Research in Astronomy and Astrophysics, 24:095019 (11pp), 2024 September https://doi.org/10.1088/1674-4527/ad74dd

Leveraging the Empirical Wavelet Transform in Combination with Convolutional LSTM Neural Networks to Enhance the Accuracy of Polar Motion Prediction

Xu-Qiao Wang[1], Lan Du[1], Zhong-Kai Zhang[1,2], Ze-Jun Liu[1], and Hao Xiang[1]
[1] College of Geospatial Information, Information Engineering University, Zhengzhou 450001, China; zhongkai@ashn.org.cn
[2] Henan Industrial Technology Academy of Spatio-Temporal Big Data, Zhengzhou 450046, China
Received 2024 April 25; revised 2024 August 9; accepted 2024 August 21; published 2024 September 18

Abstract

High-precision polar motion prediction is of great significance for deep space exploration and satellite navigation. Polar motion is affected by a variety of excitation factors, and nonlinear prediction methods are more suitable for polar motion prediction. In order to explore the effect of deep learning in polar motion prediction. This paper proposes a combined model based on empirical wavelet transform (EWT), Convolutional Neural Networks (CNN) and Long Short Term Memory (LSTM). By training and forecasting EOP 20C04 data, the effectiveness of the algorithm is verified, and the performance of two forecasting strategies in deep learning for polar motion prediction is explored. The results indicate that recursive multi-step prediction performs better than direct multi-step prediction for short-term forecasts within 15 days, while direct multi-step prediction is more suitable for medium and long-term forecasts. In the 365 days forecast, the mean absolute error of EWT-CNN-LSTM in the *X* direction and *Y* direction is 18.25 mas and 15.78 mas, respectively, which is 23.5% and 16.2% higher than the accuracy of Bulletin A. The results show that the algorithm has a good effect in medium and long term polar motion prediction.

Key words: methods: data analysis – methods: miscellaneous – astrometry – reference systems – Earth

1. Introduction

The Earth Rotation Parameters (ERP) are the main transformation parameters that link the Earth Reference Frame and the Celestial Reference Frame (Kalarus et al. 2010). These parameters include polar motion(PM), the difference between universal Time (UT1) and Coordinated Universal Time (UTC), and Length of Day (LOD) (Petit & Luzum 2010). With the continuous development of modern geodesy, leading to a significant improvement in the observation accuracy of ERP (Dow et al. 2005). Currently, there is a growing demand for real-time prediction of ERP in deep space exploration and satellite navigation fields, However, existing measurement techniques result in delays in ERP products (Bachmann et al. 2016), making high-precision ERP forecasting particularly important, the Earth Orientation Center (EOC) at Paris Observatory provides comprehensive EOP C04 series on a regular basis (Bizouard et al. 2019), while the International Earth Rotation and Reference Systems Service Rapid Service/Prediction Center in Washington regularly provides Bulletin A and Bulletin B reports, greatly facilitating research on Earth rotation.

Polar motion is the instantaneous movement of the Earth's rotation axis relative to the Earth's surface or the mean pole, caused by the interaction between mass redistribution on the Earth's surface and material movements within the Earth (Gross 2007; Sun et al. 2019; Dobslaw et al. 2010). Polar motion consists of three significant components: long-term drift, Chandler wobble (CW) (Chandler 1891; Zharkov & Molodensky 1996), and annual wobble(AW). The long-term drift rate is approximately 3.5 mas per year, Chandler wobble varies in range from 100 to 200 mas, and annual wobble is about 100 mas (Gross 2007; Su et al. 2014). Prediction models for polar motion are typically divided into linear prediction models and nonlinear prediction models, with the classic Least Squares and Autoregressive (LS+AR) prediction model belonging to linear prediction models. Nonlinear prediction methods mainly include artificial neural networks, wavelet decomposition algorithms, etc. Due to the influence of various excitation factors on ERP and its strong occasional abruptness, nonlinear prediction methods have better effects. Artificial neural networks have great potential in long-term prediction of ERP (Egger 1992), with results proving that artificial neural networks can effectively predict Earth rotation parameters (Schuh et al. 2002). Long Short-Term Memory Network (LSTM) has been widely used in time series forecasting, with some scholars applying LSTM models to ERP forecasting as well (Gou et al. 2023). Convolutional Neural Network easily captures potential features of multi-dimensional data and effectively solves uncertainties brought by original time series. IERS provides a large amount of

Figure 1. Flowchart of the EWT-CNN-LSTM Prediction Model.

historical polar motion data which requires more complex deep learning models for predictions. For example, the hybrid deep model CNN-LSTM has better prediction performance (Xue et al. 2019; Khodabakhsh et al. 2020; Mehtab et al. 2020). This paper proposes a combination method based on Empirical Wavelet Transform, Convolutional Neural Network and Long Short-Term Memory (EWT-CNN-LSTM), This model uses EWT to decompose polar motion data into clusters containing individual features, and then inputs these clusters into CNN, followed by using LSTM for predictions.

The structure of this article is as follows: Section 2 introduces the process and theoretical basis of the EWT-CNN-LSTM algorithm. Section 3 presents the data set and detailed procedures, evaluation metrics used in the experiments. In Section 4, recursive multi-step prediction and direct multi-step prediction are employed for PM prediction using PM data from IERS EOP 20C04 from 2008 to 2018. The accuracy of PM results predicted by different models such as LSTM, CNN-LSTM, and EWT-CNN-LSTM is compared with the forecast results of IERS Bulletin A. To better analyze the univariate forecasting performance of the models, this study does not consider information such as the effective angular momentum (EAM). Finally, a summary of the experiments is provided in the last section.

2. Method

2.1. Construction of the EWT-CNN-LSTM Model

This paper proposes a hybrid prediction model based on EWT-CNN-LSTM to optimize the extreme shift prediction results, as shown in Figure 1. The main steps of the EWT-CNN-LSTM algorithm are as follows:

(1) Empirical wavelet decomposition: The input PM time series is decomposed into several component signals

using empirical wavelets, thereby reducing the complexity of the original time series.

(2) K-means clustering: The component signals obtained in step 1 are clustered into high-frequency, mid-frequency, and low-frequency sub-sequences based on their entropy values using K-means clustering.

(3) CNN-LSTM model training and prediction: The three different frequency sub-sequences obtained in step 2 are separately inputted into the CNN-LSTM model for training and prediction. The predicted results are then reconstructed to obtain the final prediction result.

2.2. Empirical Wavelet Transform

The Empirical Wavelet Transform is a method for processing non-stationary signals, proposed by Gilles in 2013 (Gilles 2013). It combines the adaptive decomposition characteristics of Empirical Mode Decomposition (EMD) and the advantages of the tight support wavelet transform theory. EWT can flexibly select frequency bands, effectively solving the mode mixing problem in EMD, and has lower computational complexity.

The specific procedures of the EWT algorithm entail initially converting the original signal into the corresponding Fourier spectrum. Simultaneously, the spectrum is rationally partitioned in a prescribed manner to acquire a series of consecutive intervals. Subsequently, a group of wavelet filter banks is constructed based on the interval ranges, and this filter bank is utilized to identify the positions of the characteristic information within the spectrum, completing the filtering process of the original signal. Consequently, a set of frequency-modulated and amplitude-modulated components of the original signal is adaptively obtained.

For a given signal, the first step is to perform Fourier transform to obtain the normalized Fourier spectrum within the range of 2π. The study range is [0, π], assuming that the signal consists of N individual components. In the domain of Fourier, it is divided into N continuous segments by ω_n with ω_n representing the boundaries between each interval, totaling $N+1$ boundaries where $\omega_0 = 0$ and $\omega_n = \pi$.

Empirical wavelet is a bandpass filter defined on interval Λ_n,. When constructing bandpass filters, Gilles adopts the ideas of Littlewood-Paley and Meyer in constructing wavelets. When $n > 0$, empirical wavelet function $\hat{\psi}_n(\omega)$ and $\hat{\phi}_n(\omega)$ empirical scale function are defined as follows:

$$\hat{\psi}_n(\omega) = \begin{cases} 1 & \omega_n + \tau_n \leqslant |\omega| \leqslant \omega_{n+1} - \tau_{n+1} \\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\tau_{n+1}}(|\omega| - \omega_{n+1} + \tau_{n+1})\right)\right] & \omega_{n+1} - \tau_{n+1} \leqslant |\omega| \leqslant \omega_{n+1} + \tau_{n+1} \\ \sin\left[\frac{\pi}{2}\beta\left(\frac{1}{2\tau_n}(|\omega| - \omega_n + \tau_n)\right)\right] & \omega_n - \tau_n \leqslant |\omega| \leqslant \omega_n + \tau \\ 0 & \text{others} \end{cases} \quad (1)$$

$$\hat{\phi}_n(\omega) = \begin{cases} 1 & |\omega| \leqslant \omega_n + \tau_n \\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\tau_n}(|\omega| - \omega_n + \tau_n)\right)\right] & \omega n\tau_n \leqslant |\omega| \leqslant \omega_n + \tau_n \\ 0 & \text{others} \end{cases} \quad (2)$$

After acquiring the bandpass filter, it is possible to construct the frequency division signal of the original signal. The empirical wavelet function and scale function are internally integrated into the signal, followed by obtaining the empirical wavelet coefficient and scale coefficient, ultimately resulting in the decomposed mode.

2.3. Sample Entropy and K-means Clustering Algorithm

Sample Entropy (SampEn/SE) is a measure of the disorder of data. Sample entropy has the advantage of being computationally independent of data length and having better consistency. Numerically, a lower sample entropy value indicates higher self-similarity in a time series, while a higher sample entropy value suggests greater complexity and less obvious regularity in the sequence. In practical applications, it can be used for feature selection, classification recognition, clustering and other data processing to improve accuracy and reliability in data analysis.

The k-means algorithm is based on the idea of clustering data points in space around k centroids, classifying each point to the nearest centroid, and iteratively updating the centroids until the best clustering result is obtained. The algorithm follows these steps:

(1) Determine the number of clusters, k, and set k points as initial centroids.

(2) Calculate the distance between each element and the k centroids, assigning each element to its closest cluster.

(3) Recalculate the centroids based on the new clustering result and reassign elements to clusters.

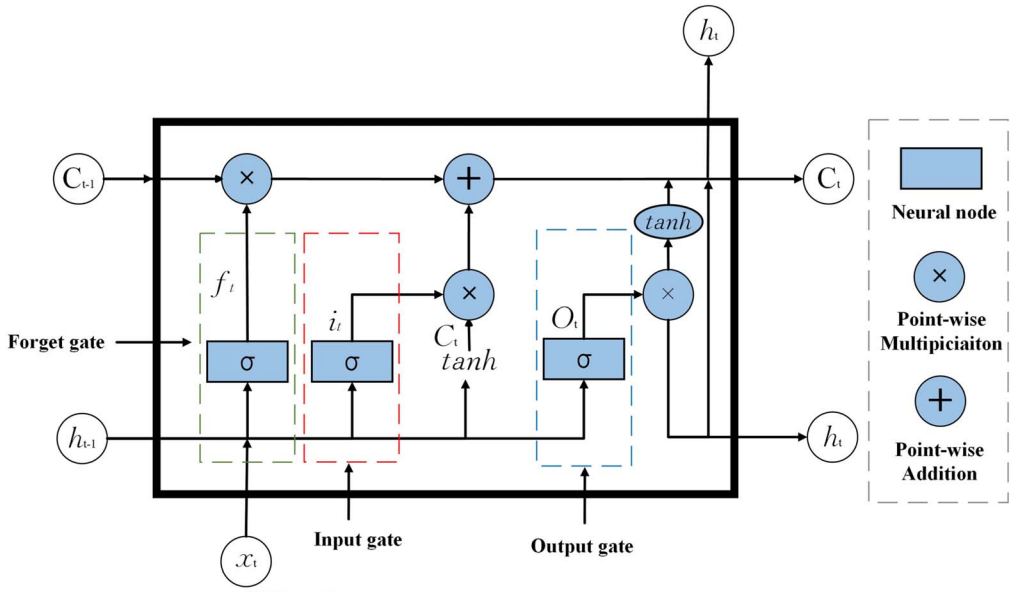

Figure 2. LSTM Structure Diagram.

(4) Iterate step (3) until no new clustering results are obtained.

2.4. CNN-LSTM

LSTM is a variant of recurrent neural network (RNN) (Hochreiter & Schmidhuber 1997), compared to other network models, the LSTM algorithm is specifically designed to handle time series data by introducing structures such as forget gates, input gates, and output gates. These structures enable the model to have long-term memory capabilities, allowing it to forget irrelevant information while effectively retaining and updating important information (Gers et al. 2000; Gou et al. 2023), thus avoiding the issues of vanishing and exploding gradients. Figure 2 illustrates the basic structure of LSTM, where each unit inherits information from the previous time step. As a result, LSTM is more sensitive to long-term patterns and has been widely applied in time series prediction tasks (Zhang et al. 2019; Karevan & Suykens 2020; Wang et al. 2021).

CNN is commonly used for image processing (Krizhevsky et al. 2012). In PM time series, CNN can identify local features hidden in the data and transform them to improve computational efficiency and model accuracy. Additionally, pooling operations can reduce the number of features and alleviate overfitting issues.

The combination of CNN-LSTM leverages the advantages of both approaches. In time series prediction, it typically achieves higher prediction accuracy than using LSTM alone by effectively extracting features and utilizing temporal dependencies. Furthermore, it demonstrates better generalization capabilities across different data sets.

3. Design of PM Prediction Scheme

3.1. Data Description

This study primarily utilizes PM data, The training data is sourced from EOP 20C04, and the results are compared using Bulletin A. The EOP 20C04 series is composed of GNSS, SLR, VLBI, and DORIS combinations and has been in existence since 1962. The Bulletin A product is provided by IERS to global users in the form of a bulletin for Earth orientation parameter forecast measurements and prediction services. It is solely composed of GNSS and VLBI combinations and is regularly published by the IERS Rapid Service Center through A and B bulletins. Bulletin A data is released weekly, including leap seconds, measured values, predicted values, among other information. Based on these characteristics, this study uses EOP 20 C04 data as training data and employs Bulletin A data for error verification in predicting results.

3.2. Design of Prediction Strategies and Period

In the time series prediction of deep learning, it is usually divided into single-step prediction and multi-step prediction. Single-step prediction means that only one future value is predicted for each input window at a time. However, in practical applications, the prediction task requires predicting values for multiple future time steps, which is known as multi-step prediction. Currently, the commonly used multi-step prediction strategies are direct multi-step prediction and recursive multi-step prediction. The direct method involves developing a separate model for each time step to predict multiple values. On the other hand, recursive multi-step

prediction essentially belongs to single-step prediction, where multiple uses of single-step models are employed to achieve the goal of predicting multiple values. In this approach, the predicted value from the previous time step serves as the input for predicting the next time step. As the range of predictions increases, using predicted values instead of observed values leads to accumulating errors in recursive.

To investigate the effectiveness of the two methods mentioned above in polar motion prediction, different strategies for predicting polar motion at different periods were determined. The LSTM direct method and recursive method were used to predict different forecast lengths, and the prediction cycle boundaries for the direct and recursive methods were identified.

For medium to long-term polar motion prediction, four schemes were designed for direct multi-step prediction using LSTM, CNN, CNN-LSTM, EWT-CNN-LSTM to forecast polar motion for 15, 30, 60, 90, 180, and 365 days. Bulletin A product was also included for comparison.

Due to the higher uncertainty of early data compared to recent data, early data was removed. The EOP 20C04 from 2008 January 1 to 2018 December 31 with a total of eleven years' worth of data was used as the base data set at a sampling rate of one day. This provided a total of 4020 time series data samples for model training and forecasting inputs.

Considering the possibility of interruptions or unavailability of final Earth Orientation Parameters products provided by IERS during the actual polar motion prediction process, it is necessary to establish the prediction model using earlier accurate data when the final products are not accessible for some reasons. Therefore, a time interval of 60 days after the training sample period is chosen for prediction. Although this approach may reduce the accuracy of the prediction algorithm, it can better test the universality of the algorithm and meet the complexity in practical polar motion predictions. Hence, the predicted time span is from 2019 February to 2021 June, in order to better validate the performance of the model. Due to the weekly update cycle of Bulletin A product, in order to conduct an accurate comparison of prediction accuracy, all 70 files within the aforementioned forecast period for Bulletin A product are saved. This includes 47 files from 2019 and 23 files from 2022. Additionally, the EWT-CNN-LSTM algorithm and comparative algorithms are used to generate 70 prediction files for the same time period, with each prediction file covering a forecasting period of 365 days.

3.3. Evaluation Criteria

The present experiment employs the common regression evaluation indicators of absolute error (AE) and mean absolute error (MAE) to analyze the accuracy of the prediction results. The formulas are as follows:

$$\mathrm{AE} = |\hat{y} - y| \tag{3}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \tag{4}$$

Here $\hat{y}$ represents the predicted value, y represents the actual value, and n represents the total number of days for prediction.

In time series forecasting of deep learning, each prediction result has randomness. To make the statistical error more representative, all errors in this study are averaged over multiple predictions.

4. Results and Discussion

4.1. Recursive and Direct Multi-step Prediction

Typically, recursive prediction is considered suitable for short-term forecasting, while direct multi-step prediction is more appropriate for medium to long-term forecasting. This is due to the fact that recursive prediction utilizes the previous time step's forecast value as input for the next time step, which can lead to error accumulation and a decline in prediction accuracy. However, in cases of strong data correlation, the effectiveness of short-term forecasting may surpass that of direct multi-step prediction. Therefore, it is crucial in PM prediction to determine which forecasting method is suitable for different time spans and to use different methods for different periods in order to ensure a scientifically sound approach. As illustrated in Figure 3 from the experimental results, overall MAE increases with the length of forecast days for both methods; PMY's MAE consistently remains smaller than PMX's. The MAE error of recursive prediction far exceeds that of direct multi-step prediction. When focusing on the initial few days' results, it can be observed that within 15 days, the MAE error of recursive prediction is smaller than that of direct multi-step prediction. Additionally, the error of direct multi-step prediction gradually decreases within 1–30 days, indicating its inferior performance in short-term forecasting. Therefore, it can be concluded that within a short-term forecast range of 1–15 days, we should utilize recursive multi-step predictions; beyond 15 days' forecasts call for direct multi-step predictions.

4.2. Recursive and Direct Multi-step Prediction

Taking PMX as an example, EWT decomposition is performed on all PMX data containing training and testing sets, due to the good adaptability of EWT decomposition, there is no problem that different wavelet basis functions in wavelet decomposition affect the decomposition results. Therefore, when using EWT decomposition, only the number of decomposition layers needs to be determined. The determination of the optimal number of decomposition layers needs to consider multiple aspects. In this paper, the Sensitive IMFs assessing index method is used to determine the initial number of decomposition layers (Liu et al. 2022), and the decomposition effect is analyzed by viewing the decomposition results

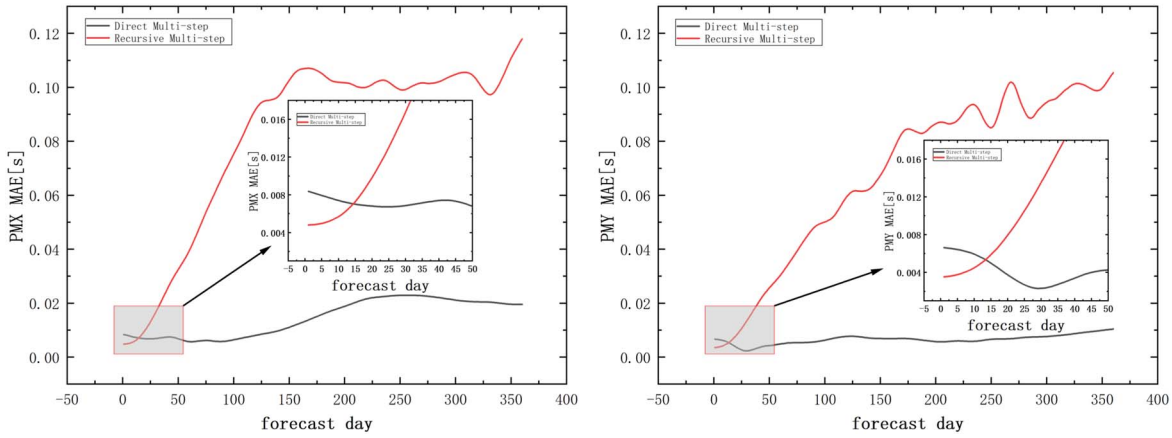

Figure 3. Comparison of MAE between direct multi-step and recursive multi-step prediction results.

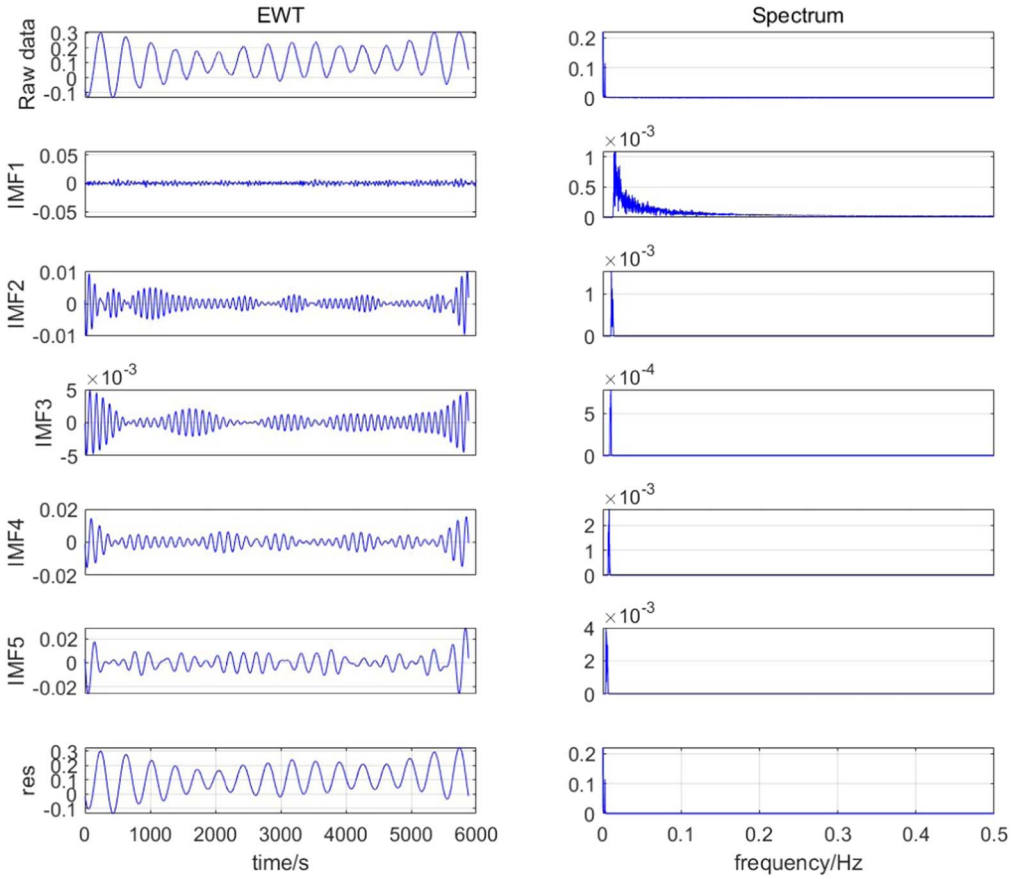

Figure 4. EWT Decomposition Diagram.

and the corresponding component spectrograms. Since the subsequent clustering algorithm will classify the signals into high-frequency, mid-frequency, and low-frequency categories, it is advisable to set a larger number of decomposition layers in order to achieve more accurate clustering results, the final determination is to set the number of decomposition layers to 6.

Figure 4 shows the results of PMX Raw data decomposed by EWT into six layers and the corresponding spectrum. Raw data is PMX original time series, and IMF1-IMF5 is different frequency components from high frequency to low frequency. It can be seen that the subsequences retain the fluctuation characteristics of the original characteristic variables, and no

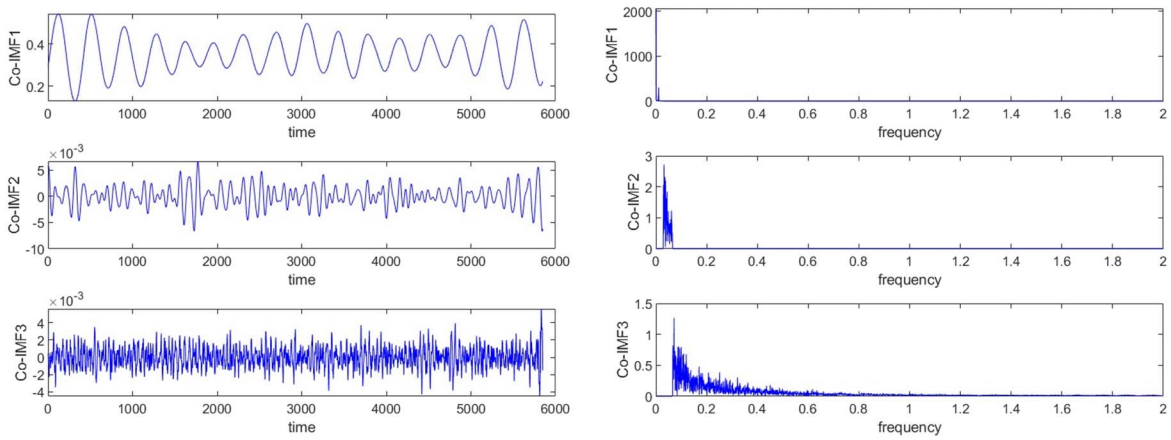

Figure 5. K-means clustering results and corresponding spectrogram.

mode aliasing occurs. The spectrogram shows that the original sequence is well divided into components of different frequencies, and the decomposition results are good. The sample entropy of each decomposed signal component will be calculated separately after decomposition, and k-means clustering will be conducted based on sample entropy. The clustering results are shown in Figure 5, where it can be observed from the frequency spectrum that the data has been effectively classified into high-, mid-, and low-frequencies.

A multi-layer CNN-LSTM network can effectively delve into target features, thereby enhancing prediction accuracy. However, an excessively complex structure may result in overfitting. Therefore, the model's structure and selection significantly impact prediction outcomes. The model parameters are as follows: Dropout rate is set at 0.1, the optimization algorithm used is Adam, the epoch is set to 100, the LSTM layer consists of 50 neurons, and the learning rate gradually decays over time.

4.3. Analysis of AE in PM Prediction Results

The use of absolute error can reveal the discrepancy between predicted values and actual values. Figure 6 shows the three-dimensional graph of absolute error for the prediction results. Specifically, (a), (c), and (e) represent the AE distribution of PMX forecast results for LSTM, CNN-LSTM, and EWT-CNN-LSTM schemes, while (b), (d), and (f) correspond to the AE distribution of PMY forecast results. Overall, the prediction results for PMY are superior to those for PMX. The variation trends of each scheme exhibit certain similarities. In the three-dimensional graph, the orange and red areas for EWT-CNN-LSTM are minimal in comparison to the other two schemes, with both average height and peak value being smaller than those of the remaining two schemes. This indicates that this scheme has a higher prediction accuracy compared to the other two schemes. In order to verify the deviation between predicted results and Bulletin A, we have compiled statistics on maximum value, minimum value, and mean value of AE for each time period as shown in Table 1.

Table 1
Statistical Information on the AE Values of Different Prediction Models(mas)

Time Span			EWT-CNN-LSTM	CNNLSTM	LSTM	Bulletin A
1–6	x	Max	17.83	18.87	21.91	7.15
		Min	0.00	0.00	0.00	0.00
		Average	5.04	5.61	5.03	1.03
	y	Max	12.75	18.71	13.51	3.50
		Min	0.00	0.00	0.00	0.00
		Average	4.65	5.05	4.15	1.04
7–90	x	Max	41.26	64.11	59.31	30.84
		Min	0.00	0.00	0.00	0.00
		Average	10.23	15.61	16.24	9.13
	y	Max	22.24	39.73	28.26	20.91
		Min	0.00	0.00	0.00	0.00
		Average	5.51	6.82	6.45	4.12
91–365	x	Max	51.16	73.14	63.91	59.41
		Min	0.00	0.00	0.00	0.00
		Average	20.41	25.43	25.51	21.63
	y	Max	37.86	44.11	48.72	45.12
		Min	0.00	0.00	0.00	0.00
		Average	12.51	14.44	14.53	13.81

From Table 1, it can be observed that in the short-term forecast results from 1 to 7 days, the prediction results of EWT-CNN-LSTM are generally similar to those of LSTM and CNN-LSTM. The absolute error of Bulletin A is smaller. In the medium-term forecast from 7 to 90 days, the average absolute errors in the X and Y directions for this approach are 10.23 mas and 5.51 mas, respectively, which are superior to the other two approaches. Compared with Bulletin A, the difference in mean values is within 1–2 mas, indicating overall similarity in

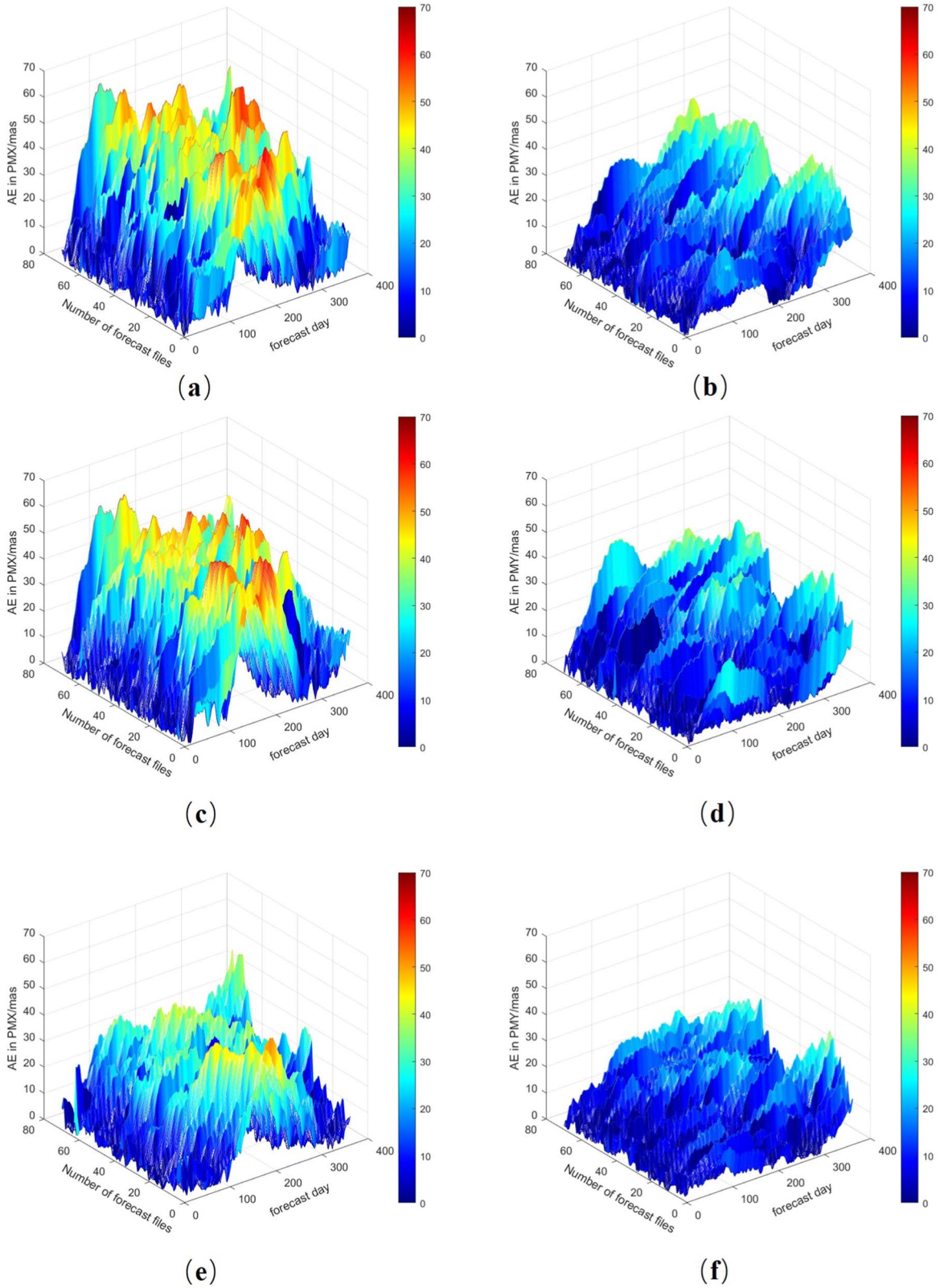

Figure 6. The 3D distribution of PM prediction result for LSTM, CNN-LSTM, and EWT-CNN-LSTM.

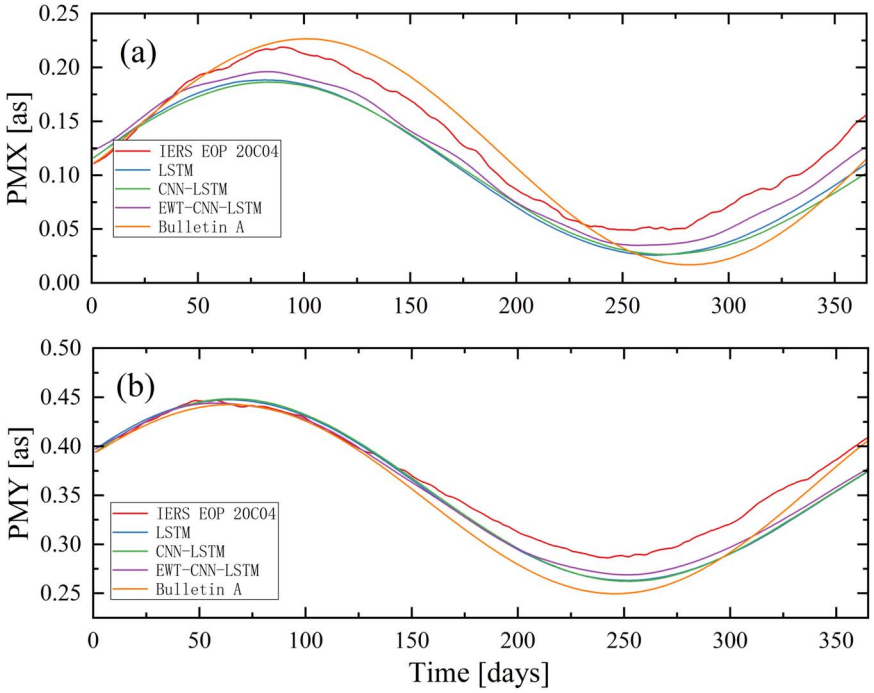

Figure 7. Comparison of 365 days forecast results from different prediction models.

Table 2
MAE Statistics for PMX and PMY Predictions Under Four Different Scenarios

No.	PMX					PMY				
	LSTM	CNN-LSTM	EWT-CNN-LSTM	Bulletin A	PCT	LSTM	CNN-LSTM	EWT-CNN-LSTM	Bulletin A	PCT
7	1.25	1.09	1.12	0.92	−19.0%	1.49	1.31	0.96	0.73	−31%
30	2.79	1.26	4.71	4.2	−12.0%	4.7	5.23	2.08	1.63	−27.6%
60	6.19	2.65	6.19	5.89	5.0%	7.9	6.84	6.03	6.14	1.7%
90	12.56	8.16	6.58	8.14	19.1%	12.6	8.71	7.04	7.11	0.9%
180	31.72	26.91	16.44	18.04	8.90%	21.45	15.31	11.73	13.78	14.8%
365	26.42	22.66	18.25	23.87	23.50%	20.58	23.04	15.78	18.84	16.2%

prediction results. In the long-term forecast from 90 days to 365 days, PM has maximum AE values of 51.16 mas and 37.86 mas in both directions, corresponding to mean values of 20.41 mas and12.51 mas respectively, all smaller than those of the other two approaches as well as Bulletin A. Therefore, we can conclude that in mid-to-long term polar motion prediction, the EWT-CNN-LSTM method outperforms the other two methods and also exhibits improved accuracy compared to Bulletin A.

4.4. Analysis of MAE in PM Prediction Results

Figure 7 presents the predicted results of PMX and PMY under different schemes, with a prediction time of 365 days and a base sequence of 11 yr. The red line in the figure represents the true values provided by IERS EOP 20C04, while the blue, green, purple, and orange lines represent the predictions of LSTM, CNN-LSTM, EWT-CNN-LSTM, and Bulletin A respectively. It can be seen from the figure that the prediction results of LSTM and CNN-LSTM are similar. The proposed EWT-CNN-LSTM method is superior to the former two methods, and it is closer to the true values in medium- to long-term predictions compared to Bulletin A.

The MAE of PMX and PMY predictions under different schemes is illustrated in Figure 8. Overall, the EWT-CNN-LSTM model exhibits the smallest error, particularly in the medium to long term. This can be attributed to the enhanced representation of overall trend and removal of interference factors by low-frequency components after EWT decomposition and K-Means clustering, enabling LSTM to make more accurate predictions. The CNN-LSTM model shows slightly

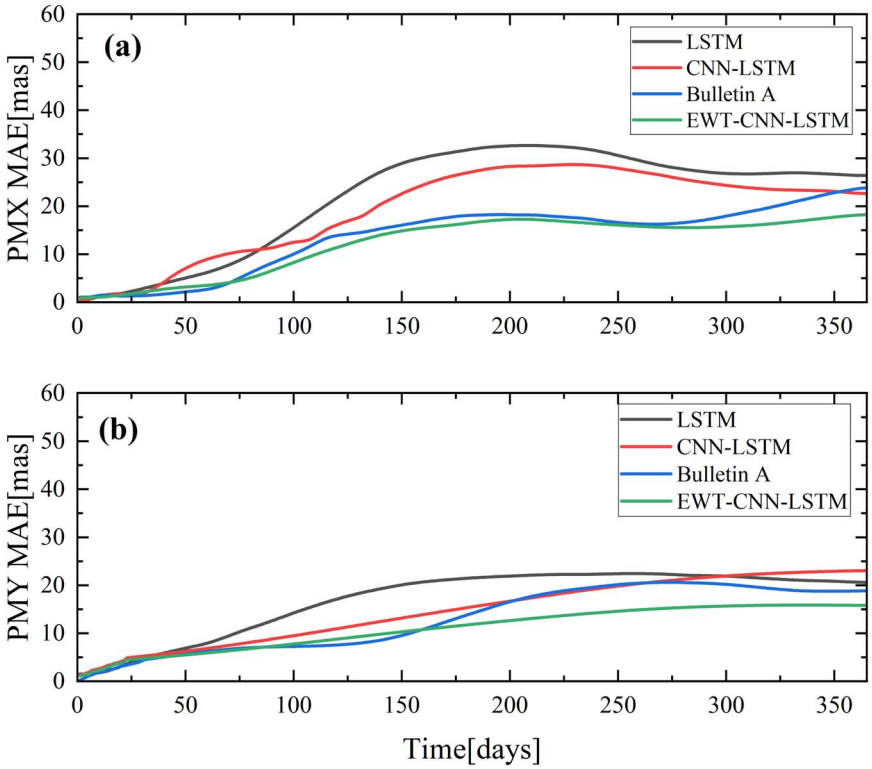

Figure 8. MAE of PMX and PMY predictions in four scenarios.

better performance than the LSTM model, but both still lag behind Bulletin A in terms of accuracy. Table 2 presents statistical MAE for various prediction periods, revealing that the EWT-CNN-LSTM model outperforms Bulletin A in both X and Y directions after approximately 60 days, with this improvement becoming more significant as the number of predicted days increases. However, within 60 days, Bulletin A still demonstrates higher predictive accuracy. In a 365 days forecast, the EWT-CNN-LSTM model achieves an MAE of 18.25 mas for PMX and 15.78 mas for PMY predictions respectively—representing improvements of 23.5% and 16.2% over Bulletin A's respective values of 23.87 mas and 18.84 mas.

5. Conclusion

The high-precision prediction of polar motion is of great significance for satellite navigation and high-precision orbit determination. According to different forecasting models, it is mainly divided into linear forecasting models and nonlinear forecasting models, with the latter mainly based on neural networks. For time series prediction, the LSTM model is often used. Based on this, we have constructed the EWT-CNN-LSTM model, which decomposes the sequence through EWT, calculates its sample entropy, obtains K-Means clustering results, and decomposes the original data into high-frequency, medium-frequency, and low-frequency data. These are then separately trained and predicted using CNN-LSTM before reconstructing and restoring the prediction results. The purpose of this approach is to reduce the complexity of the data in order to improve the prediction results of CNN-LSTM.

This study first analyzed the performance of recursive multi-step prediction and direct multi-step prediction in PM forecast within 15 days. The results showed that in short-term forecasts within 15 days, recursive multi-step predictions outperformed direct multi-step predictions. Furthermore, this study compared four types of forecast results—LSTM, CNN-LSTM, EWT-CNN-LSTM, and Bulletin A—over a two-year period using IERS EOP 20C04 as a reference point. The experimental results showed that the EWT-CNN-LSTM model had better predictive accuracy in medium-to-long term PM forecasts compared to Bulletin A forecast results; specifically improving PMX by 23.5% and PMY by 16.2%.

This study primarily focused on analyzing deep learning models with single-variable input for multiple-step output in polar motion prediction. Future research will involve using multivariable inputs considering Earth Atmospheric Motion (EAM) to enhance polar motion prediction accuracy further. Additionally, the combination of deep learning with traditional linear forecasting models such as LS will be explored more deeply in the field of high-precision PM forecast.

Acknowledgments

This work has been supported by the National Natural Science Foundation of China (NSFC) under grant No. 42304044 and the Natural Science Foundation of Henan, China under grant No. 222300420385.

ORCID iDs

Zhong-Kai Zhang https://orcid.org/0000-0001-8293-7400

References

Bachmann, S., Thaller, D., Roggenbuck, O., Lösler, M., & Messerschmitt, L. 2016, JGeod, 90, 631
Bizouard, C., Lambert, S., Gattano, C., Becker, O., & Richard, J.-Y. 2019, JGeod, 93, 621
Chandler, S. C. 1891, AJ, 11, 83
Dobslaw, H., Dill, R., GröTzsch, A., BrzezińSki, A., & Thomas, M. 2010, JGRB, 115, B10406
Dow, J. M., Gurtner, W., & Schlüter, W. 2005, JGeod, 40, 375
Egger, D. 1992, Allgemeine Vermessungsnachrichten, 31, 517
Gers, F. A., Schmidhuber, J., & Cummins, F. 2000, Neural Comput., 12, 2451
Gilles, J. 2013, ITSP, 61, 3999
Gou, J., Kiani Shahvandi, M., Hohensinn, R., & Soja, B. 2023, JGeod, 97, 52
Gross, R. S. 2007, Treatise on Geophysics, 3 (Amsterdam: Elsevier), 239
Hochreiter, S., & Schmidhuber, J. 1997, Neural Computation., 9, 1735
Kalarus, M., Schuh, R., Kosek, R., et al. 2010, JGeod, 84, 587
Karevan, Z., & Suykens, J. A. K. 2020, NN, 125, 1
Khodabakhsh, A., Ari, I., Bakır, M., & Alagoz, S. M. 2020, Forecasting Multivariate Time-Series Data Using LSTM and Mini-Batches in Data Science: From Research to Application (Berlin: Springer), 121
Krizhevsky, A., Sutskever, I., & Hinton, G. 2012, Advances in Neural Information Processing Systems, 25
Liu, Q., Yang, J., & Zhang, K. 2022, Measurement, 187, 110348
Mehtab, S., Sen, J., & Dasgupta, S. 2020, IEEE, 1, 1481
Petit, G., & Luzum, B. 2010, ITN, 36, 1
Schuh, H., Ulrich, M., Egger, D., Müller, J., & Schwegmann, W. 2002, JGeod, 76, 247
Su, X., Liu, L., Houtse, H., & Wang, G. 2014, JGeod, 88, 145
Sun, Z., Xu, T., Jiang, C., Yang, Y., & Jiang, N. 2019, AcGG, 54, 499
Wang, J., Jiang, W., Li, Z., & Lu, Y. 2021, RemS, 13
Xue, N., Triguero, I., Figueredo, G. P., & Landa-Silva, D. 2019, in 2019 IEEE Congress on Evolutionary Computation (CEC) (Wellington: IEEE), 1517
Zhang, X., Liang, X., Zhiyuli, A., et al. 2019, IOP Conf. Ser.: Mater. Sci. Eng., 569, 052037
Zharkov, V., & Molodensky, S. 1996, Plan. Space Sci., 44, 1457