

Received April 30, 2020, accepted May 15, 2020, date of publication May 19, 2020, date of current version June 2, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2995592

# The Prediction of Dam Displacement Time Series Using STL, Extra-Trees, and Stacked LSTM Neural Network

YANGTAO LI<sup>ID</sup>, TENGFEI BAO<sup>ID</sup>, JIAN GONG<sup>ID</sup>, XIAOSONG SHU<sup>ID</sup>, AND KANG ZHANG

State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210024, China  
College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing 210024, China

Corresponding author: Tengfei Bao (baotf@hhu.edu.cn)

This work was supported in part by the National Key Research and Development Program under Grant 2018YFC1508603 and Grant 2016YFC0401601, and in part by the National Natural Science Foundation of China under Grant 51579086 and Grant 51739003.

**ABSTRACT** Dam behavior prediction is a fundamental component of dam structural health monitoring. By comparing the predictions and the observations, anomalies can be detected, and then the remedial measures can be executed in time. As the most intuitive monitoring indicators, deformation is often used to evaluate dam structural health status. In this research, we propose a novel combined model for predicting the dam displacement time series. First, the seasonal-trend decomposition based on Loess (STL) method is utilized to decompose the dam displacement time series into seasonal, trend, and remainder components. Then the extremely randomized trees (extra-trees) model is used to predict seasonal components based on the causal models and influencing factors, whereas the stacked Long-Short Term Memory (LSTM) model is used to predict trend and remainder components based on the numerical models and historical observation data. Finally, the predicted results of the three components are aggregated to obtain the total predicted dam displacement. Seven state-of-the-art methods are introduced as benchmark methods to verify the effectiveness and feasibility of the proposed model. To quantitatively evaluate and compare the prediction results, three evaluation indicators, and a statistic test method are introduced. The experimental results show that the proposed model is the best-performing method compared with other benchmark methods both in prediction accuracy and stability. This indicates the proposed novel combined model STL-extra-trees-LSTM is a promising method for predicting displacement time series.

**INDEX TERMS** Dam behavior prediction, time series decomposition, deep learning, ensemble learning.

## I. INTRODUCTION

In China, more than 98000 dams in active service, and most of them were constructed in the 1950s -1970s. These dams have potential problems such as low design standards, insufficient strength of dam materials, and serious aging [1]. Moreover, they locate in the deep mountain valleys with the poor geological environments and are subject to the influence of extreme loads and weather, such as floods, earthquakes, and cold waves. Once the dams break, reservoir water will pour downstream in an instant, creating an uncontrollable flood that will cause environmental, social, and economic disasters [2].

Structural failure of most dams is not a sudden event, but a gradual process under the long-term effects of various loads [3]. If the dam behavior prediction model can be

established based on the prototype observation, the structural degradation of the dam and its trend can be detected in time. However, dam structural performance is a dynamic evolution process with the interaction of the dam material property and multiple external factors.

Dam behavior prediction is a fundamental component of dam structural health monitoring. They are utilized to calculate the dam response under internal and external loads. The anomalies can be detected timely by comparing the predictions with the observations. Then the conduct maintenance and remedial measures can be executed in time.

As the most intuitive indicators, deformation is commonly used to evaluate the safety status of the dam. Dam deformation process is a dynamic non-linear evolution with the characteristics of complexity, uncertainty, diversity, and time-varying. Traditional statistical methods, such as

The associate editor coordinating the review of this manuscript and approving it for publication was Orazio Gambino<sup>ID</sup>.

multiple linear regression (MLR), have been widely utilized to dam safety monitoring and behavior prediction. However, the regression results are only the approximate fit of the actual relationship between variables. It is hard to predict dam displacement time series with high accuracy and stability, especially for long prediction sequences. Moreover, with the development of monitoring technology, the sampling frequency of dam monitoring devices has changed from once a week to multiple times a day. It is desirable to propose advanced intelligent methods to meet the requirements of big data processing.

In recent years, many machine learning (ML)-based methods with strong capability on handling non-linear problems are introduced to predict dam displacement [1], [4]. For example, Lin *et al.* [5] proposed a Gaussian process-based prediction model for dam deformation and utilized an example analysis to verify the feasibility of the Gaussian process model. These methods obtain good performance in learning input-output relationships without requiring a detailed physical process between actions and effects [6]. But there are still many drawbacks that limit the application of ML methods, such as overfitting problems for artificial neural network and parameter tuning for support vector machine.

These limitations foster the development of combined models [7]. Various kinds of combined models are proposed, and one of them is the signal pre-processing measure model. The signal pre-processing measure is implemented to decompose the time series into a collection of stationary and regular sub-sequences. Then the models that fit the characteristics of the sub-sequences are selected as the prediction model [8]. Different decomposition methods have been utilized to process the dam prototype observation time series. For instance, in [9], wavelet analysis (WA) is used to decompose and reconstruct the residual sequence of the dam displacement time series. In [10], an improved empirical mode decomposition (EMD)-based method is presented to remove noise from prototypical observations on dam safety. Ahumada and Garegnani [11] applied the wavelet de-noising and Hodrick-Prescott (HP) filter to decompose the displacement time series of the landslide into periodic, trend terms, and random noise. Nevertheless, these decomposition methods still suffer from some problems. For example, WA can decompose a time series into both high and low frequency series at each level. But there are no standard definitions of high and low frequency in dam displacement time series decomposition. It is fully determined by the experience of the user. Moreover, EMD can decompose a time series into a collection of stationary Intrinsic Mode Functions (IMFs), but the mode mixing problem still exists. Specifically, a signal of different scales still exists in one IMF or a similar scaled resides in different IMFs [12]. Furthermore, the decomposition results of HP filter are determined by the value of smoothing parameter  $\lambda$ . In addition, HP filter can only process series with low sampling frequency, such as the yearly, quarterly or monthly data [13].

The seasonal-trend decomposition based on Loess (STL, local polynomial regression fitting) is a well-established filtering procedure for decomposing a time series into additive variation components [14]. Compared with other decomposition methods, STL has some significant advantages [15], [16]. For example, it has strong resilience to outliers in the sequence, resulting in robust sub-series. Moreover, STL can deal with any seasonal frequency that great than one [17]. Also, the modeling process of STL is purely based on numerical methods, without requiring any mathematical modeling tools.

In this study, STL is used to decompose the dam displacement time series into three components: seasonal, trend, and remainder components. According to the previous researches, the cause of each component of dam displacement is analyzed. The seasonal component represents periodic dam responses influenced by the water level and seasonal temperature changes during its lifetime [18]. Extremely randomized trees (extra-trees) are utilized to predict the seasonal displacement based on multiple influencing factors (e.g. reservoir level, and temperatures). The trend component reflects the long-term evolution process of dam behavior under the comprehensive effect of the property degradation of dam material and inherent rheological property [19]. The remainder component is caused by uncertain factors, such as structural damage, seepage coupling, and joint fissure [20]. It is hard to predict trend and remainder components accurately based on the causal models and influencing factors. Thus, the stacked Long-Short Term Memory Neural Network (LSTM NN) is utilized to predict them based on the numerical model and historical observation data [21]. Then the predicted results of the three components are aggregated to obtain the final predicted results. To effectively compare the prediction performance of the proposed model, seven state-of-the-art methods are selected as benchmark methods. Three quantitative evaluation indicators and a modified hypothesis test are utilized to compare the prediction performance between the proposed and the benchmark methods.

The structure of this paper is described as below. Section 2 presents a brief description of the STL, extra-trees, and stacked LSTM models, followed by a formulation of the proposed combined model in detail. In Section 3, the research design of the study case, input variable selection, evaluation indicators, and model implementation is introduced. In Section 4, the experimental results of the proposed model and benchmark methods for various displacement time series are illustrated and discussed. Finally, the conclusions are drawn in Section 5.

## II. METHODOLOGY

The overall process of the proposed model is introduced and described in this section. Firstly, the theoretical basis of the STL, extra-trees, and stacked LSTM models is described briefly. Then the proposed STL-extra-trees-LSTM model is formulated, and the specific steps of the proposed model are presented in detail.

### A. SEASONAL-TREND DECOMPOSITION BASED ON LOESS

The STL method is a filtering procedure for decomposing a time series into three additive components based on the Loess smoother [14]. STL is commonly used in processing a large number of long-scaled time series due to its simple design and fast computation speed.

Assuming there is a dam displacement time series  $X_t$ , STL, can disaggregate the global time series  $X_t$  into three additive components, namely seasonal ( $S_t$ ), trend ( $T_t$ ), and remainder components ( $R_t$ ).

$$X_t = S_t + T_t + R_t \quad (1)$$

As an iterative calculation method, the implementation of STL is composed of two recursive procedures: inner and outer loops. In each of the passes through the inner loop, seasonal smoothing is used to update the seasonal component, followed by a trend smoothing is used to update the trend component. The specific calculation process is divided into six steps as follows.

Assuming that  $S_t^{(k)}$ ,  $T_t^{(k)}$  are the seasonal and trend components at the end of the  $k$ th pass, respectively. The updates of the  $(K+1)_{st}$  step of  $S_t^{(k+1)}$  and  $T_t^{(k+1)}$  are calculated in the following steps.

**Step1:** Detrending. A detrended series is obtained by subtracting the original series from the estimated trend series  $T_t^{(k)}$ :  $X_t^{\text{trend}} = X_t - T_t^{(k)}$ .

**Step2:** Cycle-subseries smoothing. Each cycle-subseries of the detrended series is smoothed by a Loess smoother, and a preliminary seasonal series  $C_t^{(k+1)}$  is obtained.

**Step3:** Low-pass Filter of smoothed cycle subseries. The preliminary seasonal series obtained from **Step2** is processed using a low-pass filter, followed by a Loess smoother, to obtain the remaining trend series  $L_t^{(k+1)}$ .

**Step4:** Detrending of smoothed cycle subseries. The seasonal component  $S_t^{(k+1)}$  from the  $(K+1)_{st}$  loop is obtained by subtracting the preliminary seasonal series from the low-pass values:  $S_t^{(k+1)} = C_t^{(k+1)} - L_t^{(k+1)}$ .

**Step5:** Deseasonalizing. A deseasonalized series  $X_t^{\text{deseason}}$  is obtained by subtracting the original series  $X_t$  from the seasonal component  $S_t^{(k+1)}$ :  $X_t^{\text{deseason}} = X_t - S_t^{(k+1)}$ .

**Step6:** Trend Smoothing. The deseasonalized series  $X_t^{\text{deseason}}$  obtained in **Step5** is smoothed by a Loess smoother to obtain the trend component  $T_t^{(k+1)}$ .

After the inner iteration reaches the accuracy requirement, the inner iteration is over, and the outer loop starts. In the outer loop, the estimation of the seasonal and trend components obtained in the inner loop is used to calculate the remainder component  $R_t$ .

$$R_t^{(k+1)} = X_t - T_t^{(k+1)} - S_t^{(k+1)} \quad (2)$$

Any large values in  $R_t$  are regarded as outliers, and weight coefficients are calculated. In the further iteration of the inner loop, the weight coefficients are used to down-weight the influence of outliers when updating the seasonal and trend components.

As such, the dam displacement time series is decomposed into three additive components, seasonal, trend, and remainder components by using the STL method.

### B. EXTREMELY RANDOMIZED TREES

Ensemble learning is an important and practical method that boosts various base learners to generate a strong learner with good generalization capability [22], [23]. One of the most commonly-used base learners is the decision tree, which is easier to intercept than neural networks and does not rely on prior knowledge. But the decision tree is prone to overfitting because the sample space may be excessively divided in the recursive process of decision trees. Many methods have been proposed to overcome the drawbacks of traditional decision trees, and one of them is the extra-trees model. Compared with other tree-based ensemble methods, the extra-trees model mainly has two differences. Firstly, the extra-trees model goes further in randomness [24]. Like random forest (RF) models, a random subset of candidate features is generated in extra-trees models. But different from searching the most discriminative thresholds in RF models, it randomly generates thresholds for each candidate feature. Then the optimal randomly generated threshold is selected as the splitting rules [25]. Moreover, the whole training samples are used to grow the trees in extra-trees models rather than a bootstrap replica in RF models [26]. These improvements lead to more diversified trees and fewer splitters to evaluate when training an extra-trees model. Additionally obtained randomness allows the extra-trees model with faster computation speed and reduction of the variance of the model with slightly increased bias [25].

### C. STACKED LONG-SHORT TERM MEMORY NEURAL NETWORK

Deep learning is a sub-area of ML methods that can express much more complex relationships by adding more layers and nonlinear elements in a layer [27]. The feature learning of deep learning is realized through a general-purpose learning mechanism instead of time-consuming human manual feature extraction or expert domain knowledge [28]. Deep learning techniques have been widely utilized to solve practical problems and achieved state-of-the-art or highly competitive results [29].

Recurrent Neural Network (RNN) is a class of deep neural network, which utilizes its internal memory loops to deal with sequence data. Although traditional RNNs exhibit strong capability of solving the nonlinear problems in sequence prediction, the gradient vanishing or exploding problems still exist, which limits its application. Improved from RNN, the LSTM model was proposed by Hochreiter and Schmidhuber in 1997 [30]. Exist researches show that the LSTM model can effectively learn the temporal and long term dependencies from time series compared with traditional RNNs [31], [32]. LSTM model uses special units called memory blocks to replace the traditional neurons in the hidden layers. The memory block of the LSTM consists of one or more memory

cells and three gates (input, forget, and output gates), which is shown in FIGURE 1. The input gate determines what values from the input to update the memory state. The forget gate determines to what extent to forget the previous output results and selects the optimal time lag for the input sequence. The output gate determines what to output based on input and the memory of the block [33]. At time  $t$ , the input, forget, input modulation, and output gates can be formulated as  $i_t, f_t, \tilde{c}_t$  and  $o_t$ , respectively. The process of updating the state of the cell and calculating the output of LSTMs can be described as follows [34]:

$$f_t = \sigma(w_{f,x}x_t + w_{f,h}h_{t-1} + w_{c,i}c_{t-1} + b_f) \quad (3)$$

$$i_t = \sigma(w_{i,x}x_t + w_{i,h}h_t + w_{c,f}c_{t-1} + b_i) \quad (4)$$

$$\tilde{c}_t = \tanh(w_{\tilde{c},x}x_t + w_{\tilde{c},h}h_{t-1} + b_{\tilde{c}}) \quad (5)$$

$$o_t = \sigma(w_{o,x}x_t + w_{o,h}h_{t-1} + b_o) \quad (6)$$

where  $w_{f,x}, w_{i,x}, w_{\tilde{c},x}, w_{o,x}$  represent the weight coefficients, which is used to map the input of hidden layer to the gates and input cell state;  $w_{f,h}, w_{i,h}, w_{\tilde{c},h}, w_{o,h}$  are the weight coefficients, which is used to connect the previous cell output state to the gates;  $b_f, b_i, b_{\tilde{c}}$  and  $b_o$  are four bias vectors.  $\sigma(\bullet)$  represents the standard logistics sigmoid function,  $\sigma(x) = \frac{1}{1+e^{-x}}$ .

Based on the above-mentioned equations, at each time  $t$  iteration, the cell output state  $C_t$  and the output  $h_t$  can be denoted as

$$c_t = f_t * c_{t-1} + i_t * \tilde{c}_t \quad (7)$$

$$h_t = o_t * \tanh(c_t) \quad (8)$$

where  $*$  represents the scalar product of two vectors.

Current studies show deep LSTM models with several hidden layers can build up a progressively higher level of representations of sequence data [35]. Deep LSTM models are the networks with several LSTM hidden layers, in which the output of an LSTM hidden layer will be fed into the next level of the LSTM hidden layer as the input. This stacked-layers mechanism can effectively enhance the learning power of NN [36] and is utilized to construct the prediction model in this research.

Training any learning-based model for sequence prediction could be regarded as an optimization problem, and the aim is to minimize a loss function. Mean square error (MSE) is adopted as the loss function in this research. To optimize the loss function, a stochastic gradient descent method called Adam optimizer is utilized as the optimizer to train the stacked LSTM model. Compared with other optimizers, it requires little memory space, has high implementation efficiency, and a small number of hyperparameters [37].

#### D. THE PROPOSED COMBINED STL-EXTRA-TREES-LSTM METHOD

Based on the above-mentioned theories, a novel hybrid model for identifying and predicting the displacement time series is proposed. FIGURE 2 shows the flow chart of the proposed

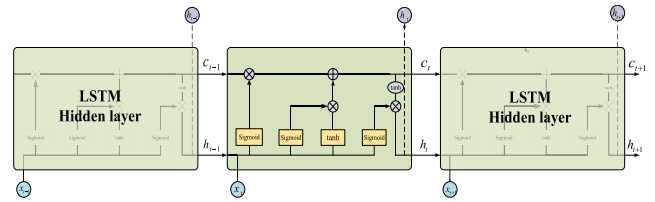


FIGURE 1. The hidden layer of LSTM architecture.

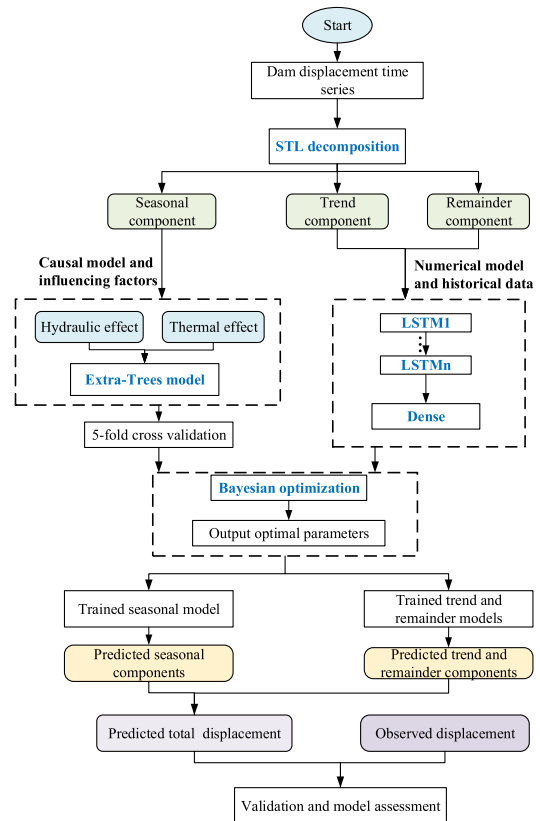


FIGURE 2. Flow chart of the proposed combined model.

novel hybrid model. The specific implementation of the proposed model is shown as follows:

**Step1: Decomposition.** The original dam displacement time series  $X_t$  is decomposed into three components including seasonal  $S_t$ , trend  $T_t$ , and remainder  $R_t$  components by using the STL method.

**Step2: Seasonal component prediction.** Extra-trees models are utilized to construct the prediction model for extracted seasonal components based on the influencing factors. The optimal parameters of the model are determined by the Bayesian optimization method and 5-fold-cross-validation.

**Step3: Trend and remainder components prediction.** Stacked LSTM NNs containing multiple hidden LSTM layers are utilized to predict the extracted trend and remainder components. The number of hidden LSTM layers and the parameter of each LSTM layer is determined by the Bayesian optimization.



FIGURE 3. The design drawing of the project in the case study.

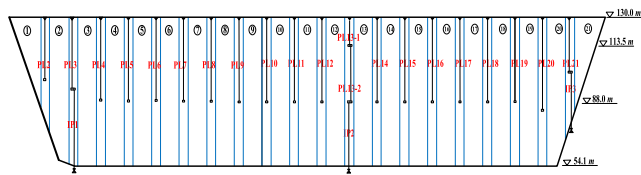


FIGURE 4. The layout of the plumb line monitoring system.

**Step4: Summation.** The total predicted dam displacement is obtained by adding such three predicted components. To verify the feasibility of the proposed method, various state-of-the-art prediction methods are used as comparison methods.

III. RESEARCH DESIGN

In this research, the research design of the case study, input variable selection, evaluation indicators, and model implementation are provided in detail.

A. CASE STUDY

A multiple-arch concrete dam is located in Anhui Province, China. It is composed of 20 arches and 21 buttresses. The maximum dam height is 75.90m, the dam crest length is 510m, the dam crest elevation is 129.96m, and the normal storage level is 125.56m.

To ensure the safe operation of the dam, the safety monitoring system was rebuilt in 2002-2005. Various monitoring projects are arranged at the surface and interior of the dam body, such as determination of dam horizontal and vertical displacement, crack opening, seepage flow, and groundwater level. Among them, the horizontal displacement monitoring project is measured by plumb lines (PL) and inverted plumb lines (IP). There are 20 PLs placed in No.2-5, No.7-12, No.14-21 buttresses, and 2 IPs placed in No.13 buttress. FIGURE 3 and FIGURE 4 show the design drawing of the project and the specific layout of the plumb line monitoring system.

Due to the thin dam body, temperature variations have a significant influence on the deformation in this case. A large number of thermometers are implemented in the dam body and foundation to monitor the temperature variations, such

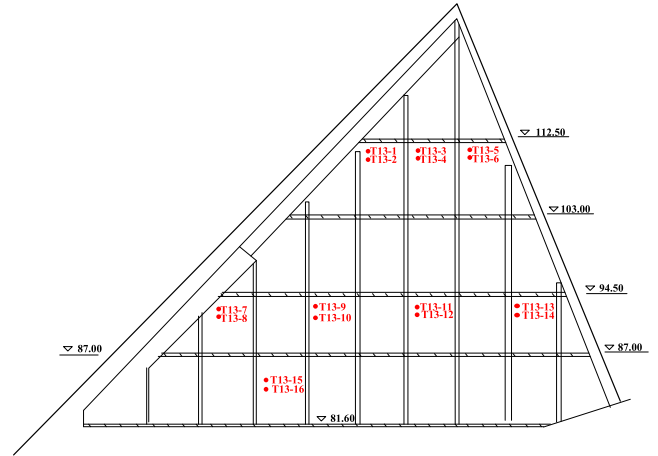


FIGURE 5. The layout of thermometers in No.13 buttress.

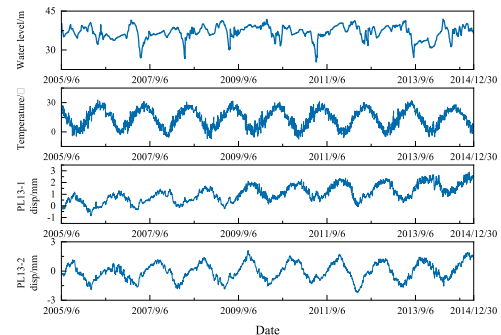


FIGURE 6. Hydrographs of water level, air temperature, and measured displacement.

as air, water, and concrete temperatures. FIGURE 5 shows the layout of thermometers in the No.13 buttress.

Compared with left and right bank slope sections and overflow dam sections, riverbed non-overflow dam sections are less affected by external environmental interference. Also, its deformation is more affected by water levels and temperatures. In this study, monitoring points PL13-1(elevation 129m) and PL13-2 (elevation 86m) placed in No.13 buttress are used for analysis. FIGURE 6 shows the hydrographs of environmental variables (reservoir level, air temperature) and measured displacement. Measured displacement time series consists of a total of 3355 sets for both two monitoring points. The first 3000 observations from September 16, 2005 to December 18, 2013 are used to construct the model and determine the unknown parameters. The rest 355 observations from December 19, 2013 to December 30, 2014 are used to test the prediction capability of the model.

B. INPUT VARIABLES SELECTION

In this study, the hydrostatic-thermal-time(HTT) statistical model is used to intercept measured dam displacement [38]. In the HTT model, the hydraulic effect is considered to be related to the reservoir level  $H$ . It could be modeled in the form of a four-order polynomial with terms of  $H, H^2, H^3,$

**TABLE 1. Contribution Ratios of each PC for temperatures.**

PCs	Air temperature	Water temperature	Concrete temperature
PC1	97.48%	98.23%	97.43%
PC2	1.99%	1.54%	2.47%

and  $H^4$ . The thermal effect is represented by the recorded thermometer data. Considering a large number of thermometers embedded in the dam, principal component analysis is utilized to extract the principal components (PCs) of the original thermometer data to represent the thermal effect. Table 1 shows the contributions ratios of each PC for air, water, and concrete temperatures.

It can be inferred that for both three temperatures, the cumulative contribution ratios of PC1 and PC2 exceeds 95%. This indicates that the first two PCs can represent the main information of the original thermometer data. Thus, a total of 6 temperature variables are selected as inputs to construct the HTT model.

The time-varying effect refers to the evolution of the dam over time and can be modeled as  $\theta$ , and  $\ln(\theta)$ . Finally, a total of 12 influencing factors are determined as input variables for predicting dam displacement, including 4 hydraulic variables, 6 temperature variables, and 2 time-varying variables. The target output of the HTT model is dam displacement.

### C. EVALUATION INDICATORS

To compare the prediction performance of the proposed and the benchmark methods, various indices, mean absolute error (MAE), MSE, and coefficient of determination ( $R^2$ ), are used as evaluation indicators. These indices are formulated as follows:

$$MAE(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (9)$$

$$MSE(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|^2 \quad (10)$$

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (11)$$

where  $\hat{y}_i$  is the predicted value of the  $i$ -th sample,  $y_i$  is the corresponding true values, and  $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ .

Apart from accuracy evaluation, the Diebold-Mariano (DM) test [39] is introduced to compare the statistical significance of two prediction models. The DM test is utilized to test the null hypothesis of equality of prediction mean square errors, and the null hypothesis is that the two predictions have the same accuracy. The DM statistic test can be formulated as

$$DM = \frac{\bar{d}}{\sqrt{[\gamma_0 + 2 \sum_{k=1}^{h-1} \gamma_k] / n}} \quad (12)$$

where  $\bar{d}$  is the sample mean loss differential,  $\gamma_t = \text{cov}(d_{t+1}, d_t)$ ,  $\gamma_0$  is the variance of  $d_t$ ,  $n$  is the length of the prediction series.

Since the DM test tend to reject the null hypothesis for small samples, a modified DM test called HLN-DM test is proposed, and the corrected statistic is obtained as

$$HLN - DM = \sqrt{\frac{T + 1 - 2h + h(h-1)}{T}} DM \quad (13)$$

where  $T$  is the size of the samples, and  $h$  is the number of steps ahead, the value of which is determined by using  $h = n^{1/3} + 1$ .

### D. MODEL IMPLEMENTATION

In this study, the STL is utilized to decompose the dam displacement time series into seasonal, trend, and remainder components. The implementation of STL is by using Python *statsmodel* module. To execute the *stl* function, the number of observations in each cycle needs to be predetermined. Since the dam displacement time series change in an annual cycle, the frequency of the series is selected as 365.

To compare the feasibility of the proposed model, some state-of-the-art methods are introduced as the benchmark methods. These methods include extra-trees, MLR, Gaussian process regression (GPR), multilayer perceptron (MLP), support vector regression (SVR), and RF.

The modeling of extra-trees, GPR, MLP, SVR, and RF for comparison is based on the HTT models and influencing factors. These ML models were implemented using the of the python *sklearn* module. Three hyperparameters of the extra-trees and the RF models (the number of estimators  $n_{estimators}$ , the maximum depth  $max\_depth$ , and the minimum number of samples to split  $min\_samples\_split$ ) made up a three-dimension space. The hyperparameters of the SVR models including the regularization parameter  $C$  and the kernel coefficient  $gamma$  made up a two-dimension space. The hyperparameters of the MLP models including the number of neurons in the hidden layers  $n_{neurons}$  made up a one-dimension space. The best parameters were selected and determined in terms of prediction accuracy (e.g., MSE) according to the Bayesian hyper-parameter optimization and 5-fold-cross-validation. TABLE 2 shows the final determined hyperparameters of the extra-trees model. The kernel of GPR was set as a square exponential kernel, and the unknown hyperparameters of the selected kernel were estimated by maximizing the log-marginal-likelihood based on the optimizer.

The modeling of stacked LSTM is based on the numerical models and historical observation data. The LSTM models were implemented by using Python *tensorflow* module. It should be noted there is no theoretical knowledge to predetermine the structure of the LSTM neural network for specific data, and the practical method is to determine the hyperparameters by the trial-and-error method. To solve the above problems, in this study, Bayesian optimization is utilized to fine-tune these parameters of stacked LSTM models. The number of layers and the neurons in each layer determines

**TABLE 2.** The parameters of the extra-trees model for seasonal component and total displacement.

Hyperparameters	Domain	PL13-1		PL13-2	
		Seasonal	Total	Seasonal	Total
<i>n_estimators</i>	[10,500]	457	112	400	170
<i>min_samples_split</i>	[2,30]	30	30	8	30
<i>max_depth</i>	[1,30]	30	30	8	30

**TABLE 3.** The parameters of comparison models.

	Hyperparameters	Domain	PL13-1	PL13-2
MLP	<i>n_neurons</i>	[10,500]	382	274
SVR	<i>C</i>	$[-10^4, 10^4]$	25.594	16.296
	<i>gamma</i>	[0.0001,0.1]	0.021	0.0026
RF	<i>n_estimators</i>	[10,500]	273	499
	<i>min_samples_split</i>	[2,30]	30	4
	<i>max_depth</i>	[1,30]	30	29

**TABLE 4.** The DETERMINED parameters of LSTM.

Hyperparameter	PL13-1		PL13-2	
	Trend	Remainder	Trend	Remainder
Number of neurons	85,115	82,105,120	80,108,120	100
Number of layers	2	3	3	1
Dropout	0.1	0.2	0.1	0.1

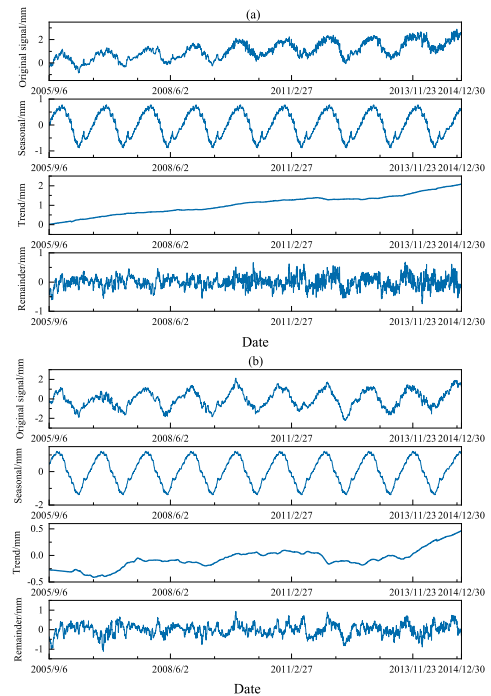
**TABLE 5.** The DETERMINED parameters of LSTM for comparison.

Hyperparameter	PL13-1	PL13-2
Number of neurons	60	55
Number of layers	1	1
Dropout	0.1	0.1

the performance of the deep learning model. Also, a dropout with each layer was introduced to regularize the network and prevent overfitting. Three parameters including the number of layers, the number of neurons in each layer, and the recurrent dropout rates made up a three-dimension space, and the domains of the parameters are set as [20], [120], {1, 2, 3}, and {0.1, 0.2, 0.3}. Then the Bayesian optimization is utilized to determine the optimal hyperparameters for the LSTM model in terms of prediction accuracy (e.g., MSE). TABLE 4 and TABLE 5 show the final determined parameters of the stacked LSTM model. The stacked LSTM networks were trained for 100 iterations, and the batch size was set as 100.

**IV. RESULTS AND DISCUSSIONS**

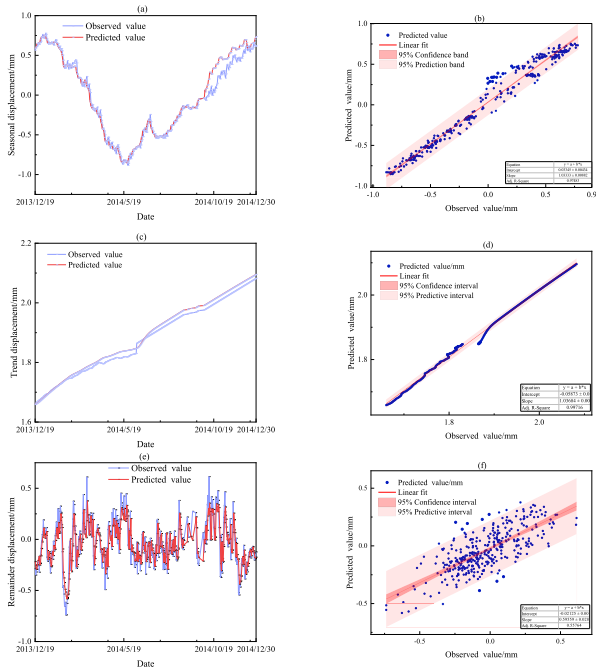
FIGURE 7 shows the decomposition results of dam displacement time series from PL13-1 and PL13-2 by using STL. All of the seasonal components of both two dam displacement time series show a one-year cycle, and the preliminary analysis of this regularity change is caused by hydrostatic loads, temperatures, and other cyclical factors. Moreover, a constant growth trend can be seen in the trend components of PL13-1, while the trend component of PL13-2 shows



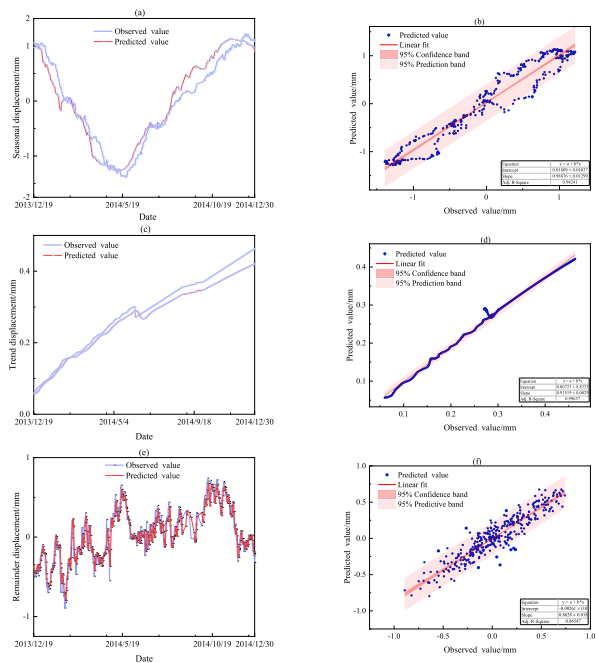
**FIGURE 7.** The decomposition results of time series from both two monitoring points via STL (a) PL13-1 (b) PL13-2.

much more gentle growth. The preliminary analysis of the trend component is caused by the property degradation of dam material and inherent rheological property. However, caused by the restraint of the dam foundation, this phenomenon is much more observable in the upper dam body than the dam body near the foundation. Since the elevation of PL13-1(129m) is higher than PL13-2(86m), the deformation time-varying effect reflected from PL13-1 is much more observable. The remainder component is the residual variability after the seasonal and trend components have been eliminated. It can be seen the remainder components of both two monitoring points are typical noise series and present a high degree of complexity of irregularity.

After the decomposition, extra-trees models are utilized to construct prediction models for seasonal components based on influencing factors, whereas stacked LSTM models are utilized to predict trend and remainder components based on the historical observation data. Taking monitoring point PL13-1 and PL13-2 as an example, the prediction results and the linear regression analysis of each component are shown in FIGURE 8 and FIGURE 9. FIGURE 8 and FIGURE 9(a) and (b) display extra-trees model achieve high accuracy when predicting the seasonal component, and the correlation coefficients between the predictions and the observations are 0.975 and 0.942. Most predictions are within the 95% prediction interval, which indicates extra-trees models have a good prediction ability for seasonal components. Once again, FIGURE 8 and FIGURE 9(c)and (d) display the stacked LSTM models show excellent performance in trend component prediction, and the correlation coefficients

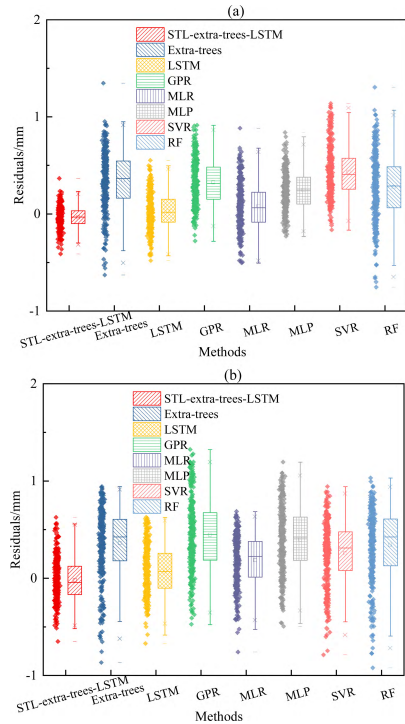


**FIGURE 8.** Predicted components by the proposed model and the linear regression analysis for PL13-1.



**FIGURE 9.** Predicted components by the proposed model and the linear regression analysis for PL13-2.

reach 0.997 and 0.996. Although the remainder component is a noisy series with the characteristics of uncertainty and irregularity, the stacked LSTM model still obtains good performance. FIGURE 8 and FIGURE 9(e) and (f) show the predicted values of the stacked LSTM model are consistent with the observed values, and most of the predictions are within the 95% prediction band.



**FIGURE 10.** Boxplot of residuals for the proposed model and other benchmark methods (a) PL13-1 (b) PL13-2.

Then the predicted results of seasonal, trend, and remainder components are aggregated as the final predicted displacement. To detect the outliers between the prediction values and the observations, a standardized method of displaying the distribution of data boxplot is proposed. The drawing of the boxplot is purely based on the observations without relying on any assumption and has strong robustness on the outliers. FIGURE 10 displays the residual boxplot for the proposed model and other benchmark models of both two monitoring points. The residual distribution of the proposed model is within 1.5 interquartile range (IQR), only a small number of mild outliers exist, and there are not any extreme outliers. Moreover, the median values of outliers of the proposed model are less than those of the other four benchmark methods, which indicates the prediction performance of the proposed model is outstanding and stable.

To quantitatively evaluate the predictive performance of the proposed and benchmark methods, MSE, MAE, and  $R^2$  of the five models are calculated for both two monitoring points. Table 6 shows the prediction performance of the proposed model and other benchmark methods. It can be inferred that the proposed STL-extra-trees-LSTM model achieves better performance than a single prediction model in all evaluation indicators. This indicates applying STL to preprocess the displacement time series, and then select the prediction models that fit the characteristics of the component itself is effective. From Table 6, it can be seen all single ML (extra-trees/GPR/MLP/SVR/RF) models do not achieve satisfactory performance on both two monitoring points, especially for



**TABLE 6. Prediction accuracy assessment of different models.**

Indicators Models	MSE		MAE		R <sup>2</sup>	
	PL13-1	PL13-2	PL13-1	PL13-2	PL13-1	PL13-2
<b>Proposed</b>	<b>0.017</b>	<b>0.047</b>	<b>0.103</b>	<b>0.177</b>	<b>0.963</b>	<b>0.941</b>
Extra-trees	0.503	0.253	0.646	0.451	-0.117	0.684
LSTM	0.063	0.066	0.202	0.21	0.861	0.917
GPR	0.646	0.333	0.719	0.48	-0.433	0.584
MLR	0.384	0.098	0.569	0.267	0.147	0.877
MLP	0.294	0.268	0.456	0.44	0.347	0.665
SVR	0.297	0.164	0.495	0.353	0.34	0.796
RF	0.584	0.253	0.693	0.442	-0.295	0.683

**TABLE 7. HLN-DM test results for different models for PL13-1.**

	Extra-trees	LSTM	GPR	MLR	MLP	SVR	RF
<b>Proposed</b>	<b>-6.78</b>	<b>-7.61</b>	<b>-7.12</b>	<b>-4.31</b>	<b>-7.6</b>	<b>-5.51</b>	<b>-4.62</b>
Extra-trees		6.00	1.59	6.32	4.02	-0.91	0.84
LSTM			-5.72	-2.36	-5.43	-4.95	-3.97
GPR				4.44	4.5	-3.29	-0.82
MLR					-2.9	-4.4	-3.82
MLP						-4.32	-2.37
SVR							1.04
RF							

PL13-1. It should be noted that even the MLR model even achieves a better performance than the ML methods. It is mainly because the time series of PL13-1 exhibits a growing trend over time. For this kind of time series, the ML methods based on the causal models cannot achieve high performance in prediction. Moreover, deep learning methods (e.g. LSTM) show a strong capability for predicting dam deformation. It is conceivable deep learning methods can mine the potential information between data through specific learning mechanisms. The experimental results verify the feasibility and the advancement of the proposed STL-extra-trees-LSTM model.

To further intuitively compare the performance of the predictive models, the HLN-DM test is introduced to test the statistical significance of the proposed and other benchmark models. Table 7 and Table 8 show the results of the HLN-DM statistics test for different models.

The following inference can be deduced from the results in Table 7 and Table 8

First, the proposed STL-extra-trees-LSTM model outperforms benchmark models (e.g., extra trees, LSTM) for monitoring point PL13-1 at a 1% level of statistical significance. Dam deformation evolution is a dynamic non-linear process with the characteristics of diversity, time-varying, and uncertainty. The single ML method cannot achieve high prediction performance. Therefore, it is desirable to utilize signal pre-processing methods to decompose the time series into serval

**TABLE 8. HLN-DM test results for different models for PL13-2.**

	Extra-trees	LSTM	GPR	MLR	MLP	SVR	RF
<b>Proposed</b>	<b>-5.95</b>	<b>-1.29</b>	<b>-3.92</b>	<b>-3.62</b>	<b>-4.15</b>	<b>-4.25</b>	<b>-5.74</b>
Extra-trees		5.71	-1.25	4.68	-0.34	3.67	-0.03
LSTM			-4.23	-2.45	-4.68	-4.45	-5.71
GPR				3.43	2.77	3.21	1.20
MLR					-3.51	-2.76	-4.74
MLP						3.05	0.31
SVR							-3.38
RF							

sub-sequences, and then the prediction model that fits the characteristics of the sub-sequence is implemented.

Second, the LSTM models are superior to the other ML methods (extra-trees/GPR/MLP/SVR/RF) model at a 1% level of statistical significance. This indicates deep learning methods achieve better performance and higher advantages in predicting the displacement time series compared with ML methods.

Third, the difference between Extra-trees and GPR models is not significant, whereas the MLR model is superior to the other two ML models. This indicates single ML methods cannot obtain satisfactory results on predicting displacement time series, especially for the long sequences in the test set.

In further validating the effectiveness and generalization of the proposed combined model, two monitoring points PL7 and PL17 from the left and right bank slope sections are used as reference points. Also, the displacement data of the other four PL monitoring points including PL2, PL3, PL20, and PL21, and the three IP monitoring points including IP1, IP2, and IP3 are utilized to verify the generalization and applicability of the proposed model. The prediction results and quantitative assessment indicators are shown in **Appendix1**. FIGURE 11 shows the decomposition result of PL7 and PL17 by using the STL method. Then the obtained seasonal, trend, and remainder components are predicted by the proposed model and then aggregated to obtain the total prediction results. The process of determining the hyperparameters can refer to the previous process for PL13-1 and PL13-2. FIGURE 11 and FIGURE 12 show the prediction results and the linear regression analysis of each component for monitoring points PL7 and PL17. Table 9 shows the prediction performance assessment of the proposed model and other benchmark methods.

According to FIGURE 12, FIGURE 13, and Table 9, it can be observed that :

- It is confirmed again the proposed STL-extra-trees-LSTM exhibits better prediction accuracy and stability than the benchmark methods.
- A single ML/statistical method shows poor performance in predicting dam displacement with high accuracy, especially for the long prediction sequence.

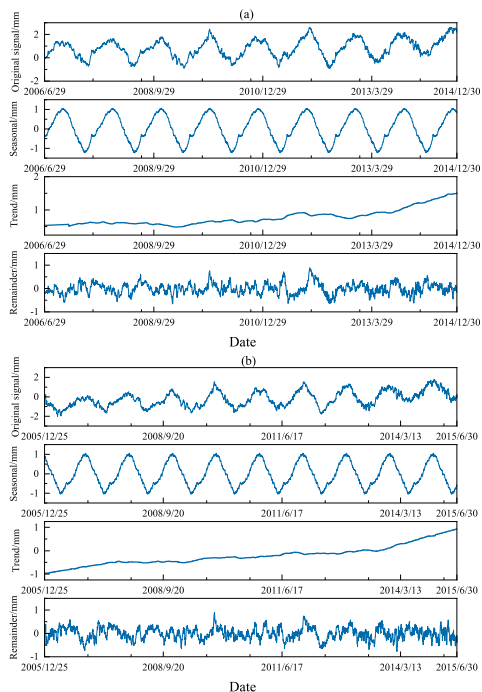


FIGURE 11. The decomposition results of time series from both two monitoring points via STL (a) PL7 (b)PL17.

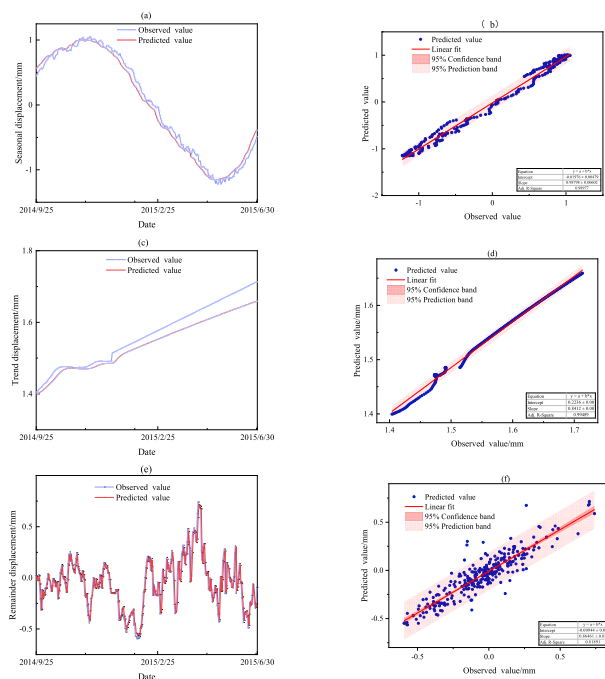


FIGURE 12. Predicted components by the proposed model and the linear regression analysis for PL7.

- Also, the deep learning methods show a good capability of predicting dam deformation accurately.
- Applying pre-processing methods to decompose a time series into several sub-sequences is desirable to improve the performance of the prediction model.

From the above comparative analysis of experimental results, it can be concluded the proposed combined model is

TABLE 9. Prediction accuracy assessment of different models.

Models	Indicators	MSE		MAE		R <sup>2</sup>	
		PL7	PL17	PL7	PL17	PL7	PL17
Proposed		0.019	0.017	0.106	0.103	0.962	0.963
Extra-trees		0.262	0.503	0.433	0.646	0.469	-0.117
LSTM		0.118	0.063	0.285	0.202	0.762	0.861
GPR		0.522	0.646	0.621	0.719	-0.057	-0.433
MLR		0.248	0.384	0.419	0.569	0.498	0.147
MLP		0.102	0.294	0.237	0.456	0.794	0.347
SVR		0.239	0.297	0.418	0.495	0.517	0.34
RF		0.271	0.584	0.433	0.693	0.452	-0.295

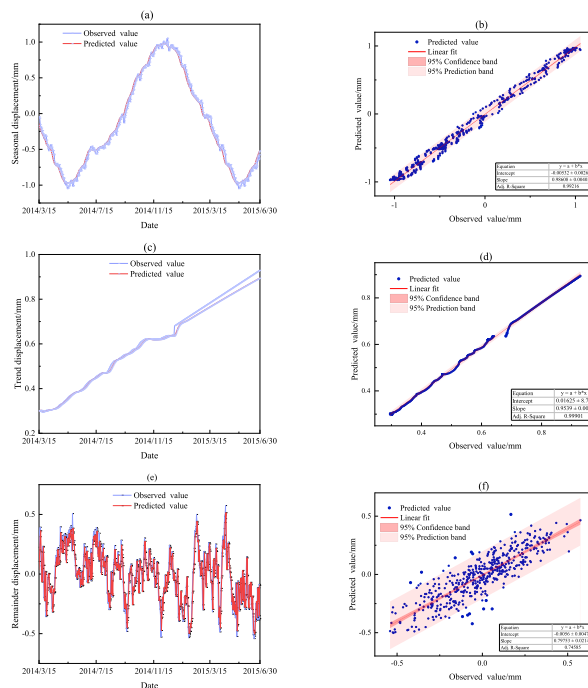
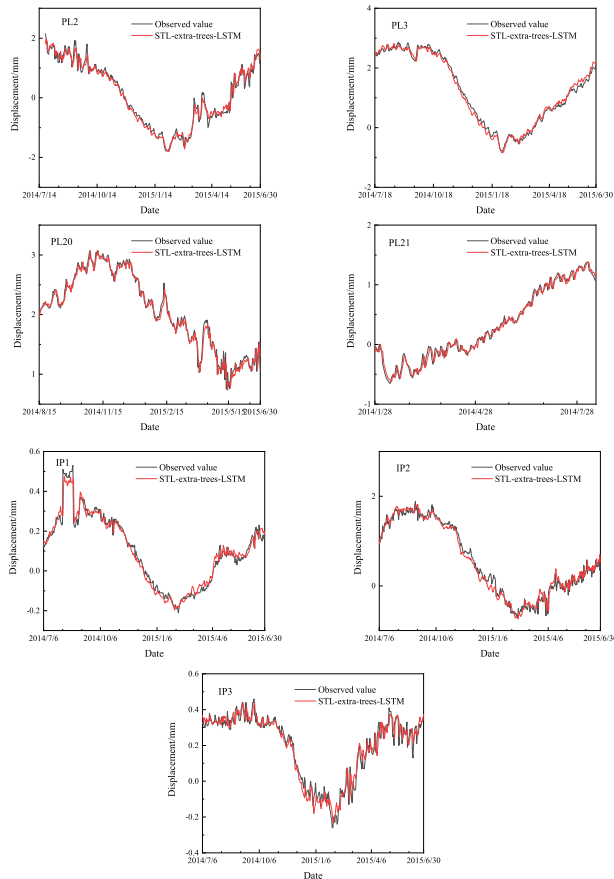


FIGURE 13. Predicted components by the proposed model and the linear regression analysis for PL17.

TABLE 10. Evaluation indicators for the prediction results of seven monitoring points.

	MSE	MAE	R <sup>2</sup>
PL2	0.027	0.123	0.976
PL3	0.012	0.094	0.991
PL20	0.009	0.067	0.978
PL21	0.005	0.055	0.987
IP1	0.001	0.022	0.968
IP2	0.016	0.099	0.975
IP3	0.002	0.033	0.942

generally the best-performing model for the displacement time series, in terms of all various quantitative evaluation indicators. Also, the results of the HLN-DM test further illustrate the proposed STL-extra-trees-LSTM model outperforms the other benchmark methods at a 1% level of statistical significance. The experimental results show that utilizing



**FIGURE 14.** The prediction result of seven monitoring points by the proposed model.

STL methods to decompose the displacement time series into various additive components, and then selecting the ML and deep learning methods to predict each component separately can enhance both the prediction capability and stability significantly.

## V. CONCLUSIONS

In this research, a novel combined model for predicting dam displacement time series with time-varying effect was proposed. In the proposed model, STL decomposition is utilized to decompose the original displacement time series into three components: seasonal, trend, and remainder components. Preliminary analysis shows seasonal component changes regularly and is mainly affected by periodic loads such as hydrostatic loads and temperature variations. Trend and remainder components show rapid growth and irregular changes, respectively, which are influenced by the time-varying effect and uncertainty factors, respectively.

By analyzing the characteristics of each component, the prediction models that fit the characteristics of the component itself is selected and applied to predict dam displacement. Extra-trees models are used to construct prediction models for the seasonal components based on the causal model and influencing factors, while the stacked LSTM

models are utilized to predict the trend and remainder components based on the numerical model and historical observation data. The predicted results of the three components are superimposed to generate the final prediction result, and four state-of-the-art methods are selected as comparison methods.

Three quantitative evaluation indicators including MSE, MAE,  $R^2$ , and a statistical test method HLN-DM test are used to verify the effectiveness of the proposed model. The experimental results show that the proposed model achieves great performance both in prediction accuracy and stability. Besides, it indicates that STL is an effective and efficient method for decomposing the dam displacement with the time-varying effect. Selecting the suitable prediction model based on the decomposition results of STL can effectively enhance both the prediction accuracy and stability.

## A. LIMITATION

The proposed method still has certain limitations in some aspects. The first limitation is about the selection of the optimal hyperparameters of LSTM models. It should be noted there is no theoretical knowledge to predetermine the structure of the LSTM neural network for specific data, and the practical method is to determine the hyperparameters by the trial-and-error method. In this research, the Bayesian optimization is introduced to determine the parameters of stacked LSTM models. However, the process is complicated and time-consuming when the number of layers of LSTM is large. Therefore, in future research, we need to find a better and simplified optimization method to determine the model parameters of LSTM. Moreover, the dam safety monitoring system is usually arranged with multiple measuring points, but the proposed model cannot handle the modeling of multiple monitoring points simultaneously. Concretely, the model will be constructed and trained separately multiple times, without considering the correlation between the multiple monitoring points.

## B. FUTURE WORK

Dam behavior prediction a fundamental component of dam structural health monitoring. As the intuitive monitoring indicators, deformation is often utilized to evaluate the safety status of a dam in service. It is desirable to develop advanced and effective methods for dam deformation prediction. If we can establish strong prediction models according to the prototype observation data, and then the structural degradation of the dam and its trend can be detected in time.

In our further work, advanced data mining techniques such as transfer learning, deep learning will be introduced into the field of dam safety monitoring to fully mine the dam structure change information to avoid dam crash accident.

## APPENDIX 1 THE PREDICTION RESULTS OF THE PROPOSED STL-EXTRA-TREES-LSTM MODEL FOR OTHER SEVEN MONITORING POINTS

See Fig. 14 and Table 10.

## REFERENCES

- [1] H. Su, Z. Wen, X. Sun, and M. Yang, "Time-varying identification model for dam behavior considering structural reinforcement," *Struct. Saf.*, vol. 57, pp. 1–7, Nov. 2015, doi: [10.1016/j.strusafe.2015.07.002](https://doi.org/10.1016/j.strusafe.2015.07.002).
- [2] H. Su, X. Li, B. Yang, and Z. Wen, "Wavelet support vector machine-based prediction model of dam deformation," *Mech. Syst. Signal Process.*, vol. 110, pp. 412–427, Sep. 2018, doi: [10.1016/j.ymsp.2018.03.022](https://doi.org/10.1016/j.ymsp.2018.03.022).
- [3] D. Zhong, Y. Sun, and M. Li, "Dam break threshold value and risk probability assessment for an Earth dam," *Natural Hazards*, vol. 59, no. 1, pp. 129–147, Oct. 2011, doi: [10.1007/s11069-011-9743-6](https://doi.org/10.1007/s11069-011-9743-6).
- [4] B. Dai, C. Gu, E. Zhao, and X. Qin, "Statistical model optimized random forest regression model for concrete dam deformation monitoring," *Struct. Control Health Monitor.*, vol. 25, no. 6, p. e2170, Jun. 2018, doi: [10.1002/stc.2170](https://doi.org/10.1002/stc.2170).
- [5] C. Lin, T. Li, S. Chen, X. Liu, C. Lin, and S. Liang, "Gaussian process regression-based forecasting model of dam deformation," *Neural Comput. Appl.*, vol. 31, no. 12, pp. 8503–8518, Dec. 2019, doi: [10.1007/s00521-019-04375-7](https://doi.org/10.1007/s00521-019-04375-7).
- [6] F. Salazar, R. Morán, M. Á. Toledo, and E. Oñate, "Data-based models for the prediction of dam behaviour: A review and some methodological considerations," *Arch. Comput. Methods Eng.*, vol. 24, no. 1, pp. 1–21, Jan. 2017, doi: [10.1007/s11831-015-9157-9](https://doi.org/10.1007/s11831-015-9157-9).
- [7] K.-T. Bui, D. Tien Bui, J. Zou, C. Van Doan, and I. Revhaug, "A novel hybrid artificial intelligent approach based on neural fuzzy inference model and particle swarm optimization for horizontal displacement modeling of hydropower dam," *Neural Comput. Appl.*, vol. 29, no. 12, pp. 1495–1506, Jun. 2018, doi: [10.1007/s00521-016-2666-0](https://doi.org/10.1007/s00521-016-2666-0).
- [8] Y. Huang, L. Yang, S. Liu, and G. Wang, "Multi-step wind speed forecasting based on ensemble empirical mode decomposition, long short term memory network and error correction strategy," *Energies*, vol. 12, no. 10, p. 1822, May 2019, doi: [10.3390/en12101822](https://doi.org/10.3390/en12101822).
- [9] A. Bruce, D. Donoho, and H.-Y. Gao, "Wavelet analysis [for signal processing]," *IEEE Spectr.*, vol. 33, no. 10, pp. 26–35, Oct. 1996, doi: [10.1109/6.540087](https://doi.org/10.1109/6.540087).
- [10] N. Rehman and D. P. Mandic, "Multivariate empirical mode decomposition," *Proc. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 466, no. 2117, pp. 1291–1302, May 2010, doi: [10.1098/rspa.2009.0502](https://doi.org/10.1098/rspa.2009.0502).
- [11] H. Ahumada and M. Garegnani, "Hodrick-Prescott filter in practice," in *Proc. Jornadas Econ. Monet. Int.*, 1999, doi: [10.1.1.121.3365](https://doi.org/10.1.1.121.3365).
- [12] T. H. E. State, B. Of, and M. P. Decision, "State bank of Pakistan," vol. 1956, pp. 9–10, Feb. 2014.
- [13] X. Zhu, Q. Xu, M. Tang, H. Li, and F. Liu, "A hybrid machine learning and computing model for forecasting displacement of multifactor-induced landslides," *Neural Comput. Appl.*, vol. 30, no. 12, pp. 3825–3835, Dec. 2018, doi: [10.1007/s00521-017-2968-x](https://doi.org/10.1007/s00521-017-2968-x).
- [14] R. Cleveland, W. Cleveland, J. McRae, and I. Terpenning, "STL: A seasonal-trend decomposition procedure based on Loess (with discussion)," *J. Off. Stat.*, vol. 6, no. 1, p. 3, 1990.
- [15] T. Xiong, C. Li, and Y. Bao, "Seasonal forecasting of agricultural commodity price using a hybrid STL and ELM method: Evidence from the vegetable market in China," *Neurocomputing*, vol. 275, pp. 2831–2844, Jan. 2018, doi: [10.1016/j.neucom.2017.11.053](https://doi.org/10.1016/j.neucom.2017.11.053).
- [16] M. Li, Y. Shen, Q. Ren, and H. Li, "A new distributed time series evolution prediction model for dam deformation based on constituent elements," *Adv. Eng. Informat.*, vol. 39, pp. 41–52, Jan. 2019, doi: [10.1016/j.aei.2018.11.006](https://doi.org/10.1016/j.aei.2018.11.006).
- [17] M. Theodosiou, "Forecasting monthly and quarterly time series using STL decomposition," *Int. J. Forecasting*, vol. 27, no. 4, pp. 1178–1195, Oct. 2011, doi: [10.1016/j.ijforecast.2010.11.002](https://doi.org/10.1016/j.ijforecast.2010.11.002).
- [18] J. Mata, A. Tavares de Castro, and J. Sá da Costa, "Constructing statistical models for arch dam deformation," *Struct. Control Health Monitor.*, vol. 21, no. 3, pp. 423–437, Mar. 2014, doi: [10.1002/stc.1575](https://doi.org/10.1002/stc.1575).
- [19] S. Wang, C. Gu, and T. Bao, "Observed displacement data-based identification method of deformation time-varying effect of high concrete dams," *Sci. China Technol. Sci.*, vol. 61, no. 6, pp. 906–915, Jun. 2018, doi: [10.1007/s11431-016-9088-9](https://doi.org/10.1007/s11431-016-9088-9).
- [20] B. Wei, D. Yuan, H. Li, and Z. Xu, "Combination forecast model for concrete dam displacement considering residual correction," *Struct. Health Monitor.*, vol. 18, no. 1, pp. 232–244, Jan. 2019, doi: [10.1177/1475921717748608](https://doi.org/10.1177/1475921717748608).
- [21] L. S. Ribeiro, V. E. Wilhelm, É. F. Faria, J. M. Correa, and A. C. P. dos Santos, "A comparative analysis of long-term concrete deformation models of a buttress dam," *Eng. Struct.*, vol. 193, pp. 301–307, Aug. 2019, doi: [10.1016/j.engstruct.2019.05.043](https://doi.org/10.1016/j.engstruct.2019.05.043).
- [22] Z.-H. Zhou, "Ensemble learning," in *Encyclopedia Biometrics*, S. Z. Li and A. K. Jain, Eds. Boston, MA, USA: Springer, 2015, pp. 411–416.
- [23] K. Chen, J. Jiang, F. Zheng, and K. Chen, "A novel data-driven approach for residential electricity consumption prediction based on ensemble learning," *Energy*, vol. 150, pp. 49–60, May 2018, doi: [10.1016/j.energy.2018.02.028](https://doi.org/10.1016/j.energy.2018.02.028).
- [24] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Mach. Learn.*, vol. 63, no. 1, pp. 3–42, Apr. 2006, doi: [10.1007/s10994-006-6226-1](https://doi.org/10.1007/s10994-006-6226-1).
- [25] E. Eslami, A. K. Salman, Y. Choi, A. Sayeed, and Y. Lops, "A data ensemble approach for real-time air quality forecasting using extremely randomized trees and deep neural networks," *Neural Comput. Appl.*, vol. 6, Mar. 2019, doi: [10.1007/s00521-019-04287-6](https://doi.org/10.1007/s00521-019-04287-6).
- [26] M. T. Uddin and M. A. Uddiny, "Human activity recognition from wearable sensors using extremely randomized trees," in *Proc. Int. Conf. Electr. Eng. Inf. Commun. Technol. (ICEEICT)*, May 2015, pp. 21–23, doi: [10.1109/ICEEICT.2015.7307384](https://doi.org/10.1109/ICEEICT.2015.7307384).
- [27] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: [10.1038/nature14539](https://doi.org/10.1038/nature14539).
- [28] J. Rico, J. Barateiro, J. Mata, A. Antunes, and E. Cardoso, *Applying Advanced Data Analytics and Machine Learning to Enhance the Safety Control of Dams*. Springer, 2019.
- [29] X. Qu, J. Yang, and M. Chang, "A deep learning model for concrete dam deformation prediction based on RS-LSTM," *J. Sensors*, vol. 2019, pp. 1–14, Oct. 2019, doi: [10.1155/2019/4581672](https://doi.org/10.1155/2019/4581672).
- [30] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: [10.1162/neco.1997.9.8.1735](https://doi.org/10.1162/neco.1997.9.8.1735).
- [31] Y. Li and H. Cao, "Prediction for tourism flow based on LSTM neural network," *Procedia Comput. Sci.*, vol. 129, pp. 277–283, 2018, doi: [10.1016/j.procs.2018.03.076](https://doi.org/10.1016/j.procs.2018.03.076).
- [32] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, "Long short-term memory neural network for traffic speed prediction using remote microwave sensor data," *Transp. Res. C, Emerg. Technol.*, vol. 54, pp. 187–197, May 2015, doi: [10.1016/j.trc.2015.03.014](https://doi.org/10.1016/j.trc.2015.03.014).
- [33] W. Liu, J. Pan, Y. Ren, Z. Wu, and J. Wang, "Coupling prediction model for long term displacements of arch dams based on long short term memory network," *Struct. Control Heal. Monit.*, vol. 2019, pp. 1–15, Mar. 2020, doi: [10.1002/stc.2548](https://doi.org/10.1002/stc.2548).
- [34] J. Li, H. Chen, T. Zhou, and X. Li, "Tailings pond risk prediction using long short-term memory networks," *IEEE Access*, vol. 7, pp. 182527–182537, 2019, doi: [10.1109/ACCESS.2019.2959820](https://doi.org/10.1109/ACCESS.2019.2959820).
- [35] A. Graves, A.-R. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2013, pp. 6645–6649, doi: [10.1109/ICASSP.2013.6638947](https://doi.org/10.1109/ICASSP.2013.6638947).
- [36] Z. Cui, R. Ke, and Y. Wang, "Deep bidirectional and unidirectional LSTM recurrent neural network for network-wide traffic speed prediction," 2018, *arXiv:1801.02143*. [Online]. Available: <https://arxiv.org/abs/1801.02143>
- [37] D. P. Kingma and J. Ba, "Adam: A method for stochastic Optimization," 2014, *arXiv:1412.6980*. [Online]. Available: <https://arxiv.org/abs/1412.6980>
- [38] Y. S. Lee, A. F. Vakakis, D. M. Mcfarland, and L. A. Bergman, "Constructing statistical models for arch dam deformation," *Struct. Control Heal. Monit.*, vol. 2, pp. 742–760, Jun. 2010, doi: [10.1002/stc](https://doi.org/10.1002/stc).
- [39] D. Harvey, S. Leybourne, and P. Newbold, "Testing the equality of prediction mean squared errors," *Int. J. Forecasting*, vol. 13, no. 2, pp. 281–291, Jun. 1997, doi: [10.1016/S0169-2070\(96\)00719-4](https://doi.org/10.1016/S0169-2070(96)00719-4).



**YANGTAO LI** received the B.S. degree in water resources and hydropower engineering from Shihzei University, Xinjiang, China, in 2014. He is currently pursuing the Ph.D. degree in hydraulic structural engineering with Hohai University. His research interests include data mining and the application of machine learning technology in the field of dam safety monitoring.



**TENGFEI BAO** received the B.S., M.S., and Ph.D. degrees in hydraulic structural engineering from Hohai University, Nanjing, China, in 1998, 2000, and 2004, respectively. He is currently a Professor and a Ph.D. Tutor with Hohai University. His research interests include safety monitoring, evaluation and feedback analysis of hydraulic buildings, and application research of optical fiber sensors in hydropower engineering.



**XIAOSONG SHU** received the B.S. degree in water resources and hydropower engineering from Hohai University, Nanjing, China, in 2014, where he is currently pursuing the Ph.D. degree in hydraulic structural engineering. His research interests include safety assessment and stability analysis of dam structure.



**JIAN GONG** received the B.S. degree in agricultural water conservancy project from Hohai University, Nanjing, China, in 2013, where he is currently pursuing the Ph.D. degree in hydraulic structural engineering. His research interests include dam safety monitoring theory and application of BIM 3D visualization technology.



**KANG ZHANG** received the B.S. degree in water resources and hydropower engineering from the Hefei University of Technology, Xuancheng, China, in 2014. He is currently pursuing the Ph.D. degree in hydraulic structural engineering with Hohai University. His research interests include dam safety monitoring theory and dam safety status assessment.

...