RESEARCH

Research on modeling and predicting of BDS-3 satellite clock bias using the LSTM neural network model

Shaofeng He[1,2] · Jiulong Liu[1,2] · Xiangwei Zhu[1,2] · Zhiqiang Dai[1,2] · Du Li[1,2]

Abstract

In the Global Navigation Satellite System (GNSS), the satellite clock bias (SCB) is one of the sources of ranging error, and its ability to predict directly affects the users' navigation and positioning accuracy. The BeiDou Navigation Satellite System (BDS) has three types of satellite orbits: Geostationary Orbit (GEO), Medium Earth Orbit (MEO), and Inclined Geosynchronous Orbit (IGSO). The BDS satellites are equipped with rubidium and hydrogen atomic clocks with different error characteristics. Establishing a reliable SCB prediction model is essential for real-time precise point positioning, precise orbit determination, and optimization of navigation message parameters. In this research, we apply a long short-term memory (LSTM) model for predicting BDS-3 SCB, which uses a multiple single-step predicting method to avoid error accumulation. Short- (0–6 h), medium- (6 h–3 days), and long-term (3–7 days) predicting is performed, and the results are compared with those of two traditional models to verify the reliability and accuracy of the LSTM method. For the BDS-3 IGSO satellites, the short-, medium-, and long-term accuracy is better than 0.5, 1.8, and 19.2 ns, respectively; for the BDS-3 MEO satellites, the short-, medium-, and long-term accuracy is better than 0.3, 1.4, and 8.6 ns. For long-term prediction of SCB, LSTM improves the accuracy by 72.0 and 64.0% compared to the autoregressive integrated moving average (ARIMA) and quadratic polynomial (QP) model, respectively.

Abbreviations

ARIMA	Autoregressive integrated moving average
BDS	BeiDou Navigation Satellite System
CODE	Center for Orbit Determination in Europe
FDSCB	First differences of satellite clock bias
GEO	Geostationary Orbit
GFZ	German Research Center for Geosciences
GM	Gray model
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
IGS	International GNSS Service
IGSO	Inclined Geosynchronous Orbit
KF	Kalman filter
LP	Linear programming
LSTM	Long short-term memory
MAD	Median absolute deviation
MEO	Medium Earth Orbit
MGEX	Multi-GNSS experiment
PRN	Pseudo-random noise ranging code
QP	Quadratic polynomial
RMSEs	Root-mean-square errors
RNN	Recurrent neural network
SA	Spectral analysis
SCB	Satellite clock bias
SL-LSTM	Supervised learning long short-term memory
TWSTF	Two-way satellite-ground time transfer
WHU	Wuhan University
WNN	Wavelet neural network

✉ Jiulong Liu
  liujlong8@mail.sysu.edu.cn

1 Shenzhen Campus of Sun Yat-Sen University, Shenzhen 518107, China

2 School of Electrical and Communication Engineering, Sun Yat-Sen University, Guangzhou 510000, China

Introduction

As a timing base for Global Navigation Satellite Systems (GNSS), the performance of satellite-based atomic clocks directly affects navigation, positioning, and timing accuracy. Since the post-processed satellite clock bias (SCB) products provided by the International GNSS Service (IGS) have

higher accuracy than those broadcasted by the broadcast ephemeris, it is important to use the IGS SCB product to establish a high-precision prediction model and improve the prediction accuracy of SCB. The commonly used prediction models include linear programming (LP) (Cernigliaro and Sesia 2012), quadratic polynomial (QP) (Huang et al. 2011; Huang et al. 2018), gray model (GM) (Liang et al. 2015; Lu et al. 2008), Kalman filter (KF) (Davis et al. 2012; Zhu et al. 2008), autoregressive integrated moving average (ARIMA) (Xi et al. 2014), and spectral analysis (SA) (Zheng et al. 2010; Zhao et al. 2021).

These prediction models have advantages and disadvantages: LP and QP models have the advantages of straightforward calculation and better short-term prediction accuracy. However, the accuracy and stability deteriorate significantly with time (Jonsson and Eklundh 2002). The GM model, ARIMA, and KF are affected by the characteristics of atomic clocks, the choice of key parameters, and the priori values of parameters (Zheng et al. 2009; Huang et al. 2014; Xu et al. 2013). The SA model requires a long time series of stable clock bias to implement its functionality (Heo et al. 2010). Since the time-varying properties of satellite-based atomic clocks are complex and are affected by many factors, neural network models have been applied to predict SCB. For example, the wavelet neural network (WNN) model (Wang et al. 2017) and supervised learning long short-term memory (SL-LSTM) model (Huang et al. 2021) have been used to predict the SCB of Global Positioning System (GPS) obtaining promising results.

The BeiDou Navigation Satellite System (BDS) differs from GPS in that its satellites operate on three types of orbits. The BDS-3 satellite constellation contains three Geostationary Orbit (GEO) satellites, three Inclined Geosynchronous Orbit (IGSO) satellites, and 24 Medium Earth Orbit (MEO) satellites. BDS provides worldwide basic navigation and satellite-based augmentation services. Two-way satellite-ground time transfer (TWSTF) is used for measuring SCB in BDS-2 broadcast ephemeris. BDS-3 increases inter-satellite link measurements compared to BDS-2 (Zhou et al. 2016; Liu et al. 2020) and generates SCB parameters in broadcast ephemeris using a combined short-term and long-term QP fit. IGS data centers, such as Wuhan University (WHU), Center for Orbit Determination in Europe (CODE), and German Research Center for Geosciences (GFZ), use observations from the multi-GNSS experiment (MGEX) to perform precise orbit determination of GPS, GLONASS, Galileo, and BDS, and provide precise orbit and clock bias products (Steigenberger et al. 2015; Zhao et al. 2013; Chen et al. 2014). The basic information on BDS satellites currently operating in orbit is listed in Table 1.

Neural networks in deep learning have developed rapidly (Dong et al. 2021), among which long short-term memory (LSTM) network has great potential and effectiveness in the prediction of nonlinear non-stationary time series. The BDS in-orbit satellite clock bias has complex characteristics whose prediction accuracy can still be improved. In this research, we apply the LSTM model for predicting the SCB of BDS-3 satellites and analyze the prediction performance of different in-orbit satellites over various time scales, to establish a reliable SCB prediction model.

In this research, the extraction process is applied to the dataset to form the input data and corresponding labels of the LSTM model. This method achieves multiple single-step predictions with the practical significance of avoiding the error accumulation problem caused by multi-step prediction. It improves processing speed by reducing matrix transposition and block division operations. In addition, the hyperparameters of the LSTM model are given for predicting 7-day BDS-3 SCB. Experiments on different types of BDS-3 satellites indicate that the model performance changes little when fine-tuning the model hyperparameters. Thus, a relatively robust LSTM model suitable for Beidou-3 SCB prediction is obtained.

Data sources and data pre-processing are introduced first. Next, the LSTM method is explained, and then, the results are shown. Furthermore, the results are compared in detail. The conclusion is given in the final part.

Data pre-processing

The 5-min interval high-precision SCB data for 70 days are selected as provided by the IGS Data Center of Wuhan University from August 30 to November 7, 2021. The SCB data of IGSO and MEO are analyzed since most IGS data centers do not currently provide the precise ephemeris data of BDS-3 GEO satellites.

The original SCB data are subject to anomalies such as data discontinuities, jumps, and gross errors (Feng 2009), which negatively impact the modeling and prediction of SCB. Data jumps are solved by using a moving window to find jump points (Zhou et al. 2016); the SCB data jump detection algorithm is based on the Hilbert transform (Guo

Table 1 BDS in-orbit satellites

Generation	Orbit	PRN (pseudo-random noise ranging code)
BDS-2	IGSO	6,7,8,9,10,13,16
	MEO	11,12,14
	GEO	1,2,3,4,5,18
BDS-3S	IGSO	31,56
	MEO	57,58
BDS-3	IGSO	38,39,40
	MEO	19,20,21,22,23,24,25,26,27,28,29,30,32,33,34,35,36,37,41,42,43,44,45,46
	GEO	59,60,61

2013). For the problem of data discontinuity, a method of dividing the time-domain Allan variance by the Barnes deviation function was proposed to deal with data discontinuity (Riley and Howe 2008), and there exist methods for processing non-equidistant data (Hackman and Parker 1996). The methods described above can deal with these two types of problems. Maximum SCB data are selected without interruptions and jumps in the 70-day cycle. Gross errors are detected and eliminated necessarily since they always exist.

As shown in Fig. 1, PRN 35 and PRN 39 are the BDS-3 MEO and IGSO satellites, respectively. The abscissa "Epoch" represents the number of epochs, and the epoch interval is 5 min, while the ordinate represents the deviation of the satellite clock in the precise ephemeris relative to the ephemeris reference clock, i.e., SCB. The overall trend can be observed through the original SCB sequence,
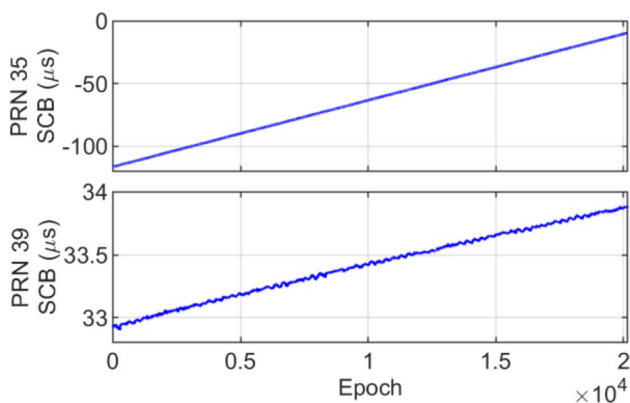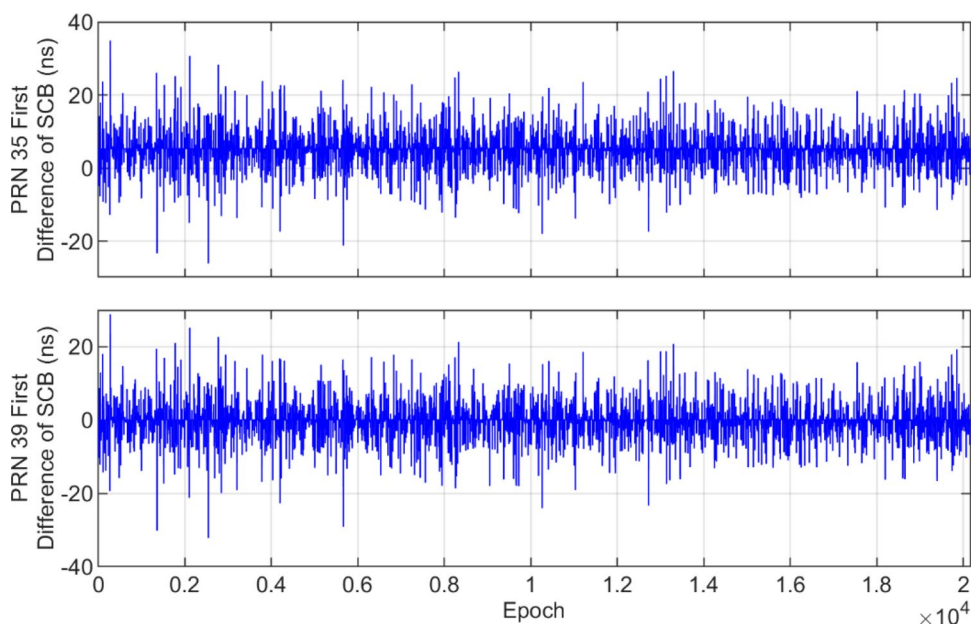

Fig. 1 Original SCB sequences of BDS-3 satellites in different orbits. Top: SCB of PRN 35 (BDS-3 MEO). Bottom: SCB of PRN 39 (BDS-3 IGSO)

and the order of magnitude is μs. But the gross errors in the sequence are in the order of ns; thus, it is not easy to be observed.

One of the most efficient ways to detect gross errors is to perform first differences on the original SCB sequence. For the SCB sequence $L = \{l_1, l_2, ..., l_j..., l_n\}$, the first-difference sequence is $\Delta L = \{\Delta l_1, \Delta l_2, ..., \Delta l_j..., \Delta l_{n-1}\}$, where $\Delta l_j = l_{j+1} - l_j$, as shown in Fig. 2. It can be seen that there are some obvious abnormal data (in the order of 10 ns) called gross errors. The existence of gross errors will have a serious negative effect on the prediction accuracy of SCB, which has to be eliminated.

The median absolute deviation (MAD) method is taken to replace the gross errors, whose mathematical model is expressed by

$$\left|\Delta l_j\right| > (k + n \times \text{MAD}), \tag{1}$$

where $k = Median\{\Delta L\}$ is the median of first differences of SCB; $\text{MAD} = Median\{|\Delta l_j - k|/0.6745\}$ is the absolute median deviation; the constant n is determined as needed. When Δl_j satisfies Eq. (1), it can be considered that Δl_j is a gross error. When gross errors are detected, the outlier can be replaced by a method such as the zeros padding or the cubic spline interpolation.

However, the traditional MAD method has limitations because the first-difference sequence still has a trend term. As shown in Fig. 3, the curve is the result of using the traditional MAD method to eliminate the gross errors, in which the first-difference sequences of PRN 35 and PRN 39 have very obvious trend terms. This can be seen from the expression:

Fig. 2 First differences of SCB sequences of BDS-3 satellites in different orbits. Top: First differences of PRN 35 (BDS-3 MEO). Bottom: First differences of PRN 39 (BDS-3 IGSO)

Fig. 3 Traditional MAD method to eliminate gross errors. Top: First differences after traditional MAD of PRN 35 (BDS-3 MEO). Bottom: First differences after traditional MAD of PRN 39 (BDS-3 IGSO)
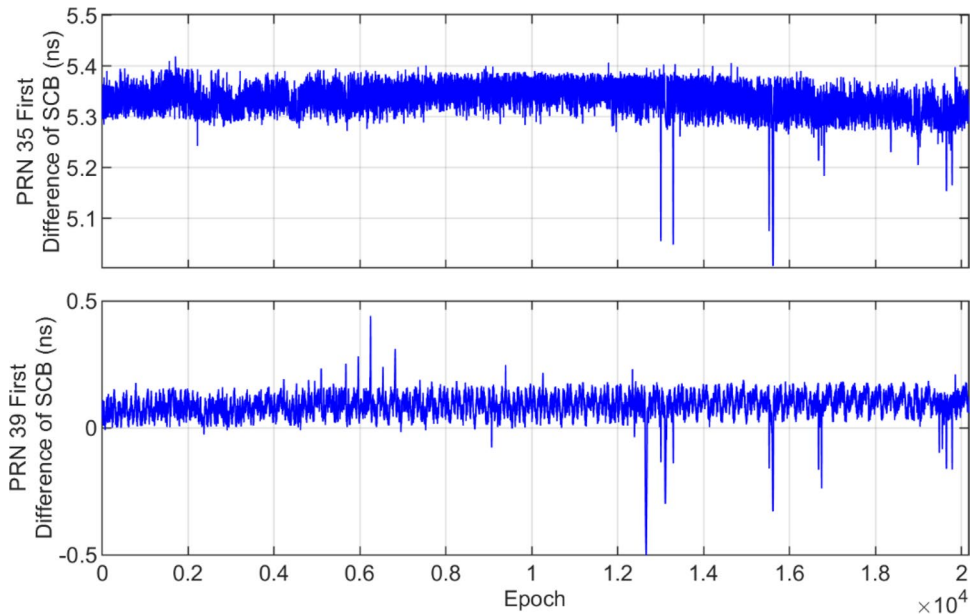
$$\Delta t^s = a_0 + a_1(t - t_0) + a_2(t - t_0)^2 + \Delta t_r, \tag{2}$$

where t_0 is the initial moment (ephemeris reference time). $a_0, a_1, and\ a_2$ are the clock bias, clock speed, and clock drift, respectively. Δt_r is the random term, whose statistical characteristics are only described by the clock stability. Ignoring the random term, the equation is a quadratic polynomial model. In practice, the magnitudes of $a_0, a_1, and\ a_2$ are about 10^{-4}, 10^{-10}, and 10^{-20} s. For one satellite, only the influence of a_1 and a_2 remains after the first differencing. Since a_2 is much smaller than a_1, the traditional MAD method does not consider the influence of a_2 so that the k value in the algorithm is a fixed constant. According to Eq. (1), data outside a certain range near the fixed k value will be eliminated and replaced, resulting in a serious loss of original data and a great increase in the prediction error, so the influence of a_2 cannot be ignored.

A straightforward idea is to make the sequence more "horizontal" by taking second differences and then take the traditional MAD method to eliminate gross errors. However, according to the principle of error propagation, it will cause the noise mean square error to be $\sqrt{2}$ times the original.

A dynamic MAD method (Huang et al. 2022) was proposed based on ridge regression (a modified least squares estimation method), where the trend term is fitted first, and then, the idea of MAD is applied to replace the gross errors. In this method, the authors redefine the expression for calculating k. When Δl_j satisfies the following conditions:

$$\begin{cases} \left| \Delta l_j - k \right| > n \times \text{MAD} \\ k = \hat{\omega} = \left(X^T X + \lambda I\right)^{-1} X^T y \end{cases}, \tag{3}$$

it can be judged as a gross error. In Eq. (3), X is the time sequence; y is the first difference; I is the unit matrix; λ is the regularization coefficient, which needs to be determined from empirical values; and $\hat{\omega}$ is the ridge regression value. k is the trend term contained in the first difference. The data outside a certain range near the k value are considered gross errors, which are eliminated and replaced by the cubic spline interpolation method, ensuring that as many original data are retained as possible, thus completing the data pre-processing step.

Methodology

Deep learning methods have been more widely used in solving nonlinear problems (Dong et al. 2021), and Recurrent Neural Network (RNN) is one of these approaches. RNN is a neural network with nodes directed to connect into a ring to find the series correlation by using the characteristics of its network structure, which is suitable for predictive filtering of time series and has been successfully applied in temporal data such as text, video, and speech.

Figure 4 shows the schematic diagram of RNN. The hidden node would have two parameters: weight and bias. However, a recurrent layer has three parameters to optimize: weight for the input, weight for the hidden unit, and bias. The parameters of the hidden layers in the RNN model inherit the output of the previous moment. It makes RNN that has the "memory" ability compared with other neural networks, so it can process the time series more effectively.

LSTM is a variant of RNN that can solve the gradient disappearance and gradient explosion problems during the training of RNN models (Yu et al. 2019). A cell in the LSTM

Fig. 4 Schematic diagram of RNN. RNN has memory ability by using neurons with self-feedback, so it is suitable for processing time-series data
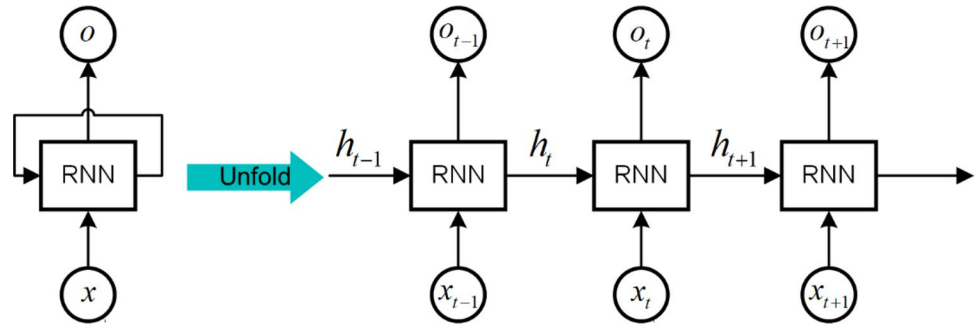
model has four gates, which enables the model to have long short-term memory and solves the gradient disappearance and gradient explosion problems. By introducing the "gate" mechanism, LSTM can give the network stronger memory ability and thus obtain better results for longer series. Figure 5 shows an LSTM neural network model. We provide a brief analysis of the applicability of the LSTM model based on the characteristics of the SCB sequence.

As the cell structure in the LSTM neural network model shown in Fig. 5, a cell has four gates: input gate i_t, forget gate f_t, select gate $\tilde{C}_t$, and output gate o_t, which are expressed as follows:
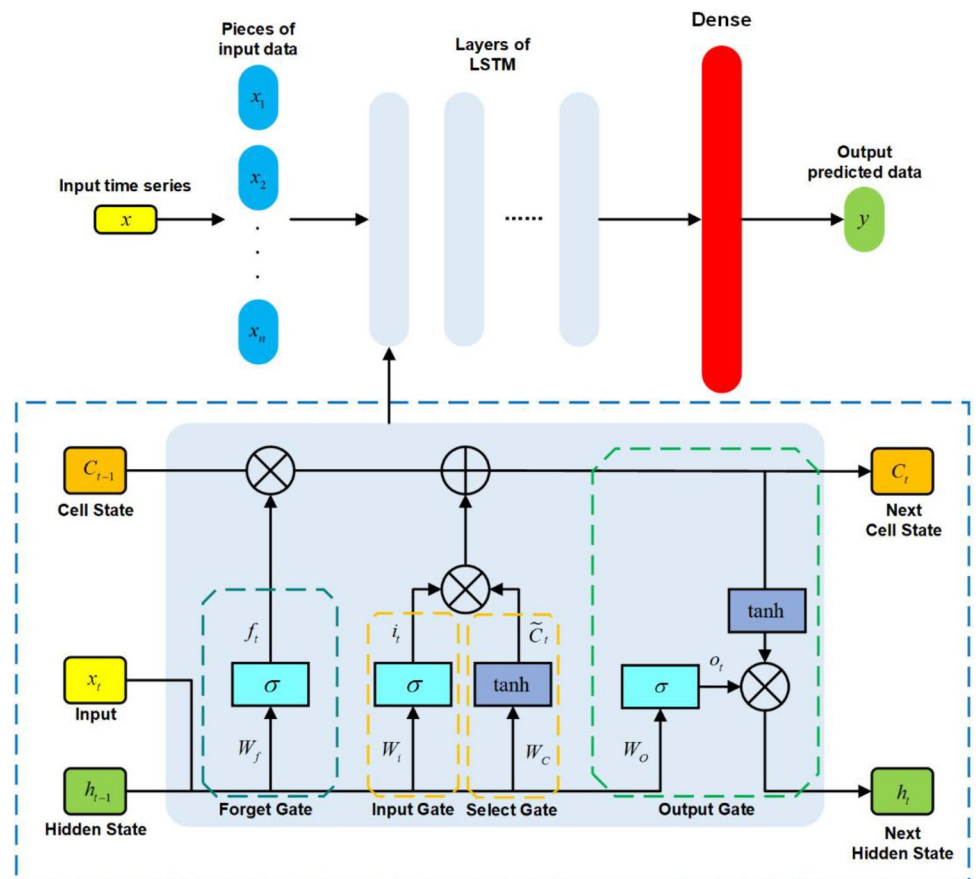
$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right), \tag{4}$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t, \tag{5}$$

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right), \tag{6}$$

$$\tilde{C}_t = \tanh\left(W_C \cdot [h_{t-1}, x_t] + b_C\right), \tag{7}$$

$$o_t = \sigma\left(W_o \cdot [h_{t-1}, x_t] + b_o\right), \tag{8}$$

Fig. 5 LSTM neural network model. Each cell of the LSTM model has four gates, which enable the model to have long short-term memory and solve the gradient disappearance and gradient explosion problems in RNN

$$h_t = o_t * \tanh\left(C_t\right), \tag{9}$$

where σ denotes the sigmoid activation function. h_t, C_t, and x_t are the hidden state, the cell state, and the input at moment t, respectively. W_* and b_* denote the weight matrix and bias of the current network layer, respectively. The information processing of LSTM is divided into three main stages:

1. Forgetting stage: In Eqs. (4) and (5), the input from the previous state node is selectively forgotten, that is, to control what needs to be remembered or forgotten in the previous state C_{t-1} by f_t.
2. Selective memory stage: In Eqs. (6) and (7), the inputs are remembered selectively, which important ones will be focused on and which unimportant ones will be remembered less. i_t is the input to the current cell, which can be selectively output by $\tilde{C}_t$.
3. Output stage: In Eqs. (8) and (9), this stage is controlled by o_t and scaled using the tanh function on C_t to determine which outputs are used as the current state.

In Eq. (2), the random term Δt_r is affected by the physical characteristics of the atomic clock itself, which leads to the nonlinear and non-stationary characteristics of the SCB. The LSTM cell contains multiparameter nonlinear functions that can be automatically adjusted during training. Considering this characteristic, an LSTM neural network model is proposed for BDS-3 SCB prediction.

It is a fact that there exists the problem of prediction error accumulation. Time-series predicting includes single-step predicting and multi-step predicting. Single-step predicting refers to predicting one future value using all true values, while multi-step predicting refers to predicting multiple future values. In this case, numerous values are predicted in the future. If the multi-step predicting method is taken, the prediction error will be gradually accumulated because one predicted value may be affected by the previously predicted ones. In contrast, the single-step predicting method is taken to predict only one value without any error accumulation problem. So, the dataset is processed to make it into a multiple single-step predicting problem to avoid error accumulation.

For the whole time series $\{x_1, x_2, ..., x_N\}$ in the training dataset, let p be the number of prediction days. In the training phase, all data are extracted with p as the sampling interval. The positive integer d is the dimension of the input data of the LSTM model and needs to satisfy:

$$d \le \frac{N}{p} - 1, \tag{10}$$

then, it produces $N - dp$ sequences:

$$\begin{bmatrix} \{x_1, x_{1+p}, x_{1+2p}, \dots, x_{1+(d-1)p}, x_{1+dp}\} \\ \{x_2, x_{2+p}, x_{2+2p}, \dots, x_{2+(d-1)p}, x_{2+dp}\} \\ \dots \\ \{x_{N-dp}, x_{N-(d-1)p}, x_{N-(d-2)p}, \dots, x_{N-p}, x_N\} \end{bmatrix}, \tag{11}$$

The first d data of each sequence are used as input data during training, and the last ones are used as labels. Before training the LSTM model, all data need to be normalized first to improve the accuracy of the training model and the speed of model convergence. In the predicting stage, the latter d data of each sequence are taken as input to the trained LSTM model for predicting, and the output sequence $\{x_{1+(d+1)p}, x_{2+(d+1)p}, ..., x_{N+p}\}$ are the values of $N - dp$ single-step prediction results. Take the last p values, add the corresponding trend terms, and get the final SCB predictions by inverse difference (generate clock bias series using first differences).

By processing the dataset this way, the LSTM model does not have the problem of prediction error accumulation during predicting SCB, making the results more accurate and reliable. The implementation flowchart is shown in Fig. 6: (1) perform first differences on the original SCB data and take the ridge regression MAD method to eliminate gross errors, then subtract the fit trend term to obtain the pre-processed first differences of satellite clock bias (FDSCB). (2) Before training, the pre-processed data need to be normalized and dimensioned. The model parameters are saved at the end of the training. (3) Use the trained model for predicting. (4) The model outputs need to go through the reverse process of data pre-processing to obtain the final prediction results.

In this research, gross errors which will decline the prediction accuracy are detected first by performing first differences on the original SCB data, and then, the MAD method based on ridge regression is taken to replace the gross errors. Then, a more stable FDSCB series is obtained by subtracting the trend term, which is suitable as the input to the LSTM model. Next, it avoids the accumulation of the prediction error problem from multi-step predicting by processing the dataset and turning it into a multiple single-step predicting problem. Finally, the average root-mean-square errors (RMSEs) of all experimented satellites are calculated, and the SCB prediction accuracy of LSTM is compared with that of two traditional models (ARIMA and QP).

Results

The original SCB data are downloaded from the IGS Data Center of Wuhan University. The FDSCB data are generated after eliminating the gross errors as the input of the

Fig. 6 Implementation flow-chart. The whole process is divided into three parts, data pre-processing (obtaining a stationary sequence), model training and predicting, and data post-processing (reconstructing the clock difference sequence)
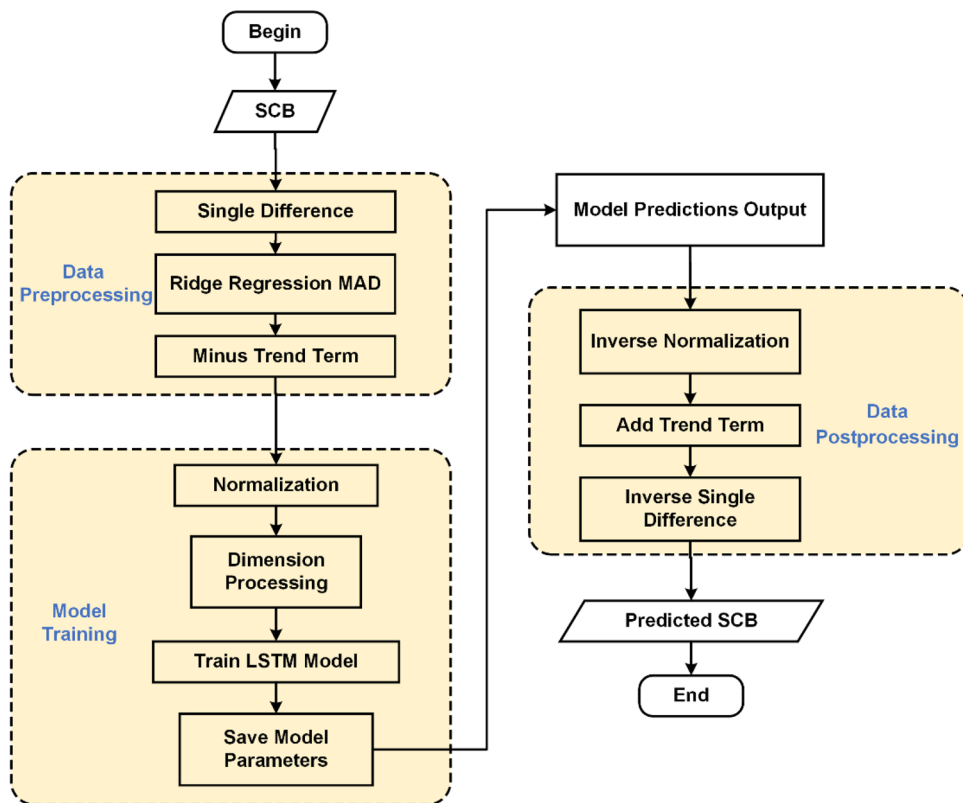
LSTM model to obtain the model output. The final SCB prediction data are obtained after data post-processing. All processing results for two satellites in different orbits during each step are shown.

Data pre-processing results

As shown in Fig. 7, the blue line represents the fitted results of the trend term k, and the red one represents the fitted residuals, which are more stable and easier to fit and suitable to input into the LSTM model

Fig. 7 Fitting the trend term by ridge regression MAD (in blue) and pre-processed FDSCB (in red). Top: The fit trend term values and fitted residuals after ridge regression MAD of PRN 35 (BDS-3 MEO). Bottom: The fit trend term values and fitted residuals after ridge regression MAD of PRN 39 (BDS-3 IGSO)
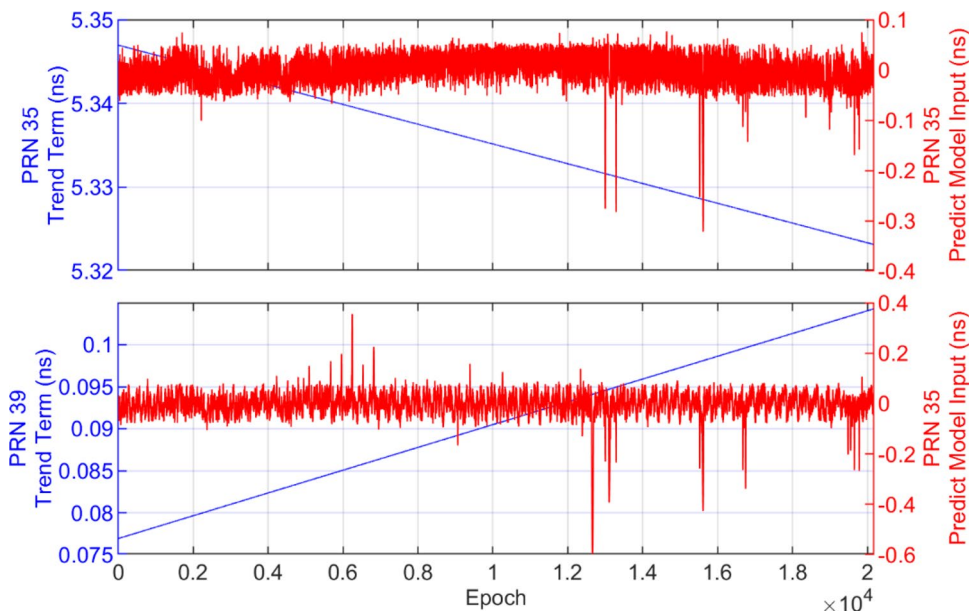
Table 2 List of the LSTM model parameters

No.	Parameter	Value
1	Loss function	MSE
2	Optimizer	Adam
3	Number of LSTM layers	2 layers (64 units + 64 units)
4	Number of dense layers	1
5	Number of training rounds	2000
6	Input dimension size	8
7	Output dimension size	1
8	Size of batch	16
9	Validation split rate	20%
10	Learning rate	0.01

Table 3 Final training error statistics based on the LSTM model for some satellites

PRN	Orbit	Training loss
35	BDS-3 MEO	2.1208×10^{-4}
39	BDS-3 IGSO	6.5023×10^{-4}

to predict. It completes the data pre-processing step. The range of FDSCB after pre-processing falls mostly between ± 0.5 ns, which means that the gross errors have almost been eliminated and replaced.

LSTM model output results

The LSTM model is established based on the Keras 2.7.0 deep learning environment with the specific parameters in Table 2. The 70-day FDSCB can be obtained through data pre-processing, after which 63 days of data are input as training data into the above-designed LSTM model for training. Some final losses during training are shown in Table 3.

Then, the trained model is used for predicting, and the output series of the LSTM model is inverse normalized to obtain the FDSCB predictions. Figure 8 shows the comparison between the FDSCB predicted and true values for the next 7 days.

### SCB prediction results based on different models

The SCB prediction accuracy of the LSTM model is compared with two baseline models, ARIMA and QP. Figure 9 shows the deviations of the SCB predictions of BDS-3 satellites in short (0–6 h), medium (6 h–3 days), and long (3–7 days) terms based on LSTM, ARIMA, and QP.

Table 4 shows the statistical results of the average RMSE ($\overline{\text{RMSE}}$) of BDS-3 SCB predictions with good data continuity and no obvious jumps using the three models (results with 4 decimal places). For a single satellite i, the RMSE is calculated as follows:

$$\text{RMSE}_i = \sqrt{\frac{1}{p} \sum_{k=1}^{p} (x_k - \hat{x}_k)}, \tag{12}$$

where $\hat{x}_k$ and x_k are the predicted and true values of SCB at the moment k, respectively, and p is the number of

Fig. 8 LSTM prediction output results. The blue curve is the true value of FDSCB, and the red one is the corresponding prediction. Top: The model output of PRN 35 (BDS-3 MEO). Bottom: The model output of PRN 39 (BDS-3 IGSO)
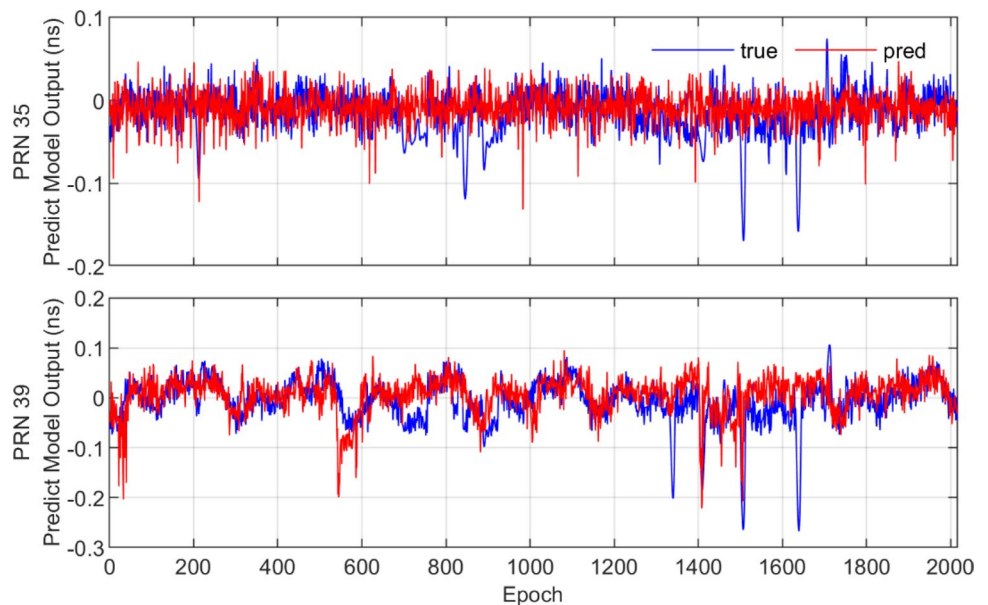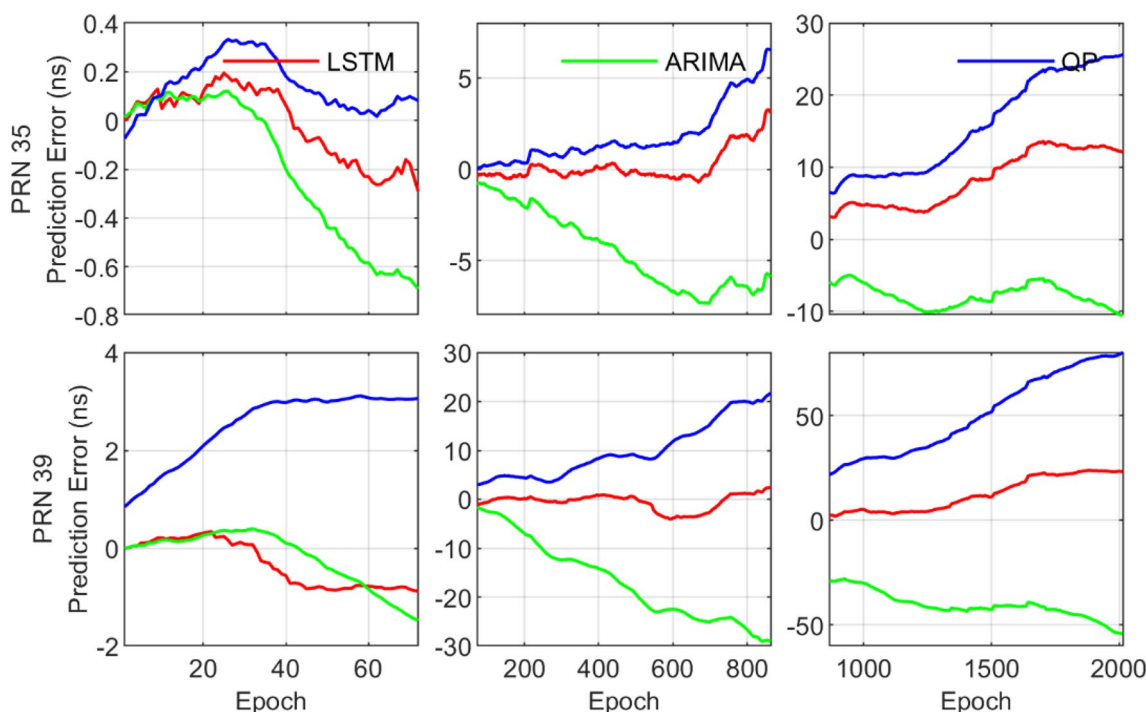
Fig. 9 Deviation diagram of short-, medium-, and long-term SCB predictions based on LSTM, ARIMA, and QP. Top: PRN 35 prediction error. Bottom: PRN 39 prediction error. Left: 0–6 h. Middle: 6 h–3 days. Right: 3–7 days. It means the prediction error at different stages

Table 4 $\overline{RMSE}$ of SCB predictions of BDS satellites based on different models

Orbit	Model	0–6 h (ns)	6 h–3 days (ns)	3–7 days (ns)
IGSO	LSTM	0.4310	1.7874	19.1598
	ARIMA	1.4985	21.1576	68.4734
	QP	2.1796	9.8175	53.2732
MEO	LSTM	0.2316	1.3899	8.5490
	ARIMA	0.4231	11.6385	60.3322
	QP	2.0192	10.4359	46.2189

predictions. For satellites of the same orbit, the $\overline{RMSE}$ is calculated as follows:

$$\overline{RMSE} = \frac{1}{m} \sum_i RMSE_i, \tag{13}$$

where m is the number of satellites in the same orbit.

Discussion

Table 3 shows that the training loss values of the trained LSTM model using the MSE loss function for the data within 63 days are all less than 10^{-3}, which tells us that the

LSTM model has successfully performed regression fitting on the training dataset.

The 63-day pre-processed FDSCB is input into the trained LSTM model to predict the data for the next 7 days. It can be seen from the comparison results of the 7-day predicted values and the true values in Fig. 8 that the predicted curve is in good agreement with the true curve and reflects the trend and cycle characteristics of the true data. Therefore, the strong nonlinear fitting ability of the LSTM model is verified.

As shown from Fig. 9, the prediction accuracy of PRN 35 (BDS-3 MEO) in the short, medium, and long term is no more than 0.8, 8, and 27 ns, respectively. When it comes to PRN 39 (BDS-3 IGSO), they are no more than 3.2, 30, and 80 ns, respectively. The LSTM model has no obvious advantage in short-term predicting and is slightly worse than QP on satellite 35 and slightly worse than ARIMA on satellite 39. In medium- and long-term predicting, LSTM has the best effect, followed by QP, and ARIMA is the worst.

In Table 4, for all orbit types and all models, the SCB prediction errors increase as the prediction time increases. Compared with the other two models, the prediction ability of the LSTM model for the satellite clock errors of the two orbit types is optimal in the short, medium, and long term. For the BDS-3 IGSO satellites, the short-term prediction accuracy of LSTM is 0.4310 ns, which is 71.2% higher than ARIMA and 80.2% higher than QP; the medium-term

prediction accuracy is 1.7874 ns, which is 91.5% higher than ARIMA and 81.7% higher than QP; and the long-term prediction accuracy is 19.1598 ns, which is 72.0% higher than ARIMA and 64.0% higher than QP. For the BDS-3 MEO satellites, the short-term prediction accuracy of LSTM is 0.2316 ns, which is 45.3% higher than ARIMA and 88.5% higher than QP; the medium-term prediction accuracy is 1.3899 ns, which is 88.1% higher than ARIMA and 86.6% higher than QP; and the long-term prediction accuracy is 8.5490 ns, which is 85.8% higher than ARIMA and 81.5% higher than QP.

Table 5 shows the accuracy improvement of the LSTM model compared with QP and ARIMA. As shown in Fig. 9, for the same type of single satellite, the short-term prediction accuracy may be slightly inferior to the other two traditional models, but after statistical processing of the experimental results, it can still show the advantage of the LSTM model in the performance of SCB prediction, especially in the long-term predicting.

Experiments show that the LSTM model is comparable to the other two in short-term (within 6 h) prediction and performs better above 6 h. The average total time to process each satellite, i.e., train the model using 63 days of data and use the fitted model to predict SCB in the next 7 days, is about 20 min. With the same input data, the LSTM model takes longer to predict. This is a common problem of basic deep learning; that is, obtaining a neural network model through data-driven offline learning requires more computing time than traditional algorithms. Our operating environment is CPU R7 5800H and graphics card RTX3060. The time it takes about 20 min to predict SCB for the next 7 days is acceptable. In long-term predicting, when real-time broadcast ephemeris and precise ephemeris cannot be obtained in some cases, it is meaningful work to use the historical data of post-processed high-precision clock products to train the model to predict SCB.

While traditional QP and ARIMA models have achieved relatively good results in short-term predicting, prediction errors will diverge over time, resulting in poor predicting effects in medium- and long-term predictions. To date, there are relatively few studies on applying the LSTM neural network model to BDS-3 SCB prediction. The LSTM model in the field of deep learning realizes long sequence processing through a gating mechanism, which has considerable advantages in medium- and long-term predicting. Undeniably, the complexity of the deep learning neural network model is higher than that of traditional methods, but the LSTM method in this research can achieve better accuracy than QP and ARIMA in medium- and long-term SCB predicting.

Conclusions

This research applies an LSTM model for predicting BDS-3 SCB, including the data pre-processing and the cross-sectional comparison and analysis with two traditional models. To address the problem that the time series in engineering are nonlinear and non-smooth, which leads to the low accuracy of the predicting models in the traditional methods, we used a deep learning method to establish a multiple single-step predicting LSTM model using the precise ephemeris provided by the IGS Data Center of Wuhan University. The results are summarized and analyzed, which show that:

1. No matter which model is used, the clock difference prediction error will disperse with time. The LSTM model performs the short-, medium-, and long-term accuracy better than 0.5, 1.8, and 19.2 ns for the BDS-3 IGSO satellites. Meanwhile, it performs the short-, medium-, and long-term accuracy better than 0.3, 1.4, and 8.6 ns for the BDS-3 MEO satellites.
2. For satellites in the same orbit, the LSTM is the best, the ARIMA is the second best, and the QP is the worst for short-term SCB predicting in most cases; the LSTM is the best, the QP is the second best, and the ARIMA is the worst for medium- and long-term SCB predicting on all experimental satellites.

Basic deep learning is equivalent to the problem of function approximation, that is, the fitting of functions. Different from polynomial, B-spline, trigonometric, and wavelet functions used in traditional mathematics, nonlinear neural network functions are used as basis functions in deep learning. Among them, the LSTM neural network model is suitable for relatively long sequence prediction tasks, which improves the prediction accuracy of BDS-3 SCB. It positively impacts the reduction in positioning errors and is also expected to be applied to precise orbit determination.

Due to the limitation of the LSTM structure, the accuracy is not stable in short-term predicting. In future studies, the more popular attention-based transformer in recent years will be considered to model SCB prediction to achieve higher prediction accuracy.

Table 5 Accuracy improvement table of LSTM model compared with QP and ARIMA in percentage

Orbit	Model	0~6 h (%)	6 h~3 days (%)	3 days~7 days (%)
IGSO	LSTM:ARIMA	71.2	91.6	72.0
	LSTM:QP	80.2	81.8	64.0
MEO	LSTM:ARIMA	45.3	88.1	85.8
	LSTM:QP	88.5	86.7	81.5

Authors' Contributions JL and XZ helped in conceptualization; JL helped in methodology; SH worked in software; JL, ZD, and DL contributed to validation; JL worked in formal analysis; JL helped in investigation; JL contributed to resources; SH helped in data curation; SH helped in writing—original draft preparation; JL helped in writing—review and editing; SH helped in visualization; XZ and ZD worked in supervision; JL worked in project administration; and XZ and DL helped in funding acquisition. All authors have read and agreed to the final version of the manuscript.

Data availability We thank the IGS Data Center of Wuhan University for providing open-source data for this research. The website is http://www.igs.gnsswhu.cn/index.php/Home/DataProduct/igs.html.

Declarations

Competing interest The authors declare no competing interests.

References

Cernigliaro A, Sesia I (2012) INRIM tool for satellite clock characterization: frequency drift estimation and removal. Mapan 27(1):41–48. https://doi.org/10.1007/s12647-012-0001-5

Chen JP, Zhang YZ, Zhou XH, Pei X, Wang JX, Wu B (2014) GNSS clock corrections densification at SHAO: from 5 min to 30 s. Sci China Phys Mech Astron 57(1):166–175. https://doi.org/10.1007/s11433-013-5181-7

Davis J, Bhattarai S, Ziebart M (2012) Development of a Kalman filter based GPS satellite clock time-offset prediction algorithm. In: 2012 IEEE European Frequency and Time Forum, pp. 152–156. https://doi.org/10.1109/EFTF.2012.6502355

Dong Shi, Wang Ping, Abbas Khushnood (2021) A survey on deep learning and its applications. Computer Science Review 40:100379. https://doi.org/10.1016/j.cosrev.2021.100379

Feng SL (2009) Study on the methods of data preprocessing and performance analysis for atomic clocks. PLA Information Engineering University, 20–21.

Guo JS (2013) Time scale steering in UTC (NIM). Beijing University of Technology, Beijing

Guo H (2006) Study on the analysis theories and algorithms of the time and frequency characterization for atomic clocks of navigation satellites. Information Engineering University.

Hackman C, Parker TE (1996) Noise analysis of unevenly spaced time series data. Metrologia 33(5):457

Heo YJ, Cho J, Heo MB (2010) Improving prediction accuracy of GPS satellite clocks with periodic variation behaviour. Measurement Science and Technology 21(7):073001. https://doi.org/10.1088/0957-0233/21/7/073001

Huang GW, Yang YX, Zhang Q (2011) Estimate and predict satellite clock error used adaptively robust sequential adjustment with classified adaptive factors based on opening windows. Acta Geodaetica Et Cartographica Sinica 40(1):15

Huang GW, Zhang Q, Xu GC (2014) Real-time clock offset prediction with an improved model. GPS Solut 18(1):95–104. https://doi.org/10.1007/s10291-013-0313-0

Huang GW, Cui BB, Zhang Q, Fu WJ, Li PL (2018) An improved predicted model for BDS ultra-rapid satellite clock offsets. Remote Sens 10(1):60. https://doi.org/10.3390/rs10010060

Huang BH, Ji ZX, Zhai RJ, Xiao CF, Yang F, Yang BH, Wang YP (2021) Clock bias prediction algorithm for navigation satellites based on a supervised learning long short-term memory neural network. GPS Solut 25(2):1–16. https://doi.org/10.1007/s10291-021-01115-0

Huang BH, Yang BH, Li MG, Guo ZK, Mao JY, Wang H (2022) An improved method for MAD gross error detection of clock error. Geomat Inf Sci Wuhan Univ 47(5):747–752. https://doi.org/10.13203/j.whugis20190430

Jonsson P, Eklundh L (2002) Seasonality extraction by function fitting to time-series of satellite sensor data. IEEE Trans Geosci Remote Sens 40(8):1824–1832. https://doi.org/10.1109/TGRS.2002.802519

Liang YJ, Ren C, Yang XF, Pang GF, Lan L (2015) Grey model based on first difference in the application of the satellite clock bias prediction. Acta Astrono Sin 56(3):264–277. https://doi.org/10.15940/j.cnki.0001-5245.2015.03.007

Liu JL, Cao YL, Hu XG, Tang CP (2020) Beidou wide-area augmentation system clock error correction and performance verification. Adv Space Res 65(10):2348–2359. https://doi.org/10.1016/j.asr.2020.02.010

Lu XF, Yang ZQ, Jia XL, Cui XQ (2008) Parameter optimization method of gray system theory for the satellite clock error predicating. Geomat Inform Sci Wuhan Univ 33(5):492–495

Riley W and Howe D (2008) Handbook of Frequency Stability Analysis, Special Publication (NIST SP), National Institute of Standards and Technology, Gaithersburg, MD, https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=50505

Steigenberger P, Hugentobler U, Loyer S, Perosanz F, Prange L, Dach R, Uhlemann M, Gerd C, Montenbruck O (2015) Galileo orbit and clock quality of the IGS Multi-GNSS experiment. Adv Space Res 55(1):269–281. https://doi.org/10.1016/j.asr.2014.06.030

Wang YP, Lu ZP, Qu YY, Li LY, Wang N (2017) Improving prediction performance of GPS satellite clock bias based on wavelet neural network. GPS Solut 21(2):523–534. https://doi.org/10.1007/s10291-016-0543-z

Xi C, Cai CL, Li SM, Li XH, Li ZB, Deng KQ (2014) Long-term clock bias prediction based on an ARMA model. Chin Astron Astrophy 38(3):342–354. https://doi.org/10.1016/j.chinastron.2014.07.010

Xu B, Wang Y, Yang XH (2013) Navigation satellite clock error prediction based on functional network. Neural Process Lett 38(2):305–320. https://doi.org/10.1007/s11063-012-9247-8

Yu Y, Si XS, Hu CH, Zhang JX (2019) A review of recurrent neural networks: LSTM cells and network architectures. Neural Comput 31(7):1235–1270. https://doi.org/10.1162/neco_a_01199

Zhao QL, Guo J, Li M, Qu LZ, Hu ZG, Shi C, Liu JN (2013) Initial results of precise orbit and clock determination for COMPASS navigation satellite system. J Geodesy 87(5):475–486. https://doi.org/10.1007/s00190-013-0622-7

Zhao L, Li N, Li H, Wang RL, Li MH (2021) BDS satellite clock prediction considering periodic variations. Remote Sens 13(20):4058. https://doi.org/10.3390/rs13204058

Zheng ZY, Chen YQ, Lu XS (2009) An improved grey model and its application research on the prediction of real-time GPS satellite clock errors. Chin Astron Astrophy 33(1):72–89. https://doi.org/10.1016/j.chinastron.2009.01.001

Zheng ZY, Dang YM, Lu XS, Xu WM (2010) Prediction model with periodic item and its application to the prediction of GPS satellite clock bias. Acta Astron Sin 51:95–102. https://doi.org/10.15940/j.cnki.0001-5245.2010.01.012

Zhou SS, Hu XG, Liu L, Guo R, Zhu LF, Chang ZQ, Tang CP, Gong XQ, Li R, Yu Y (2016) Applications of two-way satellite time and frequency transfer in the BeiDou navigation satellite system. Sci China Physi Mech Astron 59(10):1–9. https://doi.org/10.1007/s11433-016-0185-6

Zhu XW, Xiao H, Yong SW, Zhuang ZW (2008) The Kalman algorithm used for satellite clock offset prediction and its performance analysis. J Astronaut 29(3):966–970

Zhiqiang Dai is currently an assistant professor at the School of Electronics and Communication Engineering, Sun Yat-Sen University, China. He is mainly engaged in the theory, algorithm, and software development of GNSS/SBAS precise data processing, real-time PPP, and multi-sensor navigation data fusion.

Shaofeng He is currently pursuing a master's degree in communication engineering at the School of Electronics and Communication Engineering, Sun Yat-Sen University, China. His current research interests include global navigation satellite systems, brain-inspired navigation, and intelligent signal processing.

Du Li is an associate research fellow at the School of Electronic and Communication Engineering, Sun Yat-Sen University, China. His research interests include antenna miniaturization, GNSS high-precision antenna, GNSS anti-jamming antenna array, innovational theory antenna, and an acoustic-driven electrically small antenna.

Jiulong Liu is currently a postdoctoral fellow at the School of Electronic and Communication Engineering, Sun Yat-Sen University, China. He graduated from the University of Chinese Academy of Sciences with a Ph.D. in 2020. Currently, he is mainly engaged in research related to SBAS, GNSS orbit, clock error, and BeiDou integrated PNT system.

Xiangwei Zhu is currently a Professor with the School of Electronics and Communication Engineering, Sun Yat-Sen University, China. His current research interests include global navigation satellite systems, time synchronization, intelligent signal processing, and instrument design.