Acta Geodyn. Geomater., Vol. 21, No. 3 (215), 195–206, 2024
DOI: 10.13168/AGG.2024.0018
journal homepage: https://www.irsm.cas.cz/acta

ORIGINAL PAPER

GNSS ELEVATION TIME SERIES PREDICTION MODEL BASED ON MULTIPLE GATE RECURRENCE UNIT AND TEMPORAL CONVOLUTIONAL NETWORKS PARALLEL

Zengnan HOU [1)], Zhengkai HUANG [1)] *, Huijuan LIU [2)] and Wentao WANG [2)]

[11)] *School of Transportation Engineering, East China Jiao Tong University, Nanchang 330013, China*
[2)] *Hebei Water Resources and Hydropower Survey, Design and Research Institute Group Co., Ltd, Tianjin 300250, China*

**Corresponding author's e-mail:* zhkhuang@whu.edu.cn

ARTICLE INFO

Article history:
Received 9 July 2024
Accepted 11 September 2024
Available online 24 September 2024

Keywords:
Gate Recurrence Unit
Temporal Convolutional Networks
Long Short-Term Memory
Convolutional Neural Networks

ABSTRACT

The research on Global Navigation Satellite System (GNSS) time series prediction provides reliable data to support for the monitoring of crustal plate movements, and is of great significance for an in-depth understanding of the movement law and potential change trend of crustal plates. Addressing the issue that the Gate Recurrence Unit (GRU) model has a limited ability to learn the characteristics of complex time series and cannot effectively capture the spatial characteristics of GNSS time series, this paper presents a Multiple Gate Recurrence Unit (MGRU) and Temporal Convolutional Networks (TCN) parallel (MGRU+TCN) dual branch parallel prediction model. The MGRU branch consists of multiple GRU modules and multiple hidden layers. The temporal features and the spatial features extracted are fused by concatenate function, and the prediction results are obtained through full connection layer output. The validity of the model is verified by using the elevation time series of several GNSS reference stations. The experimental show some results that compared with the single model of Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), MGRU and TCN, the Root Mean Square Error (RMSE) is on average reduced by 0.73 mm, the Mean Absolute Error (MAE) is on average reduced by 0.60 mm, and the R^2 is on average increased by 0.11. Compared with the prediction model of MGRU-TCN and TCN-MGRU serial structure, RMSE decreased by 0.60 mm on average, MAE decreased by 0.47 mm on average, and R^2 increased by 0.08 on average. The results show that the parallel structure MGRU+TCN prediction model has higher prediction accuracy and good GNSS elevation time series prediction ability compared to the single prediction model and the serial structure prediction model.

1. INTRODUCTION

Over the past 20 years, space observation technology has advanced rapidly. The International GNSS Service (IGS) reference stations have accumulated time series data. This data provides support for geodesy and Earth dynamics research (Fernandes et al., 2004; Serpelloni et al., 2013; Montillet et al., 2015; Shen et al., 2019). GNSS data can reflect long-term trends and exhibit nonlinear changes (Chen et al., 2013; Wang et al., 2021). As the key technology of modern positioning, navigation and timing, it plays an indispensable role in numerous fields. These data are used for research on crustal plate movement (Serpelloni et al., 2013; Ohta et al., 2012; Kong et al., 2023; Younes, 2023). Landslide detection (Cina et al., 2015; Shen et al., 2021; Shen et al., 2022). Deformation monitoring of bridges or dams (Xi et al., 2018; Chen et al., 2018; Xin et al., 2018) and maintenance of regional or global coordinate frameworks (Lahtinen et al., 2019; Li et al., 2020; Chen et al., 2021) provide data. The changes in coordinate over continuous time points are forecasted using the time series of long-term observation data of GNSS reference stations, providing a basis for clarifying the movement trend (Li et al., 2023).

With the continuous development of artificial intelligence, machine learning has been applied in various fields. As an important branch of machine learning, deep learning is based on the concept of artificial neural network. It processes data by constructing a multi-layer neural network structure. It has strong ability of feature extraction and pattern recognition, and can automatically learn complex feature representation from the original data. It has achieved great success in image recognition, speech recognition, natural language processing and other fields, and has greatly promoted the development of artificial intelligence (Li et al., 2023). Recurrent Neural Network (RNN) is a kind of neural network used to process sequence data. Its main feature is that it can remember and transfer the information in the sequence. Through the circular structure, the output of the current time is associated with the information of the previous time, so it is suitable for processing sequential and dependent data such as natural language and time series data. However, RNN may

Cite this article as: Hou Z, Huang Z, Liu H, Wang W: GNSS elevation time series prediction model based on multiple gate recurrence unit and temporal convolutional networks parallel. Acta Geodyn. Geomater., 21, No. 3 (215), 195–206, 2024. DOI: 10.13168/AGG.2024.0018

have problems such as gradient disappearance or explosion (Li et al., 2018). In 1997, the Long Short-Term Memory (LSTM) was proposed by German computer scientists Hochreiter and Schmidhuber. Its main components include a forgetting gate, an input gate, and an output gate. Its core idea is to control the flow of information through these gates, avoiding the gradient disappearance or explosion of the RNN model. However, the more complex structure and too many parameters lead to a large amount of calculation (Yu et al., 2019; Chen et al., 2023a). Cho et al. (2014) proposed a variant of the LSTM known as the Gate Recurrence Unit (GRU), which has a simpler than the LSTM structure. It controls the flow of information through update and reset gates, thereby reducing the number of parameters (Li et al., 2022; Fu et al., 2016; Dey et al., 2017). Irie et al. (2016) used RNN, LSTM and GRU networks compare their prediction performance on voice data sets. It is found that GRU avoids the RNN gradient explosion problem and has the fastest prediction speed while ensuring the same high-precision prediction results as LSTM. Although GRU model has high prediction accuracy and fast prediction speed, its feature extraction ability is weak to process complex time series, and it cannot extract the spatial features of GNSS time series. Building on this, this paper proposes a model, MGRU, which incorporates multiple GRU modules and hidden layers, and integrates it with the Temporal Convolutional Networks (TCN) model proposed by Lea et al. in 2016, to create a parallel structure prediction model, Multiple Gate Recurrence Unit and Temporal Convolutional Networks parallel (MGRU+TCN). On the basis of the advantages of GRU and TCN single prediction models, further improve the prediction accuracy. The method extracts the time characteristics of GNSS time series through MGRU branch, and the TCN branch extracts the spatial characteristics. The features extracted from the two branches are fused through concatenate function. Finally, the results are output through the full connection layer. Taking the elevation time series of multiple GNSS reference stations as an example, the effectiveness of the MGRU+TCN parallel structure prediction model is verified by analyzing the prediction results of mainstream single prediction models and combination models with different structures.

The structure of this paper is as follows: Section 2 introduces the principles of the GRU and TCN, details the working principle and prediction process of the MGRU+TCN model, and discusses the accuracy evaluation indices for the prediction results. Section 3 introduces GNSS station data, data noise reduction and preprocessing; Determine the number of GRU modules and hidden layers in MGRU branch, and the number of residual blocks in TCN branch; The prediction results and accuracy of MGRU+TCN model and single model LSTM, CNN, MGRU and TCN, as well as the prediction results and accuracy of serial structure with TCN-MGRU and MGRU-TCN are compared and analyzed. Finally, Section 4 presents the conclusion.

2. PRINCIPLE AND METHOD

2.1. GATED RECURRENT UNIT (GRU)

The GRU, a variant of the LSTM network, operates on a similar principle. It controls the input, memory and other information through the gating mechanism, so as to make a prediction at the current time. Their main differences lie in the design of gating structure and cell state. GRU has only two gating structures, namely update gate and reset gate, while LSTM has three gates, namely input gate, forget gate and output gate. The update gate controls the degree to which the information from the previous time is transmitted to the current time, while the reset gate controls how much information from the previous memory comes from the current time. This design makes GRU more flexible in dealing with long-term dependencies, while reducing the number of model parameters and computational complexity. The structure is shown in Figure 1 (Fu et al., 2016; Dey et al., 2017).

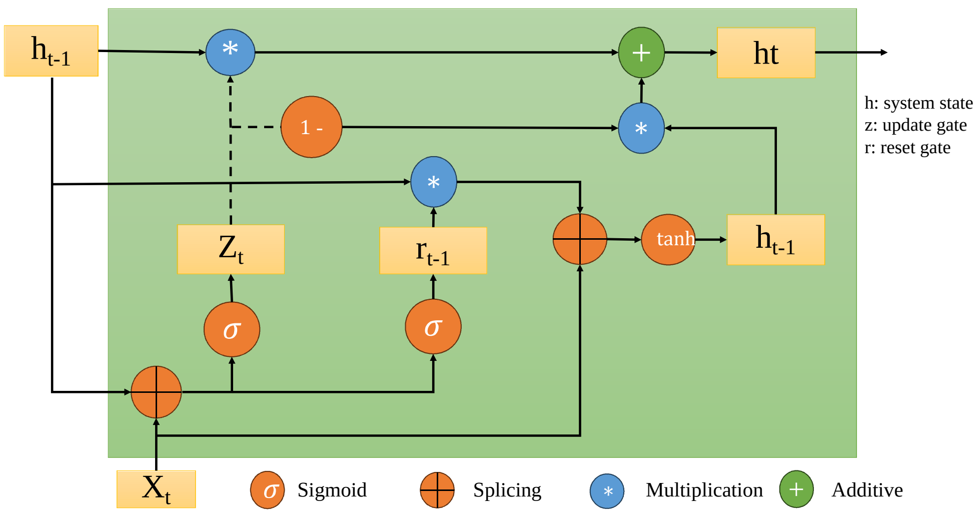

Fig. 1 GRU model structure.

2.2. TEMPORAL CONVOLUTIONAL NETWORK (TCN)

TCN is mainly based on causal convolution, hole convolution of stored data and residual block, and its structure is shown in Figure 2. Causal convolution uses only past information for predictions, allowing for a step-by-step analysis of time series data and the extraction of inherent patterns. Meanwhile, hole convolution broadens the network's receptive field, enhancing its ability to detect long-term dependencies. Void convolution can cover a larger time span with a smaller number of layers by introducing an interval in the convolution core, so that the dynamic changes of time series can be understood more comprehensively (Hewage et al., 2020). Compared with traditional Convolutional Neural Networks (CNN), TCN demonstrates significant improvements in processing time series data. It breaks through the limitation of CNN in processing variable length sequences, and can better deal with time series of different lengths. It also solves the gradient explosion problem caused by the increase of network depth in the training process through residual blocks. The residual connection structure is shown in Figure 3 (Zhang et al., 2021).

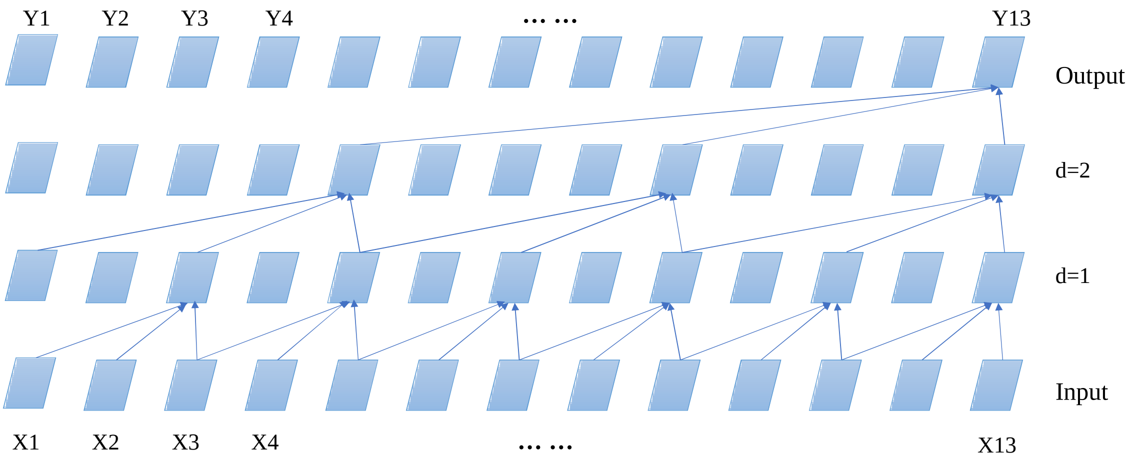

Fig. 2 TCN model structure.

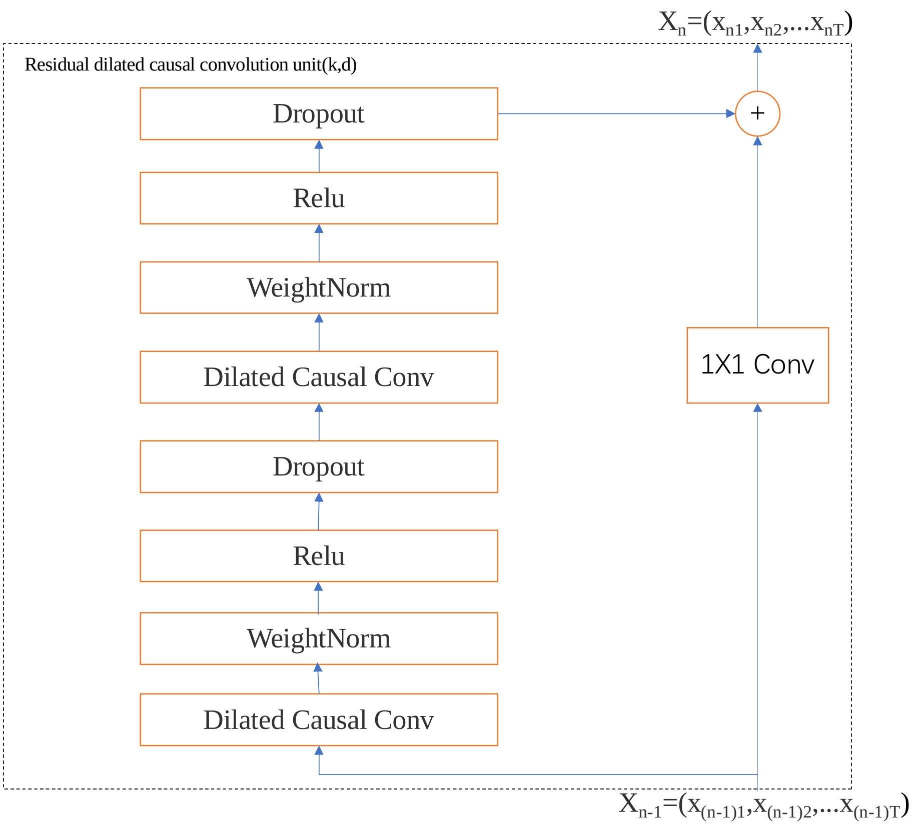

Fig. 3 Residual connection structure.

2.3. MGRU+TCN PREDICTION MODEL

This article uses GRU to extract temporal features of time series, and TCN to extract spatial features of time series. In order to enhance the ability of learning time features from complex time series, the MGRU branch adopts multiple GRU modules and multiple hidden layers, with each GRU module's hidden layer connected to a dropout layer to prevent overfitting. In addition, the concatenate function is used to fuse the features extracted from MGRU and TCN, and the output results are integrated through a fully connected layer. The structure of the MGRU+TCN parallel prediction model is shown in Figure 4. The specific prediction process is as follows.

Step1: Firstly, CEEMDAN (Complete Ensemble Empirical Mode Decomposition with Adaptive Noise) algorithm is used to de-noise the data. Secondly, fixed length sliding window and step size are used to obtain more data samples for the denoised data.

Step2: Split the data samples according to the ratio of 8:1:1 of the training set, validation set and test set, and normalize the split data.

Step3: Input the training set data into the MGRU+TCN parallel network model for training. Every 10 times of training, the model will test the prediction results of the validation set data to determine whether the Root Mean Square Error (RMSE) has decreased. If it has not decreased for 1000 consecutive times, stop training, otherwise continue training.

Step4: Input the test set into MGRU+TCN parallel prediction model to get the prediction results, calculate RMSE, MAE and R^2, and evaluate the prediction performance of the model.

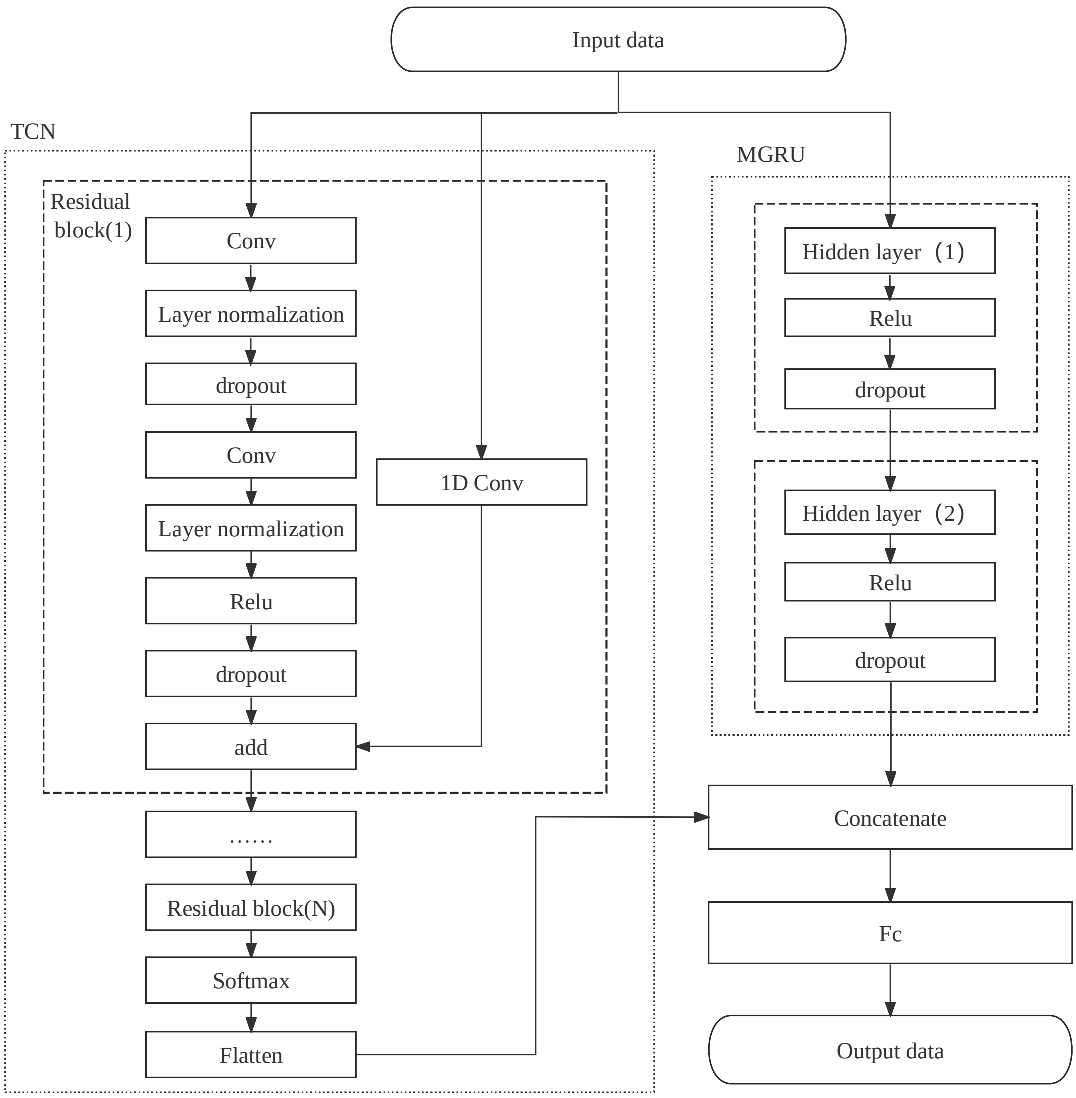

Fig. 4 MGRU+TCN parallel prediction model structure.

2.4. ACCURACY EVALUATION INDEX

In order to evaluate the prediction accuracy of the prediction model, this paper uses RMSE, Mean Absolute Error (MAE), and coefficient of determination (R^2) as evaluation indexes.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \tag{1}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\left(y_i - \hat{y}_i\right)\right| \tag{2}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \overline{y}\right)^2} \tag{3}$$

Where, y_i is the real GNSS data value, $\overline{y}$ is the average of real GNSS data values, $\hat{y}_i$ is the forecast result, n is the number of GNSS data points. The smaller the value of RMSE and MAE, the higher the prediction accuracy of the model. On the contrary, the larger the value, the lower the prediction accuracy of the model. The closer the R^2 value is to 1, the better the predictive model can explain the variability of dependent variables, and the closer it is to 0, the weaker the explanatory ability of the predictive model (Chen et al., 2023b).

3. DATA AND EXPERIMENTS

3.1. DATA SELECTION

The experimental data in this paper selects the daily elevation data from 50 GNSS reference stations within the extended solid earth science ESDR system (Es3). The data from the selected reference stations spans a period from 2010 to 2022, totaling 12 years. To ensure the reliability of model predictions, the data loss rate for the selected stations is maintained below 10 %. The distribution of GNSS reference stations is shown in Figure 5.

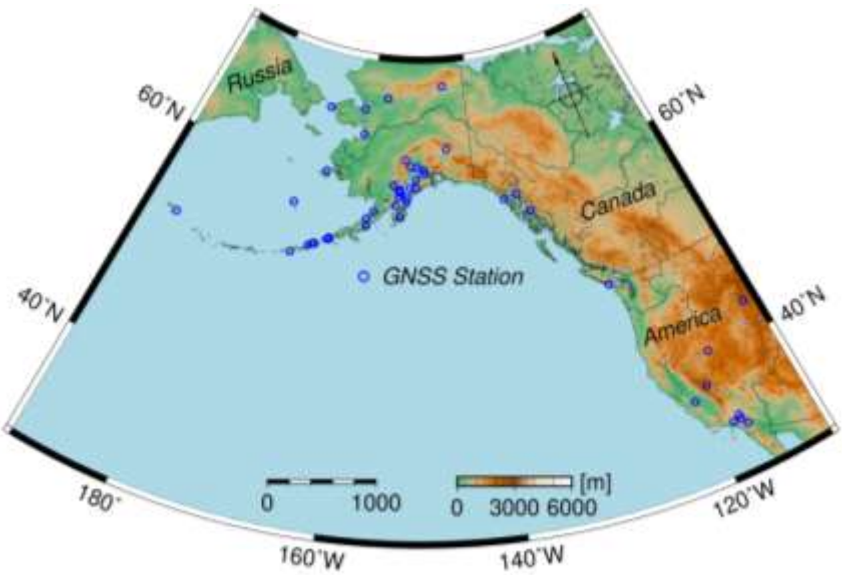

Fig. 5 Distribution of GNSS reference stations.

3.2. DATA PREPROCESSING

Select GNSS site experimental data use Hector software to remove outliers (Williams., 2008). For step discontinuities, use least squares fitting method for correction and Regularized Expectation Maximization (RegEM) algorithm for interpolation (Chen et al., 2023b). Affected by multipath effect, clock error and tropospheric delay, the complex noise in GNSS time series is not conducive to the feature learning of prediction model (Guo et al., 2022). To eliminate the noise's impact, this paper uses CEEMDAN noise reduction algorithm to denoise the data. After dividing the time series into K Intrinsic Mode Functions (IMF), it calculates the correlation coefficient between each IMF and the original time series, and removes the noise by taking the first extreme point of the correlation coefficient as the boundary. The noise reduction results are shown in Figure 6.

As can be seen from Figure 6, the GNSS time series become more stable and retain more useful signals after noise reduction using the CEEMDAN algorithm. Therefore, the denoised GNSS time will be used in subsequent experiments. In the process of sample data expansion, different sliding window lengths affect the prediction accuracy of the model (Guo et al., 2022). After trying different fixed window lengths, we finally selected the window length of 16 and the step size of 1 to expand the data sample. Figure 7 shows the sample expansion process. The

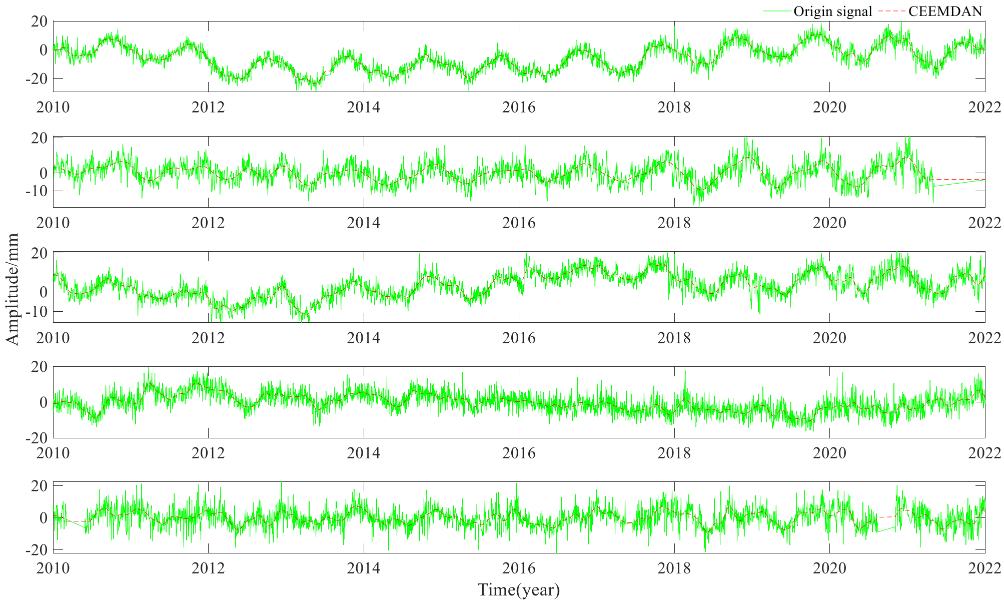

Fig. 6 Noise reduction results of CEEMDAN algorithm.

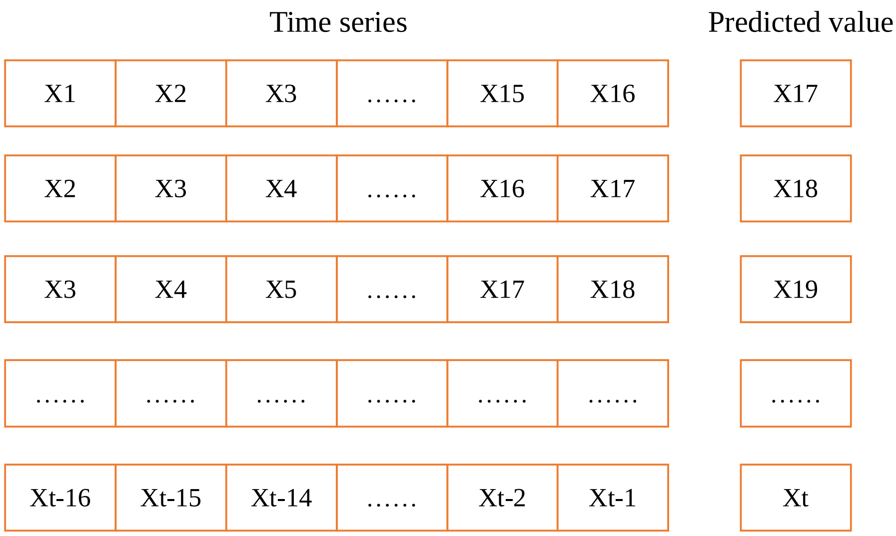

Fig. 7 Sample expansion process with fixed window length of 16 steps as 1.

data is then segmented into a training set, a validation set, and a test set in an 8:1:1 ratio. We apply the Mapminmax normalization technique to enhance the model's convergence rate and overall performance. The data preprocessing flowchart is shown in Figure 8.

3.3. PARAMETER SETTING OF PREDICTION MODEL

In order to better extract the time characteristics of complex time series, we conducted experiments on branches with 1, 2 and 3 GRU modules combined as MGRU respectively. The hidden layer is uniformly set as 256 hidden units, followed by a 0.5 discard layer after each hidden layer. The number of GRU modules of MGRU branches is determined through the comparative analysis of RMSE and MAE accuracy results. Figure 9 shows the prediction accuracy results of the number of different GRU modules of MGRU branches.

As can be seen from Figure 9, with the increase of GRU modules, the RMSE and MAE first become smaller and then larger, and the MGRU combined by two modules is 1.11 and 0.73 mm lower than the RMSE combined by one and three modules; MAE decreased by 1.31 and 0.56 mm. Therefore, we choose the MGRU branch with two GRU modules as the MGRU+TCN model. At the same time, in order to further improve the prediction accuracy of MGRU branches, we try to set the second module as 1, 2 and 3 hidden layers on the basis of the two GRU modules. We also analyzed the prediction results by calculating the RMSE and MAE index values. Table 1 shows the number of hidden units in the hidden layer and the evaluation index results.

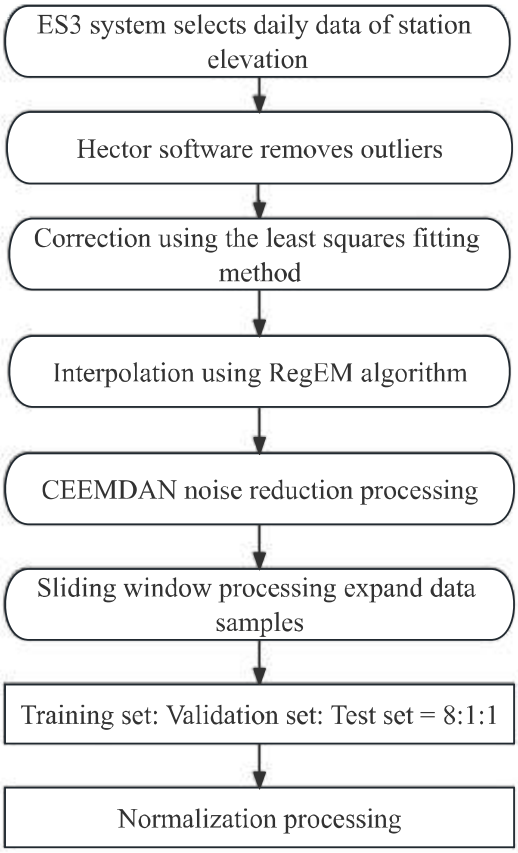

Fig. 8 GNSS station data preprocessing process.

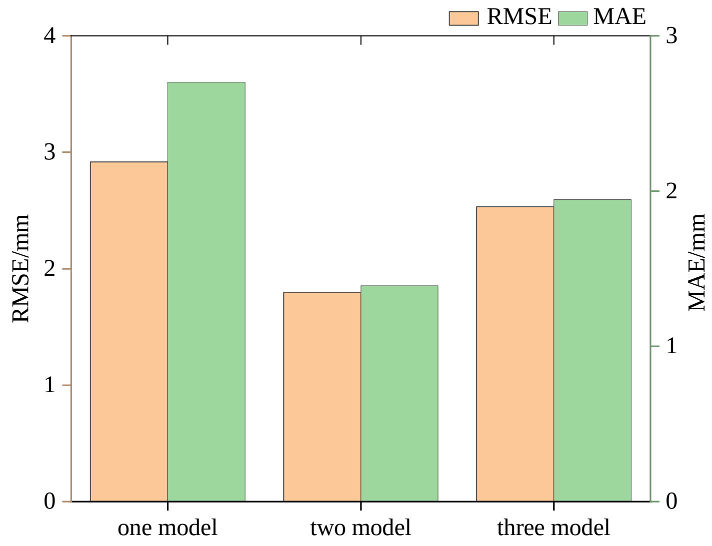

Fig. 9 Prediction accuracy results of MGRU prediction branch under different number of modules.

Table 1 Prediction accuracy results of MGRU prediction branches under different hidden layers.

MGRU	First hidden layer(unit)	Second hidden layer(unit)	Third hidden layer(unit)	RMSE (mm)	MAE (mm)
1	256	-	-	1.80	1.39
2	128	256	-	1.48	1.25
3	64	128	256	1.70	1.46

It can be seen from Table 1 that with the increase of the number of hidden layers, RMSE and MAE also decrease first and then increase. When two hidden layers are used in the second module, the accuracy values of RMSE and MAE are the lowest. Compared to one and three hidden layers, this configuration reduces the RMSE decreases by 0.32 and 0.22 mm respectively; and the MAE decreased by 0.14 and 0.21 mm. Finally, we choose two GRU modules to be combined as MGRU branches, and set two hidden layers in the second GRU module.

The more residual blocks in the TCN branch of MGRU+TCN prediction model, the more complex the model, and the greater the probability of over fitting. Therefore, in order to determine the number of residual blocks, we set N as 2, 3, 4, and 5 for experiments, and set the convolution kernel, channel number, and discard rate as 2, 32, and 0.5, respectively, to calculate the RMSE and MAE of the prediction results. Figure 10 displays the prediction accuracy results for the TCN branch with different number of residual blocks.

As can be seen from Figure 10, when the residual block N is set to 2, the RMSE and MAE values reach the minimum; When N is 5, RMSE and MAE are the largest; At N =4, the RMSE and MAE values have little difference from those at N =2. In order to reduce the complexity of MGRU+TCN parallel prediction model and avoid over fitting problem, TCN branch sets residual block N as 2. The detailed parameters of the final MGRU+TCN parallel prediction model are shown in Table 2.

In Table 2, unit in the MGRU branch represents the number of hidden units in the hidden layer, and the hidden units are lost at a deactivation rate of 0.5 after each GRU module hidden layer. In the TCN branch, the number of residual blocks is 2, and the convolution cores, channels and discard rates of residual block 1 and residual block 2 are 2, 32 and 0.5.

In order to verify the feasibility of the MGRU+TCN parallel prediction model proposed in this paper, LSTM, MGRU, CNN, TCN, TCN-MGRU and MGRU-TCN models are used for comparative analysis. In order to ensure the reliability of the experiment, the convolution kernel, channel number and discard rate of single prediction model CNN are also set to 2, 32 and 0.5. The parameters and structures of TCN, MGRU, TCN-MGRU and MGRU-TCN are the same as those of TCN and MGRU in MGRU+TCN. The number of LSTM hidden layer units is 256, followed by a lost layer with a parameter of 0.5.

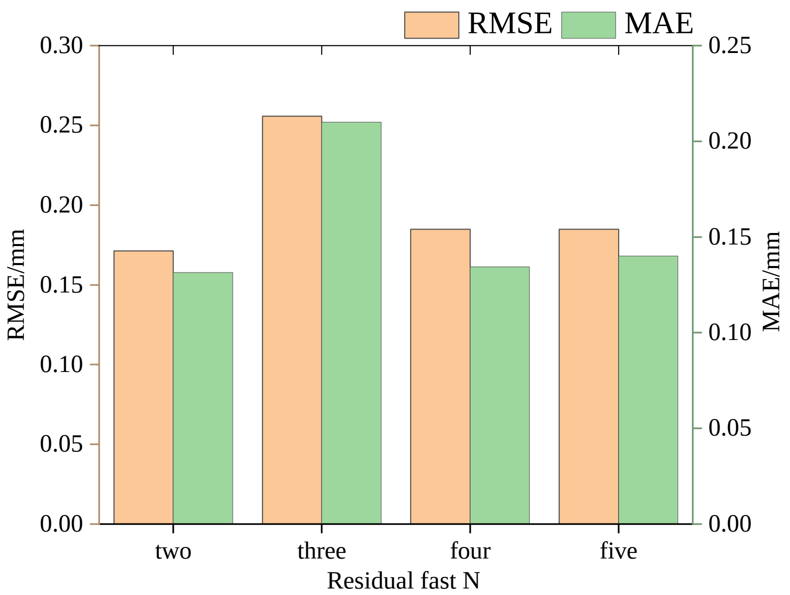

Fig. 10 Prediction accuracy results of different residual blocks N in TCN prediction branch.

Table 2 Parameters of MGRU+TCN parallel prediction model.

MGRU+TCN	Parameters	
MGRU	Module 1	Unit (Hidden layer1) = 256, Dropout=0.5
	Module 2	Unit (Hidden layer2) = 128, Unit (Hidden layer3) = 256, Dropout=0.5
TCN		Convolutional kernel size=2, Number of channels =32, Residual block N=2, Dropout=0.5

In the model training superparameters, Adam optimization algorithm is used for gradient optimization; Epochs is 50 times; The initial learning rate was 0.01. The input and output dimensions are 1. To avoid fitting problems during model training, we set L2 regularization coefficient as 0.001; In order to ensure that the model training is sufficient and there is overfitting, we evaluate the performance of the model by bringing in the validation set every 10 iterations. If the RMSE (loss value) between the predicted value and the real value does not decrease for 1000 consecutive times, stop training to get the final prediction model.

3.4. ANALYSIS OF EXPERIMENTAL RESULTS

To verify the feasibility of the MGRU+TCN prediction model, we processed the GNSS reference station time series mentioned in Chapter 3.1 through CEEMDAN denoising, sample expansion, and normalization, using LSTM, CNN, MGRU, TCN, TCN-MGRU, MGRU-TCN and MGRU+TCN prediction models for prediction. Figure 11 shows the comparison between the predicted results and the true values of different validation and testing sets of prediction models for site av34.

The LSTM model in Figure 11 (a) closely follows the real value's, fluctuation trend, although the predictions are generally higher; The prediction accuracy of CNN and TCN models in Figure 11 (b) and (d) is improved compared with LSTM, but there is still a certain gap between the predicted value and the real value at the trough; In Figure 11 (c), the MGRU prediction model has good prediction result at the peak, but the error between the predicted value and the real value is large at the trough; The TCN-MGRU and MGRU-TCN models in Figure 11 (e) and Figure 11 (f) extract features from one model as input to another model, and their prediction results deviate greatly from the true values, even worse than the single TCN model; the MGRU+TCN model proposed in this paper has the best prediction result, which is very close to the real value at the trough, peak and continuous fluctuation of the waveform. In order to further verify the effectiveness and applicability of the MGRU+TCN prediction model, we calculated the accuracy index values of the prediction results of LSTM, CNN, MGRU, TCN, TCN-MGRU, MGRU-TCN and MGRU+TCN models, and calculated the improvement I of the evaluation index of MGRU+TCN model compared with other models. Table 3 presents the prediction accuracy indices; while Figure 12 illustrates the absolute error as a box plot for each model's predictions against actual values.

Table 3 shows that compared to LSTM, CNN, MGRU, TCN, TCN-MGRU and MGRU-TCN, the RMSE of the MGRU+TCN parallel prediction model proposed in this paper decreased by 1.05, 0.65, 1.12, 0.30, 0.72 and 0.54 mm respectively; MAE decreased by 0.91, 0.58, 0.94, 0.11, 0.55, 0.43 mm respectively; R^2 increased by 0.16, 0.07, 0.12, 0.01, 0.07, 0.03 respectively. The prediction error of MGRU and TCN models with single branch is larger than that of MGRU+TCN parallel structure model, indicating that the prediction network with single branch structure has room for improvement. For TCN-MGRU and MGRU-TCN serial hybrid model, the accuracy is not significantly improved compared with that of single model structure, even compared with TCN model, RMSE increases by 26.35 %, 45.79 %, MAE increases by 44.78 %, 41.91 %, and R^2 decreases by 6 %. In

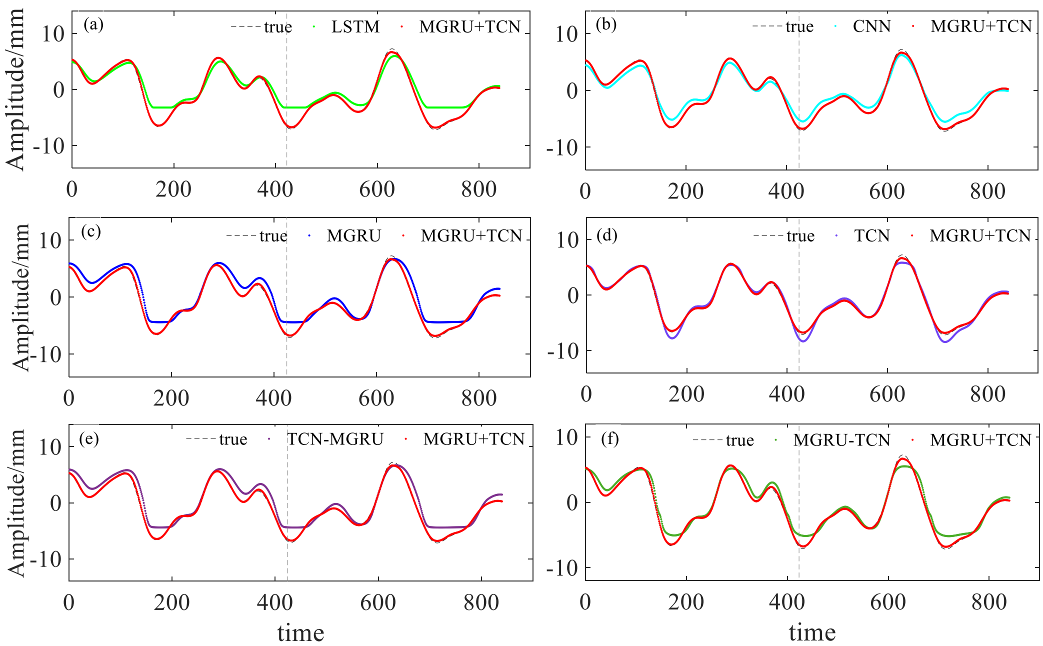

Fig. 11 Comparison of prediction results of single model LSTM, CNN, MGRU, TCN, TCN-MGRU, MGRU-TCN and MGRU+TCN parallel prediction model (av34gnss reference station as an example)

Table 3 Prediction accuracy of different models

Site	Model	RMSE(mm)	I/%	MAE(mm)	I%	R^2
	LSTM	1.53	59.48	1.28	60.94	0.94
	CNN	1.31	52.67	1.09	54.13	0.97
	MGRU	1.63	61.96	1.31	61.83	0.94
ab43	TCN	1.09	43.12	0.81	38.27	0.97
	TCN-MGRU	1.65	62.42	1.28	60.94	0.93
	MGRU-TCN	1.21	48.76	0.98	48.98	0.98
	MGRU+TCN	0.62	/	0.50	/	0.99
	LSTM	1.22	71.31	1.02	72.55	0.91
	CNN	0.85	58.82	0.69	59.42	0.96
	MGRU	1.56	77.56	1.28	78.13	0.82
ab27	TCN	0.72	51.39	0.55	49.09	0.98
	TCN-MGRU	1.20	70.83	0.91	69.23	0.91
	MGRU-TCN	1.10	68.18	0.83	66.27	0.94
	MGRU+TCN	0.35	/	0.28	/	0.99
	LSTM	1.15	60.00	1.00	64.00	0.88
	CNN	1.33	65.41	1.10	67.27	0.83
	MGRU	1.58	70.89	1.38	73.91	0.84
ac03	TCN	0.52	11.54	0.41	12.20	0.98
	TCN-MGRU	0.76	39.47	0.61	40.98	0.96
	MGRU-TCN	0.80	42.50	0.65	44.62	0.96
	MGRU+TCN	0.46	/	0.36	/	0.99
	LSTM	1.73	83.82	1.64	85.98	0.72
	CNN	0.73	61.64	0.57	59.65	0.93
	MGRU	1.37	79.56	1.16	80.17	0.88
av06	TCN	0.54	48.15	0.29	20.69	0.99
	TCN-MGRU	0.88	68.18	0.70	67.14	0.93
	MGRU-TCN	0.61	54.10	0.50	54.00	0.97
	MGRU+TCN	0.28	/	0.23	/	0.99
	LSTM	1.54	87.01	1.15	86.09	0.72
	CNN	0.96	79.17	0.98	83.67	0.90
	MGRU	1.36	85.29	1.09	85.32	0.87
av34	TCN	0.54	62.96	0.41	60.98	0.98
	TCN-MGRU	1.04	80.77	0.80	80.00	0.90
	MGRU-TCN	0.91	78.02	0.72	77.78	0.94
	MGRU+TCN	0.20	/	0.16	/	1.00

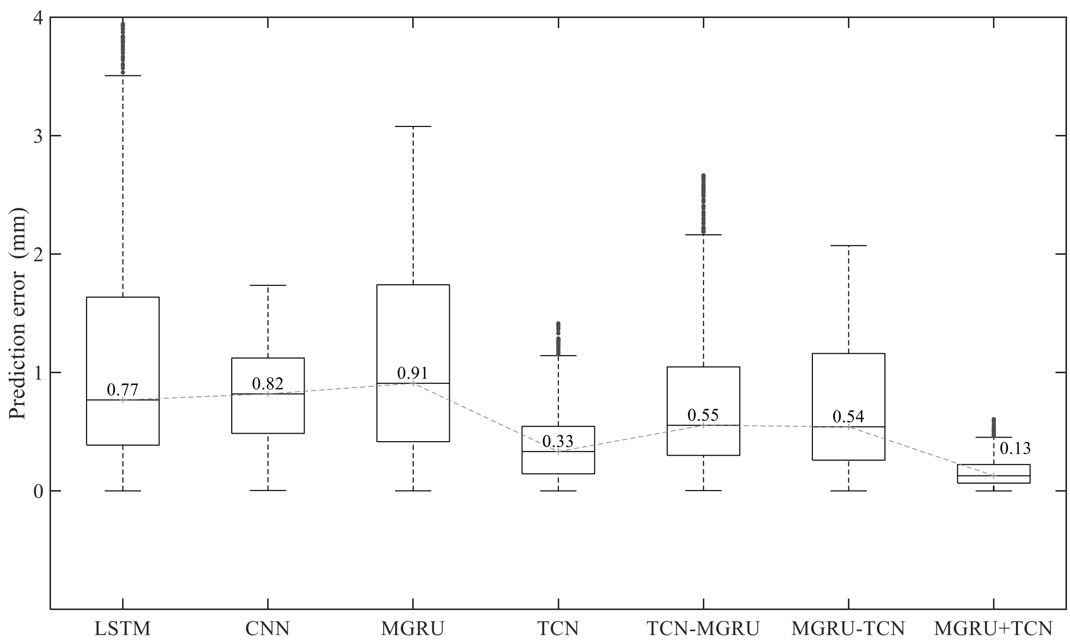

Fig. 12 Absolute error box diagram of different prediction algorithms (av34gnss reference station as an example).

Figure 12, the narrower the top and bottom width of the box chart is, the more convergent the prediction error is. The top and bottom width of the box chart of MGRU+TCN is the narrowest, and the minimum median value is 0.13, followed by the TCN model. It shows that the prediction network of MGRU+TCN parallel structure proposed in this paper is feasible. Compared with the model combination of TCN-MGRU and MGRU-TCN serial structure, MGRU+TCN improves the low prediction accuracy result of single model, and the learning characteristics are more comprehensive. The difference between the prediction result and the real value is small, so MGRU+TCN has high prediction ability.

4. CONCLUSION

Addressing the limitations of single prediction models, which include low accuracy and incomplete feature extraction, this paper designs a model combination mode, and finally proposes a prediction model with MGRU+TCN parallel structure. Experiments utilizing the elevation time series data from 50 GNSS reference stations over 2010 to 2022 have led to the following conclusions:

1. The MGRU branch of MGRU+TCN model incorporates two GRU units and includes two hidden layers within the second GRU unit. This configuration enhances the RMSE and MAE accuracy by 48.85 % and 54.11 %, respectively, compared to a single GRU model, demonstrating that the MGRU's multi-module and multi-layer design more effectively captures temporal features of the time series.
2. The MGRU+TCN prediction model leverages the MGRU branch to extract temporal features and the TCN branch to extract spatial features, fusing them using a concatenation function. Compared with the LSTM, CNN, MGRU, TCN single model and the TCN-MGRU, MGRU-TCN serial structure prediction model, the RMSE accuracy is improved by about 63.20 % on average, the MAE accuracy is improved by about 62.77 % on average, and the R^2 can reach between 0.99-1.00. The model achieves the smallest overall absolute error between predictions and actual values, indicating more comprehensive learning capabilities and superior predictive performance.

ACKNOWLEDGEMENT

This work was sponsored by the National Natural Science Foundation of China (42264001, 42204047, 42364002), the Open Research Program of Key Laboratory of Marine Environmental Survey Technology and Application, Ministry of Natural Resources (MESTA-2020-A004), the Major Discipline Academic and Technical Leaders Training Program of Jiangxi Province (20225BCJ23014), Hebei Water Conservancy Research Plan (2022-28).

REFERENCES

Altamimi, Z., Rebischung, P., Métivier, L. and Collilieux, X.: 2016, ITRF2014: A new release of the International Terrestrial Reference Frame modeling nonlinear station motions. J. Geoph. Res., Solid Earth, 121, 8, 6109–6131. DOI: 10.1002/2016JB013098

Chen, H., He, X., Lu, T., Huan, J. and Sun, X.: 2023a, An improved VMD–EEMD–LSTM time series hybrid prediction model for sea surface height derived from satellite altimetry data. J. Mar. Sci. Eng., 11, 12, 2386. DOI: 10.3390/jmse11122386

Chen, H., Lu, T., Huang, J., He, X., Yu, K., Sun, X., ... and Huang, Z.: 2023b, An improved VMD-LSTM model for time-varying GNSS time series prediction with temporally correlated noise. Remote Sens., 15, 14, 3694. DOI: 10.3390/rs15143694

Chen, M. and Zhang, Q.: 2021, Analysis of positioning deviation between Beidou and GPS based on National Reference Stations in China. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., 43, 209–214. DOI: 10.5194/isprs-archives-XLIII-B1-2021-209-2021

Chen, Q., van Dam, T., Sneeuw, N., Collilieux, X., Weigelt, M. and Rebischung, P.: 2013, Singular spectrum analysis for modeling seasonal signals from GPS time series. J. Geodyn., 72, 25–35. DOI: 10.1016/j.jog.2013.05.005

Chen, Q., Jiang, W., Meng, X., Jiang, P., Wang, K., Xie, Y. and Ye, J.: 2018, Vertical deformation monitoring of the suspension bridge tower using GNSS: A case study of the forth road ridge in the UK. Remote Sens., 10, 3, 364. DOI: 10.3390/rs10030364

Cho, K.: 2014, On the Properties of Neural Machine Translation: Encoder-decoder Approaches. arXiv preprint arXiv, 1409.1259.

Cina, A. and Piras, M.: 2015, Performance of low-cost GNSS receiver for landslides monitoring: Test and results. Geomat. Nat. Hazards Risk, 6, 5-7, 497–514. DOI: 10.1080/19475705.2014.889046

Dey, R. and Salem, F.M.: 2017, Gate-variants of gated recurrent unit (GRU) neural networks. In: 2017 IEEE 60th International Midwest Symposium on Circuits and Systems, 1597–1600. DOI: 10.1109/MWSCAS.2017.8053243

Fernandes, R.M.S., Ambrosius, B.A.C., Noomen, R., Bastos, L., Combrinck, L., Miranda, J. M. and Spakman, W.: 2004, Angular velocities of Nubia and Somalia from continuous GPS data: implications on present-day relative kinematics. Earth Planet. Sci. Lett., 222, 1, 197–208. DOI: 10.1016/j.epsl.2004.02.008

Fu, R., Zhang, Z. and Li, L.: 2016, Using LSTM and GRU neural network methods for traffic flow prediction. In: 2016 31st Youth Academic Annual Conference of Chinese Association of Automation, 324–328. DOI: 10.1109/YAC.2016.7804912

Guo, J., Wang, W., Tang, Y., Zhang, Y. and Zhuge, H.: 2022, A CNN-Bi_LSTM parallel network approach for train travel time prediction. Knowl-Based Syst., 256, 109796. DOI: 10.1016/j.knosys.2022.109796

Hewage, P., Behera, A., Trovati, M., Pereira, E., Ghahremani, M., Palmieri, F. and Liu, Y.: 2020, Temporal convolutional neural (TCN) network for an effective weather forecasting using time-series data from the local weather station. Soft Comput., 24, 16453–16482. DOI: 10.1007/s00500-020-04954-0

Irie, K., Tüske, Z., Alkhouli, T., Schlüter, R. and Ney, H.: 2016, LSTM, GRU, highway and a bit of attention: An empirical overview for language modeling in speech recognition. In: Interspeech 2016, 3519–3523. DOI: 10.21437/Interspeech.2016-491

Kong, Q., Zhang, L., Han, J., Li, C., Fang, W. and Wang, T.: 2023, Analysis of coordinate time series of DORIS stations on Eurasian plate and the plate motion based on SSA and FFT. Geod. Geodyn., 14, 1, 90–97. DOI: 10.1016/j.geog.2022.05.001

Lahtinen, S., Jivall, L., Häkli, P., Kall, T., Kollo, K., Kosenko, K., ... and Nordman, M.: 2019, Densification of the ITRF2014 position and velocity solution in the Nordic and Baltic countries. GPS Solut., 23, 1–13. DOI: 10.1007/s10291-019-0886-3

Li, C., Li, G., Wang, K. and Han, B.: 2022, A multi-energy load forecasting method based on parallel architecture CNN-GRU and transfer learning for data deficient integrated energy systems. Energy, 259, 124967. DOI: 10.1016/j.energy.2022.124967

Li, S., Li, W., Cook, C., Zhu, C. and Gao, Y.: 2018, Independently recurrent neural network (IndRNN): Building a longer and deeper RNN. In: Computer Vision and Pattern Recognition 2018, 5457–5466. DOI: 10.1109/CVPR.2018.00572

Li, Z., Lu, T., He, X., Montillet, J.P. and Tao, R.: 2023, An improved cyclic multi model-eXtreme gradient boosting (CMM-XGBoost) forecasting algorithm on the GNSS vertical time series. Adv. Space Res., 71, 1, 912–935. DOI: 10.1016/j.asr.2022.08.038

Li, Z., Chen, W., van Dam, T., Rebischung, P., Altamimi, Z.: 2020, Comparative analysis of different atmospheric surface pressure models and their impacts on daily ITRF2014 GNSS residual time series. J. Geod., 94, 42. DOI: 10.1007/s00190-020-01370-y

Montillet, J.P., Williams, S.D.P., Koulali, A. and McClusky, S.C.: 2015, Estimation of offsets in GPS time-series and application to the detection of earthquake deformation in the far-field. Geophys. J. Int., 200, 2, 1207–1221. DOI: 10.1093/gji/ggu473

Ohta, Y., Kobayashi, T., Tsushima, H., Miura, S., Hino, R., Takasu, T., ... and Umino, N.: 2012, Quasi real-time fault model estimation for near-field tsunami forecasting based on RTK-GPS analysis: Application to the 2011 Tohoku-Oki earthquake (Mw 9.0). J. Geophys. Res., Solid Earth, 117(B2). DOI: 10.1029/2011JB008750

Serpelloni, E., Faccenna, C., Spada, G., Dong, D. and Williams, S.D.: 2013, Vertical GPS ground motion rates in the Euro-Mediterranean region: New evidence of velocity gradients at different spatial scales along the Nubia-Eurasia plate boundary. J. Geophys. Res., Solid Earth, 118, 11, 6003–6024. DOI: 10.1002/2013JB010102

Shen, N., Chen, L. and Chen, R.: 2022, Displacement detection based on Bayesian inference from GNSS kinematic positioning for deformation monitoring. Mech. Syst. Signal Process., 167, 108570. DOI: 10.1016/j.ymssp.2021.108570

Shen, N., Chen, L., Liu, J., Wang, L., Tao, T., Wu, D. and Chen, R.: 2019, A review of global navigation satellite system (GNSS)-based dynamic monitoring technologies for structural health monitoring. Remote Sens., 11, 9, 1001. DOI: 10.3390/rs11091001

Shen, N., Chen, L., Wang, L., Hu, H., Lu, X., Qian, C.,….Chen, R.: 2021, Short-term landslide displacement detection based on GNSS real-time kinematic positioning. IEEE Trans. Instrum. Meas., 70, 1–14. DOI: 10.1109/TIM.2021.3055278

Wang, J., Jiang, W., Li, Z. and Lee, Y.: 2021, A new multi-scale sliding window LSTM framework (MSSW-LSTM): a case study for GNSS time-series prediction. Remote Sens., 13, 16, 3328. DOI: 10.3390/rs13163328

Williams, S.D.: 2008, CATS: GPS coordinate time series analysis software. GPS Solut., 12, 147–153. DOI: 10.1007/s10291-007-0086-4

Xi, R., Jiang, W., Meng, X., Zhou, X. and He, Q.: 2018, Rapid initialization method in real-time deformation monitoring of bridges with triple-frequency BDS and GPS measurements. Adv. Space Res., 62, 5, 976–989. DOI: 10.1016/j.asr.2018.06.018

Xin, J., Zhou, J., Yang, S.X., Li, X. and Wang, Y.: 2018, Bridge structure deformation prediction based on GNSS data using Kalman-ARIMA-GARCH model. Sensors, 18, 1, 298. DOI: 10.3390/s18010298

Younes, S.A.: 2023, Study of crustal deformation in Egypt based on GNSS measurements. Surv. Rev., 55, 391, 338–349. DOI: 10.1080/00396265.2022.2099690

Yu, Y., Si, X., Hu, C. and Zhang, J.: 2019, A review of recurrent neural networks: LSTM cells and network architectures. Neural Comput., 31, 7, 1235–1270. DOI: 10.1162/neco_a_01199

Zhang, R., Sun, F., Song, Z., Wang, X., Du, Y. and Dong, S.: 2021, Short-term traffic flow forecasting model based on GA-TCN. J. Adv. Transp., 2021, 1, 1338607. DOI: 10.1155/2021/1338607