ISSN: 2414-1895 DOI: 10.6919/ICJE.202202_8(2).0072

Dynamic Prediction of Landslide Displacement Time Series Combining Singular Spectrum Decomposition and Improved Attention Mechanism

Weijing Zhang[1], Rongwei Ma[2]

[1] Chengdu University of Technology, Chengdu 610000, China

[2] Xinjiang University, Urumqi 830000, China

Abstract

The displacement and deformation of landslides are affected by many factors, which is a typical non-linear change process. The neural network can deal with linear problems well. However, the traditional serialization network will cause gradient disappearance and sequence forgetting due to sequential calculation. This paper proposes a combined prediction model that combines singular spectrum analysis (SSA) and improved attention mechanism. First, SSA is used to decompose the displacement time series and eliminate noise, reduce random fluctuation factors and increase the amount of effective information, and then use the improved attention mechanism to predict the decomposition sub-sequence, and superimpose the results to obtain the final prediction value. Experiments show that the combined model has better effects and higher accuracy than traditional sequence models, such as RNN and LSTM.

Keywords

Landslide; Deep learning; SSA; Improved Attention.

1. Introduction

The change of landslide displacement is a non-linear change process caused by multiple factors, and its forming factors can be internal and external. Most of the internal factors are caused by the continuous changes in the internal forces of the landslide body, which are affected by the topography, rock and soil conditions, and soil moisture content, and have trend changes; the external factors are caused by external factors, mainly affected by precipitation, rain intensity, and accumulated rainfall. The internal stability of the equal sound landslide body, and the changes of these factors over time have obvious periodicity. Singular Spectrum Analysis (SSA) is a nonlinear time series decomposition and reconstruction analysis method, which can effectively deal with complex sequences. It can identify the differences in the original sequence without selecting a priori basis functions and complex calculation processes and extract trend decomposition items, period decomposition amounts, and noise decomposition amounts. They have good objectivity and adaptability, and are widely used in various engineering disciplines. However, there are few applications in the field of landslide disasters at present, and it has a good research prospect.

In recent years, many scholars have applied deep learning methods to time series analysis and forecasting and the establishment of nonlinear models. In particular, sequence neural networks can extract state features and historical information in sequences better than traditional neural networks, making long-term sequence prediction results more accurate than traditional time-series models, such as Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM). However, the calculation process of these three types of networks is sequential. The weight parameters at the current moment are completely dependent on the output at the previous moment. Parallel calculation cannot be used

ISSN: 2414-1895 DOI: 10.6919/ICJE.202202_8(2).0072

to improve efficiency. In the process of sequential calculation, the data with the farther distance has the current predicted value. The smaller the impact and the formation of RNN-type models, there will be information loss when predicting the final result, resulting in a large error in the prediction accuracy.

Google proposed the Transformer model in 2017, and for the first time used the Attention mechanism to achieve very good results. In 2021, Zhou proposed an Informer model based on Transformer to be applied to long time series prediction. Improving the attention mechanism (ProSparse Self-Attention) to solve the problems of Transformer's high computational complexity in processing long time series, large memory footprint, and slow decoder prediction speed.

In the actual data collection, the time series usually contains a lot of noise and some data points are abnormal due to various force majeure, resulting in the prediction results still containing these invalid factors, resulting in unsatisfactory final prediction results. In addition, external factors such as precipitation are cyclical, and what really induces landslides is internal, not external cyclical factors. Therefore, this paper proposes an improved attention mechanism landslide displacement prediction algorithm fused with singular spectrum analysis. By denoising the displacement data, the trend term, period term and noise sequence are decomposed, so that the model is more focused on internal factors (trend term) to improve prediction accuracy.

2. Landslide Displacement Prediction Model

2.1 Principles of Singular Spectrum Analysis

SSA is a sequence component analysis method based on singular value decomposition, which decomposes the original time series data into several sub-sequences that can gather trend items, periodic items, and noise items. The specific process is as follows.

SSA reconstructs a trajectory matrix $\mathbf{X}$ with dimension K according to the original one-dimensional time series$[z_1, z_2, z_3, \cdots, z_N]$ with length N:

$$\mathbf{X} = [X_1, X_2, X_3, \cdots, X_N] = \begin{bmatrix} z_1 & z_2 & \cdots & z_K \\ z_2 & z_3 & \cdots & z_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ z_L & z_{L+1} & \cdots & z_N \end{bmatrix} \tag{1}$$

where: $\mathbf{X}_i = [z_i, z_{i+1}, \cdots, z_{i+L-1}]^T; i = 1,2, \cdots, K; K = N - L + 1$.

Perform singular value decomposition on the trajectory matrix. Calculate the eigenvalues of the matrix $\mathbf{X}^T\mathbf{X}$ and sort them in descending order according to the size of the eigen values, denoted as $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_L \geq 0$, and denote the eigenvector corresponding to the eigenvalue as $\mathbf{U}_i$. Define $V_i = \mathbf{X}^T\mathbf{U}_i\sqrt{\lambda_i}$, then the singular value decomposition of $\mathbf{X}$ can be written as:

$$\mathbf{X} = X_1 + X_2 + \cdots + X_d \tag{2}$$

where: $\mathbf{X}_i = \sqrt{\lambda_i}\mathbf{U}_i V_i^T, \mathbf{V}_i = \mathbf{X}^T\mathbf{U}_i\sqrt{\lambda_i}$; $\mathbf{U}_i$ and $\mathbf{V}_i$ are respectively the ith singular value of the trajectory matrix X and its corresponding left singular vector and right singular vector.

Next, group on the basis of the above steps. According to the different singular values decomposed, the sub-sequences are grouped, the harmonic components of each component are filtered, and the abnormal data is eliminated.

Reconstruct the time series by diagonal averaging. Convert the grouped matrix into a series of new reconstruction matrices of length N. Let $L^* = \min(L, K)$, $K^* = \max(L, K)$, $N = L + K - 1$, if $L <$

ISSN: 2414-1895 DOI: 10.6919/ICJE.202202_8(2).0072

K, then $x_{ij}^{*} = x_{ij}$,otherwise $x_{ij}^{*} = x_{ji}$. The sequence obtained after reconstruction can be recorded as: $[y_1, y_2, \cdots, y_N]$.

$$y_k = \begin{cases} \frac{1}{k+1}\sum_{m=1}^{k+1} x_{m,k-m+2}^{*}, 1 \le k \le L^{*} \\ \frac{1}{L^{*}}\sum_{m=1}^{L^{*}} x_{m,k-m+2}^{*}, L^{*} \le k \le K^{*} \\ \frac{1}{N-n}\sum_{m=k-K^{*}+2}^{N-K^{*}+1} x_{m,k-m+2}^{*}, K^{*} \le k \le N \end{cases} \tag{3}$$

In the process of predicting the displacement of a landslide, the original time series data contains a certain amount of noise and uncertain volatility. Therefore, this paper uses the singular spectrum method to preprocess the data in order to reduce random items and uncertainty to improve the data. Signal-to-noise ratio to optimize the prediction results.

2.2 Improved Attention Mechanism

The traditional RNN, LSTM, Transformer and Informer processing time series data essentially use the decoder-encoder method. However, RNN and LSTM will forget the early information when processing data in sequence. Transformer can calculate the attention between all features at the same time, but the amount of calculation is too large, the memory occupancy rate is too high, and the long sequence cannot be predicted. Informer uses the ProSparse Self-Attention mechanism in the encoding-decoding process, and only considers the part of the attention mechanism that contributes the most. It performs better than Transformer in terms of overall operating processing efficiency and memory overhead. The algorithm unit structure is as follows:

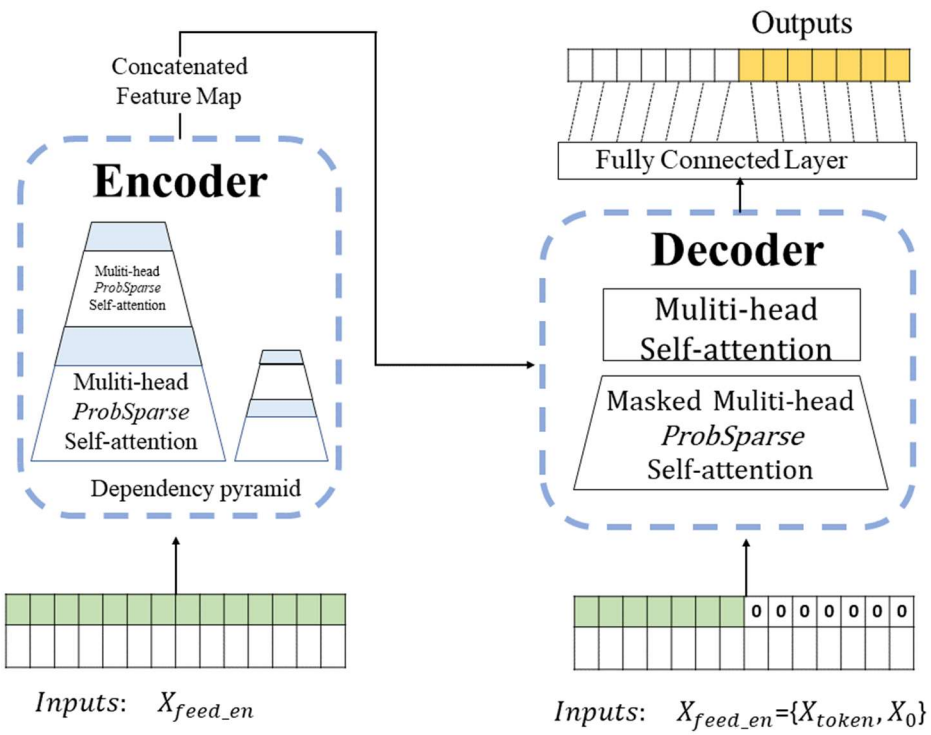

Figure 1. Module

ISSN: 2414-1895 DOI: 10.6919/ICJE.202202_8(2).0072

The input part of the encoder contains the coding position information of the sequence data, the dimension mapping feature, the sequence length and the time information of the input sequence. In the encoding part of the encoder, the attention mechanism weight calculated by the ProbSparse Self-attention mechanism is expressed by the formula (4).

$$Z = \mathrm{Softmax}\left(\frac{\bar{Q}K^{T}}{\sqrt{d}}\right)V \tag{4}$$

Among them, $\bar{Q}$, K, V are the active query, key value, and value of the input vector after different linear transformations, which represent different learning information in the sequence information. Multi-ProbSparse Self-attention can divide the input after Embedding into n according to the mapping dimension, and then merge them after doing ProbSparse Self-attention separately.

In the decoding part of the decoder, the relevant information between the data is also learned through ProbSparse Self-attention, and the ProbSparse Self-attention attention weight is calculated, and finally combined with the information passed in by the Enconder, it is predicted through the Self-Attenton layer. The connection layer obtains the predicted value.

3. Real Case Analysis

3.1 Model Prediction Process

The combined model analysis and prediction displacement process is as follows:

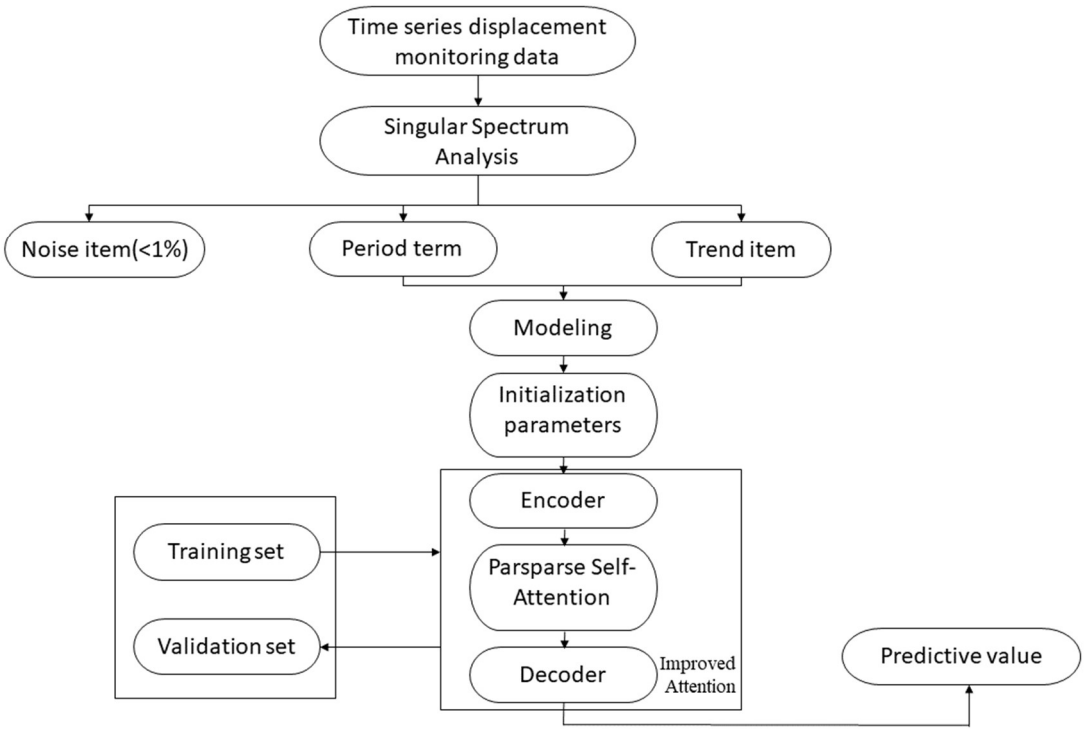

Figure 2. Processing steps

3.2 Experimental Comparison

The data used in this paper are the GXT1 monitoring point data located in the Xintan landslide in the upper section of the Xiling Gorge of the Yangtze River in Hubei Province. The daily landslide displacement from July 1, 1989 to January 31, 1985 was recorded, a total of 19,800 records were

ISSN: 2414-1895 DOI: 10.6919/ICJE.202202_8(2).0072

recorded. According to the recorded data and Figure 3-1, the landslide has experienced a leaping growth since 1979.

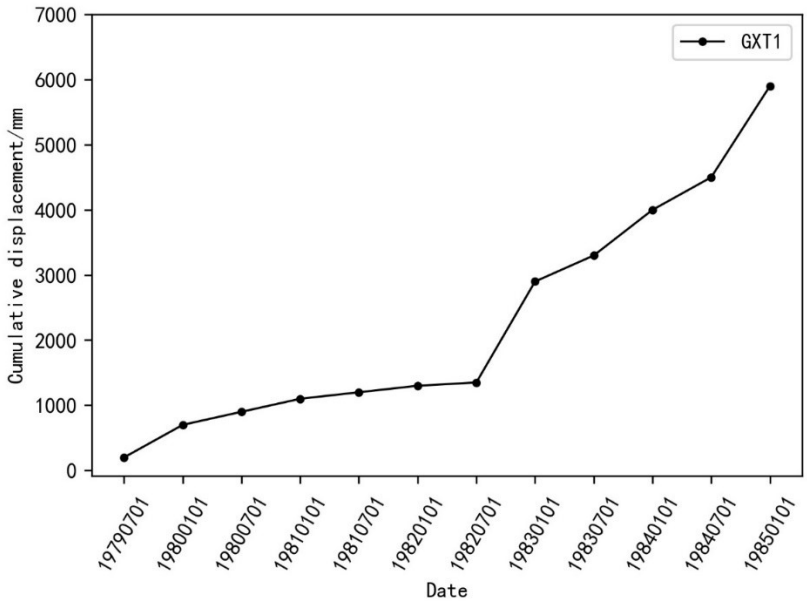

Figure 3. Landslide displacement

The data preprocessing is decomposed into trend items and period items, respectively, and the traditional LSTM and the improved attention mechanism model are used for comparative experiments. Divide the entire data into training set, validation set and prediction set. The training set data includes July 1, 1979 to December 31, 1981, and the validation set data includes January 1, 1982 to December 31, 1982. The forecast set data includes the period from January 1, 1983 to July 1, 1985, the final value of the forecast is expressed cumulatively by month, and the evaluation index is the relative error. The training results are shown in the figure below:

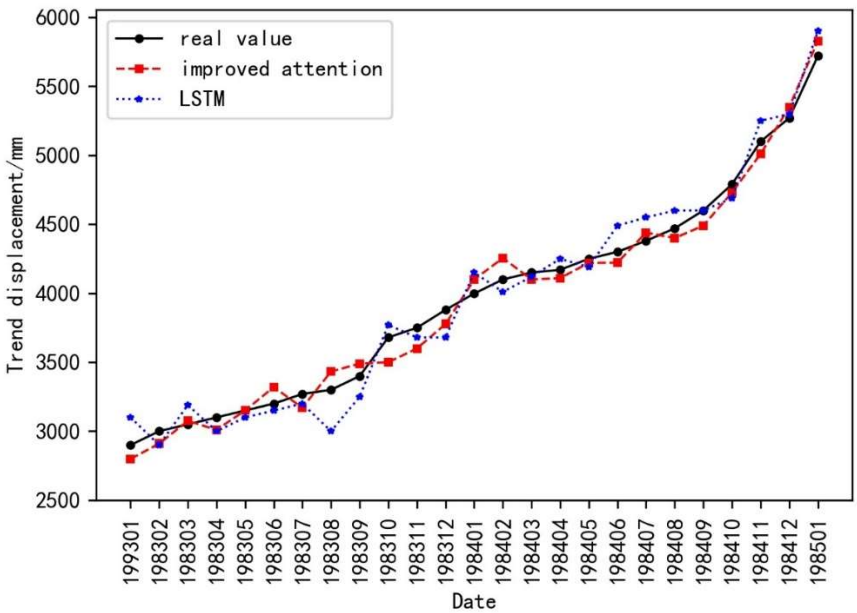

Figure 4. Trend series comparison

ISSN: 2414-1895 DOI: 10.6919/ICJE.202202_8(2).0072

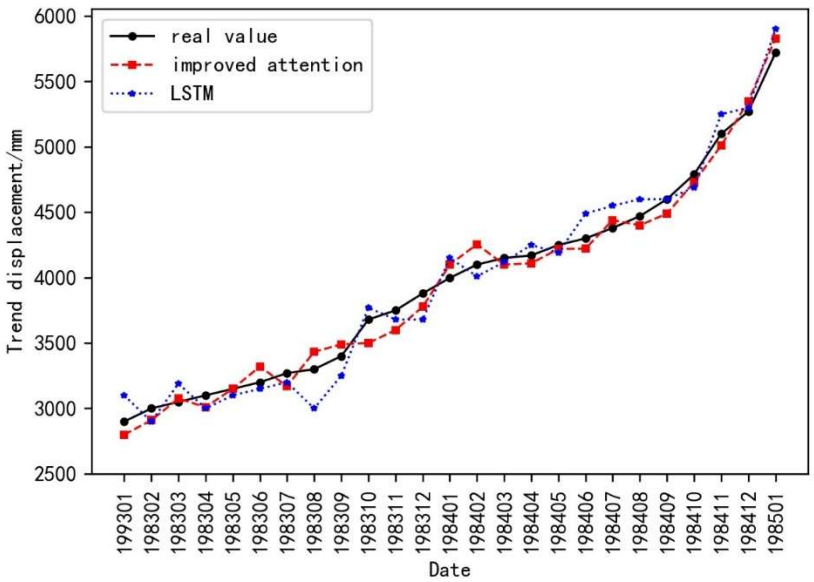

Figure 5. Periodic sequence comparison

From the above figure, it can be intuitively obtained that in the entire nonlinear process of landslide displacement and deformation prediction, the improved attention mechanism is due to the traditional serialization model, combined with the relative error (RE) of the evaluation index, it can be found that the average of the improved attention mechanism. The error is 5.512%, and the average error of LSTM reaches 7.131%. Therefore, according to the decomposition principle of singular spectrum analysis method, the predicted values of the decomposed trend sub-items and period sub-items are combined to obtain more accurate prediction results.

4. Conclusion

In the process of predicting such non-linear changes in landslides, with the rise of artificial intelligence technology, traditional time series methods can no longer meet more accurate predictions, which requires new models that can better deal with non-linear changes. In this paper, the singular spectrum analysis method is used to eliminate data noise, retain more effective data information, and use an improved attention mechanism to predict landslide displacement. At the same time, it also overcomes the drawbacks of long-term forgetting such as RNN and LSTM, and further improves the accuracy of prediction. And the accuracy of long-term series prediction.

References

[1] Liu H, Chen C. Data processing strategies in wind energy forecasting models and applications: A comprehensive review[J]. Applied Energy, 2019, 249(SEP.1):392-408.

[2] Vaswani A , Shazeer N , Parmar N , et al. Attention Is All You Need[J]. arXiv, 2017.

[3] Meng Q, Wang H , He M , et al. Displacement prediction of water-induced landslides using a recurrent deep learning model[J]. European Journal of Environmental and Civil Engineering, 2020:1-15.

[4] Mohan A , Singh A K , Kumar B , et al. Review on remote sensing methods for landslide detection using machine and deep learning[J]. Transactions on Emerging Telecommunications Technologies, 2020(4).

[5] Elsner J B. Analysis of Time Series Structure: SSA and Related Techniques[J]. Journal of the American Statistical Association, 2002, 97(460):1207-1208.

[6] Colominas M A , Schlotthauer G , Torres M E . Improved complete ensemble EMD: A suitable tool for biomedical signal processing[J]. Biomedical Signal Processing & Control, 2014, 14(nov.):19-29.

[7] Zhou H , Zhang S , Peng J, et al. Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting[J]. 2020.