IOP Conference Series: Earth and Environmental Science

PAPER • OPEN ACCESS

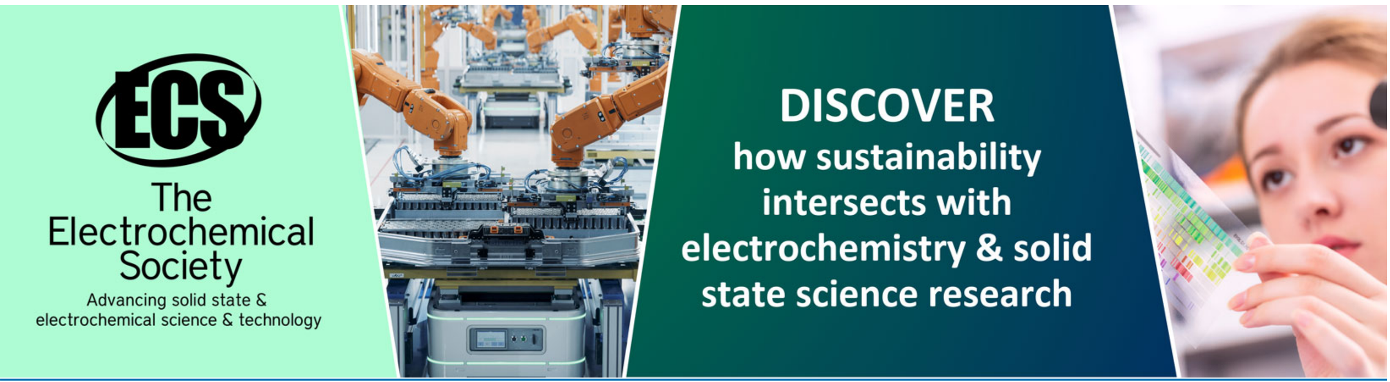
Rainfall-induced landslide displacement prediction model based on attention mechanism neural network

To cite this article: Tianjun Tang *et al* 2021 *IOP Conf. Ser.: Earth Environ. Sci.* **643** 012137

View the article online for updates and enhancements.

You may also like

- Application of fast singular spectrum decomposition method based on order statistic filter in rolling bearing fault diagnosis
  Yonggang Xu, Jinxin Cao, Jiyuan Zhao et al.
- Simulation of shallow landslides induced by typhoon in July 2018 in Gogoshima, Ehime Prefecture, Japan based on a simple model
  T V Nguyen, T Ozaki, A Wakai et al.
- An enhanced prediction model for BDS ultra-rapid clock offset that combines singular spectrum analysis, robust estimation and gray model
  Huijie Xue, Tianhe Xu, Wenfeng Nie et al.



This content was downloaded from IP address 3.141.202.54 on 06/05/2024 at 22:14

IOP Conf. Series: Earth and Environmental Science **643** (2021) 012137 doi:10.1088/1755-1315/643/1/012137

Rainfall-induced landslide displacement prediction model based on attention mechanism neural network

Tianjun Tang[1], Feifei Tang[1*] and Xifei Wang[1]

[1] School of Civil Engineering, Chongqing Jiaotong University, Chongqing, 400074, China

[*]Corresponding author's e-mail: tangfeifei@cqjtu.edu.cn

Abstract. Being influenced by fluctuating precipitation, long-term rainfall-induced landslide displacement prediction could be unstable. To solve this problem, a rainfall-induced landslide displacement prediction model based on Attention Mechanism Neural Network(AMNN) is proposed in this paper. Firstly, accumulative landslide displacement is decomposed into the trend term and the periodic term. Secondly, multivariate linear regression is adopted to fit the trend term and AMNN is used to predict the periodic term. Finally, the accumulative predicted displacement is given by summing the trend and the periodic displacement components. In this paper, one rainfall-induced landslide in Chongqing province, China was taken as an example to evaluate the designed model. Compared with some existing methods, the model proposed in this paper can captured the correlation between each feature sequence and the predicted term with higher prediction accuracy of 0.97 Goodness of Fit. The results of this paper are believed to be contributive to further rainfall-induced landslide forecast and early warning research.

1. Introduction

According to the statistical data uncovered by 2019 Chinese Geological Bulletin, among all the landslides, about 90% are rainfall-induced. The landslide displacement can directly reflect the changing trend and stability of the landslide, accurate displacement prediction can effectively reduce the damages.

In recent years, data-driven prediction model based on monitoring data as follows has been a focused areas of landslide warning[1-2]. Chen et al. adopted Recurrent Neural Network (RNN) to predict the landslide displacement in Baishui River area[3]. RNN can predict the output of the current time step based on its memory of historical information, but it has the problem of the vanishing gradient when predicting long-term sequences. To solve this problem, Long-Short Term Memory (LSTM) neural network was proposed[4]. Xu et al. applied LSTM to predict the landslide displacement in Baijiabao area[5]. This model turned out to have higher stability than RNN. These time-series analysis models have provided a method for landslide displacement prediction. However, since the shared parameters are used in RNN, LSTM and Gated Recurrent Unit (GRU), these models are easily interfered by redundant information and failed to capture key information, resulting in unstable long-term prediction.

To simplify the model, capturing key sequences of landslide time series, ensuring the long-term stability, an Attention Mechanism[6] is introduced based on the GRU time series model, a time series and AMNN rainfall-induced landslide displacement prediction model is proposed in this paper. Compared with GRU, LSTM and Support Vector Regression (SVR), the case study shows that the proposed model yields a concentrated relative error distribution for long-term prediction, the frequency of large errors is much lower and the prediction performance is much more stable.

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

IOP Conf. Series: Earth and Environmental Science **643** (2021) 012137 doi:10.1088/1755-1315/643/1/012137

2. Method

The proposed AMNN landslide displacement prediction model in this paper is composed of three layers as data pre-processing layer, data prediction layer and model evaluation layer. The data pre-processing layer decomposes the time series of accumulative landslide displacement data. The data prediction layer predicts the trend term and the periodic term respectively, summing the two components to yield the accumulative displacement prediction value. The model evaluation layer is used to verify the performance of the model according to its error gross and error distribution.

2.1. Data pre-processing layer

The data pre-processing layer decomposes the time series of accumulative displacement data into the trend term, the periodic term and the random term[7]. The trend term is mainly influenced by internal geological conditions and shows a stepped increasing trend, while the periodic term is mainly affected by external factors such as precipitation reservoir level and approximates periodic function. The random term, furthermore, can be ignored since it is mainly affected by accidental load and is relatively smaller. Hence, the cumulative displacement decomposition model is established as Equation (1):

$$\psi(t) = \rho(t) + \lambda(t) \tag{1}$$

Where $\psi(t)$ represents the accumulative displacement at time t, $\rho(t)$ represents the trend term and $\lambda(t)$ represents the periodic term at time t respectively.

In previous studies, it is found that under the influence of precipitation and internal geological conditions, rainfall-induced landslide shows a stepped increasing trend. Miao et al. suggested that the periodic term of accumulative displacement can be eliminated to obtain the trend term by using moving average method[8] as Equation (2):

$$\rho(t) = \frac{s_{t-1} + s_{t-2} + \cdots s_{t-n}}{n} \tag{2}$$

Where $\alpha(t)$ is the trend term at time t, n represents the periodic length of moving average, s_{t-1} represents the observed value at time t. The periodic term $\lambda(t)$ is given by subtracting the trend term $\rho(t)$ from the accumulative displacement $\psi(t)$.

2.2. Data prediction layer

The data prediction layer is mainly for predicting the trend term and the periodic term, the multivariate linear regression is utilized in trend term fitting and AMNN is applied in periodic term prediction. The accumulative predicted displacement is given by summing the trend and the periodic displacement components.

2.2.1. Trend term prediction. Being influenced by conditions such as internal geology and slope gravity, rainfall-induced landslide trend term in time axis shows a stepped linear increasing trend, passing an inflection point, entering into another linear increasing stage. Based on this feature, multivariate linear regression is adopted for data fitting. The fitting equation is shown in Equation (3):

$$\widetilde{\rho}(t) = \beta_1 s_1 + \beta_2 s_2 + \cdots \beta_n s_n + \varepsilon \tag{3}$$

Where $\widetilde{\rho}(t)$ represents the predicted trend term at time t, $\beta_1, \beta_2 \cdots \beta_n$ are parameters of multivariate linear regression and ε is the error value.

2.2.2. periodic term prediction. Rainfall-induced landslide periodic term is mainly affected by the fluctuation of precipitation. To uncover the relevance between precipitation fluctuations and the periodic term at each time series, AMNN is employed in prediction. The network contains two parts, the encoder and the predictor, as shown in Figure 1:

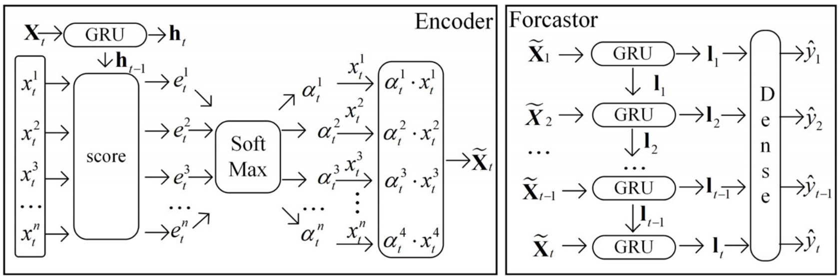


Figure 1. Flowchart of AMNN prediction model.

The time series of characteristic factors $\mathbf{X}_t = [x_t^1, x_t^2, x_t^3 \cdots x_t^n]$ is input into encoder and is input into GRU for time information iteration, ensuring the high serial correlation between the hidden states of each t and t-1 time series. t time series calculation process of GRU[9] is presented in Equation (4):

$$\left.\begin{aligned} \mathbf{r}_t &= \sigma(\mathbf{W}_{rh}\mathbf{h}_{t-1} + \mathbf{W}_{rx}\mathbf{X}_t) \\ \mathbf{z}_t &= \sigma(\mathbf{W}_{zh}\mathbf{h}_{t-1} + \mathbf{W}_{zx}\mathbf{X}_t) \\ \tilde{\mathbf{h}} &= \tanh[\mathbf{W}_{hh}(\mathbf{r}_t \circ \mathbf{h}_{t-1}) + \mathbf{W}_{hx}\mathbf{X}_t] \\ \mathbf{h}_t &= (1-\mathbf{z}_t)\circ\tilde{\mathbf{h}} + \mathbf{z}_t \circ \mathbf{h}_{t-1} \end{aligned}\right\} \tag{4}$$

Where $\mathbf{W}_{rh}$ and $\mathbf{W}_{rx}$ are the weight matrix of the forget gate, $\mathbf{W}_{zh}$ and $\mathbf{W}_{zx}$ the weight matrix of the update gate, $\mathbf{W}_{hh}$ and $\mathbf{W}_{hx}$ the weight matrix when calculating the candidate value. $\circ$ represents the Hadamard product, $\tanh(\cdot)$ and $\sigma(\cdot)$ the activation function and sigmoid respectively.

Based on the hidden state $\mathbf{h}_{t-1}$ of the former time series of GRU, scoring function of the attention mechanism presented in Equation (5) calculates the attention distribution weight at the current time series t. Then, SoftMax function presented in Equation (6) is used to normalize the attention distribution weight to yield the importance $\boldsymbol{\alpha}_t^n$ for each factor during the periodic term prediction at *t*.

$$e_t^n = \mathrm{score}(\mathbf{h}_{t-1}, \mathbf{X}_t) = \mathbf{V}_s^T \tanh(\mathbf{W}_s\mathbf{h}_{t-1} + \mathbf{U}_s\mathbf{X}_t + \mathbf{b}_s) \tag{5}$$

$$\boldsymbol{\alpha}_t^n = \mathrm{SoftMax}(e_t^n) = \frac{\exp(e_t^n)}{\sum_{i=1}^{n}\exp(e_t^i)}, \sum_{i=1}^{n}\boldsymbol{\alpha}_t^i = 1 \tag{6}$$

Where $\mathbf{W}_s$ and $\mathbf{U}_s$ represent the weight matrix of the scoring function of the attention mechanism, $\mathrm{score}(\cdot)$ is the scoring function and $\mathbf{b}_s$ the bias vector.

By multiplying $\boldsymbol{\alpha}_t^n$ given by attention mechanism layer with corresponding element $\boldsymbol{\alpha}_t^n$, the relevance of periodic term and characteristic factor is augmented, meanwhile, the redundant information in n^{th} features at time series t is decreased. The encoder's output $\tilde{\mathbf{X}}_t$ at t is shown in Equation (7).

$$\tilde{\mathbf{X}}_t = [\alpha_t^1 x_t^1, \alpha_t^2 x_t^2, \cdots \alpha_t^n x_t^n]^T \tag{7}$$

Inputting $\tilde{\mathbf{X}}_t$ into the predictor to further capture the dynamic serial correlation, inputting $\tilde{\mathbf{X}}_t$ into GRU to yield $\mathbf{l}_t$ which is then input into the Dense layer to give the prediction value $\hat{y}_t$ of the periodic term. Equation (8) demonstrates the prediction process.

$$\left.\begin{aligned} \mathbf{l}_t &= f_{gru}(\mathbf{l}_{t-1}, \tilde{\mathbf{X}}_t) \\ \hat{y}_t &= \mathbf{W}_d \times \mathbf{l}_t + b_d \end{aligned}\right\} \tag{8}$$

where $\mathbf{W}_d$ and b_d represent the weight matrix and the bias term of the fully connected layer respectively.

The weight and bias set involved in the encoder and the predictor is marked as θ . The stochastic gradient descent is used to calculate the minimum value of the loss function $J(\hat{y}_t, y_t, \theta)$ to optimize the weight and bias. The iteration process can be calculated as follows:

$$\left.\begin{aligned} J(\hat{y}_t, y_t, \theta) &= \frac{1}{2}\sum_{t=1}^{n}(\hat{y}_t - y_t)^2 \\ \theta &= \theta - \eta \times \nabla_\theta J(\hat{y}_t, y_t, \theta) \end{aligned}\right\} \tag{9}$$

where η represents the learning rate of the random gradient descending and $\nabla_\theta J(\cdot)$ represents the gradient of the loss function.

2.3. Data prediction layer

The Root Mean Square Error (RMSE) and Goodness of Fit (R^2) are adopted to evaluate the model prediction performance. Calculation of RMSE and R^2 are presented in Equation (10) respectively.

$$\left.\begin{aligned} \mathrm{RMSE} &= \sqrt{\frac{1}{M}\sum_{t=1}^{M}(y_t - \hat{y}_t)^2} \\ R^2 &= 1 - \sum_{t=1}^{M}(y_t - \hat{y})^2 \times \left(\sum_{t=1}^{M}(y_t - \overline{y})^2\right)^{-1} \end{aligned}\right\} \tag{10}$$

3. Case Study

3.1. Landslide overview

The experiment area is located on the right bank of Yangtze River. The deformation area is in the leading edge of the bedrock slope, which is stream concentrated while raining. Rainfall, therefore, is one of the major external triggering factors of landslide deformation. Monitoring devices are installed on the landslide with a three vertical and five horizontal distribution. Considering the data completeness, GPS monitoring point data and YL1 precipitation data are collected for analysis. The monitoring data are divided into two parts, data from 14th July, 2015 to 22nd April, 2018 as the training dataset, and data from 23rd April, 2018 to 31st July, 2018 as the testing dataset.

3.2. Trend term prediction

The accumulative displacement of monitoring point JC08 is used and the moving average period n is set to 10, the trend term is extracted. To avoid using future time series data for prediction, the accumulative displacement data from neighbouring monitoring points JC02, JC05 and JC07 preceding three days back are adopted. In figure 2, the visualization analysis of the trend term and the accumulative displacement of neighbouring monitoring points are presented.

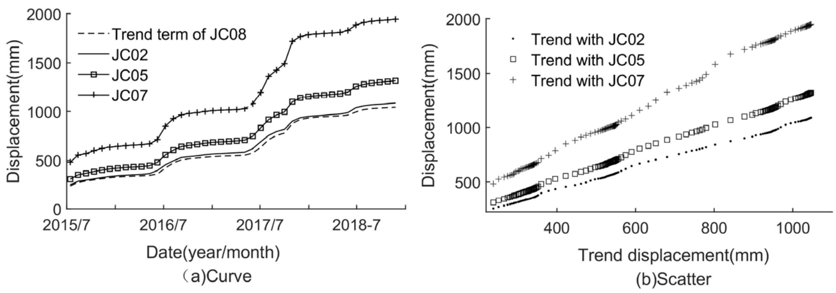


Figure 2. Relationship between trend item of JC08 and accumulative displacement of JC02, JC05 and JC07 preceding three days back.

In Figure2(a), accumulative displacement trend term and the accumulative displacement of neighbouring monitoring point (JC02, JC05 and JC07) preceding three days back are shown the same stepped trend. The trend term and three accumulative displacements preceding three days back are presented strong linear correlation in Figure2(b). Therefore, the accumulative displacement data of neighbouring points preceding three days back are input into Equation (3) as features and the least square method is adopted for parameter calculation. Fitting regression equation is then given, as shown in Equation (11):

$$\widetilde{\rho}(t)=0.251\times s^t_{\mathrm{JC02}}+0.257\times s^t_{\mathrm{JC05}}+0.212\times s^t_{\mathrm{JC07}}-13.33 \tag{11}$$

where s^t_{JC02} , s^t_{JC05} and s^t_{JC07} represents the accumulative displacement proceeding three days back of monitoring point JC02, JC05 and JC07 respectively.

The training and testing sets are input into multivariate liner regression equation. The RMSE is 4.41mm, and R^2 is 0.94. The fitting results of the testing set are shown in Figure 3.

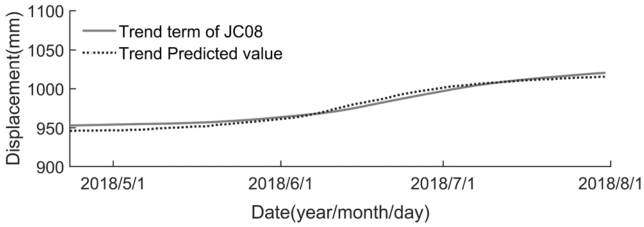


Figure 3. Prediction results of the trend terms of JC08.

3.3. Periodic term prediction

Landslide accumulative displacement is used to establish the time series decomposition model and the periodic term of accumulative displacement is acquired. The periodic term and the daily precipitation are shown in the Figure 4.

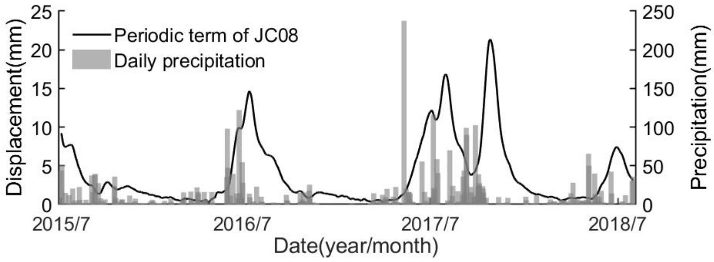


Figure 4. Relationship between periodic item of JC08 and daily precipitation.

It can be observed that when the daily precipitation is high, the period term of JC08 increases promptly and is lagged to some extent. When the daily precipitation is stable, the periodic term tends to be stable as well. It is demonstrated that precipitation is a major influencing factor for the periodic term fluctuation. Since landslide is a complex non-linear dynamics process, a non-parameter statistic method, Spearman Rank Correlation is applied to macro-extract features from driving sequences. The higher the correlation is, the stronger the correlation will be. Calculation of Spearman Rank Correlation is shown as follow:

$$\left.\begin{aligned} d_i &= c_i - c_{i-1} \\ \rho &= 1-\frac{6\sum d_i^2}{n(n^2-1)} \end{aligned}\right\} \tag{12}$$

where d_i represents the ranked sequence difference between the two monitored variants in descending order, c_i represents the ranked number of the monitored variants in descending order, ρ represents the Spearman Rank correlation, n is the sample number.

To acquire the effective precipitation value that affects the fluctuation of landslide periodic term, rolling accumulation is applied to precipitation data. Accumulative precipitations in 1, 10, 15, 30, 45 and 60 days are selected and Spearman Rank correlation is calculated, as shown in Table 1:

Table 1. The Spearman Rank correlation between accumulative precipitations and periodic displacement.

Accumulative precipitations (days)	1	10	15	30	45	60
Spearman Rank correlation	0.07	0.20	0.29	0.55	0.71	0.81

Based on Table 1, accumulative precipitations of 10, 15, 30, 45 and 60 days are selected as the model feature. Python3.8 and Tensorflow2.1r are used to build the deep learning environment to train the AMNN, GRU, LSTM and SVR. For AMNN, GRU and LSTM, the learning rate is set as 0.05, iteration times as 5000. Equation (9) is used to iterate and optimize parameters. For SVR, radical basis function is selected as the kernel function, penalty value is set as 1000. Prediction results of the four models after data iteration training are presented in Figure 5 and Table 2:

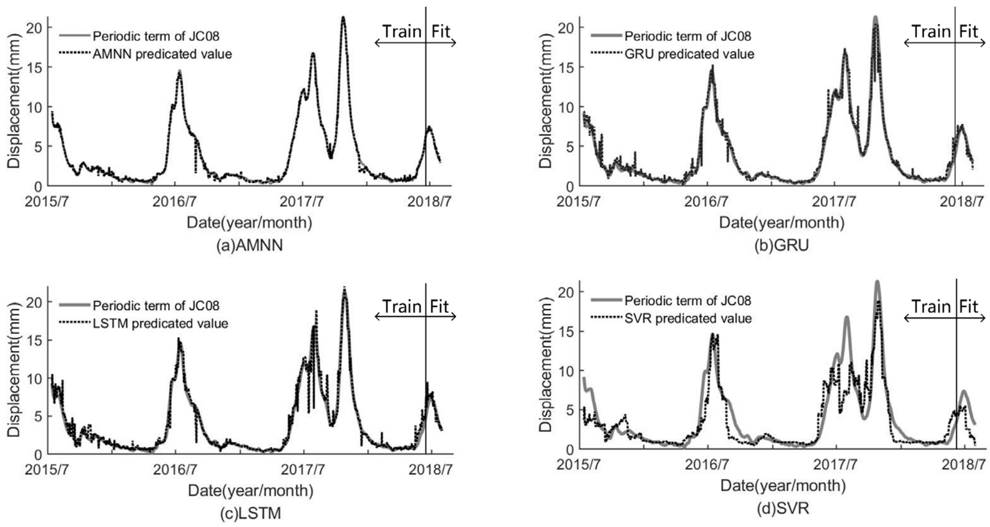


Figure 5. Prediction results of the periodic terms of JC08 using AMNN, GRU, LSTM and SVR.

Table 2. Comparison of the periodic term prediction accuracies at JC08.

Model	AMNN	GRU	LSTM	SVR
RMSE (mm)	0.30	0.71	0.91	2.07
R^2	0.97	0.90	0.84	0.20

3.4. Accumulated displacement prediction

The accumulative displacement prediction value is acquired by summing the predicted trend term of JC08 and the predicted periodic term using AMNN in accordance with time series. Results are presented in Figure 6. It is shown that the predicted accumulative displacement curve is quite close to the observation curve with the RMSE of 4.37mm and R^2 of 0.97.

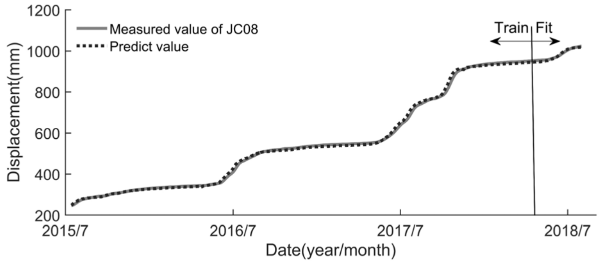


Figure 6. Prediction results of the accumulative displacement at JC08.

4. discussion

In this paper, AMNN is proposed to establish a rainfall-induced landslide prediction model. High relevant features are macro-selected from driving sequences of monitored data and input into the model. The encoder of the model captures the fluctuation of the precipitation and micro-guides and modifies driving sequences using attention distribution possibility. The modified sequence is input into the predictor to predict the periodic term to solve the unstable long-term prediction problem of recurrent neutral networks resulted from shared parameters in hidden layer.

To test the prediction stability of the attention mechanism model, relative errors of periodic term given by AMNN, GRU, LSTM and SVR are statistically analysed using Gaussian mixture distribution. Their Gaussian frequency fitting distribution curves are shown in Figure 7 for direct observation.

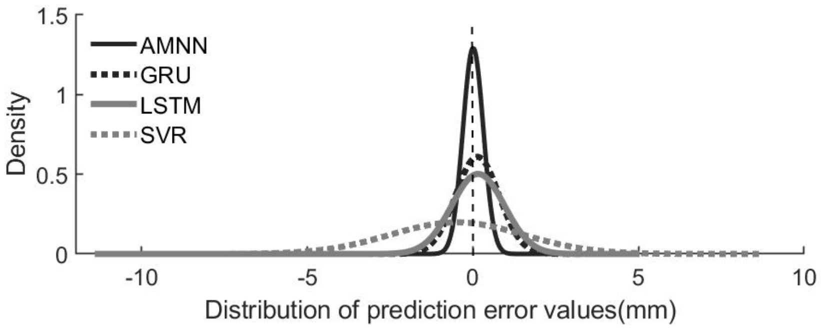


Figure 7. Gaussian frequency fitting distribution curves.

It can be seen that relative error distribution curves of GRU, LSTM and SVR show a bias distribution with their mode, presenting a large distribution range. Relative errors of AMNN, on the other hand, show a symmetrical distribution with its error range varying within [-1,1], which means AMNN has smaller relative error and is more concentrated error range, and therefore, the proposed model performs better compared with other three models. In Figure 5, it can be observed that though the training and testing sets fitting curves of AMNN show small fluctuations, they can return to stable state within a short period. While GRU and LSTM can record and forget information timely, their fitting curves show quite a number of small fluctuations since driving sequences make same contribution to the predicted value at each time series. SVR, as a static model, cannot capture the time series correlation of driving sequences and thus shows a lagged fitting.

To validate the effectiveness of the AMNN model in attention extraction, testing data from 19st June, 2016 to 23rd June, 2016 and training data from 16th May, 2018 to 20th May, 2018 are input into the model to extract the importance degree α_t^n of each feature at this time series to the periodic term to be predicted. The visualization results are shown in the figure 8.

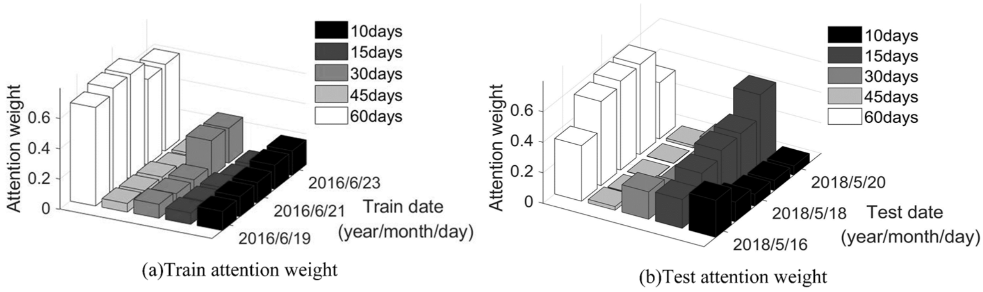


Figure 8. The importance degree of each feature.

Each column in the three-dimensional histogram represents a prediction time series, each row represents the importance degree. It can be observed that when micro-modifying occurred in the feature sequence, the model lays priority on the accumulative precipitation in the last 60 days, which is consistent with the rank correlation of macro-selected parameters. Therefore, it can be concluded that AMNN is able to predict long-term stability.

Since AMNN is a deep leaning model, its accuracy is largely influenced by the data size and quality and time-effectiveness of the data should be maintained by updating dataset timely.

5. Conclusion

To improve the long-term prediction stability of rainfall-induced landslide displacement prediction model, deeply extract the augment information between driving sequences and landslide displacement, decreasing redundant information of driving sequences, a rainfall-induced landslide displacement prediction model based on attention mechanism neural network is proposed in this paper and a case study is conducted taking one rainfall-induced landslide in Chongqing province, China as an example. Based on the experiment results, the following conclusions are drawn:

(1) During the model training and prediction phase, the attention mechanism neutral network can capture the correlation between precipitation driving sequences and the periodic term, producing the importance degree for each feature at each time series to the predicted periodic term. Driving sequences are micro-modified and stability for long term prediction is improved.

(2) Comparing with GRU, LSTM and SVR, the proposed model yields a more concentrated relative error distribution, a lower frequency of large errors and more stable prediction performance, making technological contributions to rainfall-induced landslide early warning research.

Acknowledgments

This research project was supported by The Scientific and Technological Research Program of Chongqing Municipal Education Commission under Grant No. KJQN201900728 and The Basic Research and Frontier Exploration Project of Chongqing Province (Natural Science Foundation of Chongqing Province) under Grant No. cstc2018jcyjAX0515.

References

[1] Huang, F., Huang, J., Jiang, S., Zhou, C. (2017) Landslide displacement prediction based on multivariate chaotic model and extreme learning machine. Engineering Geology., 218: 173-186.
[2] Zhu, X., Ma, S., Xu, Q., Liu, W. (2018) A WD-GA-LSSVM model for rainfall-triggered landslide displacement prediction. J. Mt. Sci., 15: 156–166.
[3] Chen, H., Zeng, Z., Tang, H. (2012) Study on landslide deformation prediction based on recurrent neural network under the function of rainfall. Neural Processing Letters., 41: 169-178.
[4] Gers, F., Schmidhuber, J., Cummins F. (2000) Learning to forget: continual prediction with LSTM. Neural Computation., 12: 2451－2471.

[5] Xu, S., Niu, R. (2018) Displacement prediction of Baijiabao landslide based on empirical mode decomposition and long short-term memory neural network in Three Gorges area, China. Computers & geosciences., 111: 87－96.
[6] Cho, K., Courville, A., Bengio, Y. (2015) Describing multimedia content using attention-based encoder-decoder networks. IEEE Transactions on Multimedia., 17: 1875-1886.
[7] Du, J., Yin, K., Lacasse, S. (2013) Displacement prediction in colluvial landslides, three gorges reservoir, china. Landslides., 10: 203-218.
[8] Miao, F., Wu, Y., Xie, Y., Li, Y. (2018) Prediction of landslide displacement with step-like behaviour based on multialgorithm optimization and a support vector regression model. Landslides., 15: 475-488.
[9] Umut, U., Ilkay, O., Oktay, T. (2018) Electricity price forecasting using recurrent neural networks. Energies., 11: 1255-1278.