ε greater than 0.

$$\sum_{j=1}^Q ||t_j - y_j|| < \varepsilon \quad (17)$$

The connection weight β between the hidden layer and the output layer can be obtained by finding the least squares solution of $\min_{\beta} ||H\beta - T'||$, then $\hat{\beta} = H^+ T'$. Where H^+ is the generalized inverse Moore-Penrose matrix of matrix H .

2.7. Research and Design

2.7.1. Project Overview

Taking a roller compacted concrete gravity dam as an example, which is located in Fujian Province, China, the design flood level is 633.0 m. The elevation of the foundation surface is 562.0 m, and the maximum dam height is 73.4 m. The horizontal displacement

monitoring of the concrete dam crest adopts the tension line method, and a total of 11 measuring points are arranged, of which 9 working measuring points are located on the top of each dam section; 2 check base points are located on the left and right ends of the tension line. The deformation monitoring system of the dam includes horizontal displacement and vertical displacement. This article mainly models and analyzes the horizontal displacement of the dam crest. Taking into account the unevenness and local differences of the dam deformation, the monitoring data of the middle measuring point EX5 and the left bank measuring point EX2 of the fourth dam section are taken for analysis, as shown in Figure 5. The deformed data obtained by processing the outliers of these two sets of data are shown in Figure 6. It can be seen that the data is highly volatile and EX5 shows periodicity on the whole, but EX2 looks messy and has no rules to follow. The numbers of deformation monitoring data of the two measuring points are 739 and 417, respectively. The division results of their training set, validation set and prediction set are shown in Table 1.

Figure 5. Dam layout plan.

(a)

(b)

Figure 6. Deformation monitoring data graph((a) EX5, (b) EX2).

Table 1. Division of measuring point data set.

Measurement Points	Total	Training Set	Validation Set	Prediction Set
EX5(739)	739	539	100	100
EX2(417)	417	317	50	50

2.7.2. Determination of K Value Based on VMD

Since the choice of decomposition modulus K will greatly affect the decomposition effect of VMD, it is necessary to determine the decomposition modulus K before proceeding

to decompose the data. In this section, EX5 is taken as an example, the data are decomposed nine times using the VMD model, and the instantaneous frequency mean value change line chart of the component under different K values is calculated, as shown in Figure 7. It can be seen from the figure that when $K = 6$, the instantaneous frequency mean change of the two components is significantly reduced. When $K > 6$, this trend is more obvious, indicating that there is modal confusion in the decomposition results at this time, which will make the sub-sequences obtained by decomposition not able to express the dam deformation characteristics well. Therefore, $K = 5$ is selected as the best decomposition modulus, and the measurement point of EX2 is selected as $K = 6$ according to the above method; the specific decomposition results of EX5 and EX2 are shown in Figure 8.

Figure 7. EX5 k value graph.

It can be seen from the decomposition results that the frequency characteristics of all components are obvious, and there are no undesirable phenomena such as modal confusion. In the decomposition results of the corresponding deformation data of EX5, the frequency of IMF1–IMF3 is lower, IMF1 is basically consistent with the dam deformation trend, and the IMF2–IMF3 component process line is relatively smooth; IMF4–IMF5 has obvious periodicity, and ER has high volatility. Therefore, IMF1–IMF3 will be used as the input of RF; IMF4–IMF5 and ER will be the input of LSTM. Similarly, in the decomposition result of the corresponding deformation data of EX2, IMF1–IMF4 are used as the input of RF, and IMF5–IMF6 and ER are used as the input of LSTM.

2.7.3. Evaluation Index

In order to evaluate the proposed model, six evaluation indicators are introduced for comparison with other models in this study. MAE and RMSE are used to indicate the degree of deviation between the predicted value and the true value, and R^2 is used to express the correlation between the true value and the predicted value; P_{MAE} , P_{RMSE} and P_{R^2} are used to measure the degree of improvement of the proposed model compared with other models. Their definitions and formulas are shown in Table 2, in which y_i is the observed value and $\hat{y}_i$ is the predicted value.

Table 2. Evaluation index formula.

Metric	Equation	Definition
MAE	$MAE = (\sum_{i=1}^n y_i - \hat{y}_i) / n$	Mean absolute error
RMSE	$RMSE = \sqrt{(\sum_{i=1}^n (y_i - \hat{y}_i)^2) / (n - 1)}$	Root mean square error
R^2	$R^2 = 1 - \frac{MSE(\hat{y}, y)}{Var(y)}$	Determination correlation coefficient
$P_{MAE}(\%)$	$P_{MAE} = (MAE_1 - MAE_2) / MAE_1 $	Boosting percentage of MAE
P_{RMSE}	$P_{RMSE} = (RMSE_1 - RMSE_2) / RMSE_1 $	Boosting percentage of RMSE
P_{R^2}	$P_{R^2} = (R^2_1 - R^2_2) / R^2_1 $	Boosting percentage of P_{R^2}

Figure 8. Cont.

Figure 8. Decomposition result of VMD measuring point deformation sequence ((a) EX5, (b) EX2).

2.7.4. Model Realization

After decomposition, the high frequency components and ER will be predicted by LSTM. The realization of the LSTM model needs to consider the hidden layer, the number of neurons in each layer, the size of the training batch, the activation function and the optimizer. Too many hidden layers and neurons will increase the computer load, increase the workload of the computer, and reduce work efficiency. On the contrary, too little can easily lead to underfitting of training data, which in turn affects the prediction accuracy. In order to prevent overfitting, a dropout layer is added for each hidden layer. In addition, during the training process, the Adam optimizer is currently considered to be the best performing optimizer. See Table 3 for detailed parameter information.

Using the decomposed low-frequency components as the input to the RF. Here, Bootstrap, max_depth and n_estimators are mainly considered. Generally speaking, if n_estimators is too small, it is easy to underfit, and if it is too large, it will be time-consuming. Therefore, it is necessary to select a moderate value for parameter tuning. Meanwhile, due to the large amount of data, max_depth needs to be set. The above pa-

rameters are determined by grid search and 5-fold cross-validation. The traversal range is determined based on historical experience and manual experiments. Relevant parameters are shown in Table 4.

Table 3. LSTM parameters.

Measurement Point	Hidden Layer	Units and Dropout	Parameters	Setting
EX5	LSTM Layer1	Units = 65, Dropout = 0.2	Batch_size	48
	LSTM Layer2	Units = 90, Dropout = 0.25	Optimizer	Adam
	LSTM Layer3	Units = 80, Dropout = 0.4	Activation Function	Relu

Table 4. RF parameters.

	Bootstrap	Max_Depth	n_Estimators
Domain	[True, False]	[1, 20]	[2, 30]
EX5	True	10	18

3. Results and Discussion

In this section, the sub-sequences of different frequencies obtained after the decomposition of the two measuring points are input into LSTM and RF, respectively, and obtain the corresponding predicted values. To evaluate the applicability of the model, the prediction effects of LSTM and RF on high and low frequency data are discussed through two monitoring points EX5 and EX2, analyzing and explaining them based on the correlation coefficient. Finally, comparing the proposed model with the traditional prediction model, it is found that the model has excellent practicability for dam deformation prediction.

As shown in Figures 9a and 10a, LSTM has great advantages in capturing the inflection point of high-frequency data, but occasionally there may be situations where it cannot be captured. After our analysis, the reason may be the data's volatility is too strong. Figures 9b and 10b show that the correlation between the predicted value and the true value at the two measuring points reached 95.3% and 95.0%, respectively. As can be seen, most of the prediction points are within the 95% confidence interval, which shows that LSTM has a strong ability to mine high-frequency data. Figures 9c and 10c indicate that RF also has a good effect in realizing the prediction of stable data, which is not only reflected in the trend of the data, but also in the time lag. It can be seen from Figures 9d and 10d, although the residual sequence is an uncertain component, the predicted correlation coefficient values at the two measurement points also reached 99.4% and 98.1%, respectively, which proved that the input of low frequency components into RF is effective. Therefore, it is considered that RF has a positive significance for the prediction of low-frequency components.

Then, the high-frequency, low-frequency, and residual prediction results are reconstructed, and finally the final prediction value of dam deformation is obtained. This section selects a single LSTM, RF and Extreme Learning Machine (ELM) as the comparison model. Compared with the traditional single hidden layer neural network, ELM has the advantages of fewer parameters, fast training speed and strong generalization ability [16]. The comparison result is shown in Figure 11. From Figure 11a, it can be seen that the predicted trends of each model are roughly the same as the actual measured trends, but the predicted results of the proposed model are closer to the real results, and at the same time, it can better capture the mutation points of the dam deformation. From Figure 11b, it can be found that the absolute error of the model is the smallest, which illustrates the accuracy of the model's prediction and its reliability in the prediction of concrete dam deformation. In order to detect the outliers between the observed value and the predicted value, this study uses box plots to further illustrate. The box plots do not rely on any assumptions and are very robust to outliers. Figure 11c shows the box line drawing results of the proposed model and other comparison models on measuring point EX5. It can be found that the residuals of the model are distributed within 1.5 quartiles (IQR), and there are almost no extreme

outliers, with only a few mild outliers. In order to illustrate the prediction effect of each model more intuitively, six evaluation indicators are selected, the corresponding values are shown in Table 5, among which proposed model is selected as the comparison model.

Figure 9. Measurement point EX5 high and low frequency prediction and correlation coefficient. ((a) High frequency prediction; (b) Linear regression analysis of LSTM; (c) Low frequency prediction; (d) Linear regression analysis of RF).

Table 5. Comparison of evaluation indexes of measuring point EX5.

	MAE	RMSE	R^2	P_{MAE} (%)	P_{RMSE} (%)	P_{R^2} (%)
Proposed Model	0.174	0.214	0.993	/	/	/
LSTM	0.402	0.512	0.972	56.7	58.2	2.2
RF	0.355	0.441	0.969	51.0	51.5	2.5
ELM	0.641	0.764	0.906	72.9	72.0	9.6

Table 5 shows that the model proposed in this study is the lowest in all comparison models, whether it is MAE or RMSE, and its MAE and RMSE values are 0.174 and 0.214, respectively. The correlation coefficient between the real value and the predicted value is the closest to 1, but due to the denseness of the data, the correlation coefficient is increased by 2.2% at the minimum and 9.6% at the maximum. It can be seen from the table that

the prediction effect of a single LSTM model and RF model on the overall data is not as good as the proposed model. The MAE and RMSE of a single LSTM are 0.402 and 0.512, respectively. The MAE and RMSE values of a single RF are 0.355 and 0.441, respectively. Based on LSTM, the proposed model has increased the corresponding MAE and RMSE values by 56.7% and 58.2%, respectively. It also increased by 51.0% and 51.5%, respectively, on the basis of RF. Compared with ELM, the improved performance is more significant, reaching 72.9% and 72.0%. In order to further verify the validity of the model, the model is then applied to the deformation data of the EX2 measuring point, and it was also compared with the other three models. The prediction results are shown in Figure 12 and Table 6.

Figure 10. Measurement point EX2 high and low frequency prediction and correlation coefficient. ((a) High frequency prediction; (b) Linear regression analysis of LSTM; (c) Low frequency prediction; (d) Linear regression analysis of RF).

Through the prediction results of each model at the EX2 measurement point, it can be found that the proposed model is still the closest to the true value, as shown in Figure 12a. Figure 12b shows that the residual is the smallest and most stable. Figure 12c exhibits that there are almost no extreme outliers between the observed value and the predicted value, which shows that the proposed model improves the performance of a single LSTM and RF to a certain extent. According to Table 6, it can be seen that the degree of improvement of each index exceeds 70%, and the correlation coefficient has increased the most. This is because the EX2 measurement point is located on the left bank of the dam, and the data is

discrete and volatile, and the decomposed data become orderly, which makes the proposed model easy to predict; however, the data that are without decomposition are difficult to mine for other models, and the law is not easy to capture, resulting in a small correlation coefficient. In summary, the accuracy and practicability of the proposed model for dam deformation prediction can be further explained.

Figure 11. Comparison of total prediction results of measuring point EX5. ((a) prediction results; (b) prediction residuals; (c) boxplot of residuals).

Table 6. Comparison of evaluation indexes of measuring point EX2.

	MAE	RMSE	R ²	P _{MAE} (%)	P _{RMSE} (%)	P _{R²} (%)
Proposed Model	0.041	0.747	0.953	/	/	/
LSTM	0.163	2.978	0.251	74.8	74.9	279.7
RF	0.149	2.52	0.464	72.5	70.4	105.4
ELM	0.388	6.60	0.125	89.4	88.7	662.4

Figure 12. Comparison of total prediction results of measuring point EX2. ((a) prediction results; (b) prediction residuals; (c) boxplot of residuals).

The prediction results of each measurement point show that the model proposed in this paper solves the problem of inaccurate prediction caused by the strong volatility of dam deformation data to a certain extent. At the same time, the proposed model is significantly better than other advanced models. The proposal of this model provides theoretical knowledge for the analysis of the deformation prediction of important dam types such as RCC dams, and also provides prior knowledge for the construction of the dam safety monitoring system.

4. Conclusions

Based on the strong volatility and non-linear deformation monitoring data of the dam, combined with the data decomposition technology of VMD and the LSTM-RF multi-scale prediction model, this research develops a hybrid framework called VMD-LSTM-RF. This article innovatively combines signal decomposition technology with deep learning, and uses the advantages of LSTM and RF to capture and reconstruct different frequencies after decomposition, which perfectly realizes the accurate prediction of dam deformation. Through the analysis and comparison of two measuring points, the proposed model can more accurately predict the deformation of the dam, and has a more obvious improvement

effect on the basis of the existing model. The MAE improvement of the measuring point EX5 can be increased by 51.0% at least. The highest can reach 72.9%; the measuring point EX2 is even as high as 662.4%, which also shows that the proposed model can better achieve the expected goal and the following conclusions are drawn:

- (1) The components obtained by VMD decomposition of the original sequence of dam deformation can effectively solve the impact of strong volatility deformation data on the prediction results. In the decomposition process, the instantaneous frequency average method can be used to determine the best decomposition modulus.
- (2) Building a multi-dimensional and multi-scale model for high-frequency components and low-frequency components; establishing a prediction model and superimpose the prediction results of each component, which can effectively excavate the change law of dam deformation, accurately capture the change trend of deformation, and greatly improve the prediction performance.
- (3) Case analysis shows that the model has good robustness and excellent prediction accuracy in different areas of the dam, which plays an important role in the long-term safe operation and management of RCC dams.

Author Contributions: Data curation, E.C. and R.Y.; Formal analysis, R.Y. and C.S.; Methodology, E.C. and R.Y.; Writing—original draft, R.Y. and S.H.; Writing—review and editing, C.S., R.Y.; Supervision, H.Z. and S.H. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the National Natural Science Foundation of China (grant no. 51579089).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Shao, C.F.; Gu, C.S.; Yang, M.; Xu, Y.X.; Su, H.Z. A novel model of dam displacement based on panel data. *Struct. Control Health Monit.* **2018**, *25*, e2037. [\[CrossRef\]](#)
2. Su, H.Z.; Li, X.; Yang, B.B.; Wen, Z.P. Wavelet support vector machine-based prediction model of dam deformation. *Mech. Syst. Signal Process.* **2018**, *110*, 412–427. [\[CrossRef\]](#)
3. Salazar, F.; Morán, R.; Toledo, M.A.; Oñate, E. Data-based models for the prediction of dam behavior: A review and some methodological considerations. *Arch. Comput. Methods Eng.* **2017**, *24*, 1–21. [\[CrossRef\]](#)
4. Shi, Y.Q.; Yang, J.J.; Wu, J.L.; He, J.P. A statistical model of deformation during the construction of a concrete face rock-fill dam. *Struct. Control Health Monit.* **2018**, *25*, 2. [\[CrossRef\]](#)
5. Li, Y.; Bao, T.; Shu, X.; Chen, Z.; Gao, Z.; Zhang, K. A Hybrid Model Integrating Principal Component Analysis, Fuzzy C-Means, and Gaussian Process Regression for Dam Deformation Prediction. *Arab. J. Sci. Eng.* **2020**, *46*, 4293–4306. [\[CrossRef\]](#)
6. Li, Y.; Zhang, H.; Wen, L.; Shi, N. A Prediction Model for Deformation Behavior of Concrete Face Rockfill Dams Based on the Threshold Regression Method. *Arab. J. Sci. Eng.* **2021**, *46*, 5801–5816. [\[CrossRef\]](#)
7. Chen, B.; Hu, T.; Huang, Z.; Fang, C. A Spatio-Temporal Clustering and Diagnosis Method of Deformation Monitoring Data in Concrete Arch Dam. *Struct. Health Monit.* **2019**, *18*, 1355–1371. [\[CrossRef\]](#)
8. Chen, B.; Gu, C.; Bao, T.; Wu, B.; Su, H. Failure analysis method of concrete arch dam based on elastic strain energy criterion. *Eng. Fail. Anal.* **2016**, *60*, 363–373. [\[CrossRef\]](#)
9. Cao, E.H.; Bao, T.F.; Gu, C.S.; Li, H.; Liu, Y.T.; Hu, S.P. A Novel Hybrid Decomposition-Ensemble Prediction Model for Dam Deformation. *Appl. Sci.* **2020**, *10*, 5700. [\[CrossRef\]](#)
10. Ren, Q.B.; Li, M.C.; Song, L.G.; Liu, H. An optimized combination prediction model for concrete dam deformation considering quantitative evaluation and hysteresis correction. *Adv. Eng. Inform.* **2020**, *46*, 101154. [\[CrossRef\]](#)
11. Qu, X.D.; Yang, J.; Chang, M. A Deep Learning Model for Concrete Dam Deformation Prediction Based on RS-LSTM. *J. Sens.* **2019**, *2019*, 4581672. [\[CrossRef\]](#)
12. Wu, Z.R. *Safety Monitoring Theory and Its Application of Hydraulic Structures*; Higher Education: Beijing, China, 2003.
13. Wu, Z.R.; Gu, Y.C.; Gu, C.S.; Guo, H.Q.; Su, H. Establishing time-dependent model of deformation modulus caused by bedrock excavation rebound by inverse analysis method. *Sci. China Ser. E Technol. Sci.* **2008**, *51*, 1–7. [\[CrossRef\]](#)

14. Dimitriadis, P.; Koutsoyiannis, D.; Iliopoulou, T.; Papanicolaou, P. A global-scale investigation of stochastic similarities in marginal distribution and dependence structure of key hydrological-cycle processes. *Hydrology* **2021**, *8*, 59. [\[CrossRef\]](#)
15. Cheng, L.; Liu, Y.R.; Yang, Q.; Pan, Y.W.; Lv, Z. Mechanism and numerical simulation of reservoir slope deformation during impounding of high arch dams based on nonlinear FEM. *Comput. Geotech.* **2017**, *81*, 143–154. [\[CrossRef\]](#)
16. Liu, H.F.; Ren, C.; Zheng, Z.T.; Liang, Y.J.; Lu, X.J. Study of a Gray Genetic BP Neural Network Model in Fault Monitoring and a Diagnosis System for Dam Safety. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 4. [\[CrossRef\]](#)
17. Kang, F.; Liu, J.; Li, J.; Li, S. Concrete dam deformation prediction model for health monitoring based on extreme learning machine. *Struct. Control Health Monit.* **2017**, *24*, e1997. [\[CrossRef\]](#)
18. Zhang, H.; Shi, X.; Lai, F.L. Research on Time Series Analysis Based Deformation Prediction Model. In *Advanced Materials Research*; Trans Tech Publications Ltd.: Stafa-Zurich, Switzerland, 2011; Volume 1270, pp. 2888–2891.
19. Dimitriadis, P.; Koutsoyiannis, D. Stochastic synthesis approximating any process dependence and distribution. *Stoch. Environ. Res. Risk Assess.* **2018**, *32*, 1493–1515. [\[CrossRef\]](#)
20. Graves, A.; Mohamed, A.; Hinton, G. Speech recognition with deep recurrent neural networks. In Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 26–31 May 2013; pp. 6645–6649. [\[CrossRef\]](#)
21. Shen, M.L.; Lee, C.F.; Liu, H.H.; Chang, P.Y.; Yang, C.H. Effective multinational trade forecasting using LSTM recurrent neural network. *Expert Syst. Appl.* **2021**, *182*, 115199. [\[CrossRef\]](#)
22. Zhang, T.; Zheng, X.Q.; Liu, M.X. Multiscale attention-based LSTM for ship motion prediction. *Ocean Eng.* **2021**, *230*, 109066. [\[CrossRef\]](#)
23. Su, Y.; Weng, K.; Lin, C.; Chen, Z. Dam Deformation Interpretation and Prediction Based on a Long Short-Term Memory Model Coupled with an Attention Mechanism. *Appl. Sci.* **2021**, *11*, 6625. [\[CrossRef\]](#)
24. Wu, Y.X.; Gu, Y.C. Metabolic grey early warning model for dam deformation based on wavelet denoising. In *MATEC Web of Conferences*; EDP Sciences: Les Ulis, France, 2018; Volume 246. [\[CrossRef\]](#)
25. Liu, B.C.; Yu, X.G.; Wang, Q.S. Air pollution concentration forecasting based on wavelet transform and combined weighting forecasting model. *Atmos. Pollut. Res.* **2021**, 101144. [\[CrossRef\]](#)
26. Wang, J.; Tang, L.Y.; Luo, Y.Y.; Ge, P. A weighted EMD-based prediction model based on TOPSIS and feed forward neural network for noised time series. *Knowl.-Based Syst.* **2017**, *132*, 167–178. [\[CrossRef\]](#)
27. Duan, W.Y.; Han, Y.; Huang, L.M.; Zhao, B.B.; Wang, M.H. A hybrid EMD-SVR model for the short-term prediction of significant wave height. *Ocean Eng.* **2016**, *124*, 54–73. [\[CrossRef\]](#)
28. Dragomiretskiy, K.; Zosso, D. Variational mode decomposition. *IEEE Trans. Signal Process.* **2014**, *62*, 531–544. [\[CrossRef\]](#)
29. Feng, Z.K.; Niu, W.J.; Tang, Z.Y.; Jiang, Z.P.; Xu, Y.; Liu, Y.; Zhang, H.R. Monthly runoff time series prediction by variational mode decomposition and support vector machine based on quantum-behaved particle swarm optimization. *J. Hydrol.* **2020**, *583*, 124627. [\[CrossRef\]](#)
30. Niu, W.J.; Feng, Z.K.; Chen, Y.B.; Zhang, H.R.; Cheng, C.T. Annual streamflow time series prediction using extreme learning machine based on gravitational search algorithm and variational mode decomposition. *Hydrol. Eng.* **2020**, *25*, 04020008. [\[CrossRef\]](#)
31. Seo, Y.; Kim, S.; Singh, V.P. Machine Learning Models Coupled with Variational Mode Decomposition: A New Approach for Modeling Daily Rainfall-Runoff. *Atmosphere* **2018**, *9*, 251. [\[CrossRef\]](#)
32. Xie, T.; Zhang, G.; Hou, J.; Xie, J.; Lv, M.; Liu, F. Hybrid forecasting model for non-stationary daily runoff series: A case study in the Han River Basin, China. *J. Hydrol.* **2019**, *577*, 123915. [\[CrossRef\]](#)
33. Zhang, Y.G.; Pan, G.F.; Chen, B.; Han, J.Y.; Zhao, Y.; Zhang, C.H. Short-term wind speed prediction model based on GA-ANN improved by VMD. *Renew. Energy* **2020**, *156*, 1373–1388. [\[CrossRef\]](#)
34. Zan, T.; Pang, Z.L.; Wang, M.; Gao, X.S. Research on Early Fault Diagnosis of Rolling Bearing Based on VMD. In Proceedings of the 2018 6th International Conference on Mechanical, Automotive and Materials Engineering (CMAME), Hong Kong, China, 10–12 August 2018.
35. Niu, M.F.; Hu, Y.Y.; Sun, S.L.; Liu, Y. A novel hybrid decomposition-ensemble model based on VMD and HGWO for container throughput forecasting. *Appl. Math. Model.* **2018**, *57*, 163–178. [\[CrossRef\]](#)
36. Li, Q.Q.; Li, M.; Guo, L.; Zhang, Z. Random Forests Algorithm Based Duplicate Detection in On-Site Programming Big Data Environment. *J. Inf. Hiding Priv. Prot.* **2020**, *2*, 199–205. [\[CrossRef\]](#)