Proc. of the 8th International Symposium on Geotechnical Safety and Risk (ISGSR)
Edited by Jinsong Huang, D.V. Griffiths, Shui-Hua Jiang, Anna Giacomini and Richard Kelly

doi: 10.3850/978-981-18-5182-7_08-014-cd

ISGSR 2022

High Arch Dam Displacement Prediction Model Based on Long Short-Term Memory Networks with Attention Mechanism

Fei Kang[1], Ben Huang[1], Junjie Li[1,2], Sizeng Zhao[1]

[1]School of Hydraulic Engineering, Faculty of Infrastructure Engineering, Dalian University of Technology, Dalian 116024, P. R. China
[2]College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing 210098, PR China
E-mail: kangfei@dlut.edu.cn, kangfei2009@163.com (F.Kang);
huangben@mail.dlut.edu.cn (B.Huang); lijunjie@dlut.edu.cn (J.J. Li).

Abstract: Dam behavior prediction that can evaluate the operational states and provide important information for safety control of dams, is an essential component of dam health monitoring. Statistical models based on regression methods have been successfully established and applied in structural health monitoring of practical engineering. However, these conventional models cannot capture the time series patterns and rely on manual parameter design. To address these problems, considering that displacement prediction is a typical time series problem, this study proposes a displacement prediction model of concrete dams using long short-term memory network (LSTM) based on deep learning techniques. The attention mechanism is adopted to capture key characteristics that influence displacement significantly. Performance of the proposed model is verified on a high arch dam. Results show that the LSTM based model outperforms the stepwise regression, back propagation neural networks, and multiple linear regression models for dam health monitoring, indicating that the proposed method is powerful and promising for arch dam behavior prediction.

Keywords: dam behavior prediction; structural health monitoring; machine learning; long short-term memory network.

1 Introduction

Dams are important infrastructure for the national economy. However, due to different reasons, dams have safety problems which affect the dam operation status and seriously threaten the safety of lives and properties of people downstream (Kang et al. 2017). Therefore, it is necessary to monitor the dam behaviors using different instruments, and identify abnormal behaviors in time (Mata et al. 2013). The observed dam displacement data which can reflect global behavior of a concrete dam is used to create a model for dam behavior prediction (Mata 2011).

Traditionally, concrete dam health monitoring models are mainly divided into deterministic, statistical and hybrid models. These mathematical models are established using displacement monitoring data, water pressure, dam temperatures and time effects to analyze the relationship between observed information and structural response of concrete dams (Kang et al. 2019a). Statistical models are established using various regression methods, including the multiple linear regression (MLR) method (Mata 2011; Salazar et al. 2015), the stepwise regression (SR) method (Xi et al. 2011) and principal component analysis regression (Yu et al. 2010), etc. However, statistical models are difficult to model complex nonlinear characteristics between input variables and dam responses, and are susceptible to the interference of uncertain variables.

Owing to the combined effects of internal and external factors, structural behavior of dams shows complex nonlinear characteristics. Recently, various machine learning methods are adopted to enrich the conventional models and establish new hybrid models for concrete dam health monitoring. Mata (2011) demonstrated that the ANN is effective to assessment concrete dam behavior. Wei et al. (2019) applied back propagation (BP) neural network to establish dam displacement monitoring model considering residual correction. The radial basis function (RBF) networks are used to establish monitoring models to simulate the concrete dam displacement (Kang et al. 2019a). However, ANN models are easy to fall into a local minimum when applied to complex time series prediction. Deep learning is a branch of the machine learning field. It has powerful learning and generalization capabilities, and has obvious advantages in implicit information mining.

Long short-term memory (LSTM) is an improved recurrent neural network (RNN) proposed by Hochreiter and Schmidhuber (1997), which is efficient to learn the long-term dependence of information. LSTM can mine hidden rules in time series data to increase the dimension of information, and was successfully used to predict water table depth (Zhang et al. 2018), short-term building energy (Li et al. 2021) and dam displacement (Qu, Yang and Chang 2019; Liu et al. 2020). However, the conventional LSTM converts the input sequence into a fixed-length vector and saves all information, which limits the model memory, and it is easy to lose information when dealing with long sequence problems. The proposed attention mechanism can solve this problem by enhancing the focus on important variables, and ignoring irrelevant information (Bahdanau et al. 2014).

In this study, a health monitoring model based on the LSTM and attention mechanism (LSTM-ATT) was established to predict concrete arch dam displacement. To prove efficiency of the proposed model, a real arch dam was used as an example to study. The performance was compared with that of MLR, SR, RBF networks, and ordinary LSTM models.

2 Mathematic theory of dam health monitoring

The HTT (Hydrostatic, Temperature, Time) model, using measured temperature data to calculate thermal effect, is a commonly used statistical model for dam health monitoring (Leger and Leclerc 2007). The HTT model is formulated as follows:

$$\delta = y_h + y_T + y_t + \varepsilon \tag{1}$$

where δ denotes the monitored dam displacement, y_h is the hydrostatic component, y_T is temperature component caused by temperature changes, y_t is irreversible time component related to the concrete creep, and ε is residuals.

The three components of Eq. (1) are calculated as Eqs. (2)-(4):

$$y_h = a_0 + a_1 h^1 + a_2 h^2 + a_3 h^3 + a_4 h^4 \tag{2}$$

$$y_T = b_0 + \sum_{i=1}^{m} b_i T_i \tag{3}$$

$$y_t = c_0 \left(1 - e^{-c_1 \theta}\right) \tag{4}$$

where h represents upstream water depth (Su et al. 2015), a_i (i=0,1,2,3,4) and b_i are unknown coefficients, T_i are measured dam temperatures, m denotes the number of temperature measuring points, $\theta = t/365$, where t denotes the days from the initial monitoring date to the observation date, c_0 and c_1 are coefficients.

3 LSTM with attention mechanism for dam health monitoring

3.1 LSTM networks

The hidden layer nodes of RNN are connected to make the output of the previous hidden layer as part of input of the next layer, therefore, RNN was used to solve time series problems. However, traditional RNN is prone to the problem of long-term dependency during training (Ma et al. 2015). Therefore, as an extension of RNN, the LSTM network was proposed to overcome the aforementioned problems.

The LSTM network adds memory cells and gate units, and prevents earlier information from disappearing during processing. These properties are especially beneficial for dam displacement prediction. The LSTM cell structure is shown in Figure 1, A denotes that three LSTM cells have same structure, and each LSTM cell is comprised of three gate units, namely input gate i_t, forget gate f_t, and output gate o_t, that collectively control the updating and discarding of information. Calculation details of LSTM cell are shown as:

$$\boldsymbol{f}_t = \sigma\left(\mathbf{W}_f \cdot \left[\boldsymbol{h}_{t-1}, \boldsymbol{x}_t\right] + \boldsymbol{b}_f\right) \tag{5}$$

$$\boldsymbol{i}_t = \sigma\left(\mathbf{W}_i \cdot \left[\boldsymbol{h}_{t-1}, \boldsymbol{x}_t\right] + \boldsymbol{b}_i\right) \tag{6}$$

$$\tilde{\boldsymbol{C}}_t = \tanh\left(\mathbf{W}_C \cdot \left[\boldsymbol{h}_{t-1}, \boldsymbol{x}_t\right] + \boldsymbol{b}_C\right) \tag{7}$$

$$\boldsymbol{C}_t = \boldsymbol{f}_t * \boldsymbol{C}_{t-1} + \boldsymbol{i}_t * \tilde{\boldsymbol{C}}_t \tag{8}$$

$$\boldsymbol{o}_t = \sigma\left(\mathbf{W}_o \cdot \left[\boldsymbol{h}_{t-1}, \boldsymbol{x}_t\right] + \boldsymbol{b}_o\right) \tag{9}$$

$$\boldsymbol{h}_t = \boldsymbol{o}_t * \tanh\left(\boldsymbol{C}_t\right) \tag{10}$$

where $\boldsymbol{x}_t$ is the current input vector, $\boldsymbol{h}_{t-1}$ is the output vector of preceding cell, $\boldsymbol{C}_t$ is current LSTM cell state, $\tilde{\boldsymbol{C}}_t$ is hidden memory cell state, $\boldsymbol{C}_{t-1}$ is the previous cell state. $\mathbf{W}_f$, $\mathbf{W}_C$, $\mathbf{W}_i$, and $\mathbf{W}_o$ are weights of the forget gate, cell state, input gate, and output gate, respectively; $\boldsymbol{b}_f$, $\boldsymbol{b}_C$, $\boldsymbol{b}_i$, and $\boldsymbol{b}_o$ are bias vectors. The σ and tanh represent sigmoid and tanh activation functions.

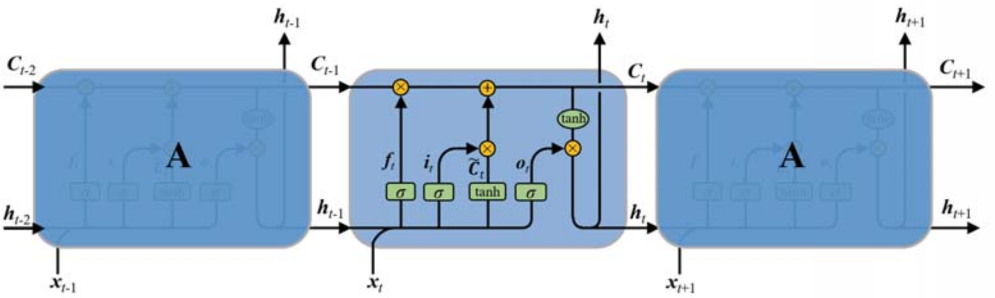


Figure 1. LSTM cell structure.

3.2 Attention Mechanism

Attention mechanism is essentially an attention allocation algorithm that highlight important feature information by mining data features. In recent years, neural networks combined with attention mechanism have been extensively studied, and has been used in the prediction of traffic flow (Zheng et al. 2021) and photovoltaic power generation (Zhou et al. 2019).

LSTM combined with attention mechanism can make the network attend to information which is more influential in current output and filter redundant irrelevant information. The output vector $\boldsymbol{H} = \{h_1, h_2, \cdots, h_n\}$ of the LSTM module is used as input to the attention mechanism for learning, and attention mechanism will automatically calculate attention weight α_i of h_i, and the calculation is shown as:

$$e_i = \tanh\left(\mathbf{W}_h \boldsymbol{h}_i + \boldsymbol{b}_h\right),\ e_i \in [-1,1] \tag{11}$$

$$\alpha_i = \mathrm{softmax}(e_i) = \frac{\exp(e_i)}{\sum_{i=1}^{n} \exp(e_i)},\ \sum_{i=1}^{n} \alpha_i = 1 \tag{12}$$

where $\mathbf{W}_h$ is weight matrix, and $\boldsymbol{b}_h$ is bias. The output feature vector $\boldsymbol{c}_i$ of attention layer is computed as a weighted sum of h_i as follows:

$$\boldsymbol{c}_i = \sum_{i=1}^{n} \alpha_i h_i \tag{13}$$

3.3 LSTM-ATT model for dam displacement prediction

In this study, considering the nonlinear characteristics of dam displacement data, LSTM-ATT network is proposed to predict arch dam displacement. The LSTM-ATT model consists of a LSTM layer, an attention layer, and a fully connected layer as shown in Figure 2. The procedure of the model is described as follows:

Step 1: Select temperature component, hydrostatic component, and time component as input variables, and select horizontal displacement as output variable.

Step 2: Obtain and process data from dam safety monitoring system and make training and test data sets.

Step 3: Set model parameters, such as the number of LSTM cells, and batch size.

Step 4: Establish and train the LSTM-ATT model using training data set until the loss value converge.

Step 5: Determine whether the loss value has converged. If the loss value has met the requirement, the process goes to Step 6, otherwise reset the parameters of Step 3.

Step 6: Evaluate the effectiveness of the model trained in the previous step by importing test sets into the model to predict horizontal displacement.

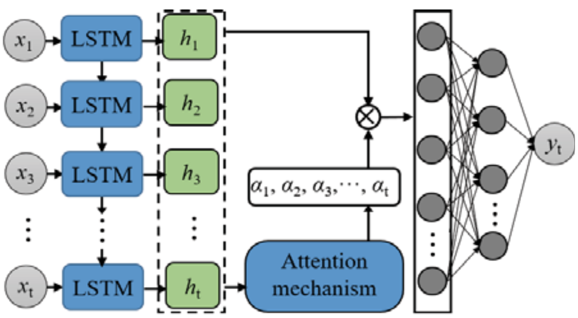


Figure 2. Architecture of the LSTM-ATT model.

The mean absolute error (MAE), maximum absolute error (AE_{max}), root mean square error ($RMSE$), and determination coefficient (R^2) were used as evaluation criteria to test the performance of different health monitoring models. The evaluation criterions can be calculated as follows:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_D(i) - y(i)\right| \tag{14}$$

$$AE_{\max} = \max\left(y_D(i) - y(i)\right), \quad i = 1,\ 2,\ \ldots\ ,\ N \tag{15}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_D(i) - y(i)\right)^2} \tag{16}$$

$$R^2 = \frac{\sum_{i=1}^{N}\left(y_D(i) - \bar{y}\right)^2}{\sum_{i=1}^{N}\left(y(i) - \bar{y}\right)^2} \tag{17}$$

where y is the monitored displacement values, $\bar{y}$ denotes average value, y_D is the predicted value, and N denotes the number of observations. The determination coefficient R^2 has a range of [0, 1], and the closer R^2 is to 1, the better the model performance. In addition, the MAE, AE_{max}, and $RMSE$ values of the optimal model should be minimized.

4 Case study

To demonstrate predictive reliability and engineering applicability of the proposed model, a case study was carried out on a real arch dam, and a comparison study on the performance of the traditional models was conducted. The simulation of SR, MLR, and RBF models were performed in Matlab R2019b, and the LSTM and LSTM-ATT models were established using the Keras framework (Gulli and Pal 2017) with version of 2.3.1.

4.1 Arch dam overview and data set

The monitoring data are taken from a 300m-level extra-high arch dam in China, with a maximum height of 285.5 meters. To assess and monitor the dam operation status, an advanced safety monitoring system was established to monitor the water level, dam temperature, vertical and horizontal displacement.

The monitoring data of horizontal displacement of the 15th dam section were used to evaluate the effectiveness of the proposed LSTM-ATT model, and the selected section is shown in Figure 3. The displacements observed by measuring point PL15_1 that is on the crest of 15th dam section are studied and the horizontal displacement curve is illustrated in Figure 4. Considering that there are too many sensors, the measured temperatures of 15 representative sensors of 16th dam section were selected as the temperature component. The data set is created with a total of 723 groups of measured data from July 4, 2014 to December 31, 2018. The training set contains 80% of the data set and the remaining 20% of data are taken to establish test set.

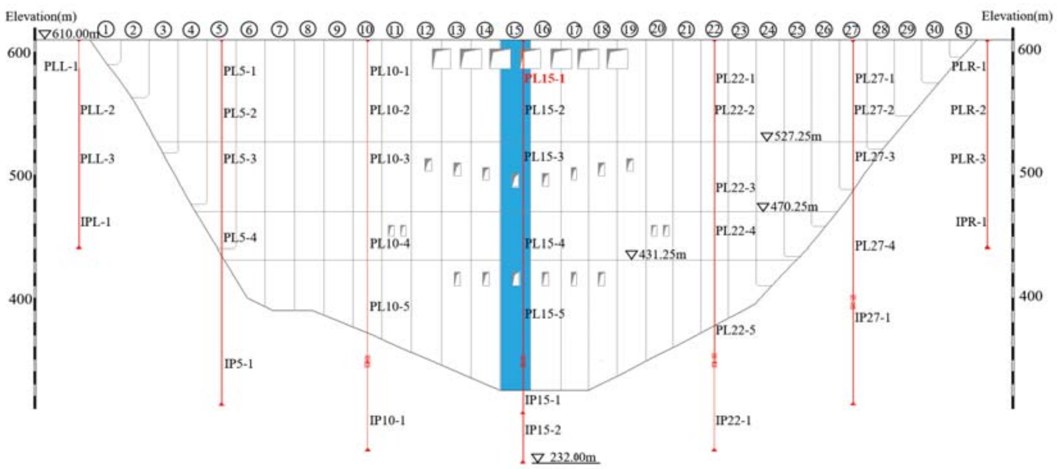


Figure 3. Dam schematic diagram, and the studied section is highlighted in blue. (PL is direct pendulums; IP is inverted pendulums).

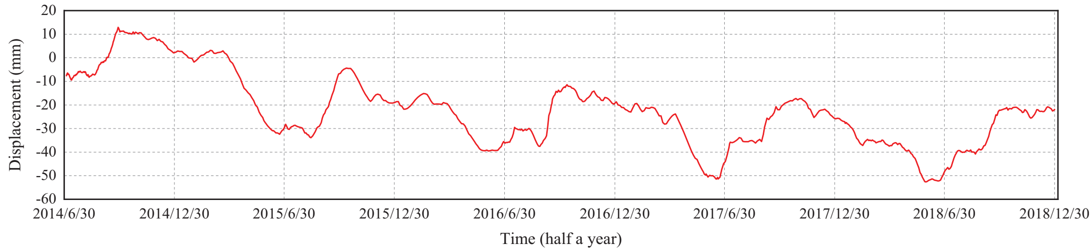


Figure 4. Measured horizontal displacements of PL15_1.

4.2 Experimental results of LSTM-ATT model

According to the theory of Section 2, 20 effective factors {H, H^2, H^3, H^4, T_1, T_2, T_3, T_4, T_5, T_6, T_7, T_8, T_9, T_{10}, T_{11}, T_{12}, T_{13}, T_{14}, T_{15}, $(1\text{-}e^{-c\theta})$} are selected to import the model of arch dam. According to experiments, when the coefficient c in $(1\text{-}e^{-c\theta})$ is set as 0.5, the HTT model is optimal. In this study, the model performs the best when the number of LSTM cells is set as 30, the batch size is 30, and the epoch is 500.

The predicted data of the LSTM model and LSTM-ATT model, and measured displacement data are shown in the Figure 5. As shown in Figure 5 (a) and (c), both LSTM and LSTM-ATT models perform well on the training set and can simulate the dam displacement. However, the prediction effect on the test set of the model with attention mechanism is better than that of the original model as shown in Figure 5 (b) and (d). The prediction accuracy on the test set of both models is shown in Table 1. As shown in Table 1, *MAE*, *RMSE* and R^2 of the LSTM-ATT model are better than that of the original model, which proves that the attention mechanism improves the overall prediction performance.

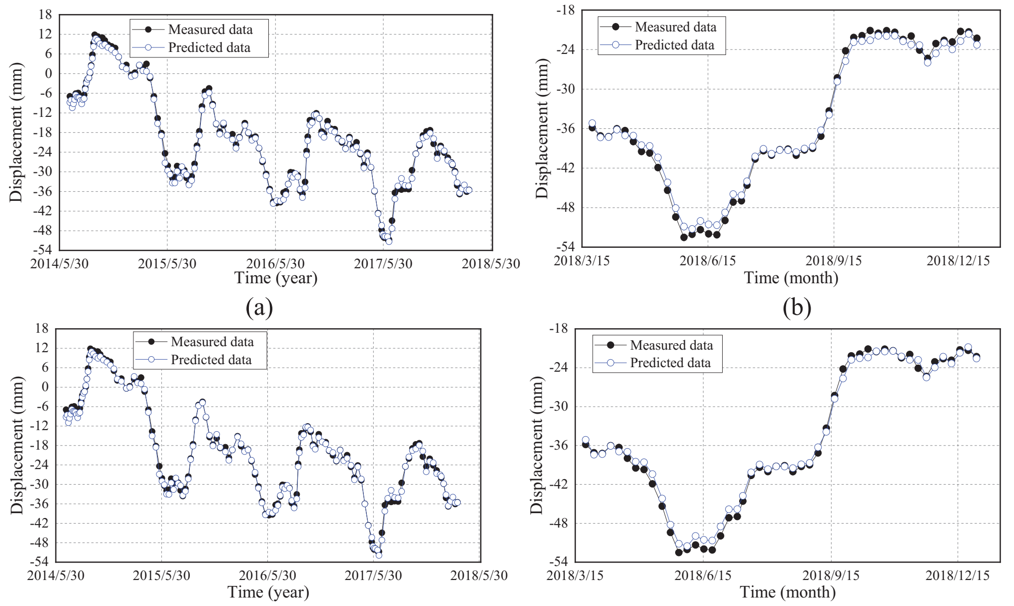


Figure 5. Performance comparison of LSTM and LSTM-ATT models: (a) LSTM performance of training set; (b) LSTM performance of test set; (c) LSTM-ATT performance of training set; (d) LSTM-ATT performance of test set.

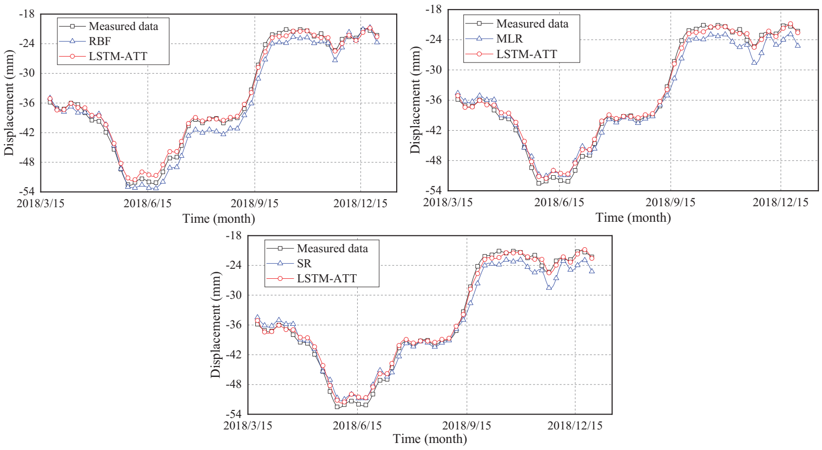
Table 1. Accuracy of dam health monitoring models on test set.

Algorithm	*MAE*	AE_{max}	*RMSE*	R^2
LSTM	0.7751	2.0365	0.9411	0.9921
LSTM-ATT	0.7063	2.1590	0.8648	0.9933

4.3 Comparative study on performance of different models

In this section, a comparative study of the performance between LSTM-ATT model and statistical models MLR (Mata 2011) and SR (Xi et al. 2011), and machine learning model RBF (Kang et al. 2019a) was conducted. The best spread of RBF is set as 294. The prediction curves of different monitoring models for the test set are shown in Figure 6. As shown in Figure 6, all models for dam health monitoring can predict the development trend of dam displacement. The predicted curve of the proposed LSTM-ATT model is closest to the measured value curve.

The performance evaluation results of different models are listed in Table 2. As shown in Table 2, it is obvious that LSTM-ATT has the smallest *MAE*, AE_{max}, *RMSE*, and largest R^2 on the test set, which demonstrates that LSTM-ATT model performs better than MLR, SR, and RBF models. Therefore, the proposed LSTM-ATT is more efficient than RBF, MLR, and SR to predict arch dam displacement for dam health monitoring.



(a) (b) (c)

Figure 6. Comparison of prediction performance for different models: (a) RBF and LSTM-ATT models; (b) MLR and LSTM-ATT models; (c) SR and LSTM-ATT models.

Table 2. Comparison of prediction performance on the data set for RBF, MLR, SR, and LSTM-ATT models.

Algorithm	Training				Test			
	MAE	AE_{max}	*RMSE*	R^2	*MAE*	AE_{max}	*RMSE*	R^2
RBF	0.4489	2.7531	0.5808	0.9986	1.2273	2.8418	1.3851	0.9828
MLR	0.7351	3.6120	0.9587	0.9961	1.3523	3.4147	1.6097	0.9767
SR	0.7421	3.5483	0.9626	0.9961	1.3408	3.3352	1.5925	0.9772
LSTM-ATT	0.7459	3.7687	0.9650	0.9960	**0.7063**	**2.1590**	**0.8648**	**0.9933**

5 Conclusion

In this study, a dam health monitoring model based on deep learning algorithm LSTM network combined with attention mechanism to predict dam behavior was proposed. The LSTM network was used to mine hidden rules in time series data. Attention mechanism was adopted to assign weights to different variables. To verify the effect of the proposed LSTM-ATT model, a 300m-level extra-high arch dam was taken as an example. Measured temperatures of this arch dam are selected as temperature component of HTT mathematic theory. Numerical experiments show that the LSTM-ATT network can predict the horizontal displacement of the arch dam effectively, and the performance is better than that of MLR, SR, RBF models and ordinary LSTM network.

Acknowledgments

This work was supported by the National Key R & D Program of China (2016YFC0401600 and 2017YFC0404900), the National Natural Science Foundation of China (52079022, 51779035, 51769033 and 51979027), and the Fundamental Research Funds for the Central Universities (DUT19LK14).

References

Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv*:1409.0473, 2014.

Gulli, A., and Pal, S. (2017). Deep learning with Keras, *Packt Publishing Ltd.*

Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. *Neural Comput*, 9(8), 1735-1780.

Kang, F., Liu, J., Li, J. J., and Li, S. J. (2017). Concrete dam displacement prediction model for health monitoring based on extreme learning machine. *Struct Control Hlth*, 24(10), e1997.

Kang, F., Li, J. J., Zhao, S. Z., and Wang, Y. J. (2019). Structural health monitoring of concrete dams using long-term air temperature for thermal effect simulation. *Eng Struct*, 180, 642-653.

Leger, P., and Leclerc, M. (2007). Hydrostatic, temperature, time-displacement model for concrete dams. *J Eng Mech-Asce*, 133(3), 267-277.

Li, G. N., Zhao, X. W., Fan, C., Fang, X., Li, F., and Wu, Y. B. (2021). Assessment of long short-term memory and its modifications for enhanced short-term building energy predictions. *J Build Eng*, 43, 103182.

Liu, W. J., Pan, J. W., Ren, Y. S., Wu, Z. G., and Wang, J. T. (2020). Coupling prediction model for long-term displacements of arch dams based on long short-term memory network. *Struct Control Hlth*, 27(7), e2548.

Ma, X. L., Tao, Z. M., Wang, Y. H., Yu, H. Y., and Wang, Y. P. (2015). Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. *Transport Res C-Emer*, 54, 187-197.

Mata, J. (2011). Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models. *Eng Struct*, 33(3), 903-910.
Mata, J., de Castro, A. T., and da Costa, J. S. (2013). Time-frequency analysis for concrete dam safety control: Correlation between the daily variation of structural response and air temperature. *Eng Struct*, 48, 658-665.
Qu, X. D., Yang, J., and Chang, M. (2019). A Deep Learning Model for Concrete Dam Deformation Prediction Based on RS-LSTM. *J Sensors*, 2019.
Su, H. Z., Wen, Z. P., Sun, X. R., and Yang, M. (2015). Time-varying identification model for dam behavior considering structural reinforcement. *Struct Saf*, 57, 1-7.
Wei, B. W., Yuan, D. Y., Li, H. K., and Xu, Z. K. (2019). Combination forecast model for concrete dam displacement considering residual correction. *Struct Health Monit*, 18(1), 232-244.
Xi, G. Y., Yue, J. P., Zhou, B. X., and Tang, P. (2011). Application of an artificial immune algorithm on a statistical model of dam displacement. *Comput Math Appl*, 62(10), 3980-3986.
Yu, H., Wu, Z. R., Bao, T. F., and Zhang, L. (2010). Multivariate analysis in dam monitoring data with PCA. *Sci China Technol Sc*, 53(4), 1088-1097.
Zhang, J. F., Zhu, Y., Zhang, X. P., Ye, M., and Yang, J. Z. (2018). Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas. *J Hydrol*, 561, 918-929.
Zheng, H. F., Lin, F., Feng, X. X., and Chen, Y. J. (2021). A Hybrid Deep Learning Model With Attention-Based Conv-LSTM Networks for Short-Term Traffic Flow Prediction. *Ieee T Intell Transp*, 22(11), 6910-6920.
Zhou, H. X., Zhang, Y. J., Yang, L. F., Liu, Q., Yan, K., and Du, Y. (2019). Short-Term Photovoltaic Power Forecasting Based on Long Short Term Memory Neural Network and Attention Mechanism. *Ieee Access*, 7, 78063-78074.