



**ARTICLE**

## Comparative Study on Deformation Prediction Models of Wuqiangxi Concrete Gravity Dam Based on Monitoring Data

Songlin Yang<sup>1,2</sup>, Xingjin Han<sup>1,2</sup>, Chufeng Kuang<sup>1,2</sup>, Weihua Fang<sup>3</sup>, Jianfei Zhang<sup>4</sup> and Tiantang Yu<sup>4,\*</sup>

<sup>1</sup>Hunan Wuling Power Technology Corporation Ltd., Changsha, 410004, China

<sup>2</sup>Wuling Power Corporation Ltd., Changsha, 410004, China

<sup>3</sup>Nanjing Research Institute of Hydrology and Water Conservation Automation, Ministry of Water Resources, Nanjing, 210012, China

<sup>4</sup>Department of Engineering Mechanics, Hohai University, Nanjing, 211100, China

\*Corresponding Author: Tiantang Yu. Email: tiantangyu@hhu.edu.cn

Received: 28 July 2021 Accepted: 14 October 2021

### ABSTRACT

The deformation prediction models of Wuqiangxi concrete gravity dam are developed, including two statistical models and a deep learning model. In the statistical models, the reliable monitoring data are firstly determined with Lahitte criterion; then, the stepwise regression and partial least squares regression models for deformation prediction of concrete gravity dam are constructed in terms of the reliable monitoring data, and the factors of water pressure, temperature and time effect are considered in the models; finally, according to the monitoring data from 2006 to 2020 of five typical measuring points including J<sub>23</sub> (on dam section 24<sup>#</sup>), J<sub>33</sub> (on dam section 4<sup>#</sup>), J<sub>35</sub> (on dam section 8<sup>#</sup>), J<sub>37</sub> (on dam section 12<sup>#</sup>), and J<sub>39</sub> (on dam section 15<sup>#</sup>) located on the crest of Wuqiangxi concrete gravity dam, the settlement curves of the measuring points are obtained with the stepwise regression and partial least squares regression models. A deep learning model is developed based on long short-term memory (LSTM) recurrent neural network. In the LSTM model, two LSTM layers are used, the rectified linear unit function is adopted as the activation function, the input sequence length is 20, and the random search is adopted. The monitoring data for the five typical measuring points from 2006 to 2017 are selected as the training set, and the monitoring data from 2018 to 2020 are taken as the test set. From the results of case study, we can find that (1) the good fitting results can be obtained with the two statistical models; (2) the partial least squares regression algorithm can solve the model with high correlation factors and reasonably explain the factors; (3) the prediction accuracy of the LSTM model increases with increasing the amount of training data. In the deformation prediction of concrete gravity dam, the LSTM model is suggested when there are sufficient training data, while the partial least squares regression method is suggested when the training data are insufficient.

### KEYWORDS

Wuqiangxi concrete gravity dam; deformation prediction; stepwise regression model; partial least squares regression model; LSTM model



## 1 Introduction

During the service period, dams not only bear various cyclic loads and sudden disasters, but also suffer from erosion and corrosion from harsh environment, and it leads to the gradual decline of local and overall safety performance. Once the dam is wrecked, it will yield disastrous consequences. Therefore, it is very important to identify the potential risk and evaluate dam safety behavior in time based on the dam monitoring data collected by prototype observation instruments.

The traditional analysis models for dam deformation monitoring data include statistical model [1,2], deterministic model [3] and hybrid model [4–6] and combination model [7,8]. The idea of statistical model is to establish the relationship between environmental variables (such as, water level, temperature, aging, etc.) and the effect variables (such as, deformation, seepage, cracking, etc.) by modeling and analyzing the monitoring data based on probability theory and mathematical statistics theory. The statistical model can simulate structural changes and predict its development trend, its implementations are simple, and its accuracy can meet engineering requirements, so the statistical model has been widely used in engineering. In the deterministic model, the displacement, stress, seepage fields, etc. of the dam and foundation under load are calculated with numerical methods such as finite element method (FEM) [9], extended finite element method (XFEM) [6,10], meshfree methods [11], etc.; and then some parameters are adjusted by optimizing the fitting between the calculated values and the measured values; thus the expression based on the essence of physics and mechanics can be established. The shortcomings of the deterministic model are uncertainty of dam and foundation material parameters, inaccuracy in setting boundary conditions and inaccuracy of model simplification. Hybrid model uses numerical methods to calculate water pressure component, and other components are obtained with the statistical model, and then the obtained values are optimally fitted with the measured values. Although the hybrid model improves the calculation accuracy from the mechanics concept, it also inherits the inherent defects and assumptions of statistical model and deterministic model. Combined model performs nonlinear optimization combination on multiple single models by integrating various useful information to achieve a more reasonable and comprehensive description of mapping relationship, and it can effectively improve fitting and prediction accuracy. The shortcomings of the combination model include: (1) the linear combination model may get unrealistic negative weights for dealing with nonlinear problems; (2) it is very difficult to construct combinatorial functions.

Li et al. [12] reviewed the dam monitoring data analysis methods including the monitoring models, monitoring indices, and anomaly value detection methods. Tonini [13] first applied a cubic polynomial to describe the water pressure and temperature components using statistical analysis. Wu [14] investigated the selection of environmental variables and the corresponding mechanisms by combining dam theory, statistical theory and engineering mechanics, and deduced mathematical models of the water pressure and temperature components. Chen [15] proposed exponential function, hyperbolic function, and logarithmic function models for time-dependent component of a concrete dam based on creep theory. Léger et al. [16] presented frequency domain solution algorithms of the one-dimensional transient heat transfer equation for describing temperature variations in arch dam cross sections, thus the temperature variations are not required to be specified at the upstream and downstream faces. Tatin et al. [17] developed a hybrid physico-statistical model to improve the assessment of displacements due to the temperature field in a dam, and cases show that the improved assessment of thermal effects on reversible phenomena leads to a reduced uncertainty on residuals. Mata et al. [18] presented a methodology for the selection of the thermometers that best represent the thermal effect in the statistical model.

Hu et al. [1] proposed the special statistical models for the displacements of high arch dams during their initial impoundment periods by improving estimations of the non-stationary thermal and the non-monotonic time-dependent effects. Wang et al. [19] presented a shape feature-based spatial clustering method for the dam temperature field, and established a displacement monitoring model of concrete dams using the shape feature clustering-based temperature principal component factor.

Stepwise regression and partial least squares regression are widely used in statistical models. In stepwise regression model, environmental variables are added to the model one by one, and the significance of environmental variables to the model is sequentially assessed to obtain the optimal variable set. Hu et al. [20] established the stepwise regression model of Bikou earth-rockfill dam deformation displacement based on the least squares method of statistics principles, and the results show that the model has higher fitting precision and longer predict cycle. Shen et al. [21] studied the influences of factors on gravity dam open crack with multiple stepwise regression method. The partial least squares method extracts the principal components from the set of independent variables, and the extracted principal components are linearly independent. However, when the extracted components contain significant amounts of information that is unrelated to the dependent variables, the results of the partial least squares method are not satisfactory [22]. Cheng et al. [23] proposed an improved observation method and improved partial least squares data analysis methods to overcome the shortcomings of the traditional methods of external deformation monitoring and data analysis of high rock-fill dams. Yin et al. [24] proposed a novel nonlinear component separation method for the effect quantities by combining kernel partial least squares and pseudosamples, and the separated displacement components of a super-high arch dam conform to the general deformation law. Considering the characteristics of complex nonlinear and multiple response variables of a super-high dam, a universal unified optimization algorithm was developed to select the kernel partial least squares parameters and achieve the optimal kernel partial least squares [25]. With the development of statistical regression, some regression models have been gradually introduced into dam safety monitoring, such as threshold regression [26], logistic regression [27] and random forest regression [28,29], etc. The multicollinearity among environmental variables and the effective optimization of the model are two key issues which affect the quality of a statistical model.

In recent years, with the development of dam safety monitoring, computer, big data, artificial intelligence and other theories and technologies, more and more data mining methods have been applied to the dam safety monitoring modeling, and many intelligent algorithm monitoring models have emerged, and these models show unique advantages in solving the problems of uncertainty and nonlinearity of monitoring model factors, prediction accuracy and generalization. Qu et al. [30] established single-point and multipoint concrete dam deformation prediction models based on long short-term memory (LSTM) network, and proposed a new evaluation system and quantitative evaluation indexes. Yang et al. [31] proposed a concrete dam deformation prediction method based on LSTM with attention mechanism, which can effectively avoid the gradient disappearance and gradient explosion problems in the recurrent neural network. Liu et al. [32] predicted long-term displacements of arch dams by combining the LSTM network and dimension reduction methods, and the results reveal that the coupling prediction models have higher accuracy and can capture the long-term characteristics of the arch dam deformation. Han et al. [33] predicted the horizontal displacement of concrete-face rockfill dams using the statistically optimized back-propagation neural network model, which can overcome the shortcomings of the statistical model and back-propagation neural network model. Chen et al. [34] proposed a novel deformation prediction model of arch dam via correlated multi-target stacking.

Shi et al. [35] developed a safety monitoring model for concrete face rockfill dam seepage with cracks considering the lagging effect using the radial basis function neural network. Yang et al. [36] presented an intelligent singular value diagnostic method based on convolutional neural network for concrete dam deformation monitoring. Liu et al. [37] investigated the applicability of the kernel-extreme learning machines-based model considering the thermal effect on the behavior prediction of concrete arch dams. Chen et al. [38] builded the structural health monitoring framework of concrete dam displacement and mined the effects of hydrostatic, seasonal and irreversible time components on dam deformation by combining with relevance vector machine, multi-kernel technique, hydrostatic-season-time statistical model and parallel Jaya algorithm. Shu et al. [39] proposed a novel prediction model based on variational autoencoder and temporal attention-based long short-term memory network for the long-term deformation of arch dams. Chen et al. [40] developed an integrated displacement prediction method based on the spatiotemporal clustering and machine learning algorithms.

This study aims to establish the deformation prediction models of Wuqiangxi concrete gravity dam based on the monitoring data. According to the monitoring data from 2006 to 2020 of measuring points J<sub>23</sub> (on dam section 24<sup>#</sup>), J<sub>33</sub> (on dam section 4<sup>#</sup>), J<sub>35</sub> (on dam section 8<sup>#</sup>), J<sub>37</sub> (on dam section 12<sup>#</sup>), and J<sub>39</sub> (on dam section 15<sup>#</sup>) located on the crest of Wuqiangxi concrete gravity dam, the stepwise regression model, the partial least squares regression model and the LSTM model of the settlement curve of the measuring points are developed. The three models are analyzed and compared, and the results show that the LSTM model can obtain excellent predictions when there are enough training data, while the partial least squares regression method is suggested for the insufficient training data.

After the introduction, the statistical models and LSTM model for deformation prediction of concrete gravity dam are described in Sections 2 and 3, respectively. Case studies and discussion are given in Section 4. Some conclusions are obtained in Section 5.

## 2 Statistical Models for Deformation Prediction of Concrete Gravity Dam

### 2.1 Data Preprocessing

The monitoring data are independent, and are easily affected by environmental factors in the process of observation, resulting in some monitoring data that do not conform to the regular changes, which are unreliable data. In order to improve the prediction performance of statistical model, the unreliable data should be removed.

Generally, the reliability of data is judged by Lahitte criterion (also called  $3\sigma$  criterion). The Lahitte criterion is to judge whether the data is reliable according to the jump characteristics and variation rules of the measured values. There are  $n$  monitoring data, i.e.,  $y_1, y_2, \dots, y_n$ , the pulsation characteristics  $d_i$  of the  $i^{\text{th}}$  monitoring data can be expressed as

$$d_i = 2y_i - (y_{i+1} + y_{i-1}) \quad (1)$$

where  $y_i$  is the  $i^{\text{th}}$  monitoring data.

Assume that  $n$  monitoring data are known,  $d_2, d_3, \dots, d_{n-1}$  can be obtained with Eq. (1). The mean value  $\bar{d}$  and mean square deviation  $\sigma$  of  $d_i$  ( $i = 2, 3, \dots, n-1$ ) are

$$\bar{d} = \frac{\sum_{i=2}^{n-1} d_i}{n-2} \quad (2a)$$

$$\sigma = \sqrt{\frac{\sum_{i=2}^{n-1} (d_i - \bar{d})^2}{n-3}} \quad (2b)$$

The ratio of absolute value of runout deviation and mean square deviation of the  $i^{th}$  monitoring data is defined as

$$q_i = \frac{|d_i - \bar{d}|}{\sigma} \quad (3)$$

As  $q_i > 3$ , it indicates that the monitoring data is abnormal or unreliable, and it should be removed.

## 2.2 Statistical Models

The displacement vector at one point in dam can be divided into horizontal displacement  $\delta_x$ , lateral horizontal displacement  $\delta_y$  and vertical displacement  $\delta_z$ . According to the cause of formation, displacement is composed with hydraulic, temperature and aging components, i.e.,

$$\delta(\delta_x \text{ or } \delta_y \text{ or } \delta_z) = \delta_H + \delta_T + \delta_\theta \quad (4)$$

where  $\delta_H$ ,  $\delta_T$  and  $\delta_\theta$  denote hydraulic, temperature and aging components, respectively.

The displacement component at one point in the concrete gravity dam under water pressure and reservoir water weight can be described as [14]

$$\delta_H = \sum_{i=1}^3 (a_{1i}H^i + a_{2i}h^i) \quad (5)$$

where  $a_{1i}$  and  $a_{2i}$  are the regression coefficients of upstream and downstream water pressure factors, respectively.  $H$  is the upstream depth, and  $h$  is the downstream water depth.

After some years of dam operation, the temperature inside the dam reaches the quasi-stable temperature field, so it can be assumed that the internal temperature of the dam body is only affected by the water temperature and air temperature, the water temperature and air temperature change harmoniously, and the deformation is linearly related to the temperature of concrete. The temperature component is expressed as [14]

$$\delta_T = b_1 \sin \frac{2\pi t}{365} + b_2 \cos \frac{2\pi t}{365} \quad (6)$$

where  $b_1$  and  $b_2$  are the regression coefficients of temperature factor,  $t$  is the cumulative number of days from the corresponding monitoring day to the initial monitoring day.

At the beginning of impoundment, the aging displacement generally changes violently, and then tends to be stable. The aging component can be expressed as [14]

$$\delta_\theta = c_1\theta + c_2 \ln\theta \quad (7)$$

where  $c_1$  and  $c_2$  are the regression coefficients of aging factor, and  $\theta = \frac{t}{100}$ .

Substituting Eqs. (5)–(7) into Eq. (4) yields

$$\delta(\delta_x \text{ or } \delta_y \text{ or } \delta_z) = \sum_{i=1}^3 (a_{1i}H^i + a_{2i}h^i) + b_1 \sin \frac{2\pi t}{365} + b_2 \cos \frac{2\pi t}{365} + c_1 \theta + c_2 \ln \theta \quad (8)$$

### 2.2.1 Stepwise Regression Method

The stepwise regression method [20,21] starts with an independent variable, then independent variable is introduced into the regression equation one by one based on the effect of the independent variable, and the independent variable which has large effect is first introduced into regression equation. On the other hand, when the first introduced independent variables become insignificant due to the introduction of the latter independent variables, they are removed from the regression equation at any time. Therefore, the so-called step-by-step is to introduce independent variables at some steps, and eliminate independent variables at some steps. At each step,  $F$  test is done to ensure that the regression equation only contains independent variables with significant effect before introducing new significant independent variables. When all the independent variables with significant effect are contained in the regression equation, the equation is the final regression equation.

There are  $k$  independent variables, and  $n$  groups of monitoring data series about the independent variables  $\mathbf{x}$  and dependent variables  $\mathbf{y}$ , i.e.,  $(x_{11}, x_{21}, \dots, x_{k1}; y_1), \dots, (x_{1n}, x_{2n}, \dots, x_{kn}; y_n)$ . The regression equation is expressed as [14]

$$\hat{y} = b_0 + \sum_{i=1}^k b_i x_i \quad (9)$$

with

$$\mathbf{b}_i = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix} = \mathbf{S}_{ij}^{-1} \mathbf{S}_{iy} \quad (10a)$$

$$b_0 = \bar{y} - \sum_{i=1}^k b_i \bar{x}_i \quad (10b)$$

where

$$\mathbf{S}_{ij} = \begin{bmatrix} S_{11} & S_{12} & \dots & S_{1k} \\ \vdots & \vdots & \vdots & \vdots \\ S_{k1} & S_{k2} & \dots & S_{kk} \end{bmatrix} \quad (11a)$$

$$\mathbf{S}_{iy} = \begin{bmatrix} S_{1y} \\ S_{2y} \\ \vdots \\ S_{ky} \end{bmatrix} \quad (11b)$$

$$\bar{x}_i = \sum_{t=1}^n \frac{x_{it}}{n}, \quad \bar{y} = \sum_{t=1}^n \frac{y_t}{n} \quad (11c)$$

$$S_{iy} = \sum_{t=1}^n x_{it}y_t - \frac{1}{n} \sum_{t=1}^n x_{it} \sum_{t=1}^n y_t \quad (11d)$$

$$S_{ij} = S_{ji} = \sum_{t=1}^n x_{it}x_{jt} - \frac{1}{n} \sum_{t=1}^n x_{it} \sum_{t=1}^n x_{jt} \quad (11e)$$

In order to solve regression coefficients  $\mathbf{b}_j$ , the normal equation of stepwise regression method is expressed as [14]

$$\mathbf{a}_{ij}\mathbf{b}_j = \mathbf{c}_{ij}\mathbf{S}_{iy} \quad (12)$$

where  $\mathbf{a}_{ij}$  is a matrix of independent variables with  $k \times k$  orders, the initial values of  $\mathbf{a}_{ij}$  are  $\mathbf{S}_{ij}$ , then  $\mathbf{c}_{ij}$  is an identity matrix of order  $k$ .

The process of solving Eq. (12) using stepwise regression method is to change  $\mathbf{a}_{ij}$  into  $\delta_{ij}$  and  $\mathbf{c}_{ij}$  into  $\mathbf{a}_{ij}^{-1}$  by elimination transformation of  $\mathbf{b}_j$  step by step.

In order to objectively evaluate the application of stepwise regression method to dam deformation prediction, the multiple correlation coefficient (MCC) and residual standard deviation (RSD) are used as the evaluation indexes of prediction accuracy. The MCC and RSD are defined as

$$MCC = \sqrt{1 - r_{nm}^m} \quad (13)$$

where  $r^m$  is the normalized correlation matrix after the  $m^{\text{th}}$  eliminating-introducing.

$$RSD = S_{nm} \sqrt{\frac{r_{nm}^m}{n - m - 1}} \quad (14)$$

where  $m$  is the number of eliminating-introducing.

The bigger the value of MCC, the better the prediction accuracy. The smaller the value of RSD, the better the prediction accuracy.

### 2.2.2 Partial Least Squares Regression Method

Partial least squares regression [23,25] is the regression of multiple dependent variables to multiple independent variables. There are  $q$  dependent variables, i.e.,  $\mathbf{X} = \{x_1, x_2, \dots, x_q\}$ ,  $p$  independent variables, i.e.,  $\mathbf{Y} = \{y_1, y_2, \dots, y_p\}$  and  $n$  sample points. Extract components  $t_1$  and  $u_1$  ( $t_1$  is linear combination of  $x_1, x_2, \dots, x_q$ ,  $u_1$  is linear combination of  $y_1, y_2, \dots, y_p$ ) from  $\mathbf{X}$  and  $\mathbf{Y}$ , respectively. In the process of extracting components, we need to meet the following requirements: (1)  $t_1$  and  $u_1$  contain as much variation information in  $\mathbf{X}$  and  $\mathbf{Y}$  as possible, respectively; (2) the correlation between  $t_1$  and  $u_1$  is the highest. After extracting the first component, perform regression of  $\mathbf{X}$  about  $t_1$ , and  $\mathbf{Y}$  about  $u_1$ . If the regression equation does not achieve satisfactory accuracy, the second component extraction is performed with the interpreted residual information  $\mathbf{X}$  by  $t_1$  and the interpreted residual information  $\mathbf{Y}$  by  $u_1$ , until the satisfactory accuracy



is achieved. If the  $m$  components  $t_1, t_2, \dots, t_m$  are extracted in  $X$ , the regression of  $y_k$  about  $t_1, t_2, \dots, t_m$  is performed, and the regression equation of  $y_k$  ( $k = 1, 2, \dots, p$ ) about the original variable  $x_1, x_2, \dots, x_q$  is established.

### 3 LSTM Model for Deformation Prediction of Concrete Gravity Dam

#### 3.1 LSTM Network

Long short-term memory (LSTM) network is a kind of back propagation recurrent neural network. In the LSTM [31], the concept of time series is introduced into the network structure to mine the association between long-term and short-term deformation monitoring data. By improving the simple node of traditional neural network into a storage unit, the problem of gradient disappearance and gradient explosion when learning the association of long-term and long-term data is avoided, and the prediction accuracy is improved. The LSTM can remember information, also known as a cell. The main excellent features of the LSTM are as follows: (1) it can learn the complex nonlinear relationship between effect sets and load sets through training; (2) it can selectively retain the previous deformation and the corresponding load when calculating the deformation under the current load; (3) it has the memorability and can learn the laws between the deformation and the previous load set.

#### 3.2 Selection of the Hyper-Parameters

In general, the hyper-parameters involved in the training of recurrent neural network are selected according to experience. However, the LSTM recurrent neural network is sensitive to the selection of hyper-parameters in training. In order to achieve better results, the optimization algorithm is adopted to optimize the hyper-parameters. The optimization of network hyper-parameters mainly focuses on three parameters: sample length, the number of hidden neurons in feedforward network layer (also known as state vector size), the learning rate which controls the adjustment range of network parameter. The grid search algorithm and random search algorithm are widely used to optimize the hyper-parameters of LSTM recurrent neural network.

Grid search determines the spatial dimension of grid search according to the number of hyper-parameters, divides the grid on each dimension, then determines the best hyper-parameters according to the results given by grid intersections. The search process of grid search is mainly divided into three steps: first, the values of less important hyper-parameters are fixed; second, the range of three hyper-parameters are set; finally, the target function LSTM recurrent neural network test set is set to the highest recognition accuracy. The existing studies show that the method based on grid search can obtain high precision result, but the computing cost increases exponentially with the increase of the number of hyper-parameters.

The search space of random search cannot be discrete, which allows random search to try more hyper-parameter combinations under the same computing resources. Grid search wastes a lot of computing resources in the hyper-parameters which have little impact on the network performance, while random search tests the unique value of each hyper-parameter which has an impact on the results almost every time, that is to say, random search tries more beneficial hyper-parameter combinations. Random search can greatly shorten the search time and improve the computational efficiency on the premise of ensuring a certain accuracy. After selecting the optimized super parameters, the recognition accuracy of LSTM recurrent neural network is greatly improved.

In this study, the LSTM recurrent neural network coupled with random search is used to construct the deformation prediction model of concrete gravity dam.



In order to objectively evaluate the application of LSTM model to dam deformation prediction, the root mean square error (RMSE) is used as the evaluation index of prediction accuracy. The RMSE is defined as

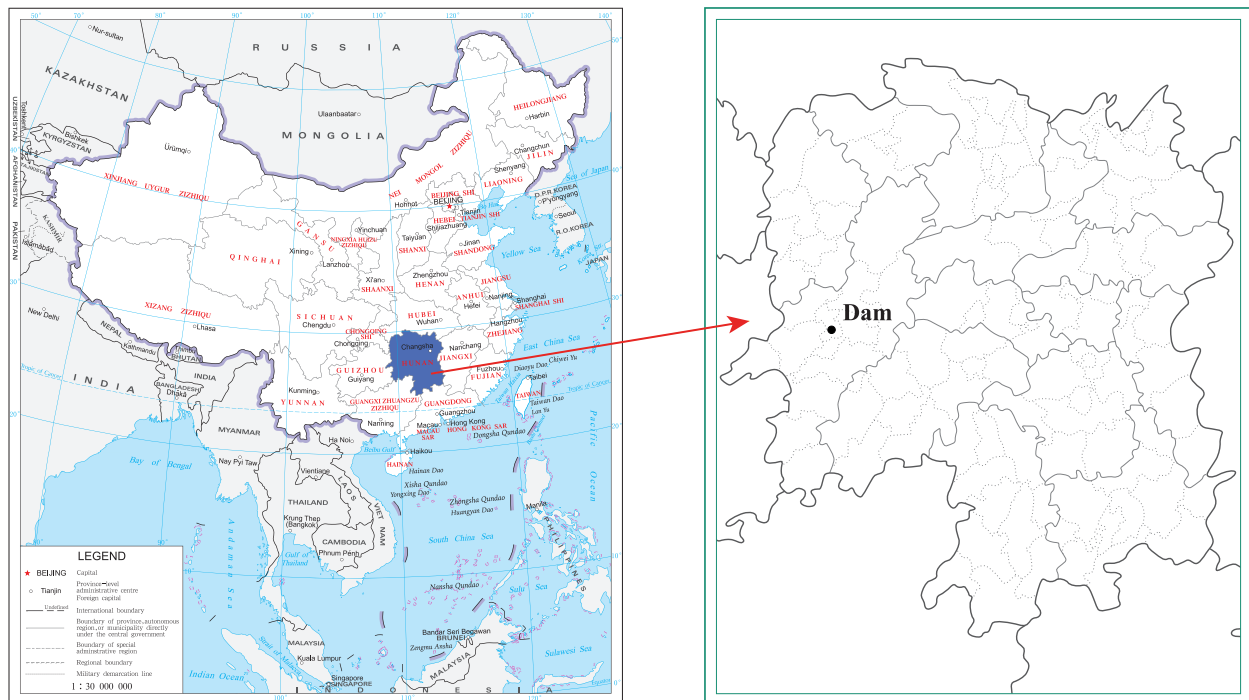
$$RMSE = \sqrt{\frac{\sum_{i=1}^n |y_i - \hat{y}_i|^2}{n}} \tag{15}$$

where  $n$  is the number of forecast data;  $y_i$  is the true value of the  $i^{th}$  data of the prediction group;  $\hat{y}_i$  is the corresponding predicted value. The smaller the value of RMSE, the better the prediction accuracy.

#### 4 Case Study and Discussion

##### 4.1 Project Overview

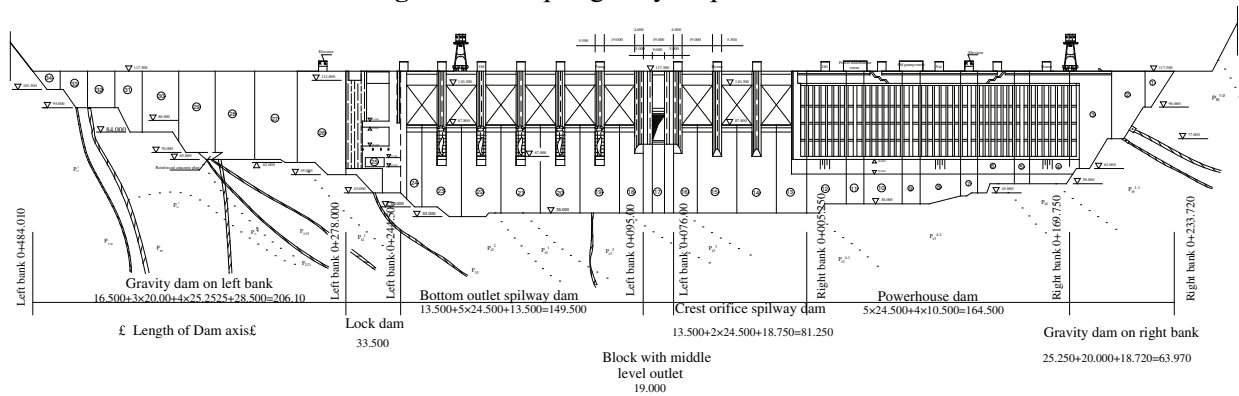
Wuqiangxi hydropower station which was completed in 1999 is located in the middle and upper reaches of the main stream of the Yuan River in Hunan province, China, as shown in Fig. 1. The project is mainly composed of three parts: the river blocking dam, the powerhouse behind the dam on the right bank and the three-stage ship lock on the left bank. Fig. 2 presents the schematic diagram of layout of Wuqiangxi hydropower station, and Figs. 3 and 4 are the upstream and downstream views.



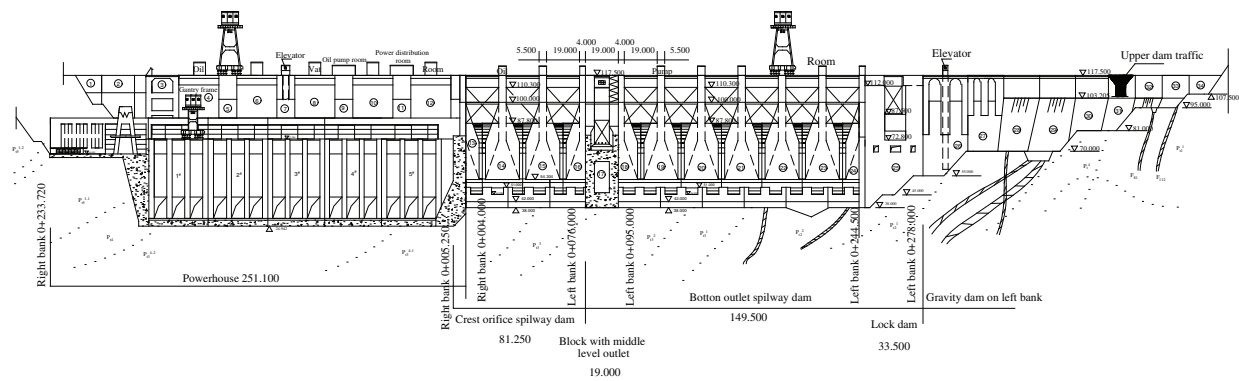
**Figure 1:** Location of Wuqiangxi hydropower station (the longitude and latitude of the dam are 110.6294 and 28.5783, respectively)



**Figure 2: Wuqiangxi hydropower station**



**Figure 3: View of the upstream of the junction (scale:1:5000)**



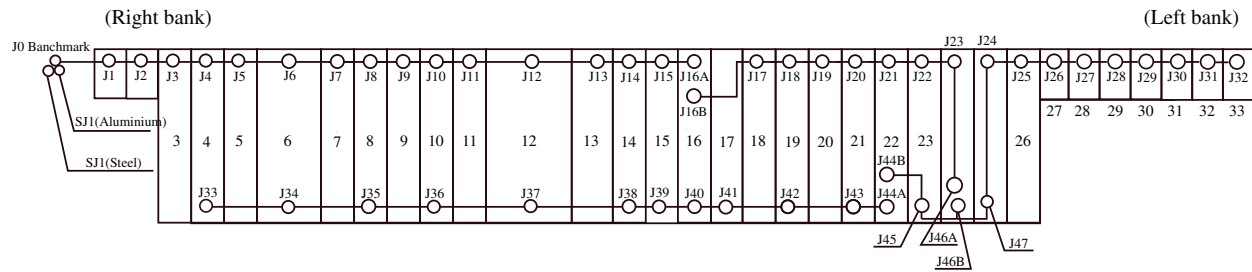
**Figure 4: View of the downstream of the junction (scale:1:5000)**

The dam is a concrete gravity dam. The crest elevation is 117.5 m, the highest dam height is 85.83 m, and the total crest length is 719.7 m. The main dam is divided into 34 dam sections, including the right bank retaining dam sections 1#~3#, 5#, 7#, 9#, and 11#; the diversion dam sections 4#, 6#, 8#, 10#, and 12#; the surface outlet spillway dam sections 13#~16#; the middle

outlet spillway dam section 17<sup>#</sup>; the overflow dam sections of the surface and bottom outlets 18<sup>#</sup>~24<sup>#</sup>; the lock dam section 25<sup>#</sup>; the left bank retaining dam sections 26<sup>#</sup>~34<sup>#</sup>.

There are 51 measurement points of the hydrostatic leveling instrument on the dam crest, numbered J<sub>0</sub>~J<sub>47</sub>, as shown in Fig. 5. Two measurement points of the hydrostatic leveling instrument are arranged at dam sections 16<sup>#</sup>, 44<sup>#</sup> and 46<sup>#</sup>, respectively. The hydrostatic pressure balance method based on connected pipe is used in the present monitoring system. A current-sharing scheme with communication by field bus is adopted. The used equipment is hydrostatic level, and the accuracy of the leveling measurements is 0.5 mm.

The hydrostatic leveling system on the dam crest is based on J<sub>0</sub> on the right bank, and double metal markers SJ<sub>1</sub> and SJ<sub>2</sub> are embedded in this part. Based on the monitoring data of measuring points J<sub>23</sub> (on dam section 24<sup>#</sup>), J<sub>33</sub> (on dam section 4<sup>#</sup>), J<sub>35</sub> (on dam section 8<sup>#</sup>), J<sub>37</sub> (on dam section 12<sup>#</sup>), and J<sub>39</sub> (on dam section 15<sup>#</sup>) from 2006 to 2020, two regression models and one LSTM model of vertical displacement of the measurement points are constructed, and the positive value is subsidence according to the regulation of Wuqiangxi dam monitoring system.



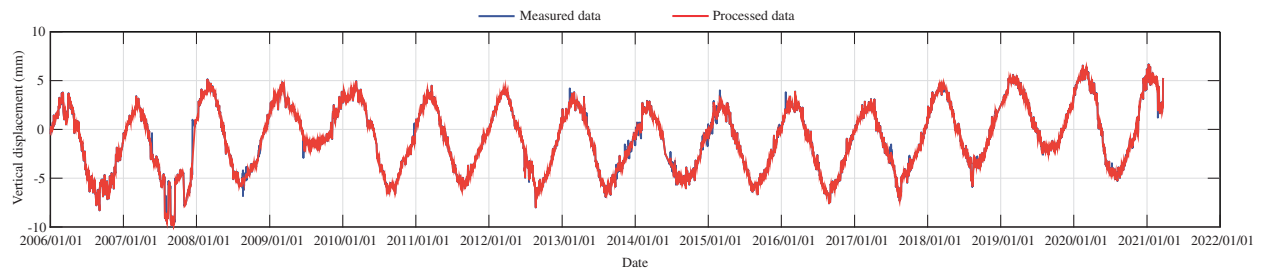
**Figure 5:** Schematic diagram of hydrostatic leveling layout of dam crest

#### 4.2 Data Preprocessing

The unreliable monitoring data of vertical displacements are determined with the Lahitte criterion. Table 1 shows the unreliable subsidence monitoring data of J<sub>23</sub>. Fig. 6 presents the process line of measured vertical displacements of J<sub>23</sub>. After removing the unreliable monitoring data, the subsidence process line of J<sub>23</sub> is smoother.

**Table 1:** Unreliable data of measured vertical displacements of J<sub>23</sub> (unit: mm)

Date	Subsidence	Date	Subsidence	Date	Subsidence
2007/5/24	-0.37	2007/7/27	-9.11	2007/9/16	-5.05
2007/7/26	-4.85	2007/7/31	-6.79	2007/12/11	-4.65
2007/7/27	-8.84	2007/9/15	-9.44	2007/12/12	1



**Figure 6:** The process line of measured vertical displacements of J<sub>23</sub> (unit: mm)

### 4.3 Statistical Regression Models

For convenience of expression, Eq. (8) can be rewritten as

$$\delta = a_{11}H + a_{12}H^2 + a_{13}H^3 + a_{21}h + a_{22}h^2 + a_{23}h^3 + b_1d + b_2e + c_1\theta + c_2f \quad (16)$$

where  $d = \sin\frac{2\pi t}{365}$ ,  $e = \cos\frac{2\pi t}{365}$ ,  $f = \ln\theta$ .

#### 4.3.1 Partial Least Squares Regression Model

In the cross-validity analysis of measuring point J<sub>23</sub> from 2006 to 2020,  $Q_1^2$ ,  $Q_2^2$  and  $Q_3^2$  are 1.0000, 0.2483 and 0.0168, respectively. Therefore, the accuracy of the three-component model can meet the requirements. The component  $t_h$  can be expressed as a function of the standardized variable  $x^*$ , i.e.,

$$t_1 = 0.089142x_1^* + 0.088813x_2^* + 0.088420x_3^* + 0.066707x_4^* + 0.070757x_5^* + 0.074168x_6^* + 0.646441x_7^* - 0.705975x_8^* - 0.162630x_9^* - 0.136581x_{10}^* \quad (17a)$$

$$t_2 = -0.345531x_1^* - 0.343128x_2^* - 0.340533x_3^* - 0.364465x_4^* - 0.365979x_5^* - 0.366488x_6^* + 0.484464x_7^* - 0.389352x_8^* + 0.055768x_9^* + 0.111455x_{10}^* \quad (17b)$$

$$t_3 = -0.109896x_1^* - 0.125921x_2^* - 0.141690x_3^* + 0.058062x_4^* + 0.071781x_5^* + 0.085863x_6^* - 0.268652x_7^* + 0.122680x_8^* - 0.549068x_9^* - 0.773957x_{10}^* \quad (17c)$$

The regression equation of standardized displacement variable is written as

$$y^* = -0.024433x_1^* - 0.025175x_2^* - 0.025882x_3^* + 0.003533x_4^* + 0.001003x_5^* - 0.001143x_6^* - 0.613302x_7^* + 0.641574x_8^* + 0.102413x_9^* + 0.063590x_{10}^* \quad (18)$$

By reducing standardized variables to original variables, the regression equation of partial least squares method is obtained as follows

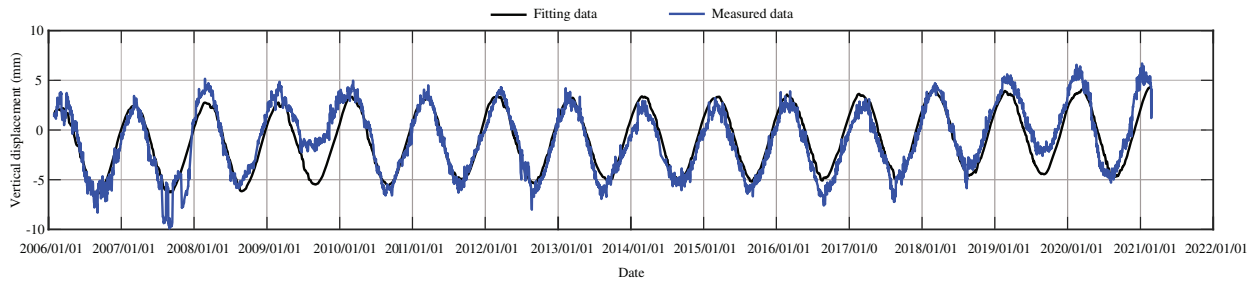
$$y = -0.341214 - 0.020904H - 0.000156H^2 - 0.000002H^3 + 0.007222h + 0.000047h^2 - 0.000002h^3 - 2.934919\sin\frac{2\pi t}{365} + 3.031150\cos\frac{2\pi t}{365} + 0.021690\theta + 0.371438\ln\theta \quad (19)$$

The multiple correlation coefficients (MCCs) and the residual standard deviations (RSDs) of the different measuring points for the partial least squares model are given in Table 2. The MCCs of the five measuring points on different dam sections exceed 0.91, and the RSDs do not exceed 0.0831. It shows that the high accuracy can be obtained with the partial least squares model.

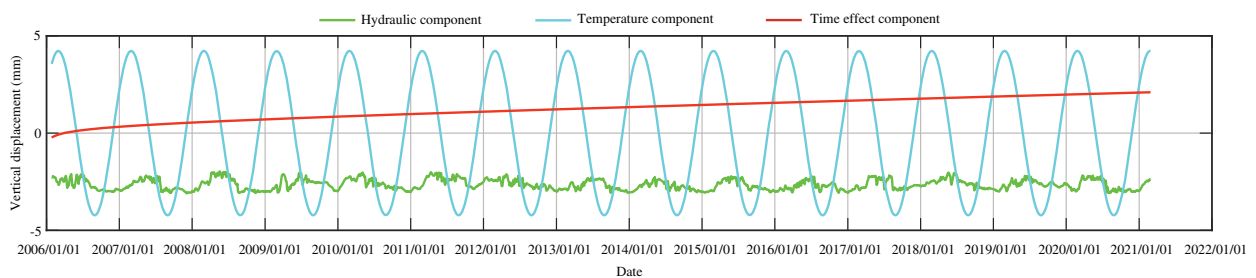
**Table 2:** The MCCs and RSDs of the different measuring points for the partial least squares model

Measuring point	J <sub>23</sub>	J <sub>33</sub>	J <sub>35</sub>	J <sub>37</sub>	J <sub>39</sub>
MCC	0.9118	0.9280	0.9273	0.9310	0.9445
RSD	0.0830	0.0786	0.0831	0.0778	0.0778

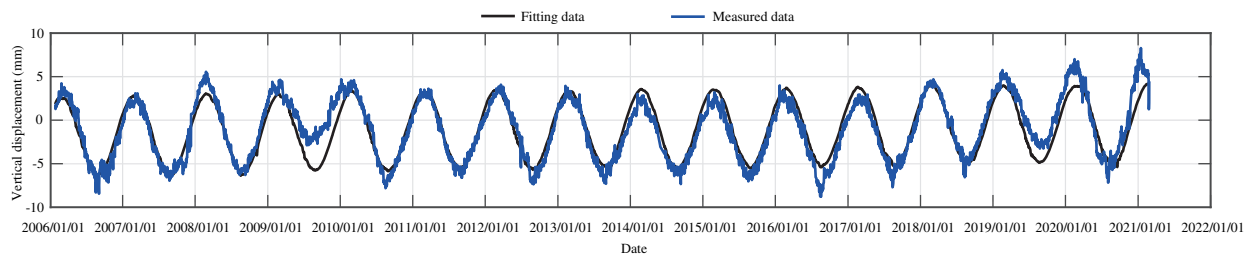
Figs. 7, 9, 11, 13, and 15 show the fitting curve of  $J_{23}$ ,  $J_{33}$ ,  $J_{35}$ ,  $J_{37}$  and  $J_{39}$  obtained with the partial least squares model, respectively. Figs. 8, 10, 12, 14 and 16 present the process line of each component of  $J_{23}$ ,  $J_{33}$ ,  $J_{35}$ ,  $J_{37}$  and  $J_{39}$ , respectively. It can be seen that this model can better reflect the variation law of dam crest settlement.



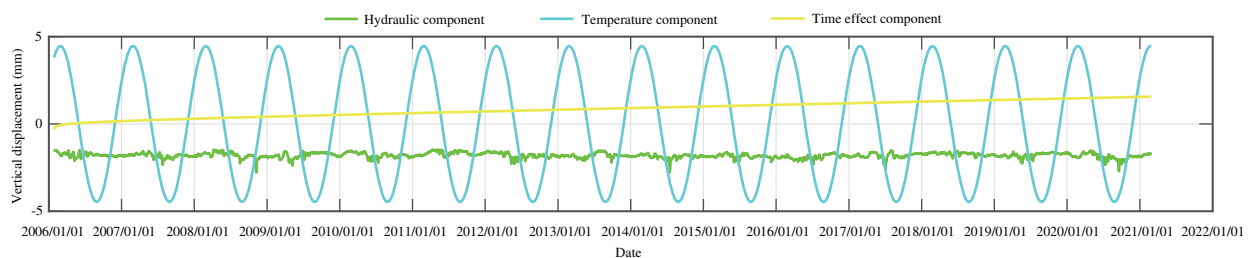
**Figure 7:** The fitting curve of vertical displacement for measuring point  $J_{23}$  by partial least squares method



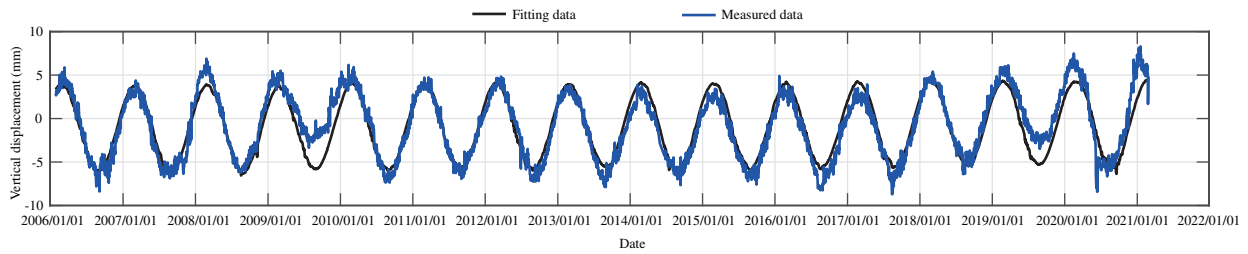
**Figure 8:** The process line of each component of vertical displacement for measuring point  $J_{23}$  by partial least squares method



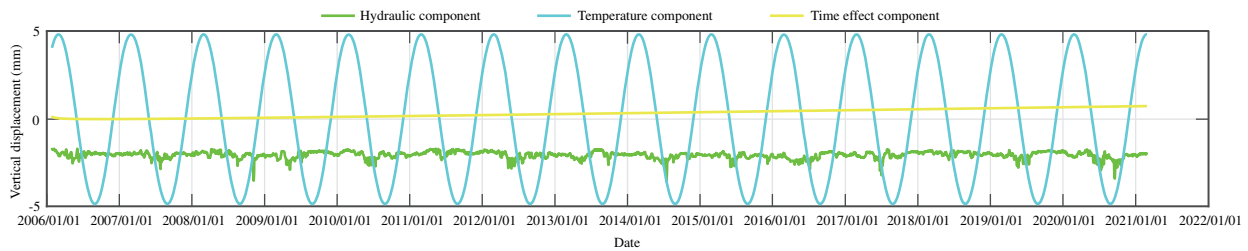
**Figure 9:** The fitting curve of vertical displacement for measuring point  $J_{33}$  by partial least squares method



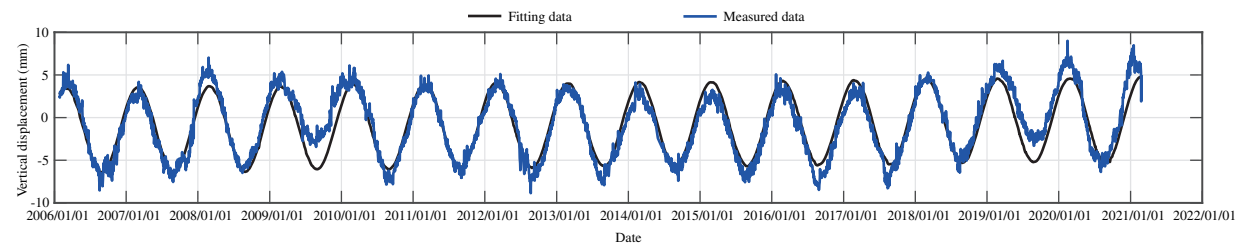
**Figure 10:** The process line of each component of vertical displacement for measuring point  $J_{33}$  by partial least squares method



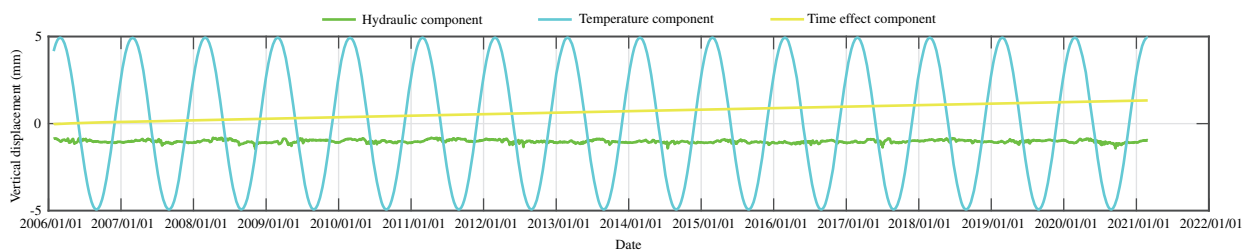
**Figure 11:** The fitting curve of vertical displacement for measuring point J<sub>35</sub> by partial least squares method



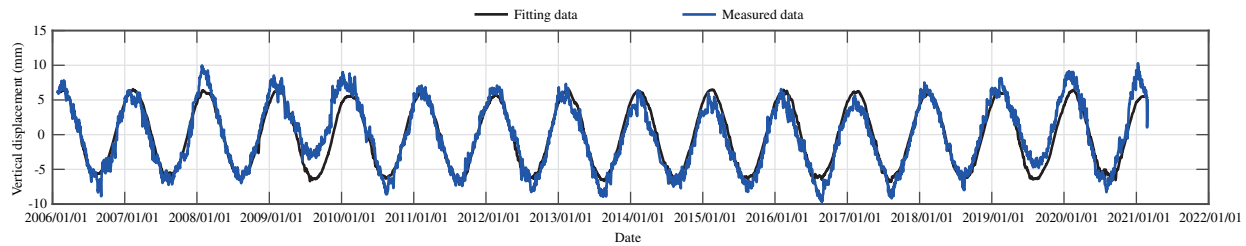
**Figure 12:** The process line of each component of vertical displacement for measuring point J<sub>35</sub> by partial least squares method



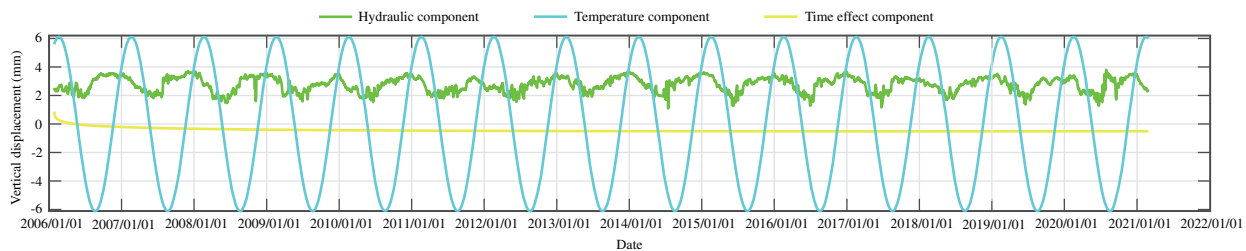
**Figure 13:** The fitting curve of vertical displacement for measuring point J<sub>37</sub> by partial least squares method



**Figure 14:** The process line of each component of vertical displacement for measuring point J<sub>37</sub> by partial least squares method



**Figure 15:** The fitting curve of vertical displacement for measuring point J<sub>39</sub> by partial least squares method



**Figure 16:** The process line of each component of vertical displacement for measuring point J<sub>39</sub> by partial least squares method

#### 4.3.2 Stepwise Regression Model

Table 3 shows the F-significance test of each factor for J<sub>23</sub> with stepwise regression method, showing the significance level of each factor. According to the significance level of the factors, the stepwise regression equation can be obtained by introducing variables in turn.

**Table 3:** F-significance test of each factor for J<sub>23</sub> in stepwise regression

Factor	$H$	$H^2$	$H^3$	$h$	$h^2$
F-significance test	0.007	0.00702	0.0069	0.0039	0.0044
Factor	$h^3$	$d$	$e$	$\theta$	$f$
F-significance test	0.0048	0.3720	0.4437	0.0235	0.0166

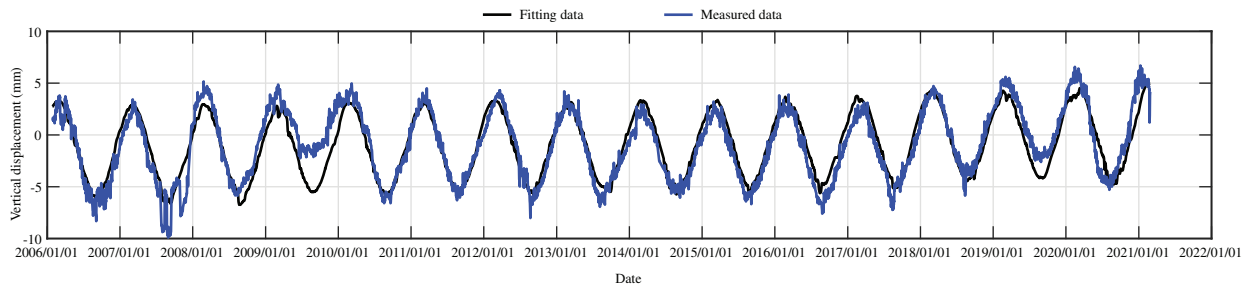
The multiple correlation coefficients (MCCs) and the residual standard deviations (RSDs) of the different measuring points for the stepwise regression model are given in Table 4. The MCCs of the five measuring points on different dam sections exceed 0.91, and the RSDs do not exceed 0.4013. It shows that the high accuracy can be also obtained with the stepwise regression model.

**Table 4:** The MCCs and RSDs of the different measuring points for the stepwise regression model

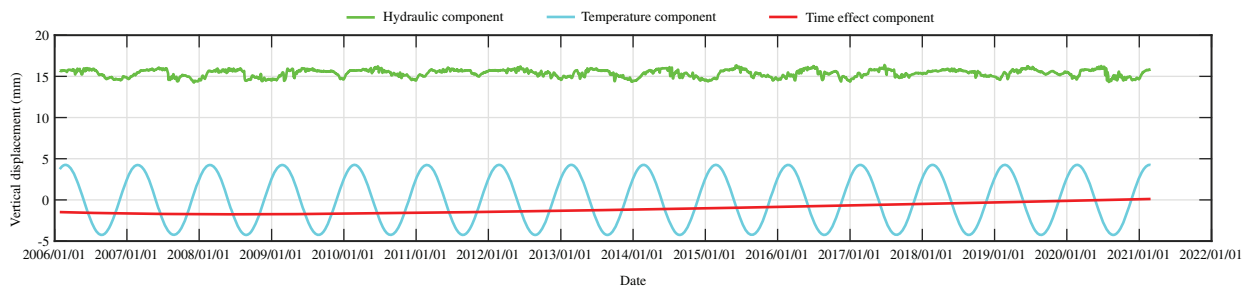
Measuring point	J <sub>23</sub>	J <sub>33</sub>	J <sub>35</sub>	J <sub>37</sub>	J <sub>39</sub>
MCC	0.9161	0.9258	0.9293	0.9336	0.9523
RSD	0.4013	0.3782	0.3695	0.3587	0.3055



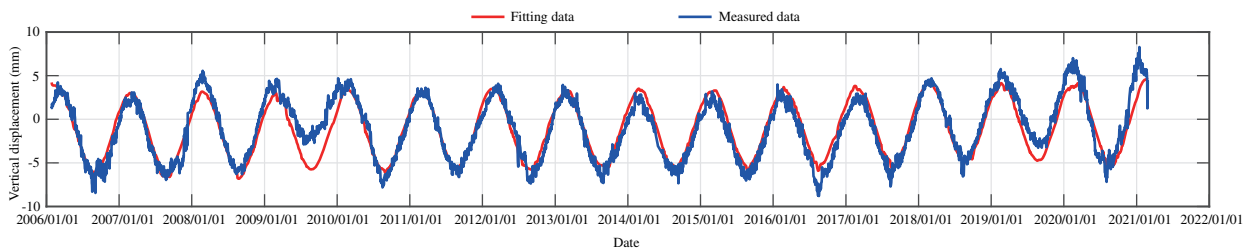
Figs. 17, 19, 21, 23 and 25 show the fitting curve of  $J_{23}$ ,  $J_{33}$ ,  $J_{35}$ ,  $J_{37}$  and  $J_{39}$  obtained with stepwise regression model, respectively. Figs. 18, 20, 22, 24 and 26 show the process line of each component of  $J_{23}$ ,  $J_{33}$ ,  $J_{35}$ ,  $J_{37}$  and  $J_{39}$ , respectively.



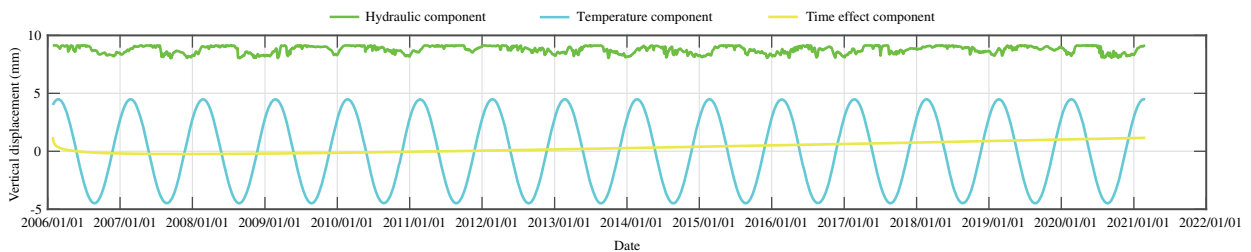
**Figure 17:** The fitting curve of vertical displacement for measuring point  $J_{23}$  by stepwise regression model



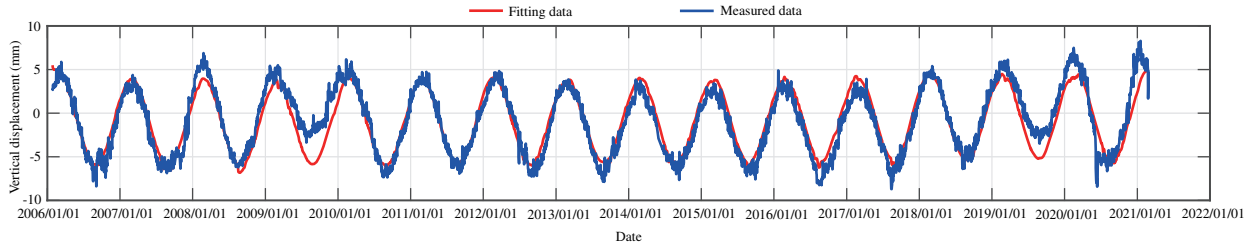
**Figure 18:** The process line of each component of vertical displacement for measuring point  $J_{23}$  by stepwise regression model



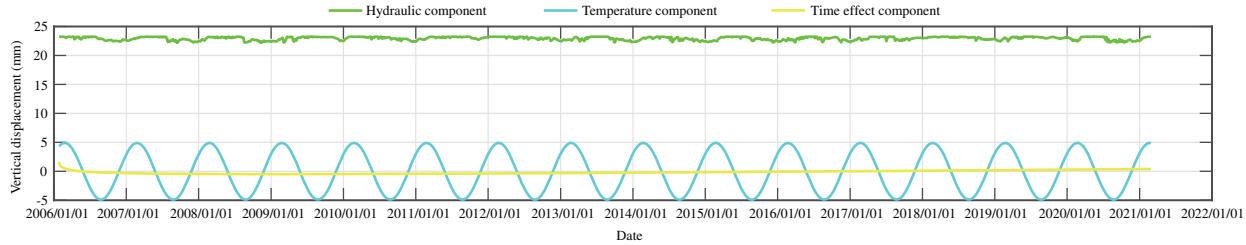
**Figure 19:** The fitting curve of vertical displacement for measuring point  $J_{33}$  by stepwise regression model



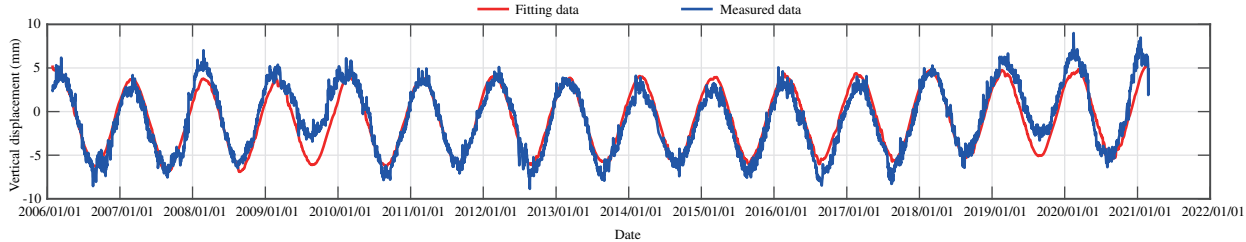
**Figure 20:** The process line of each component of vertical displacement for measuring point  $J_{33}$  by stepwise regression model



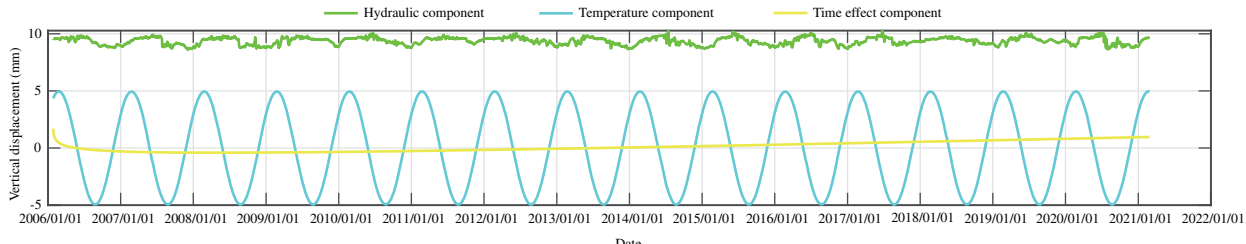
**Figure 21:** The fitting curve of vertical displacement for measuring point J<sub>35</sub> by stepwise regression model



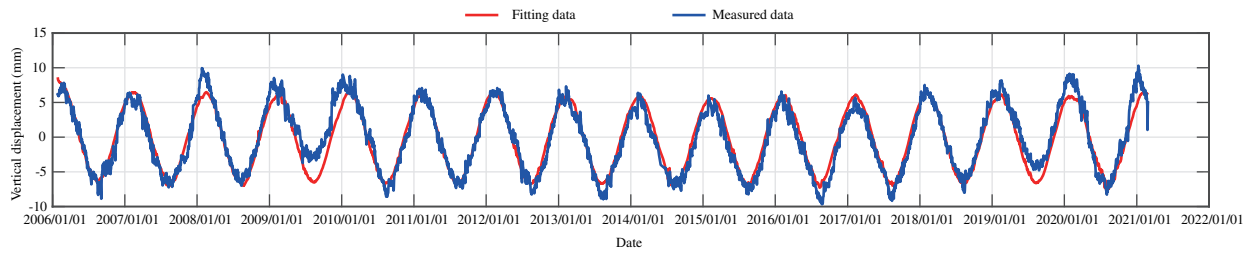
**Figure 22:** The process line of each component of vertical displacement for measuring point J<sub>35</sub> by stepwise regression model



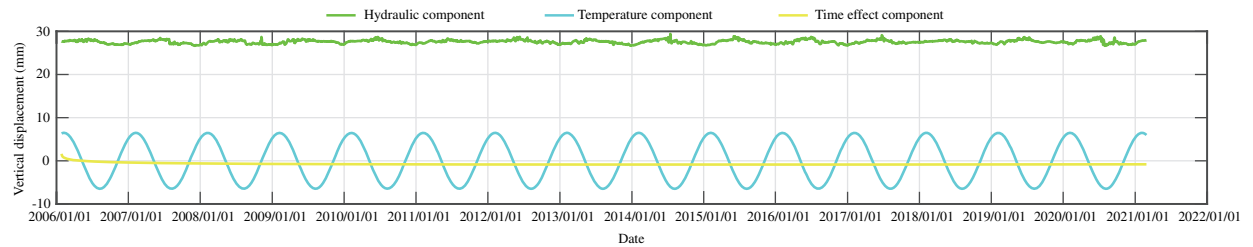
**Figure 23:** The fitting curve of vertical displacement for measuring point J<sub>37</sub> by stepwise regression model



**Figure 24:** The process line of each component of vertical displacement for measuring point J<sub>37</sub> by stepwise regression model



**Figure 25:** The fitting curve of vertical displacement for measuring point J<sub>39</sub> by stepwise regression model



**Figure 26:** The process line of each component of vertical displacement for measuring point J<sub>39</sub> by stepwise regression model

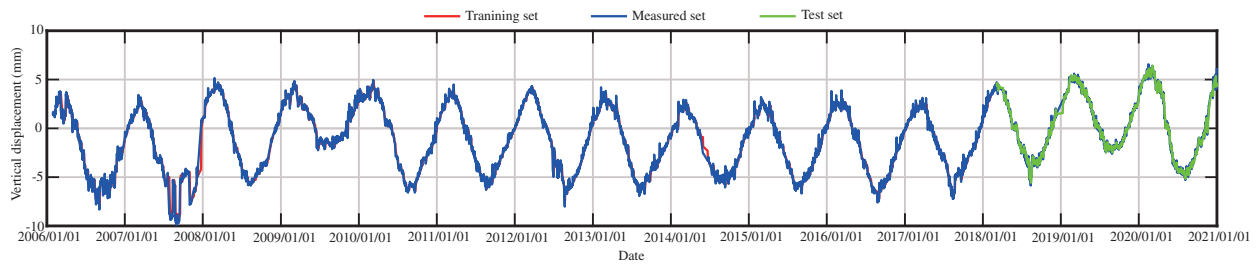
From Figs. 8 and 18, an offset can be found in the hydraulic component when comparing the different statistical models. The stepwise regression and partial least squares regression have no physical meaning, they only fit the data with different ways. The variable substitution is conducted in the partial least squares method, so the components in the partial least squares model are not equivalent to those in the stepwise regression model. In sum, there is no equivalence between the two models.

#### 4.4 LSTM Model

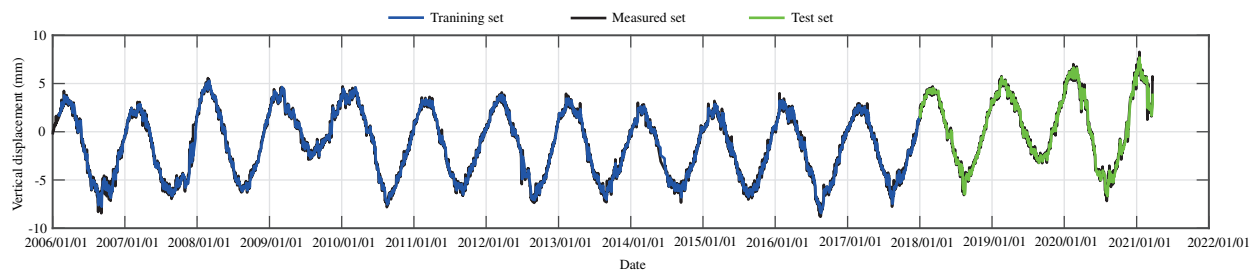
According to the numerical experiments, two LSTM layers are used, the rectified linear unit function is adopted as the activation function, and the input sequence length is 20, that is, the subsidence data of the first 20 days are used to predict the subsidence of the 21st day. The monitoring data of the training set are from 2006 to 2017, and the monitoring data from 2018 to 2020 are taken as the test set. The RMSEs of the trained model in the training set and in the test set for different measuring points are shown in Table 5. Figs. 27–31 present the training set and the test set of vertical displacement for measuring points J<sub>23</sub>, J<sub>33</sub>, J<sub>35</sub>, J<sub>37</sub> and J<sub>39</sub> by LSTM method. The RMSEs of measuring point J<sub>23</sub> in the training set and in the test set for different training datasets are shown in Table 6, and the corresponding training set and test set of vertical displacement from the trained model are shown in Figs. 32 and 33. The accuracy of the model increases with the increase of the training dataset. In addition, it is found that the accuracy of the predictions are very high when the training data are enough.

**Table 5:** The RMSEs of the trained model for different measuring points

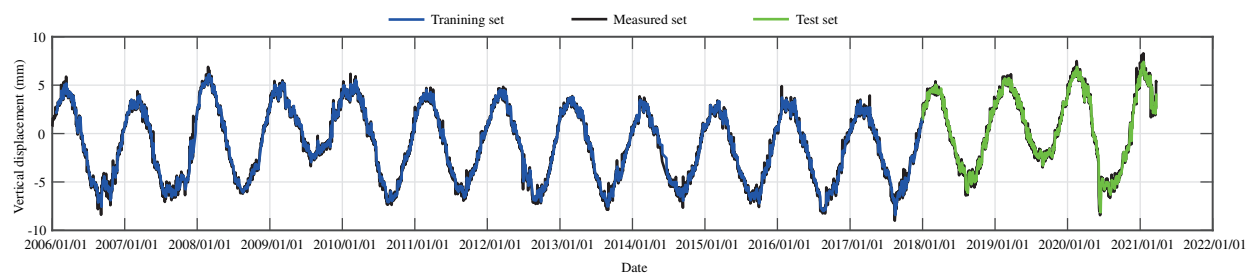
Measuring point	J <sub>23</sub>	J <sub>33</sub>	J <sub>35</sub>	J <sub>37</sub>	J <sub>39</sub>
Training set	0.39	0.39	0.48	0.47	0.47
Test set	0.40	0.44	0.54	0.50	0.49



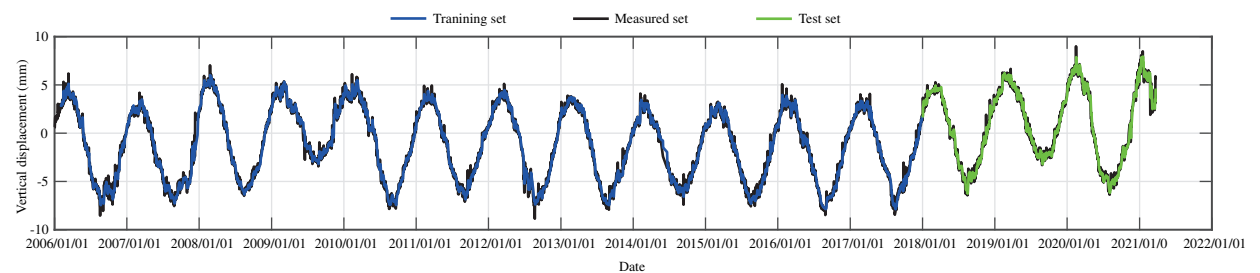
**Figure 27:** The training set and the test set of vertical displacement for measuring point J<sub>23</sub> by LSTM method



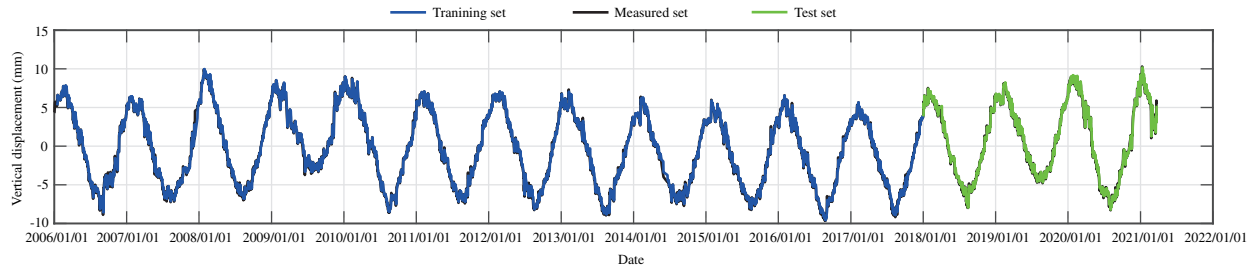
**Figure 28:** The training set and the test set of vertical displacement for measuring point J<sub>33</sub> by LSTM method



**Figure 29:** The training set and the test set of vertical displacement for measuring point J<sub>35</sub> by LSTM method



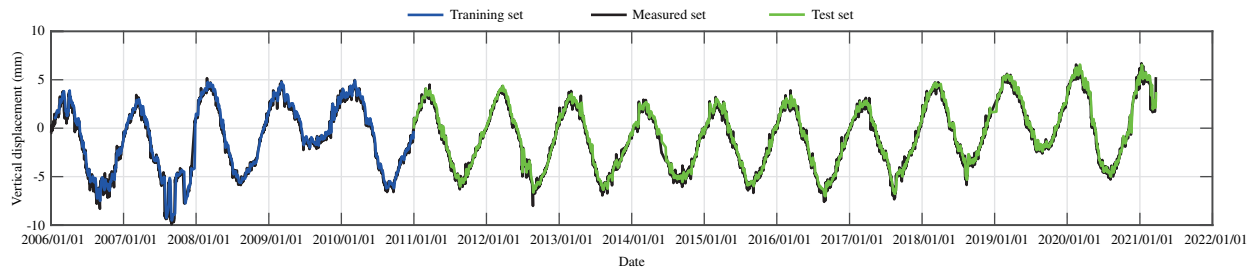
**Figure 30:** The training set and the test set of vertical displacement for measuring point J<sub>37</sub> by LSTM method



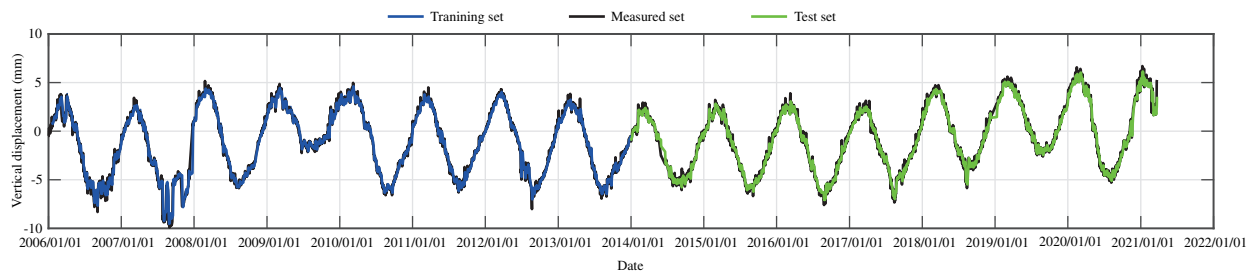
**Figure 31:** The training set and the test set of vertical displacement for measuring point J<sub>39</sub> by LSTM method

**Table 6:** The RMSEs of measuring point J<sub>23</sub> for different training datasets

Training dataset	2006–2010	2006–2013	2006–2017
Training set	0.53	0.43	0.39
Test set	0.43	0.41	0.40



**Figure 32:** The training set and the test set of vertical displacement for measuring point J<sub>23</sub> from the trained model (training dataset is from 2006 to 2010)



**Figure 33:** The training set and the test set of vertical displacement for measuring point J<sub>23</sub> from the trained model (training dataset is from 2006 to 2013)

#### 4.5 Discussion

The fitting results of the stepwise regression model and partial least square regression model are very similar, and the model quality is good, which can reflect the subsidence change law of measuring points. The multiple correlation coefficient of the stepwise regression method is slightly

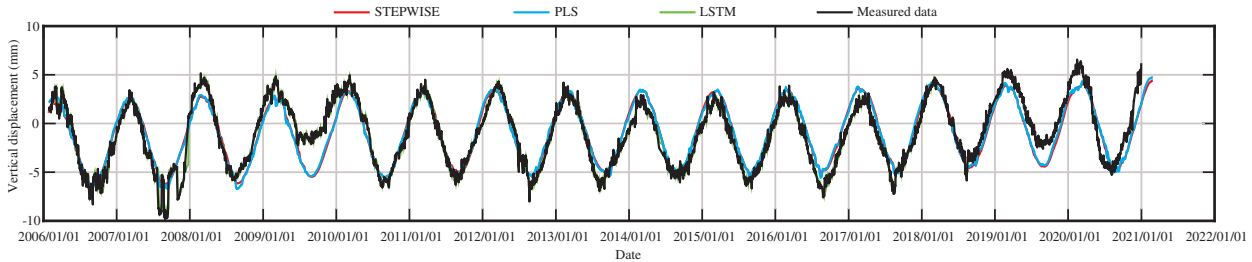
bigger than that of the partial least square method, while the residual standard deviation of the stepwise regression method is much bigger than that of the partial least square method.

From the process lines of the stepwise regression model and the partial least square model, it is found that the temperature component changes most obviously, and the change of the water pressure component and aging component is small. In other words, the displacements of measuring points are greatly affected by temperature change and less affected by water level and aging. This can be found from the introduction sequence of independent variables of stepwise regression method, which is in accord with the actual situation of dam operation. The dam body rises or sinks when the temperature rises or drops, and the changes are periodic. This is consistent with the actual situation of the dam.

Because of the high linear correlation of factors in water pressure component, stepwise regression method removes the factors of primary term of upstream water level and primary and tertiary terms of downstream water level, which will affect the regression analysis accuracy to some extent. In this regard, the partial least square regression method is more advantageous.

In the aging component, the fitting of partial least square regression method shows that the dam has settlement effect, and the subsidence will gradually slow with the time. The fitting of stepwise regression method shows that the dam has the effect of rising, and the rising amount will decrease with the time. The measuring point is not the starting time of statistical model based on the time of dam construction, so the time component has displacement at the beginning. Generally speaking, the change law of aging displacement of normal operation dam is sharp change in the initial stage and tends to be stable in the later period. However, the real displacement curves of measuring points show that the measuring points have obvious subsidence trend, so the aging components of both models are not stable, which is in agreement with the actual situation. On the whole, the aging component obtained by the partial least square model is more fit for the real situation.

Fig. 34 presents the fitting results of vertical displacement for measuring point J<sub>23</sub> obtained by the partial least squares model (PLS) and the stepwise regression model (STEPWISE) and the test results obtained by the LSTM model. From Fig. 34, in combination with Figs. 7, 17 and 27, it is found that the accuracy of the LSTM model is higher than those of the statistical models. However, the LSTM model cannot separate the influence of each component, in addition, the training data must be sufficient.



**Figure 34:** The prediction results of vertical displacement for measuring point J<sub>23</sub> by different methods

## 5 Conclusion

Two statistical models and a deep learning model of vertical displacement at five typical measuring points located on the crest of Wuqiangxi dam are constructed by partial least square method, stepwise regression method and LSTM recurrent neural network. The fitting results and the influence of each component on displacement value in the statistical models are compared and analyzed, the test curve in the LSTM model is given and compared with the fitting curves of the statistical models. The following conclusions are drawn:

- The prediction accuracy of the LSTM model is higher than the statistical models when there are enough training data, so the LSTM model is suggested when there are enough training data.
- From the multiple correlation coefficient, the fitting results of partial least squares regression model and stepwise regression model are similar, and the residual standard deviation obtained by partial least squares regression model is lower.
- The stepwise regression model removes some factors, and has a large residual standard deviation. The partial least squares regression model considers factors comprehensively and explains each component more strictly.
- It is more appropriate to use partial least squares regression model or LSTM model to predict the subsidence of measuring points in Wuqiangxi dam.

In the deformation prediction of concrete gravity dam, the LSTM model is suggested when there are sufficient training data, and the partial least squares regression model is suggested when the training data are insufficient. In addition, all deformation, water level, and temperature data in this study can be accessed at: <http://www.idmes.cn/data.html>.

**Acknowledgement:** The authors wish to express their appreciation to the reviewers for their helpful suggestions which greatly improved the presentation of this paper.

**Funding Statement:** The authors received no specific funding for this study.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

1. Hu, J., Ma, F. (2020). Statistical modelling for high arch dam deformation during the initial impoundment period. *Structural Control and Health Monitoring*, 27(12), e2638. DOI 10.1002/stc.2638.
2. Yuan, D., Wei, B., Xie, B., Zhong, Z. (2020). Modified dam deformation monitoring model considering periodic component contained in residual sequence. *Structural Control and Health Monitoring*, 27(12), e2633. DOI 10.1002/stc.2633.
3. Beaupré, L., St-Hilaire, A., Daigle, A., Bergeron, N. (2020). Comparison of a deterministic and statistical approach for the prediction of thermal indices in regulated and unregulated river reaches: Case study of the fourchue river (Quebec, Canada). *Water Quality Research Journal of Canada*, 55(4), 394–408. DOI 10.2166/wqrj.2020.001.
4. Wei, B., Liu, B., Yuan, D., Mao, Y., Yao, S. (2021). Spatiotemporal hybrid model for concrete arch dam deformation monitoring considering chaotic effect of residual seriesn. *Engineering Structures*, 228, 111488. DOI 10.1016/j.engstruct.2020.111488.
5. Pei, L., Wu, Z., Cui, M., Zhang, Q., Chen, J. (2012). Research and application on the displacement hybrid-model of high earth dam. *Journal of Sichuan University (Engineering Science Edition)*, 44(Supl), 42–47. DOI 10.15961/j.jsuese.2012.s1.051.



6. Zhen, D., Huo, Z., Li, B. (2011). Arch-dam crack deformation monitoring hybrid model based on XFEM. *Science China Technological Sciences*, 54(10), 2611–2617. DOI 10.1007/s11431-011-4550-6.
7. Ren, Q., Li, M., Song, L., Liu, H. (2020). An optimized combination prediction model for concrete dam deformation considering quantitative evaluation and hysteresis correction. *Advanced Engineering Informatics*, 46, 101154. DOI 10.1016/j.aei.2020.101154.
8. Wei, B., Yuan, D., Li, H., Xu, Z. (2019). Combination forecast model for concrete dam displacement considering residual correction. *Structural Health Monitoring*, 18(1), 232–244. DOI 10.1177/1475921717748608.
9. Tatin, M., Briffaut, M., Dufour, F., Simon, A., Fabre, J. (2018). Statistical modelling of thermal displacements for concrete dams: Influence of water temperature profile and dam thickness profile. *Engineering Structures*, 165, 63–75. DOI 10.1016/j.engstruct.2018.03.010.
10. Ding, J., Yu, T., Bui, T. (2020). Modeling strong/weak discontinuities by local mesh refinement variable-node XFEM with object-oriented implementation. *Theoretical and Applied Fracture Mechanics*, 106, 102434. DOI 10.1016/j.tafmec.2019.102434.
11. Liu, G. (2009). *Meshfree methods: Moving beyond the finite element method (Second Edition)*. USA: CRC Press.
12. Li, B., Yang, J., Hu, D. (2020). Dam monitoring data analysis methods: A literature review. *Structural Control and Health Monitoring*, 27, e2501. DOI 10.1002/stc.2501.
13. Tonini, D. (1956). Observed behavior of several Italian arch dams. *Journal of the Power Division*, 82(6), 1–26. DOI 10.1061/JPWEAM.0000062.
14. Wu, Z. (2003). *Safety monitoring theory and its application of hydraulic structures (First Edition)*. Beijing: Higher Education Press.
15. Chen, J. (1980). The nonlinear parameter estimation of aging deformation of concrete dam. *Dam Observation and Geotechnical Testing*, (2), 3–13. DOI CNKI:SUN:DBGC.0.1980-02-000.
16. Léger, P., Leclerc, M. (2007). Hydrostatic, temperature, time-displacement model for concrete dams. *Journal of Engineering Mechanics*, 133(3), 267–277. DOI 10.1061/(ASCE)0733-9399(2007)133:3(267).
17. Tatin, M., Briffaut, M., Dufour, F., Simon, A., Fabre, J. (2015). Thermal displacements of concrete dams: Accounting for water temperature in statistical models. *Engineering Structures*, 91, 26–39. DOI 10.1016/j.engstruct.2015.01.047.
18. Mata, J., T. de Castro, A., Sá da Costa, J. (2014). Constructing statistical models for arch dam deformation. *Structural Control and Health Monitoring*, 21, 423–437. DOI 10.1002/stc.1575.
19. Wang, S., Xu, C., Gu, C., Su, H., Hu, K. et al. (2020). Displacement monitoring model of concrete dams using the shape feature clustering-based temperature principal component factor. *Structural Control and Health Monitoring*, 27, e2603. DOI 10.1002/stc.2603.
20. Hu, D., Zhou, Z., Li, Y., Wu, X. (2011). Dam safety analysis based on stepwise regression model. *Advanced Materials Research*, 204, 2158–2161. DOI 10.4028/www.scientific.net/AMR.204-210.2158.
21. Shen, W., Ren, J. (2014). Multiple stepwise regression analysis crack open degree data in gravity dam. *Applied Mechanics and Materials*, 477, 888–891. DOI 10.4028/www.scientific.net/AMM.477-478.888.
22. Chen, H., Sun, Y., Gao, J., Hu, Y., Yin, B. (2018). Solving partial least squares regression via manifold optimization approaches. *IEEE Transactions on Neural Networks and Learning Systems*, 30, 588–600. DOI 10.1109/TNNLS.2018.2844866.
23. Cheng, X., Li, Q., Zhou, W., Zhou, Z. (2020). External deformation monitoring and improved partial least squares data analysis methods of high core rock-fill Dam (HCRFD). *Sensors*, 20(2), 444. DOI 10.3390/s20020444.
24. Yin, W., Zhao, E., Gu, C., Huang, H., Yang, Y. (2019). A nonlinear method for component separation of Dam effect quantities using kernel partial least squares and pseudosamples. *Advances in Civil Engineering*, 2019, 1958173. DOI 10.1155/2019/1958173.
25. Huang, H., Chen, B., Liu, C. (2015). Safety monitoring of a super-high Dam using optimal kernel partial least squares. *Mathematical Problems in Engineering*, 2015, 571594. DOI 10.1155/2015/571594.
26. He, X., Leng, Y., Zhao, S. (2010). Threshold regression forecast model for dam safety monitoring and its application. *Applied Mechanics and Materials*, 36, 182–186. DOI 10.4028/www.scientific.net/AMM.36.182.

27. Lan, S., Fang, C. (2012). Application of least square method based logistic regression model to settlement prediction of earth-rock fill dam. *Water Resources and Hydropower Engineering*, 43(3), 16–18. DOI 10.13928/j.cnki.wrahe.2012.03.011.
28. Dai, B., Gu, C., Zhao, E., Qin, X. (2018). Statistical model optimized random forest regression model for concrete dam deformation monitoring. *Structural Control and Health Monitoring*, 25, 1–15. DOI 10.1002/stc.2170.
29. Li, X., Wen, Z., Su, H. (2021). An approach using random forest intelligent algorithm to construct a monitoring model for dam safety. *Engineering with Computers*, 37, 39–56. DOI 10.1007/s00366-019-00806-0.
30. Qu, X., Yang, J., Chang, M. (2019). A deep learning model for concrete dam deformation prediction based on RS-LSTM. *Journal of Sensors*, 2019, 4581672. DOI 10.1155/2019/4581672.
31. Yang, D., Gu, C., Zhu, Y., Dai, B., Zhang, K. et al. (2020). A concrete dam deformation prediction method based on LSTM with attention mechanism. *IEEE Access*, 8, 185177–185186. DOI 10.1109/Access.6287639.
32. Liu, W., Pan, J., Ren, Y., Wu, Z., Wang, J. (2020). Coupling prediction model for long-term displacements of arch dams based on long short-term memory network. *Structural Control and Health Monitoring*, 27, e2548. DOI 10.1002/stc.2548.
33. Han, B., Geng, F., Dai, S., Gan, G., Liu, S. et al. (2020). Statistically optimized back-propagation neural-network model and Its application for deformation monitoring and prediction of concrete-face rockfill dams. *Journal of Performance of Constructed Facilities*, 34(4), 04020071. DOI 10.1061/(ASCE)CF.1943-5509.0001485.
34. Chen, S., Gu, C., Lin, C., Hariri-Ardebili, M. (2021). Prediction of arch dam deformation via correlated multi-target stacking. *Applied Mathematical Modelling*, 91, 1175–1193. DOI 10.1016/j.apm.2020.10.028.
35. Shi, Z., Gu, C., Zhao, E., Xu, B. (2020). A novel seepage safety monitoring model of CFRD with slab cracks using monitoring data. *Mathematical Problems in Engineering*, 2020, 1641747. DOI 10.1155/2020/1641747.
36. Yang, J., Qu, X., Chang, M. (2019). An intelligent singular value diagnostic method for concrete dam deformation monitoring. *Water Science and Engineering*, 12(3), 205–212. DOI 10.1016/j.wse.2019.09.006.
37. Liu, X., Kang, F., Ma, C., Li, H. (2021). Concrete arch dam behavior prediction using kernel-extreme learning machines considering thermal effect. *Journal of Civil Structural Health Monitoring*, 11, 283–299. DOI 10.1007/s13349-020-00452-x.
38. Chen, S., Gu, C., Lin, C., Zhang, K., Zhu, Y. (2021). Multi-kernel optimized relevance vector machine for probabilistic prediction of concrete dam displacement. *Engineering with Computers*, 37, 1943–1959. DOI 10.1007/s00366-019-00924-9.
39. Shu, X., Bao, T., Li, Y., Gong, J., Zhang, K. (2021). VAE-Talstm: A temporal attention and variational autoencoder-based long short-term memory framework for dam displacement prediction. *Engineering with Computers*, DOI 10.1007/s00366-021-01362-2.
40. Chen, W., Wang, X., Cai, Z., Liu, C., Zhu, Y. et al. (2021). DP-Gmm clustering-based ensemble learning prediction methodology for dam deformation considering spatiotemporal differentiation. *Knowledge-Based Systems*, 222, 106964. DOI 10.1016/j.knosys.2021.106964.