Wu et al. Earth, Planets and Space (2023) 75:179
https://doi.org/10.1186/s40623-023-01937-x

Earth, Planets and Space

FULL PAPER

Open Access

Precise prediction of polar motion using sliding multilayer perceptron method combining singular spectrum analysis and autoregressive moving average model

Kezhi Wu[1], Xin Liu[1*], Xin Jin[1], Xiaotao Chang[2], Heping Sun[3] and Jinyun Guo[1]

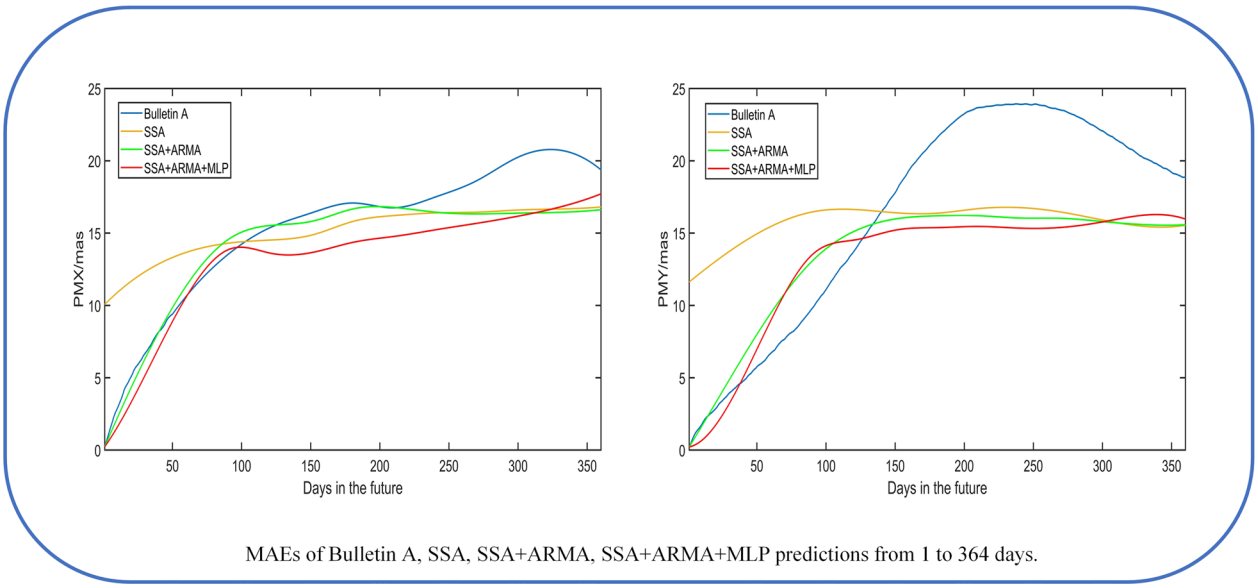
Abstract

The precise prediction of polar motion parameters is needed for the astrogeodynamics, navigation and positioning of the deep space probe. However, the current prediction methods are limited to predicting polar motion for specific periods, either short- or long-term. In this study, a sliding multilayer perceptron (MLP) method combined singular spectrum analysis (SSA) and autoregressive moving average (ARMA) for short- and long-term polar motion prediction was proposed. MLP was introduced into PM prediction due to its automatic learning characteristics and its ability to effectively process nonlinear and multi-dimensional data. The SSA was used to extract and predict the principal components of polar motion, while the remaining components were predicted using ARMA. In the meantime, SSA and ARMA were used to provide training data and target learning data for the MLP model. MLP input data were constructed by sliding processing with a window of 7 days, composed of n series of the same length (18 years). Finally, MLP was employed to predict the residuals generated during SSA and ARMA prediction. To evaluate the accuracy of the proposed method, the polar motion prediction was applied for a 364-day lead time based on the IERS EOP 14C04 product. The method outperformed the IERS Bulletin A, as demonstrated by the mean-absolute errors of the x and y components of polar motion on the 30th day, which were lower (5.14 mas and 3.37 mas, respectively) than those predicted by IERS Bulletin A (6.66 mas and 3.94 mas). Similarly, the mean-absolute errors on the 364th day were 17.79 mas and 16.29 mas, respectively, compared to the 19.24 mas and 18.81 mas predicted by IERS Bulletin A.

Keywords Polar motion prediction, Singular spectrum analysis, Autoregressive moving average model, Multilayer perceptron

*Correspondence:
Xin Liu
xinliu1969@126.com
Full list of author information is available at the end of the article

Graphical Abstract



MAEs of Bulletin A, SSA, SSA+ARMA, SSA+ARMA+MLP predictions from 1 to 364 days.

Introduction

Polar motion (PM) refers to the movement of the Earth's rotation axis on its surface (Jin et al. (2021a); Wu et al. 2021; Zhu 1982; Lambeck 1980). The accuracy of real-time PM parameters is crucial for high-precision satellite navigation and positioning systems and spacecraft tracking (Stamatakos 2017). To obtain precise PM parameters, various technologies such as Global Navigation Satellite System (GNSS), Satellite Laser Ranging (SLR), Very Long Baseline Interferometry (VLBI), Doppler Orbitography and Radiopositioning Integrated by Satellite (DORIS) can be employed (Dill and Dobslaw 2010). However, in satellite navigation technology, interruption the communication link between the satellite and ground are common, necessitating the conversion between reference frames using earth orientation parameters (EOP). This emphasizes the importance of accurately predicting PM and obtaining real-time PM parameters (Gambis and Luzum 2011; Wang et al. 2017).

Polar motion parameters are playing an increasingly crucial role in the field of space science. The rapid development of celestial dynamics has led to increasingly high requirements for the accuracy of polar motion parameters, and the accuracy of polar motion prediction directly affects the accuracy of satellite orbit determination. Due to the complexity of polar motion data processing, obtaining real-time polar motion parameters is not feasible and they can only be obtained through high-precision prediction methods (Su et al. 2014; Shen et al. 2018). The excitation source of polar motion is relatively complex, which also requires high accuracy of prediction methods. At the same time, different prediction methods exhibit varying levels of skill across different PM prediction time scales, and establishing relatively stable PM prediction methods is a challenge in the current PM prediction process.

Various methods were proposed and utilized, such as least squares (LS) (Kosek et al. 2007; Yao et al. 2013; Wu et al. 2018; Akulenko et al. 2002), singular spectrum analysis (SSA) (Shen et al. 2017; Jin et al. (2021b)), autoregressive model (AR) (Schuh et al. 2002), wavelet analysis (Zhao and Lei 2019), copula-based analysis (Modiri et al. 2018), and artificial neural networks (ANN) (Liao et al. 2012). LS and AR exhibit satisfactory performance in short-term PM prediction. However, the presence of redundant data and inappropriate orders can result in overfitting, thereby significantly diminishing their efficacy in mid-long-term PM prediction. Therefore, both LS and AR are not suitable for mid- to long-term PM prediction alone. Wavelet analysis proves effective in segregating high and low frequency components within PM series. As a result, it is commonly employed in conjunction with other models to enhance the accuracy of mid- to long-term PM prediction. Although existing prediction methods majority provide good ultra short-term (< 10 days), short-term (< 30 days) prediction results, they do not perform satisfactorily enough for long-term predictions (Sun et al. 2019). There are many PM prediction

methods performing better in the short term or only in the mid-long term.

To overcome the limitations and instability of existing prediction methods, IERS organized the Earth Orientation Parameters Combination of Prediction Pilot Project (EOPC PPP) in July 2010. The LS + AR prediction method officially recognized by IERS, has high accuracy in short-term forecasting but substantial error accumulation with increasing predicted time (Kosek et al. 2011). The LS + AR method also has a high requirement on data quality, affecting the mid-long-term PM prediction accuracy to a certain extent (Su et al. 2014; Guo and Han 2009).

SSA has been successfully applied in various fields, including geophysics, meteorology, and oceanography, to extract and reconstruct signals with different periods from the original signal and predict their principal components (Kong et al. 2020; Marques et al. 2006; Briganti and Beltrami 2008; Vautard and Ghil 1989). However, the use of SSA alone for polar motion prediction can result in distortion at the end of the fitted series (end effect) and overfitted. By combining SSA with other models, the overall PM predicted accuracy can be improved (Zhang et al. 2011; Zotov 2010; Heng and Suetsugi 2013). Shen et al. (2018) combined the SSA and ARMA model, demonstrating the applicability of this method, and improved the end effects and overfitting problems. But its short-term prediction accuracy needs improvement. Furthermore, the MLP model with its nonlinear-mapping and self-accommodating characteristics (Zhang 2003; Singh 2018), has been used to predict in several fields, including medicine, meteorology, economics, and geology (Choi et al. 2019). Polar motion is influenced by various excitation sources, which have a significant impact on the accuracy of long-term polar motion prediction. And although the PM series was fitted by SSA + ARMA, there still are complex residual changes. As a nonlinear neural network model, MLP has the processing ability to dispose complex characteristic of PM time series and improve the accuracy of PM prediction. Therefore, using MLP as a nonlinear model for predicting polar motion is theoretically appropriate.

This study proposed a new sliding method for predicting PM by combining the SSA + ARMA model with the MLP model. This method takes the characteristics of the PM series into consideration, mainly including the periodic characteristics of PM series and the 'nonlinearity' of PM series caused by complex excitation sources. The meaning of 'nonlinear' is that the polar motion series is not a linear combination of signals due to its complex excitation. It is obvious that this excitation caused the complexity of polar motion prediction. By decomposing and reconstructing the PM time series using the SSA + ARMA model, the obtained data are input into the MLP model. This integration of deep learning techniques enables the leveraging of their inherent advantages to improve the accuracy and stability of PM prediction. The sliding SSA + ARMA + MLP method proposed in this study has achieved pretty prediction accuracy in ultra-short term, short term, and mid-long term (364 days).

Methodology

Data

The PM time series from January 11, 1987, to December 26, 2019, was analyzed and used to predict PM from January 5, 2017, to December 26, 2019 by utilizing the sliding SSA + ARMA + MLP model. The original time series (PMX, PMY) covering the period of 1982 to 2019 is shown in Fig. 1. The PM data used in this study were sourced from the IERS 14 C04 product (https://www.iers.org/IERS/EN/DataProducts/EarthOrientationData/eop.html), which provides EOP data, since 1962 with a sampling interval of 1 day (Bizouard et al. 2019). The IERS also publishes fast products and data to meet engineering application criteria. Bulletin A is a prediction file regularly issued by the IERS Fast Service and Forecast Center of the U.S. Naval Observatory (USNO). Bulletin A provides the prediction of polar motion parameters for the last week and the next year, with a sampling interval of 1 day and updated once a week, with products delayed by one week upon release.

Overview of the SSA + ARMA + MLP

As shown in Fig. 2, SSA and ARMA were used to provide training data for the MLP model while predicting PM. MLP training data were constructed by sliding with a window of 7 days, composed of n rows of the same length. The n-th series was obtained by sliding the (n-1)-th series backward for 7 days, as outlined in Sect. "Prediction of MLP". PM time series were predicted using SSA, with the remaining components (RC) and SSA predictions obtained. The difference between observations and SSA predictions was used to derive the SSA predicting residuals. ARMA was then used to predict the RCs, providing ARMA modeling residuals and predictions. The difference between SSA predicting residuals and ARMA predictions was used to obtain ARMA predicting residuals, with ARMA modeling residuals serving as the training data Q, and ARMA predicting residuals as the target learning data T. The input data for MLP comprised the ARMA modeling and predicting residuals.

The PM prediction process comprised three main parts. Firstly, the principal components (PCs) were predicted using sliding SSA. To better compare Bulletin A and construct the MLP input data, the sliding window was set to 7 days. After SSA completed one prediction,

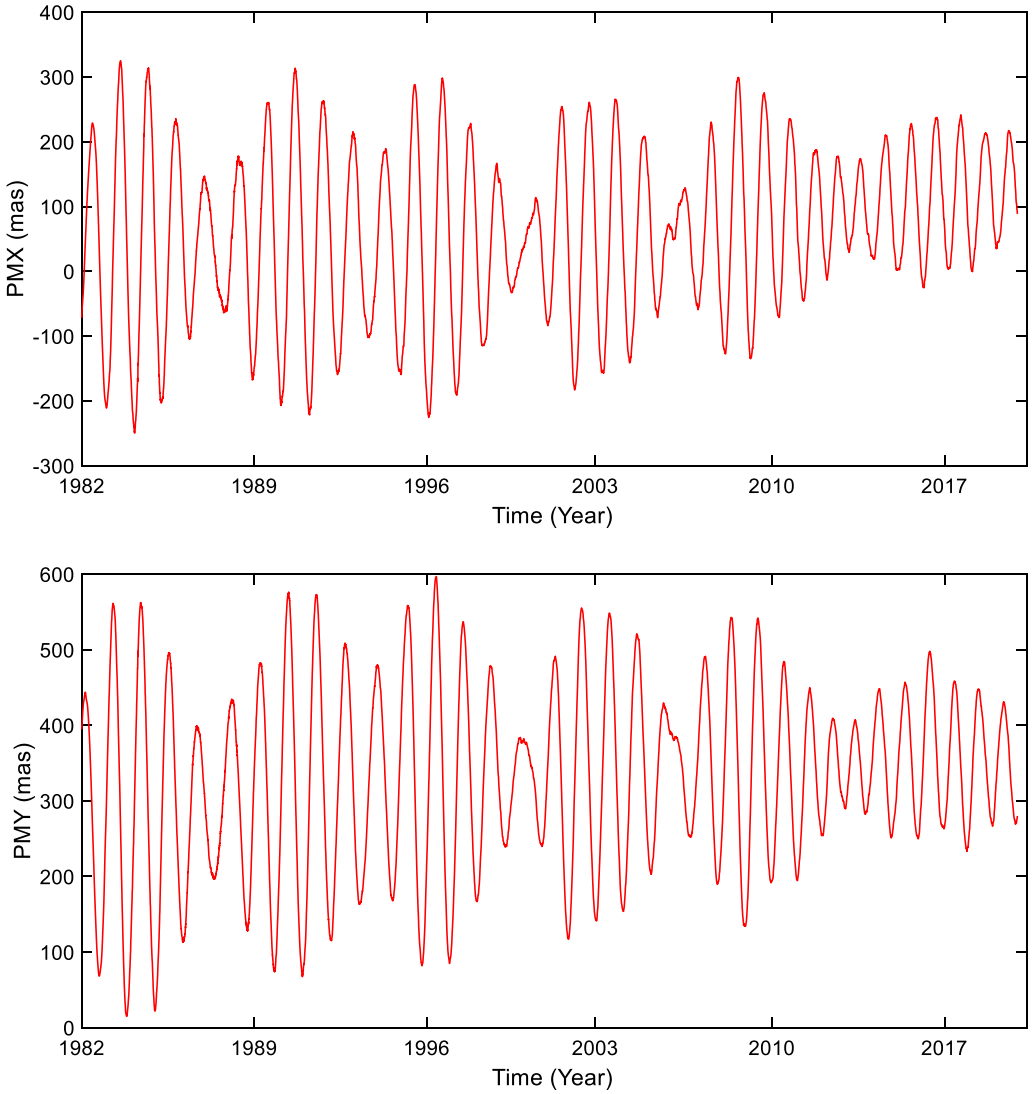


Fig. 1 The original time series data of PMX and PMY

it slid back 7 days to complete the next prediction. The predicted part was considered as the missing part, and the PCs were repeatedly filled iteratively after filling the PM data to obtain the SSA predictions and PM residuals. In the second part, the residual components (RCs) were predicted using the ARMA model, which obtained the ARMA predictions. Lastly, the MLP model was used to predict the ARMA modeling residuals. After training, the MLP model completed a 364-day PM prediction. Finally, the PM predictions were derived by adding the predicted PCs from the SSA model, the predicted RCs from the ARMA model, and the predicted ARMA residuals from the MLP model.

Prediction of SSA

Singular spectrum analysis constructs a one-dimensional nonlinear time series into a multi-dimensional trajectory matrix, which can be further decomposed and reconstructed to extract various signals (Elsner and Tsonis 1996). For the PM time series, SSA can be used to decompose different signals, such as long-term trends, periodic terms, or noise, which can reconstruct the original time series in PM prediction.

As one of the critical parameters in SSA time series analysis, the window size affects the results. A reasonable window size setting can better separate the PCs of the time series, and an appropriate window size should be a multiple of the PC period. For long-term PM time series, the days of Chandler wobble and annual oscillation period were both considered in the window size M,

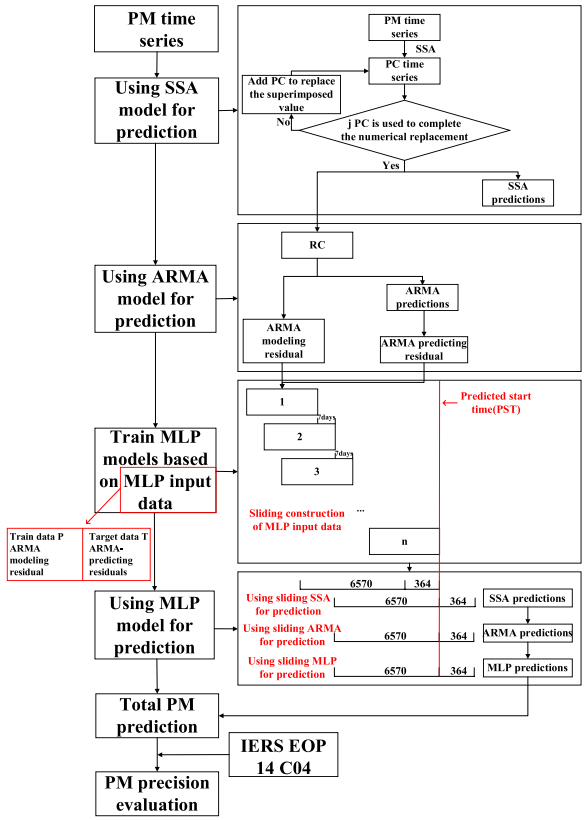


Fig. 2 Flowchart of the sliding SSA + ARMA + MLP method for PM prediction (singular spectrum analysis (SSA), autoregressive moving average model (ARMA), multilayer perceptron (MLP), principal components (PC), remaining components (RC), 6570 is the train data P volume, 364 is the target data T volume)

which was set to 2190 (Zotov 2010). After SSA decomposed and reconstructed the time series, the original PM time series could be replaced by the superposition of the first j reconstructed components. The size of j needs to be determined by SSA analysis. In Sect. "Singular spectrum analysis of PM time series analysis", the determination process of the j value is introduced in detail.

The analysis of filling the PM time series and repeating iterations during the SSA prediction period was based on previous publications (Shen et al. 2018; Schoelhamer 2001). In this study, the following are the main steps:

(1) The original PM time series length was set to $N = 6570$, the PM prediction data length was set to $u = 364$ and a new PM time series $(N + u)$ was constructed by adding n zeros at the end of the original PM time series, where u is the number of zeros.

(2) SSA decomposed the new PM time series $(N + u)$. The last n values of the new PM time series were replaced those at the end of the first PC (PC1), and this process was repeated until PC1 converged.

(3) After the first cycle, a second PC (PC2) was added to reconstruct the prediction data. PC1 and PC2 were superimposed to obtain the prediction data. Step (2) was repeated until the PC1 + PC2 series converged.

(4) The above process was repeated iteratively until the j PC series complete the numerical replacement. Finally, the u values at the end of the new PM time series were the j PC predictions of the PM time series.

Prediction of ARMA

After obtaining the RC through SSA prediction, they still contained various other stationary periods or noises. Therefore, the RCs were further modeled and predicted using the ARMA. According to the principles of multiple regression, the ARMA model used in this study can be expressed as:

$$Y_N = \beta_1 Y_{N-1} + \beta_2 Y_{N-2} + \ldots + \beta_p Y_{N-p} - \theta_1 \alpha_{N-1} - \theta_2 \alpha_{N-2} - \ldots - \theta_q \alpha_{N-q} + \alpha_N, \quad (1)$$

where $\beta_i (i = 1, 2, \ldots, p)$ and $\theta_i (i = 1, 2, \ldots, q)$ represent the autoregressive (AR) and moving average (MA) coefficients, respectively, and α_N is white noise. To determine the appropriate orders of p and q for the model, the extended autocorrelation function (EACF) analysis was selected in this study. Based on the EACF analysis, the ARMA (2, 1) model was suitable for PM prediction, where $p = 2$ and $q = 1$.

Prediction of MLP

After ARMA completed the RC prediction and output the residuals, the MLP model used the ARMA residuals as the training data. Considering the complex characteristics of the PM time series, the MLP model leveraged its deep learning capability to enhance overall PM prediction accuracy. MLP is a typical feedforward artificial neural network, where each neuron is only connected to the previous layer of neurons, making is well suited for classifying extensive data and establishing complex problems (Pal and Mitra 1992). By selecting a limited number of input and output variables, the MLP model can effectively simulate the relationship between the input and output data, thereby facilitating the establishment of a prediction method based on the MLP model (Kuremoto

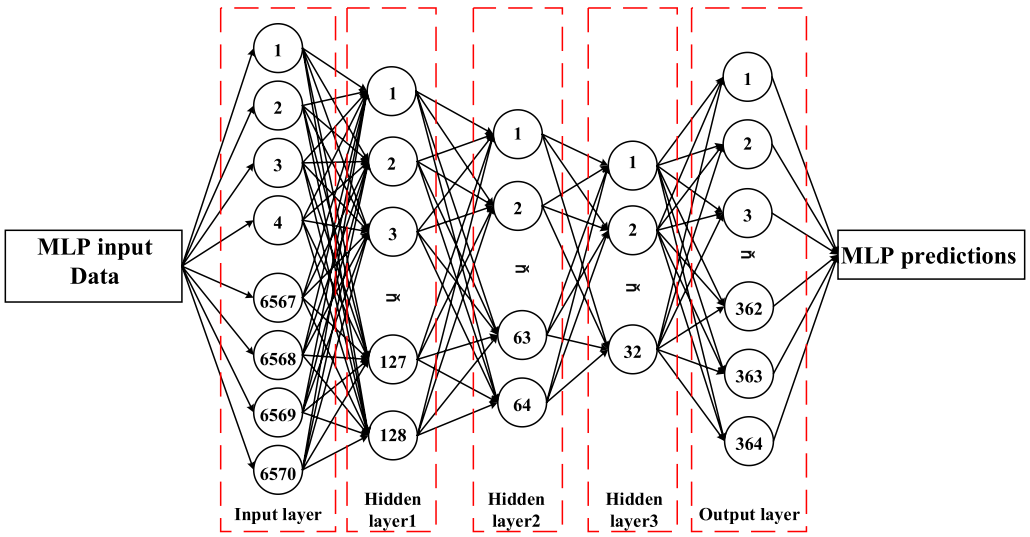


Fig. 3 Structure of the multilayer perceptron (MLP)

et al. 2014). Given that the PM prediction method employed in this study entails mapping multiple input variables to output variables, the MLP model is utilized.

As depicted in Fig. 3, the MLP model represents a network comprising densely connected information processing nodes, known as neurons. Neurons serve as the fundamental units constituting neural networks, functioning as mathematical models that emulate biological neurons for information processing and transmission purposes. The MLP model encompasses an input layer, one or more hidden layers, and an output layer (Ghasemloo et al. 2022). The input layer serves as the neuron responsible for data input, with its number aligning with the count of data variables. The hidden layers and output layer consist of neurons and output variables, respectively. Adjacent layers are fully interconnected, while neurons within the same layer remain unconnected.

The MLP model consists of three basic elements: weights, biases, and activation functions. The training process consists of two steps: from input to hidden layers and from hidden to output layers. From input to hidden layers, $I = h(W_1 x + r_1)$, where I is the output vector of the hidden layer, h is the activation function, W_1 is the weight, x is the input vector, and r_1 is the bias. From hidden to output layers, $O = g(W_2 y + r_2)$, where O is the output vector of the output layer, g is the activation function, W_2 is the weight, y is the input vector, and r_2 is the bias.

The MLP accomplishes the mapping from input to output by processing and transmitting information via neurons. The core of MLP learning lies in the iterative adjustment of the activation function and bias, achieved through information transmission and error backpropagation within the network. This iterative process aims to minimize the discrepancy between the desired output of the MLP and the actual output.

Lippmann (1988) proved that a two hidden-layer neural network can generate arbitrarily complex decision regions and approximate continuous nonlinear functions with any precision. In this study, considering the number of input neurons, three hidden layers were set to achieve a better training effect.

The ARMA residuals were used as the training data and target learning data by sliding, respectively. The MLP model was trained and used to predict PM using the following steps:

(1) Network initialization involved weight and bias initialization to determine the activation function simultaneously. The window size of $L = 6570$ was set to the length of the 18-year historical data, which is the length of the training data, and the target step size $N = 364$ was set to the size of the prediction period.
(2) For data normalization, selecting the maximum and minimum values of the PM series in the PMX and PMY, respectively. Then, min–max normalization was used to normalize the PMX and PMY of the MLP model training dataset. x was mapped to x':

$$x' = (x - x_{\min})/x_{\max} - x_{\min}. \tag{2}$$

(3) To construct the training data $\boldsymbol{Q}$ by sliding, the ARMA modeling residuals were utilized, taking into account the training window size L and the

prediction period size N. The normalized data were also incorporated, the training data $\boldsymbol{Q}$ are expressed as:

$$\mathbf{Q} = \begin{pmatrix} x_{1,1} & \cdots & x_{1,L} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \cdots & x_{n,L} \end{pmatrix} \tag{3}$$

The target learning data $\boldsymbol{T}$ were generated based on the ARMA predicting residuals and can be expressed as:

$$\mathbf{T} = \begin{pmatrix} y_{1,1} & \cdots & y_{1,N} \\ \vdots & \ddots & \vdots \\ y_{\mathrm{n},1} & \cdots & y_{n,N} \end{pmatrix}, \tag{4}$$

where, $x_1 = \{ x_{1,1}\ x_{1,2}\ x_{1,3}\ \cdots\ x_{1,L} \}$ is the ARMA modeling residuals between January 11, 1987 and January 6, 2005, the data length is L, and $x_2 = \{ x_{2,1}\ x_{2,2}\ x_{2,3}\ \cdots\ x_{2,L} \}$ is the ARMA modeling residuals obtained by sliding x_1 backward for 7 days. The sliding window was maintained as remains 7 days, the data length L remained unchanged, and $x_n = \{ x_{n,1}\ x_{n,2}\ x_{n,3}\ \cdots\ x_{n,L} \}$ is the ARMA modeling residuals sliding to the predicted start time (PST). $y_1 = \{ y_{1,1}\ y_{1,2}\ y_{1,3}\ \cdots\ y_{1,N} \}$ is the ARMA predicting residuals corresponding to ARMA modeling residuals x_1 and the data length is N. $y_2 = \{ y_{2,1}\ y_{2,2}\ y_{2,3}\ \cdots\ y_{2,N} \}$ is the ARMA predicting residuals corresponding to x_2 and $y_n = \{ y_{n,1}\ y_{n,2}\ y_{n,3}\ \cdots\ y_{n,N} \}$ corresponds to x_n ARMA predicting residuals.

(4) The MLP model was established and trained. The window size L for the long-term PM series was set to 6570, which considered the Chandler wobble and the annual oscillations. During the PM prediction with a period of 364 days, the number of neurons was 6570, 128, 64, 32, and 364 in sequence. Figure 3 shows the structure of the MLP model.

(5) Using the trained MLP model for prediction, the ARMA modeling residuals were used as the input data, and the prediction was $y_{n+1} = \{ y_{n+1,1}\ y_{n+1,2}\ y_{n+1,3}\ \cdots\ y_{n+1,N} \}$. The denormalized formula is:

$$y'_{n+1} = y_{n+1}(x_{\max} - x_{\min}) + x_{\min}. \tag{5}$$

(6) The iterative prediction process was repeated 20 times, and the final MLP predictions were obtained by taking the average value, which is a common practice to reduce the effects of random fluctuations and obtain more stable and reliable results.

Results and discussion

Singular spectrum analysis of PM time series analysis

SSA is a method for identifying and analyzing signals based on their time and frequency domain characteristics (Ghil et al. 2002). To determine the principal components of the PM series, the PM time series from January 11, 1987, to December 26, 2019, was analyzed by SSA. As described in the prediction of SSA, considering the Chandler wobble and annual wobble periods, the window size L generally needed to meet $1 < L < N/2$ (N is the length of the data), L was set to 2190.

The correlation between the PM time series was analyzed using the W-correlation method (Hassani 2007). Different signals with greater correlation can be divided into a group by SSA, and PCj is the j-th PC obtained by SSA. To determine the optimal value of j, the variance contribution rate of PM is shown in Fig. 4, while Fig. 5 displays the results of the W-correlation analysis of the PM time series.

As illustrated in Figs. 4 and 5, in the PMX, the variance contribution rates of PC1 and PC2 were relatively large, at 38.80% and 37.51%, respectively. PC3 had a variance contribution rate of 7.83%, and those of PC4 and PC5 were 7.34% and 7.27%, relatively. All other PCs had low variance contribution rates, which decreased with increasing order. For the PMY, PC1 had the largest variance contribution rate, reaching 85.54%, while PC2, PC3, PC4, and PC5 contributed 6.16%, 6.11%, 0.99%, and 0.96%, respectively. Figure 5 demonstrates strong independence among the first seven PCs, the first seven PCs account for 99.42% and 99.88% of the total variance contribution rates in both the X and Y directions of PM. Therefore, the first seven PCs of PMX and PMY were able to more accurately represent the original PM time series.

Figure 6 shows the PM components decomposed by SSA. For PMX, the Chandler wobble, the annual term, and a periodic term of approximately 489 days were three relatively prominent periodic components for PMX, corresponding to PC1 and PC2, PC4 and PC5, and PC6 and PC7, respectively. PC3 represented the trend term, while other components were not well separated due to their small correlation and were considered residual terms. For the Y direction of the PM series, PC2 and PC3 represented the Chandler wobble, PC4 and PC5 represented the annual term, and PC6 and PC7 represented the periodic term of approximately 489 days. The trend term was represented by PC1.

Prediction of PM

This study was based on historical data. The fundamental data used for the prediction were the historical observations from January 11, 1987, a period of 18 years before

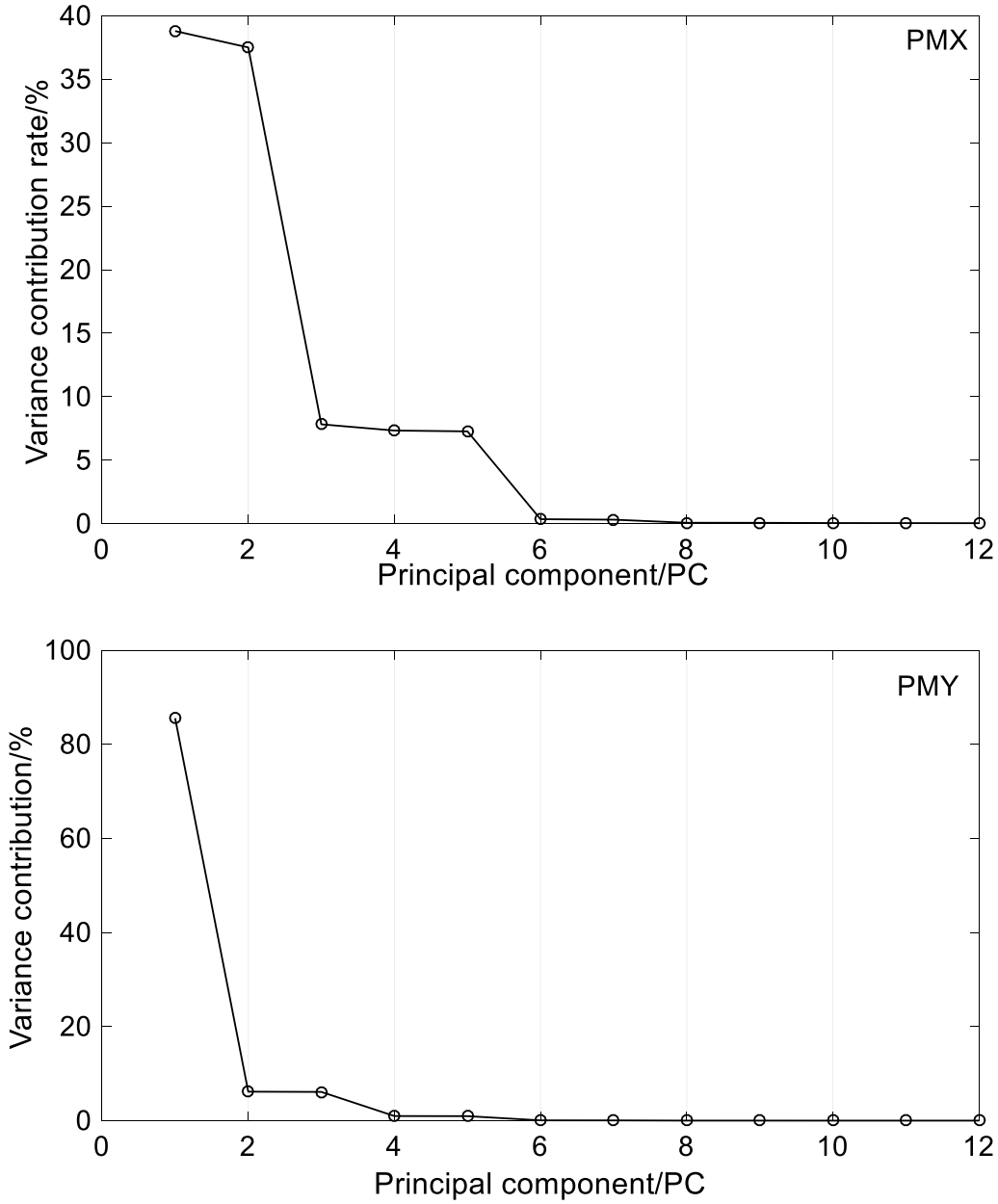


Fig. 4 Variance contribution rates of PM series determined by SSA

PST. The basic time series was updated with each prediction, and the historical length of 18 years remained unchanged. For each year, 52 time series were selected, and a total of 156 time series were selected for PM prediction. The PM prediction was compared to that of the IERS EOP 14 C04 product. In addition, the mean absolute error (MAE) was selected as the evaluation standard (Kalarus et al. 2010):

$$MAE = \sum_{i=1}^{t} |P_i - O_i|/t, \tag{6}$$

where P_i is the predicted value corresponding to period i, O_i is the observed value corresponding to period i, and t is the prediction length.

Bulletin A (https://datacenter.iers.org/eop/-/somos/5Rgv/%20getTX/6) releases a PM prediction series once

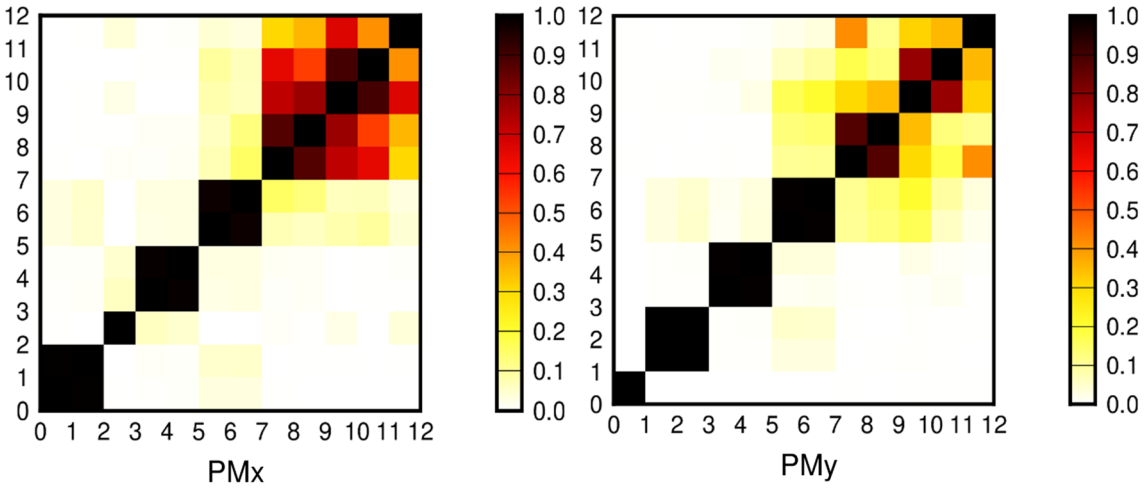


Fig. 5 W-correlation analysis of the first 12 PCs

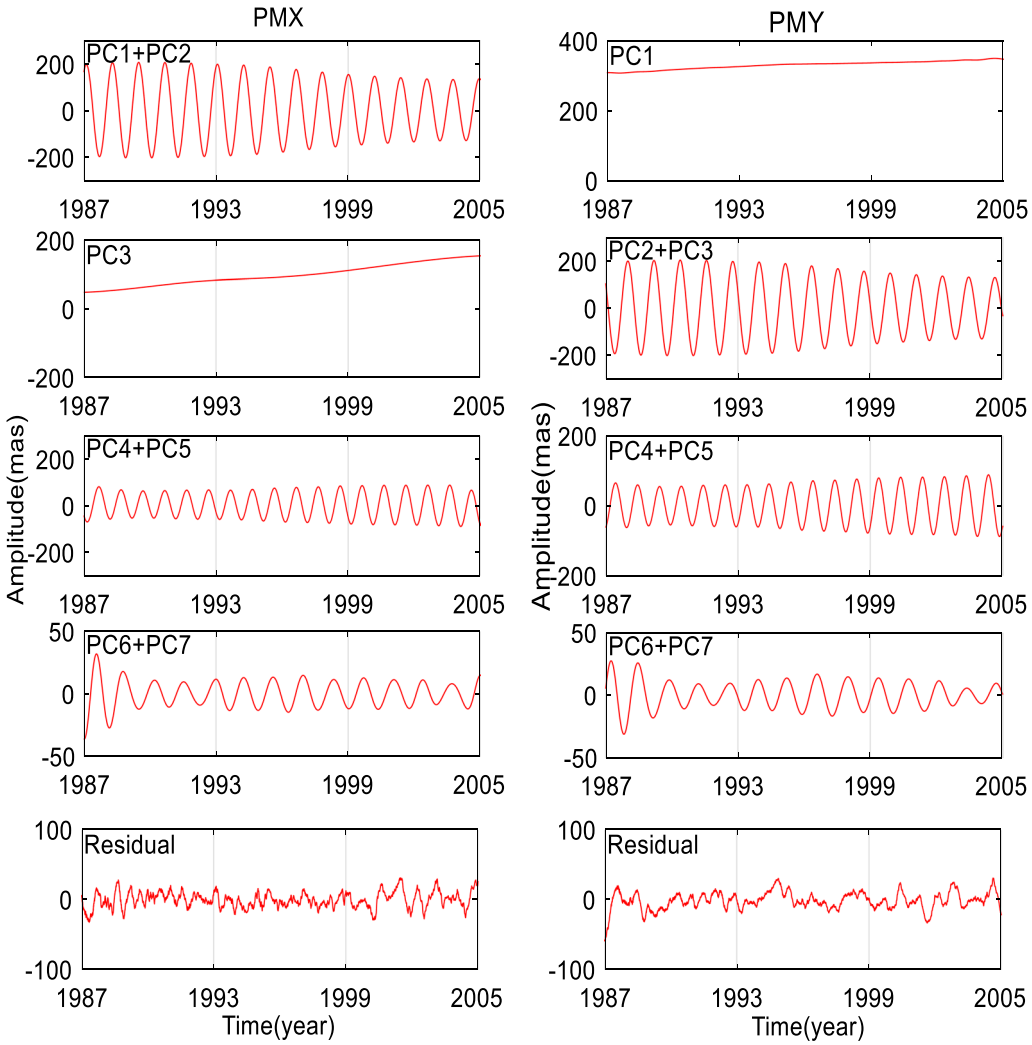


Fig. 6 PMX and PMY components obtained by SSA decomposition

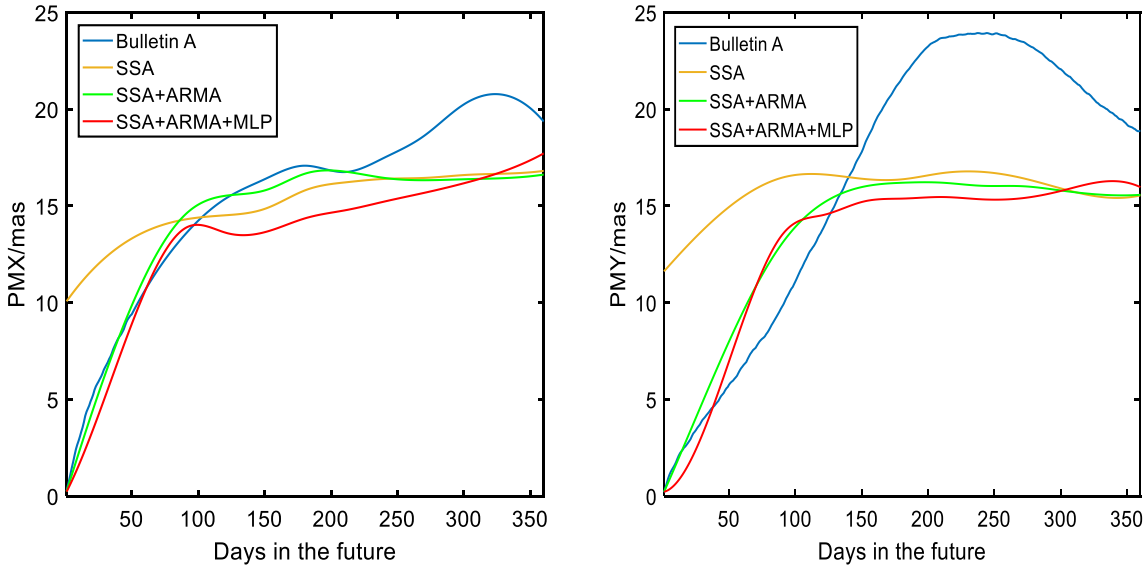


Fig. 7 MAEs of Bulletin A, SSA, SSA + ARMA, SSA + ARMA + MLP predictions from 1 to 364 days

Table 1 Comparison of the MAEs (mas) of the Bulletin A, SSA, SSA + ARMA and SSA + ARMA + MLP model

Lead time (day)	Bulletin A		SSA		SSA + ARMA		SSA + ARMA + MLP	
	PMX	PMY	PMX	PMY	PMX	PMY	PMX	PMY
5	1.49	1.09	10.44	11.93	1.12	0.84	0.80	0.35
10	2.85	1.75	10.88	12.31	2.20	1.63	1.57	0.66
15	4.15	2.39	11.29	12.68	3.26	2.44	2.40	1.12
20	5.06	2.85	11.66	13.04	4.30	3.24	3.28	1.71
25	5.92	3.41	12.01	13.39	5.31	4.05	4.20	2.41
30	6.66	3.94	12.38	13.71	6.29	5.02	5.14	3.37
60	10.64	6.69	13.70	15.50	11.50	9.61	10.59	8.95
120	15.35	13.72	14.52	16.63	15.53	15.15	13.62	14.53
180	17.08	21.56	15.82	16.38	16.68	16.20	14.37	15.38
240	17.49	23.90	16.40	16.76	16.44	16.03	15.25	15.35
300	20.25	22.07	16.60	15.89	16.38	15.79	16.20	15.79
360	19.37	18.88	16.82	15.56	16.63	15.58	17.71	15.94

a week, with approximately 52 PM prediction series per year. To enable better comparison with Bulletin A, a sequence of 52 PM series was selected each year, resulting in a total of 156 PM time series analyzed in this study. The accuracy of PM prediction was compared with Bulletin A, which was released by IERS. Figure 7 shows a comparison of MAE (364-day lead prediction), and Table 1 lists the relevant statistics. The reliability of the prediction method proposed in this study is illustrated in Fig. 7 and Table 1, with the Bulletin A solution represented in blue, the SSA-predicted data in orange, the combined results of SSA + ARMA in green, and the combined results of SSA + ARMA + MLP in red.

Based on the results presented in Fig. 7, it can be observed that the proposed method outperformed Bulletin A predictions for the PM prediction period. Specifically, for PMX, our method exhibited higher accuracy than Bulletin A for days 1–57 and 99–364. Furthermore, for the 364-day-lead prediction, the MAE of the SSA + ARMA + MLP model for the first 312-day predictions was superior to that of the SSA + ARMA model. Similarly, for PMY, the MAE of SSA + ARMA + MLP was significantly smaller than that of Bulletin A for days 1–37 and 127–364. Additionally, the SSA + ARMA + MLP model achieved better predictions than the SSA + ARMA model for days 1–70 and 106–301. The MAE for the first

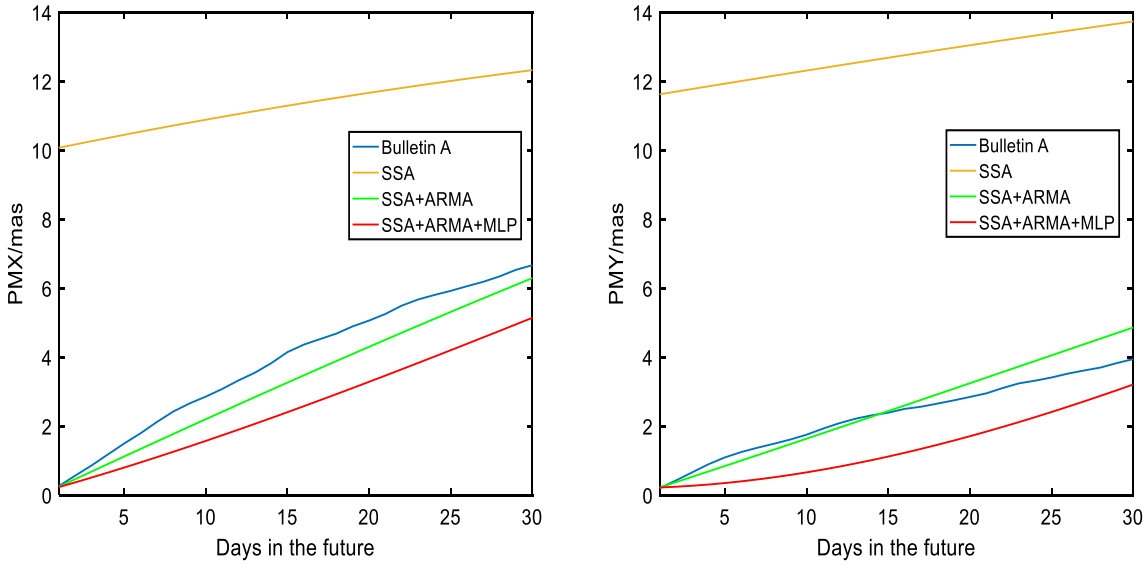


Fig. 8 MAEs of Bulletin A, SSA, SSA + ARMA, SSA + ARMA + MLP predictions from 1 to 30 days

Table 2 Comparison of the MAEs (mas) of the SSA + ARMA (Shen et al. 2018) and SSA + ARMA + MLP model

Lead time (day)	SSA + ARMA		SSA + ARMA + MLP	
	PMX	PMY	PMX	PMY
30	4.26	3.87	5.14	3.37
180	16.84	14.86	14.37	15.38
360	20.67	20.42	17.71	15.94

30-day-lead prediction were separately compared, and the results presented in Fig. 8 showed that our method outperformed Bulletin A predictions in terms of MAE for both PMX and PMY, with MAE ranging from 0.23 mas to 5.14 mas and 0.22 mas to 3.37 mas, respectively.

The MAE values are shown in Table 1. The MAEs of SSA + ARMA + MLP in the 30-day, 180-day, and 360-day predictions for PMX were 5.14mas, 14.37mas and 17.76mas, respectively, which are lower than the 6.66 mas, 17.08 mas, and 19.37 mas obtained for the Bulletin A predictions. For PMY, the MAEs from SSA + ARMA + MLP in the 30-day, 180-day, and 360-day were 3.37mas, 15.38mas, and 15.94mas, respectively, lower than those of Bulletin A (3.94mas, 21.56mas, and 18.88mas).

Because the SSA model can effectively extract the Chandler and annual wobbles of the PM series, it has an advantage in reconstructing and predicting the PC. At the same time, the ARMA model in predicting the stationary RC of the PM series after removing the trend and periodic terms were leveraged in this study. Additionally, the MLP model combined with the SSA + ARMA model enhanced the overall adaptability of the prediction model, leading to improved precision of mid-long-term PM prediction. However, it should be noted that the MAE predicted by our proposed method for days 38–126 was still greater than that of Bulletin A, particularly for PMY. The complex variation in high-frequency signals during this period may have been the cause, which needs further investigation.

The prediction and accuracy of the SSA + ARMA + MLP model in this study are compared with Shen et al (2018), and the results are shown in Table 2. As depicted in Table 2, in the Y direction, the prediction accuracy of the SSA + ARMA + MLP model surpasses that of the SSA + ARMA model (Shen et al. 2018) on the 30th and 360th days of the prediction period. Similarly, in the X direction, this study achieves significantly superior results compared to the SSA + ARMA model (Shen et al. 2018) on the 180th and 360th days of the prediction period.

Figure 9 clearly shows the improvement analysis of SSA + ARMA + MLP model. Yellow color indicates improvement, the MAE of SSA + ARMA + MLP is lower than that of Bulletin A. Green color indicates failure, the prediction accuracy of SSA + ARMA + MLP is higher than that of Bulletin A. SSA + ARMA + MLP improved the accuracy of short-term, mid-term and long-term polar motion prediction.

Figure 10 displays the distribution of absolute errors for the entire 364-day period. The results showed that for the SSA + ARMA + MLP model, the percentage of absolute errors for PMX and PMY that were smaller than 30 mas were 89.71% and 89.55%, respectively, which

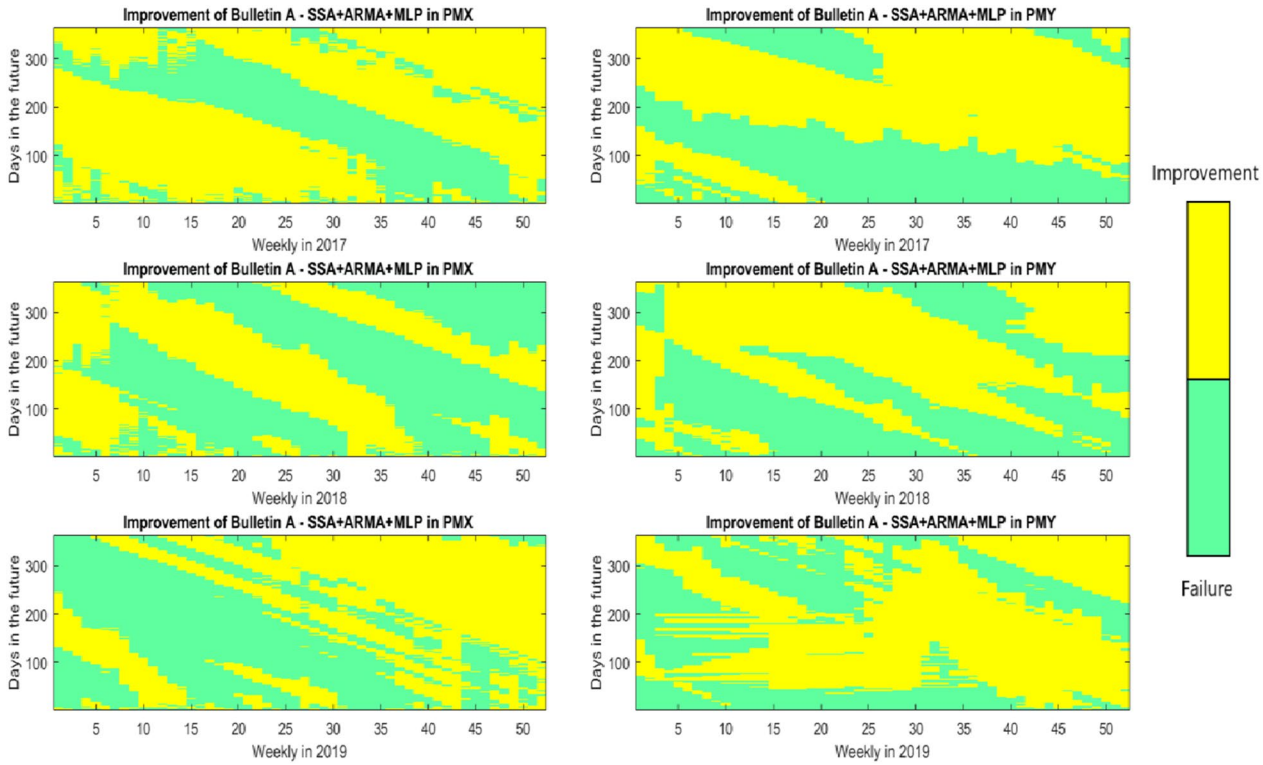


Fig. 9 Comparison of improvement in PM prediction

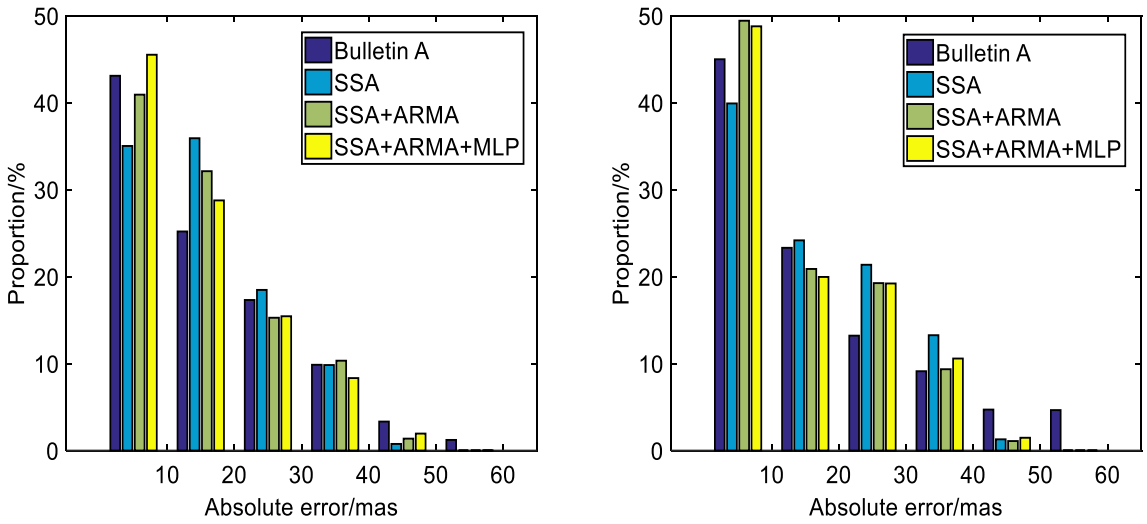


Fig. 10 Absolute error distribution of PM prediction from 1 to 364 days

are higher than the corresponding values for Bulletin A of 85.59% and 81.51%. These findings indicate that the SSA + ARMA + MLP model can significantly reduce the overall range of PM prediction errors, and the incorporation of the MLP model can further enhance the precision of PM prediction.

Conclusion

In this study, we analyzed the PM time series using SSA and demonstrated that it can effectively separate and extract the principal components of PM. Subsequently, we incorporated ARMA into the SSA to mitigate overfitting and the end effects. Considering the 'nonlinear'

characteristics of PM time series, which are caused by complex excitation source, and processing complex SSA + ARMA predicting residuals, we further introduced MLP into the SSA + ARMA model. By leveraging the nonlinear processing capabilities and deep learning techniques offered by MLP, we aimed to enhance the accuracy of mid-long term PM prediction. Summary, we considered that different predicting method have different predicting skills on different time scales, and then we proposed the sliding SSA + ARMA + MLP model for short- and long-term PM prediction and conducted 156 experiments to compare the performance of our method with that of Bulletin A. Our results show that the SSA + ARMA + MLP model outperforms Bulletin A in the 30-day-lead and 364-day-lead prediction, achieving better accuracy on days 1–57 and 99–364 for PMX and days 1–37 and 127–364 for PMY, with smaller MAEs compared to Bulletin A. This improvement can be attributed to the introduction of the MLP model, whether compared with the SSA + ARMA model proposed by other scholars or the SSA + ARMA model in this study, the introduction of MLP has improved the short-term and long-term accuracy of pole shift prediction, which enhances the overall accuracy of PM prediction. Overall, our proposed method takes the advantages of SSA, ARMA and MLP models, exhibiting high reliability and accuracy and is well suited for the prediction of PM in both short- and long-term period.

Acknowledgements
We express our sincere gratitude to the International Earth Rotation and Reference Systems Service (IERS) for generously providing the polar motion data, rapid data and Bulletin A, which were essential to the success of our research. We also would like to extend our heartfelt thanks to the editor and the reviewers for their invaluable feedback and suggestions, which significantly enhanced the quality and clarity of our manuscript. This research was supported by the National Natural Science Foundation of China (Grant Nos. 442274006, 42192535 and 42174041), Autonomous and Controllable Project for Surveying and Mapping of China (grant No. 816-517), and SDUST Research Fund (Grant No. 2014TDJH101).

Author contributions
XL, KW, and JG designed the research. KW, and XJ developed the algorithm. KW, XC, and HS analyzed data. KW wrote the manuscript with the contributions from XL. All authors were involved in discussions throughout the development.

Funding
The National Natural Science Foundation of China (Grant Nos. 42274006, 42192535 and 42174041), Autonomous and Controllable Project for Surveying and Mapping of China (Grant No. 816–517) and SDUST Research Fund (Grant No. 2014TDJH101) for supporting this study.

Availability of data and materials
The polar motion data are available at https://www.iers.org/IERS/EN/DataProducts/EarthOrientationData/eop.html.

Declarations

Competing interests
Competing interests All authors have no competing interests.

Author details
[1]College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao 266590, China. [2]Land Satellite Remote Sensing Application Center of Ministry of Natural Resources, Beijing 100048, China. [3]State Key Laboratory of Geodesy and Earth's Dynamics, Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Science, Wuhan 430077, Hubei, China.

Received: 16 June 2023 Accepted: 19 November 2023
Published online: 29 November 2023

References

Akulenko LD, Kumakshev SA, Markov YG, Rykhlova LV (2002) Forecasting the polar motions of the deformable Earth. Astron Rep 46(10):858–865. https://doi.org/10.1134/1.1515097
Bizouard C, Lambert S, Gattano C, Becker O, Richard JY (2019) The IERS EOP 14C04 solution for Earth orientation parameters consistent with ITRF 2014. J Geodesy 93(5):621–633. https://doi.org/10.1007/s00190-018-1186-3
Briganti R, Beltrami GM (2008) Singular spectrum analysis of storm surges and wave climates. J Geodesy 46(2):271–279. https://doi.org/10.1080/00221686.2008.9521960
Choi H, Moon J, Jung DH, Son JE (2019) Prediction of air temperature and relative humidity in greenhouse via a multilayer Perceptron using environmental factors. Protected Hortic Plant Factory. 28:95–103. https://doi.org/10.12791/KSBEC.2019.28.2.95
Dill R, Dobslaw H (2010) Short-term polar motion forecasts from earth system modeling data. J Geodesy 84(9):529–536. https://doi.org/10.1007/s00190-010-0391-5
Elsner JB, Tsonis A (1996) Singular spectrum analysis: a new tool in time series analysis, vol 1283. Springer berlin, Berlin, pp 932–942. https://doi.org/10.1007/978-1-4757-2514-8
Gambis D, Luzum B (2011) Earth rotation monitoring, UT1 determination and prediction. Metrologia 48(4):165–170. https://doi.org/10.1088/0026-1394/48/4/S06
Ghasemloo N, Matkan A, Alimohammadi A, Aghighi H, Mirbagheri B (2022) Estimating the agricultural farm soil moisture using spectral indices of landsat 8, and sentinel-1, and artificial neural networks. J Geovis Spat Anal 6(2):19. https://doi.org/10.1007/s41651-022-00110-4
Ghil M, Allen R, Dettinger D, Detting M, Idel K, Kondrashov D, Mann M, Robertson W, Sauunders A, Tian Y, Varadi F, Yiou P (2002) Advanced spectral methods for climatic time series. Rev Geophys. https://doi.org/10.1029/2000RG000092
Guo JY, Han YB (2009) Seasonal and inter-annual variations of length of day and polar motion observed by SLR in 1993–2006. Chin Sci Bull 54(1):46–52. https://doi.org/10.1007/s11434-008-0504-1
Hassani H (2007) Singular spectrum analysis: methodology and comparison. J Data Sci 5(2):239–257. https://doi.org/10.6339/JDS.2007.05(2).396
Heng S, Suetsugi T (2013) Coupling singular spectrum analysis with artificial neural network to improve accuracy of sediment load prediction. J Water Res Prot 5(4):395–404. https://doi.org/10.4236/jwarp.2013.54039
Jin X, Liu X, Guo JY, Shen Y (2021a) Analysis and prediction of polar motion using MSSA method. Earth Planet Space 73(1):147. https://doi.org/10.1186/s40623-021-01477-2
Jin X, Guo JY, Shen Y, Liu X, Zhao CM (2021b) Application of singular spectrum analysis and multilayer perceptron in the mid-long-term polar motion prediction. Adv Space Res 68(9):3562–3573. https://doi.org/10.1016/j.asr.2021.06.039
Kalarus M, Schuh H, Kosek W, Akyilmaz O, Bizouard CH, Gambis D, Gross R, Jovanovićet B, Kumakshev S, Kutterer H, Mendes Cerveira PJ, Pasynok S, Zotov L (2010) Achievements of the Earth orientation parameters prediction comparison campaign. J Geodesy 84(10):587–596. https://doi.org/10.1007/s00190-010-0387-1
Kong QL, Zhang LG, Han LT, Guo JY, Zhang DZ, Fang WH (2020) Analysis of 25 years of polar motion derived from the DORIS space geodetic technique using FFT and SSA methods. Sensors 20(10):2823. https://doi.org/10.3390/s20102823
Kosek W, Luzum B, Kalarus M, Wnęk A, Zbylut M (2011) Analysis of pole coordinate data predictions in the earth orientation parameters combination

of prediction pilot project. Artif Satell 46(4):139–150. https://doi.org/10.2478/v10018-012-0006-x
Kosek W, Kalarus M, Niedzielski T (2007) Forecasting of the Earth orientation parameters-comparison of different algorithms. Proc. Journèes Systèmes de Deréférence Spatio-temporels Observatoire de Paris (Paris, France, 17–19 September 2007) ed N Capitaine 155–8.
Kuremoto T, Kimura S, Kobayashi K, Obayashi M (2014) Time series forecasting using a deep brief network with restricted Boltzmann machines. Neurocomputing 137:47–56. https://doi.org/10.1016/j.neucom.2013.03.047
Lambeck K (1980) The earth's variable rotation: geophysical causes and consequences. Cambridge University Press, New York
Liao DC, Wang QJ, Zhou YH, Liao XH, Huang CL (2012) Long-term prediction of the Earth orientation parameters by the artificial neural network technique. J Geodyn 62:87–92. https://doi.org/10.1016/j.jog.2011.12.004
Lippmann RP (1988) An introduction to computing with neural nets. ACM SIGARCH Computer Architecture News 16(1):7–25. https://doi.org/10.1145/44571.44572
Marques CAF, Ferreira JA, Rocha A, Castanheira JM, Melo-Gonçalves P, Vaz N, Dias JM (2006) Singular spectrum analysis and forecasting of hydrological time series. Phys Chem Earth 31(18):1172–1179. https://doi.org/10.1016/j.pce.2006.02.061
Modiri S, Belda S, Heinkelmann R, Hoseini M, Ferrándiz JM, Schuh H (2018) Polar motion prediction using the combination of SSA and copula-based analysis. Earth Planets Space 70(1):115. https://doi.org/10.1186/s40623-018-0888-3
Pal S, Mitra S (1992) Multilayer perceptron, fuzzy sets, and classification. IEEE Trans Neural Netw 3(5):683–697. https://doi.org/10.1109/72.159058
Schoelhamer DH (2001) Singular spectrum analysis for time series with missing data. Geophys Res Lett 28(16):3187–3190. https://doi.org/10.1029/2000GL012698
Schuh H, Ulrich M, Egger D, Müller J, Schwegmann W (2002) Prediction of Earth orientation parameters by artificial neural networks. J Geodesy 76(5):247–258. https://doi.org/10.1007/s00190-001-0242-5
Shen Y, Guo JY, Liu X, Wei XB, Li WD (2017) One hybrid model combining singular spectrum analysis and LS+ARMA for polar motion prediction. Adv Space Res 59(2):513–523. https://doi.org/10.1016/j.asr.2016.10.023
Shen Y, Guo JY, Liu X, Kong QL, Guo L, Li W (2018) Long-term prediction of polar motion using a combined SSA and ARMA model. J Geodesy 92(3):333–343. https://doi.org/10.1007/s00190-017-1065-3
Singh S (2018) Backpropagation networks and multilayer Perceptron. Int J Sci Eng Res 9(7):1800–1805. https://doi.org/10.14299/ijser.2018.07.06
Stamatakos N (2017) IERS rapid service prediction center products and services: Improvement, changes, and challenges. In: Proceedings of the Journées systèmes de référence spatio-temporels pp. 2012–2017.
Su XQ, Liu LT, Hsu H, Wang GC (2014) Long-term polar motion prediction using normal time–frequency transform. J Geodesy 88(2):145–155. https://doi.org/10.1007/s00190-013-0675-7
Sun ZZ, Xu TH, Jiang CH, Yang YG, Jiang N (2019) An improved prediction algorithm for Earth's polar motion with considering the retrograde annual and semi-annual wobbles based on least squares and autoregressive model. Acta Geod Geoph 54(4):499–511. https://doi.org/10.1007/s40328-019-00274-4
Vautard R, Ghil M (1989) Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. Physica D 35(3):395–424. https://doi.org/10.1016/0167-2789(89)90077-8
Wang QX, Hu C, Xu TH, Chang GB, Moraleda AH (2017) Impacts of Earth rotation parameters on GNSS ultra-rapid orbit prediction: derivation and real-time correction. Adv Space Res 60(12):2855–2870. https://doi.org/10.1016/j.asr.2017.09.022
Wu F, Deng KZ, Chang GB, Wang QX (2018) The application of a combination of weighted least-squares and autoregressive methods in predictions of polar motion parameters. Acta Geod Geoph 53(4):247–257. https://doi.org/10.1007/s40328-018-0214-3
Wu F, Liu ZP, Deng KZ, Chang GB (2021) A polar motion prediction method considering the polar coordinates. Adv Space Res 68(3):1318–1328. https://doi.org/10.1016/j.asr.2021.03.020
Yao YB, Yue SQ, Chen P (2013) A new LS+AR model with additional error correction for polar motion forecast. Sci China Earth Sci 56(5):818–828. https://doi.org/10.1007/s11430-012-4572-3
Zhang GP (2003) Time series forecasting using a hybrid Arima and neural network model. Neurocomputing 50(1):159–175. https://doi.org/10.1016/S0925-2312(01)00702-0
Zhang Q, Wang B, He B, Peng Y, Ren M (2011) Singular spectrum analysis and Arima hybrid model for annual runoff forecasting. Water Resour Manage 25(11):2683–2703. https://doi.org/10.1007/s11269-011-9833-y
Zhao D, Lei L (2019) Possible enhancement of Earth's polar motion predictions using a wavelet-based preprocessing procedure. Stud Geophys Geod 63:83–94. https://doi.org/10.1007/s11200-018-1026-1
Zhu SY (1982) Prediction of polar motion. Bull Géodésique 56(3):258–273. https://doi.org/10.1007/BF02525586
Zotov L (2010) Dynamical modeling and excitation reconstruction as fundamental of earth rotation prediction. Artif Satell 45(2):95–106. https://doi.org/10.2478/v10018-010-0010-y

Publisher's Note