

Fusion-based Satellite Clock Bias Prediction Considering Characteristics and Fitted Residue

Jicang Lu, Chao Zhang, Yong Zheng and Ruopu Wang

(Zhengzhou Information Science and Technology Institute, China)

(E-mail: lujicang@sina.com)

As Satellite Clock Bias (SCB) prediction may be affected by various factors such as periodic items, sampling length, and stochastic items, a fusion-based prediction method is proposed by considering characteristics of SCB and fitted residue. On this basis, an instance algorithm is presented by fusing four typical prediction models. First, we use Empirical Mode Decomposition (EMD) to pre-process and decompose the SCB series into multiple components with various characteristics. Then, we analyse the fitting performance of each model for different components and prediction length, namely short-, mid- and long-term prediction, and select models with the best performance. Next, we analyse fitted residue of the reconstructed SCB, and select the model with the best fitting results. Finally, we fuse the multiple selected models for SCB prediction. The method is tested using Global Positioning System (GPS) precise clock products provided by the International Global Navigation Satellite System Service (IGS). Experimental results show that, compared with single prediction models and existing combination models, the proposed fusion-based prediction method improves accuracy and stability. In particular, the proposed method is more stable and has better performance for mid- and long-term prediction.

KEY WORDS

1. Satellite clock bias (SCB) prediction.
2. EMD (Empirical Mode Decomposition)
3. Residue.
4. Multiple model fusion.

Submitted: 13 April 2017. Accepted: 22 December 2017. First published online: 5 February 2018.

1. INTRODUCTION. As a major component of satellite navigation technology, it is important to analyse the characteristics and bias of satellites' on board atomic clocks. The clocks may experience relativity and gravitational redshift effects due to height and relative motion (Cacciapuoti et al., 2007). In addition, compared with ground atomic clocks, the environment of an on board atomic clock is more complex. The atomic clock is sensitive to various outside and intrinsic factors, such as temperature, humidity, collision, relative motion, etc. Therefore, atomic clocks could suffer large deviations without effective corrections (Tang et al., 2015). It is well known that the fundamental principle of

satellite navigation and positioning is the measurement of time. Therefore, high stability and accuracy of on board atomic clocks is a key issue for precise applications such as accurate navigation and positioning (Shi et al., 2015), long baseline navigation (Batista, 2015), precise point positioning (Li and Zhao, 2012), precise timing (Griggs et al., 2014) and so on.

Satellite clock monitoring and bias prediction for satellite navigation systems such as the Global Positioning System (GPS), Beidou, GLONASS and Galileo has been achieved by capturing real-time clock information from observing stations from wide areas. Analysis and prediction results are released after several days or weeks of processing. For example, the International Global Navigation Satellite System Service (IGS) provides clock bias and precise ephemeris products for GPS and GLONASS periodically. The International GNSS Monitoring and Assessment System (iGMAS) can provide precise orbit and clock bias products for GPS, Beidou, GLONASS and Galileo. However, these results and products are generally time-late and may not satisfy some real-time precise applications when satellite signals are lost. In addition, the prediction length is not long enough for some long-term precise applications. To date, a number of Satellite Clock Bias (SCB) analysis and prediction models or algorithms, such as the Quadratic Polynomial (QP) model (Huang et al., 2014; Wang et al., 2016), Spectrum Analysis (SA) (Heo et al., 2010; Senior et al., 2008), Grey system Model (GM) (Yuan et al., 2008; Zheng et al., 2008), Auto Regressive Integrated Moving Average (ARIMA) model (Xi et al., 2014; Stein and Evans, 1990), Kalman Model (Galleani and Tavella, 2010; Davis et al., 2012; Pratt et al., 2013), Least Square Support Vector Machine (LS-SVM) model (Liu et al., 2013), Wavelet Neural Network (WNN) model (Ai et al., 2016; Wang et al., 2017), etc. In addition, combination prediction methods were also proposed by combining typical multiple models (Lei et al., 2014; Wang et al., 2011; Xu et al., 2016). However, existing models or algorithms generally do not take fully into account the influences of factors such as periodic items, sampling length, stochastic items and so on, which result in low precision and usability. SCB is caused by multiple factors, which result in the various characteristics. Therefore, in order to research and design more precise and reasonable prediction methods or models for particular applications, the influences of periodic and stochastic effects as well as sampling length should be considered.

For more precise and stable prediction of SCB, we propose a fusion-based prediction method by considering various characteristics and fitted residue. First, an Empirical Mode Decomposition (EMD) algorithm is used to pre-process and decompose the observed SCB series into multiple components with various periodic characteristics. Then, existing typical prediction models are used to fit each component with different characteristics and sampling length. The fitting residue is analysed and the model with best fitting results is selected to reconstruct the SCB. Next, the synthetic residue between the original and reconstructed data (referred to as primary fitted error) is computed and fitted by multiple prediction models, and the model with best fitting results is selected. Finally, the selected models are fused to predict future SCB in the short-, mid- and long-term. The proposed fusion-based prediction method is tested and analysed using the GPS SCB from the IGS.

2. RELATED WORKS. In this section, some typical SCB prediction models (QP, SA, GM and ARIMA) will be briefly explained, using their mathematical expressions. Their performance will also be briefly summarised.

2.1. Typical prediction model.

2.1.1. QP model. The multinomial model is one of the most popular models in SCB analysis. Under this scheme, QP is a more typical model, which can be expressed as

$$x_i = \alpha_0 + \alpha_1 \Delta t_i + \alpha_2 \Delta t_i^2 + \int_{t_0}^{t_i} f(t)dt \tag{1}$$

where x_i is the observed SCB at time t_i , for the SCB dealt with differential processing, it could be the first or higher order difference of SCB. α_0, α_1 and α_2 are the fitting parameters to be computed. $\Delta t_i = t_i - t_0$ is the difference between observed time t_i and reference time t_0 . $\int_{t_0}^{t_i} f(t)dt$ is the stochastic error with a normal distribution, which is denoted as e_i . Generally, the stochastic term is the clock noise, which is typically composed of five noise processes, that is White Phase Modulation (WPM), Flicker Phase Modulation (FPM), White Frequency Modulation (WFM), Flicker Frequency Modulation (FFM) and Random Walk Frequency Modulation (RWFM). Thus, the noise is non-stationary.

Denote

$$\alpha = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 1 & \Delta t_1 & \Delta t_1^2 \\ 1 & \Delta t_2 & \Delta t_2^2 \\ \dots & \dots & \dots \\ 1 & \Delta t_n & \Delta t_n^2 \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}, \quad \mathbf{e} = \begin{bmatrix} e_1 \\ e_2 \\ \dots \\ e_n \end{bmatrix},$$

then

$$\mathbf{X} = \mathbf{A}\alpha + \mathbf{e} \tag{2}$$

where $\hat{\alpha}$ can be estimated by the least squares method as

$$\hat{\alpha} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{X}. \tag{3}$$

With the estimated parameters, we can use

$$\hat{x}_j = \hat{\alpha}_0 + \hat{\alpha}_1(t_j - t_0) + \hat{\alpha}_2(t_j - t_0)^2 \tag{4}$$

to obtain the fitted value of the observed data ($j \leq n$) and predict the future data ($j > n$).

2.1.2. SA model. For SCB with periodic characteristics, Spectrum Analysis (SA) can describe it better. The model can be described as

$$x_i = \alpha_0 + \beta_0 \Delta t_i + \sum_{k=1}^m A_k \sin(2\pi f_k \Delta t_i + \phi_k) + e_i \tag{5}$$

where α_0 and β_0 are constant and coefficient terms during long-term variations, respectively. m is the number of the main periodic characteristics. f_k is the frequency corresponding to the periodic characteristics and A_k and ϕ_k are amplitude and phase, respectively. e_i is the residue of x_i . m and f_k can be determined by fast Fourier transform. Denote

$\alpha_k = A_k \cos \phi_k, \beta_k = A_k \sin \phi_k$, we then have

$$x_i = \alpha_0 + \beta_0 \Delta t_i + \sum_{k=1}^m (\alpha_k \sin(2\pi f_k \Delta t_i) + \beta_k \cos(2\pi f_k \Delta t_i)) + e_i \tag{6}$$

Denote

$$\mathbf{H} = \begin{bmatrix} 1 & \Delta t_1 & \sin(2\pi f_1 \Delta t_1) & \cos(2\pi f_1 \Delta t_1) & \dots & \dots & \sin(2\pi f_m \Delta t_1) & \cos(2\pi f_m \Delta t_1) \\ 1 & \Delta t_2 & \sin(2\pi f_1 \Delta t_2) & \cos(2\pi f_1 \Delta t_2) & \dots & \dots & \sin(2\pi f_m \Delta t_2) & \cos(2\pi f_m \Delta t_2) \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & \Delta t_n & \sin(2\pi f_1 \Delta t_n) & \cos(2\pi f_1 \Delta t_n) & \dots & \dots & \sin(2\pi f_m \Delta t_n) & \cos(2\pi f_m \Delta t_n) \end{bmatrix}^T,$$

$\mathbf{a} = [\alpha_0 \quad \beta_0 \quad \alpha_1 \quad \beta_1 \quad \dots \quad \alpha_k \quad \beta_k]^T, \mathbf{e} = [e_1 \quad e_2 \quad \dots \quad e_n]^T$, then, there is

$$\mathbf{X} = \mathbf{H}\mathbf{a} + \mathbf{e} \tag{7}$$

where $\hat{\mathbf{a}}$ can be estimated by the least squares method as

$$\hat{\mathbf{a}} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{X}. \tag{8}$$

With the estimated parameters, we can use

$$\hat{x}_j = \hat{\alpha}_0 + \hat{\beta}_0 \Delta t_j + \sum_{k=1}^m (\hat{\alpha}_k \sin(2\pi f_k \Delta t_j) + \hat{\beta}_k \cos(2\pi f_k \Delta t_j)) \tag{9}$$

to obtain the fitted value of the observed data ($j \leq n$) and predict the future data ($j > n$).

2.1.3. *Grey system Model (GM).* The Grey system Model (GM) is a prediction system with incomplete definite information, that is, part of the information is known and part unknown, and some relationships between various factors are generally uncertain. The principle is: first, process the original data series with accumulated additive or subtractive factors to generate a new data series. This process decreases the randomisation of the original data and results in some regular characteristics. Then, establish a model based on differential equations to predict the future data. On the basis of this principle, multiple types of GM models have been presented by setting various parameters. Among these models, GM(1, 1), which is constructed by a first order differential equation with only one variable, is the most popular.

We denote the observed data series as $\{x_i^{(0)}, i = 1, 2, \dots, n\}$ and the corresponding times are $\{t_i^{(0)}, i = 1, 2, \dots, n\}$. Next we calculate the cumulative sum of the observed data series and the resulting series is $\{x_i^{(1)}, i = 1, 2, \dots, n\}$. Then, for first order GM $\frac{dx^{(1)}}{dt} + ax^{(1)} = u$, the differential equation after derivative discretisation can be expressed as follows

$$\mathbf{X} = \mathbf{A}\mathbf{U} \tag{10}$$

where $\mathbf{X} = \begin{bmatrix} x_2^{(0)} \\ x_3^{(0)} \\ \dots \\ x_n^{(0)} \end{bmatrix}, \mathbf{A} = \begin{bmatrix} -[x_1^{(1)} + x_2^{(1)}]/2 & 1 \\ -[x_2^{(1)} + x_3^{(1)}]/2 & 1 \\ \dots & \dots \\ -[x_{n-1}^{(1)} + x_n^{(1)}]/2 & 1 \end{bmatrix}, \mathbf{U} = \begin{bmatrix} a \\ u \end{bmatrix}.$

Then, the value \hat{U} can be estimated by the least squares method as

$$\hat{U} = (\mathbf{A}^T \mathbf{A})^{-1} (\mathbf{A}^T \mathbf{X}) = \begin{bmatrix} \hat{a} \\ \hat{u} \end{bmatrix}. \tag{11}$$

With the estimated parameters, we can use

$$\hat{x}_j^{(0)} = \left(1 - e^{\hat{a}}\right) \left[x_1^{(0)} - \frac{\hat{u}}{\hat{a}}\right] e^{-\hat{a}(j-1)} \tag{12}$$

to obtain the fitted value of the observed data ($j \leq n$) and predict the future data ($j > n$).

2.1.4. *Auto Regressive Integrated Moving Average (ARIMA) model.* ARIMA is derived from the Auto Regression Moving Average (ARMA) model by pre-processing the original data series with d -th order difference, and can be denoted as ARIMA(p, d, q). When $d = 0$, ARIMA(p, d, q) is ARMA(p, q). ARMA is a popular model used to fit stable sequences and includes Auto Regression (AR), Moving Average (MA) and ARMA. For most non-stable sequences such as SCB series, it always becomes stable after dealing with first or higher order differences. Therefore, ARIMA can be used to fit and predict SCB.

Denote the SCB series after d -th order difference as $\{x_i^{(-d)}, i = 1, 2, \dots, n\}$, then, ARIMA(p, d, q) can be expressed as

$$x_i^{(-d)} = \phi_1 x_{i-1}^{(-d)} + \dots + \phi_p x_{i-p}^{(-d)} + \varepsilon_i - \theta_1 \varepsilon_{i-1} - \dots - \theta_q \varepsilon_{i-q} \tag{13}$$

where ϕ_1, \dots, ϕ_p are parameters of auto regression, $\theta_1, \dots, \theta_q$ are parameters of a moving average and p and q are orders of auto regression and a moving average, respectively. $\varepsilon_i, \varepsilon_{i-1}, \dots, \varepsilon_{i-q}$ are white noise with normal distribution $(0, \sigma_\varepsilon^2)$. When $p = 0$ or $q = 0$, the model is MA(q) or AR(p), respectively.

The construction and application of ARIMA includes two parts: estimation of parameters and determination of model orders. The former part is a complex procedure. The popular methods used to estimate parameters are moment estimation, maximum likelihood estimation, or least squares estimation. Moment estimation is a rough method with lower computation complexity, but the accuracy is lower than the other two methods. The least squares method is a special case of maximum likelihood method with some specific conditions. Performance of the least squares method is the best, but the procedure is more complex. For a data series, there may be multiple groups of p and q satisfying the requirement. The model should be optimised by determining the best p and q . At present, Akaike Information Criterion (AIC) norm and Bayesian Information Criterion (BIC) norm are two widely used methods for parameter optimisation.

The performance of the above four typical models have been researched and analysed (Wang et al., 2016; Senior et al., 2008; Yuan et al., 2008; Xi et al., 2014; Wang et al., 2015), and can be described as follows.

- (1) For short-term prediction, all of the four models perform well. However, SA considers the periodic variation, while the exponential coefficients of GM(1, 1) are related to the number of elements in the observed SCB series. Sufficient sampling data should be provided, otherwise, prediction precision cannot be guaranteed.
- (2) For mid- and long-term prediction, SA and GM(1, 1) will perform well with sufficient sampling data. QP and ARIMA($p, 1, q$) are not suitable for mid- and long-term prediction because of error accumulation.

- (3) For periodic characteristics analysis, SA is more suitable. However, SA performs better only when there is sufficient sampling data.
- (4) As to satellite clock types, ARMA($p, 1, q$) is more easily affected by clock type and prediction condition. The results from Rubidium atomic clocks will be more accurate than those of Caesium atomic clocks (Heo et al., 2010; Wang et al., 2015).

2.2. *Existing combination prediction method.* For a particular prediction model, the performance can be influenced by various factors, such as sampling length, periodic characteristics, prediction length and type of atomic clock. For example, some models are appropriate for short-term prediction, while some are appropriate for mid- and long-term prediction. Some models are appropriate for periodic items, while some are appropriate for stochastic ones. Therefore, for particular applications, a combination prediction model that is superior to a single model can be obtained by combining multiple models, making the best use of the advantages of each. Existing combination prediction models can be divided into two categories: one is to combine the results of each model by weighted linear or nonlinear combination, such as the algorithms proposed by Wang et al. (2011) and Wang (2010). The other category is to select models for different components, and then superimpose them as the final prediction results. For example, Xu et al. (2016) combined QP and AR. Xu et al. (2013a; 2013b) combined QP and a functional network. Lei et al. (2014) combined QP and a least squares support vector machine. The principle of these two categories are illustrated in Figure 1, the left panel denotes the former category, and the right panel denotes the latter.

It can be seen from Figure 1 and existing research (Xu et al., 2016; 2013a; 2013b; Lei et al., 2014) that the combination methods generally did not take fully into account the characteristics of the single models under different conditions, as well as various characteristics

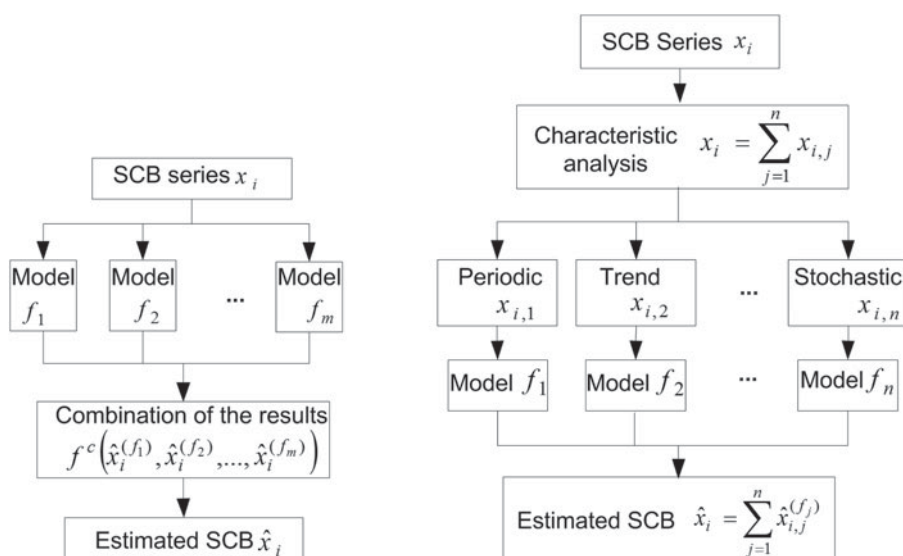


Figure 1. Principle of existing combination prediction model for SCB.

of observed SCB. Therefore, it has been difficult to obtain a combination prediction model with a superior and more stable performance under a variety of conditions.

3. FUSION-BASED SCB PREDICTION METHOD. First, this section analyses the characteristics of SCB. Then, an appropriate pre-process method will be presented to decompose an SCB series into multiple components with various characteristics. Finally, a fusion-based prediction method will be presented.

3.1. *Characteristic analysis of the SCB.* Due to influences of intrinsic physical characteristics and outside environments, the SCB series will show non-stable characteristics (Heo et al., 2010). In addition, the on board atomic clocks are generally caesium or rubidium atomic clocks, and the frequency drift over an extended period of time is usually nonlinear (Galleani et al., 2003; Zucca and Tavella, 2015). The factors above will influence stability and accuracy of satellite clocks, and lead to clock offset. Generally, the offset consists of systematic and stochastic offsets.

The systematic offset is caused by initial phase offset, frequency offset and frequency drift after a long period of working. Initial phase offset is a constant. The intrinsic frequencies are different for different atomic clocks. The difference will increase linearly as time elapses, which results in a frequency offset. Frequency drift is the intrinsic characteristic of all standard atomic frequencies and will increase non-linearly as time elapses. Frequency offset and drift display as trend characteristics in SCB, which leads to nonlinear characteristics after a long period. In addition, on board atomic clocks are influenced by multiple factors such as temperature, humidity, vibration, radiation and gravitation. Accompanied with the rotation and revolution of the Earth, and periodic variation of the orbit, these factors lead to the periodic variation characteristics of atomic clocks. For bias with this characteristic, the periodic variation function is generally used for better fitting.

The stochastic offset is caused by noise inside the electronic equipment. Although the physical mechanism of the noise is not presently clear, it can be described as multiple independent energy spectrum noises. There are five modulations of noise, namely random walk frequency modulation, flicker frequency modulation, white frequency modulation, flicker phase modulation and white phase modulation. In addition, GPS rubidium atomic clocks are also influenced by very low frequency noise flicker walk frequency modulation and random run frequency modulation. The superposition of these noises gives a stochastic model of SCB. Research (Guo, 2006) shows that the energy spectrum of most stochastic noise emanates at a low Fourier frequency. In other words, the noise will show non-stationary characteristics after long-term working, which leads to non-stationary SCB.

In summary, SCB is generally the superposition of multiple types of data such as periodic data, trend data, and stochastic data. At the same time, SCB shows non-stationary and nonlinear characteristics.

3.2. *EMD and the usability for SCB analysis.* EMD was proposed in 1998 by Huang et al. (1998) of the National Aeronautics and Space Administration (NASA). EMD aims to provide a new adaptive processing method for time-frequency signals. The principle is firstly to decompose a time series signal as finite number of Intrinsic Mode Functions (IMFs) with different time-scales, followed by processing each IMF with a Hilbert transform to obtain the time-frequency spectrogram. The signal can be reconstructed by superimposing the IMFs.

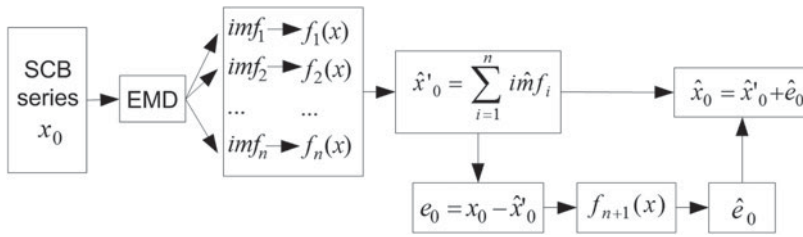


Figure 2. Procedure of fusion and EMD-based SCB prediction.

The IMFs must satisfy the following requirements:

- (1) During the whole time range, the number of the local extreme points must be equal to that of the zero-crossing point for the function, or differ by one only.
- (2) At any time, an average of the local maximum value envelope (upper envelope solid) and the local minimum value envelope (lower envelope solid) must be zero.

EMD is excellent in processing nonlinear and non-stationary signals and for satellite clocks, can reflect the characteristics of SCB more exactly.

First, the SCB series is pre-processed, and decomposed by EMD into multiple IMFs with different characteristics. Then, each IMF is analysed and the related future IMFs are predicted. Finally, the residue is analysed and the predicted SCB are obtained by reconstruction. For the IMFs, when the periodicity is not obvious, the data may vary like a stochastic sequence. Then, IMFs with this property will be added as one component for processing. For example, the minimum satellite period is about 1.5 hours, and there will be six clock biases when the sampling interval is 15 minutes. If an IMF fluctuates frequently within six consecutive datasets, then, it may be more appropriate to consider it as a stochastic sequence. Note that this paper aims at selecting an appropriate prediction model for SCB with various characteristics. Therefore, when there are two or more IMFs performing stochastic characteristics, they will be added as one component for further analysis and prediction. Other IMFs stay unchanged.

The procedure of SCB decomposition and reconstruction based on EMD is illustrated in Figure 2. Firstly, select n prediction models $\{f_j(x), j \in [1, n]\}$ for the fitting and prediction of n IMFs $\{imf_i, i \in [1, n]\}$. Then, select an appropriate model to fit and predict the residue e_0 between reconstructed signal \hat{x}'_0 and the original signal. The estimated original signal can be obtained by adding the results \hat{e}_0 to \hat{x}'_0 .

3.3. Construction of fusion-based prediction model and SCB prediction. SCB is a superposition of data series with various characteristics. At the same time, performance of different prediction models varies for data with different characteristics, and the accuracy and stability of a single model are usually poor. Therefore, for SCB prediction, an appropriate data analysis algorithm such as EMD should first be used to decompose the observed data into multiple components (such as IMFs). Then, the best prediction model should be selected to fit and estimate each component. Finally, the most accurate and stable prediction results can be obtained by fusing each component predicted by the selected models. According to this principle, a fusion-based SCB prediction method is proposed by considering SCB characteristics and fitted residue. The proposed method includes two stages: the

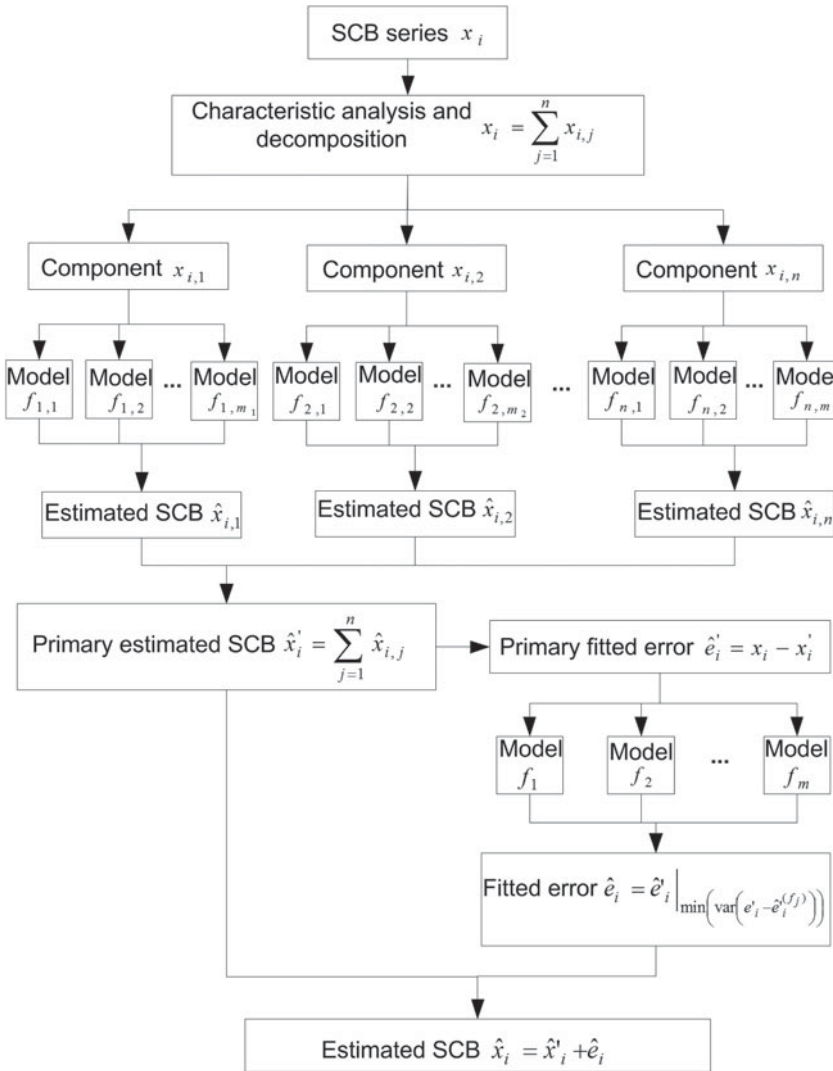


Figure 3. Framework of the proposed fusion-based SCB prediction method.

first is construction of the fusion-based prediction model, and the second is prediction of future SCB. The procedure of model construction is illustrated in Figure 3.

The framework in Figure 3 is described as follows:

- (1) Determine the length of observed SCB according to future SCB to be predicted. If the length of SCB to be predicted is p , then the previous observed SCB with length $1, 2, \dots, p$ will be used for analysis. The i -th data is denoted as x_i .
- (2) Characteristic analysis and data decomposition. For the previous observed SCB with length $1, 2, \dots, p$, pre-process it with first order difference and decompose it by methods such as EMD. For observed data of each length, a set $\{x_{i,j}, j \in [1, n]\}$ with n components such as IMFs can be obtained, where $x_i = \sum_{j=1}^n x_{i,j}$.

- (3) *Data component fitting and prediction.* Analyse the characteristics of each component, such as periodic, trend and stochastic characteristics. Select the prediction model set for each component, and denote them as $\{f_{i,j}, j \in [1, m_1]\}, \{f_{i,j}, j \in [1, m_2]\}, \dots, \{f_{i,j}, j \in [1, m_n]\}$, respectively. Then, fit the previously observed SCB with length 1, 2, \dots , p and predict the future corresponding components.
- (4) *Prediction model selection and primary fitted residue computation.* According to the above fitting results, take the prediction model with the minimum fitted variance as the selected model for the specific length of the observed data. The fitted components are denoted as $\hat{x}_{i,n} = \hat{x}_{i,n}^{(f_{nj})} \Big|_{\min(\text{var}(x_{i,n} - \hat{x}_{i,n}^{(f_{nj})}))}$. The primary fitted results of the observed data are obtained by superimposing the fitted results of each component, that is $\hat{x}'_i = \sum_{j=1}^n \hat{x}_{i,j}$. The primary fitted residue is calculated with $\hat{e}'_i = x_i - \hat{x}'_i$.
- (5) *Residue prediction and fusion model determination.* For the primary fitted residue of observed data of each length, analyse the characteristics and select the prediction model with minimum fitted variance. The estimation result of the primary fitted residue is $\hat{e}_i = \hat{e}'_i \Big|_{\min(\text{var}(e'_i - \hat{e}'_i))}$. Finally, take the observed data of the length with the minimum synthetic fitted variance as the reference for SCB prediction.

After the length of observed data is determined, the fusion-based prediction model can be determined. This consists of the selected prediction models for each component and the primary fitted residue. Then, the expected future SCB is calculated by $\hat{x}_i = \hat{x}'_i + \hat{e}_i$.

4. APPLICATION EXAMPLE AND EXPERIMENTAL RESULTS ANALYSIS. This section will present an applied example of the proposed prediction method. The method will then be tested on an actual SCB series to analyse its performance.

4.1. *Fusion-based prediction model application example.* According to the framework illustrated in the last section, four typical models (QP, SA, GM(1,1) and ARIMA($p,1,q$)) will be taken as examples to test the proposed fusion-based prediction method (referred to as MergM). Two existing combination prediction methods mentioned in Section 2.2 will be compared with the proposed method. One is the combination of QP and AR proposed by Xu et al. (2016) (referred to as LS+AR), the other one is the weight-based combination model proposed by Wang (2010) (referred to as ComT).

The actual SCB series is from GPS "final" precise SCB data published by IGS, which were sampled at 15 minute intervals from 2 November 2014 to 1 March 2015. There are a total of 120 days from the first day of GPS week 1817 to the first day of GPS week 1834 without SCB hops or interruption. The first half of the data is used for construction and parameters determination of the fusion model, and the other half is used for prediction and performance analysis. For coverage of various types of satellites and on board atomic clocks, the satellites be analysed are BLOCK IIA PRN10, BLOCK IIR PRN02, BLOCK IIR-M PRN05 and BLOCK IIF PRN27. At the same time, the corresponding clock types of PRN02, PRN05, PRN10 and PRN27 are IIR Rb (built in 2004), IIR-M Rb (built in 2009), IIA Cs (built in 1996) and IIF Cs (built in 2013), respectively.

The Root Mean Square (RMS) and maximum absolute difference (Range) of prediction error are taken as criteria to evaluate the performance. RMS is calculated by $\sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2}$, where n is the number of elements in the predicted SCB series, e_i is the prediction error.

Range is calculated by $e_{\max} - e_{\min}$, where e_{\max} and e_{\min} are the maximum and minimum prediction errors, respectively.

4.2. *Prediction results and comparison analysis.* We use four typical models (QP, SA, GM(1,1) and ARMA($p,1,q$)), two existing combination models (LS+AR and ComT) and the proposed fusion-based prediction method MergM to predict the future SCB: one day (1d, short-term), three days (3d, short-term), 15 days (15d, mid-term), 30 days (30d, mid-term) and 60 days (60d, long-term), respectively. For existing models, the length of used observed data is equal to the prediction length, that is, we use the previous 1 d, 3 d, 15 d, 30 d and 60 d observed data to predict the future 1 d, 3 d, 15 d, 30 d and 60 d SCB, respectively. RMS and Range of the results are shown in Table 1, where the results in italic and bold are the best two. Taking 3 d, 30 d and 60 d predictions of PRN05 and PRN10 as examples, the prediction errors of each method are plotted in Figure 4 for detailed comparison.

The comparison results in Table 1 show that, compared with existing models, the proposed fusion-based prediction model possess better stability and accuracy. The results of the proposed method are the best under almost all conditions. In addition, most existing methods will accumulate errors over time, which is directly reflected in Figure 4. In comparison, the prediction results of the proposed method are more accurate and stable for all short-term, mid-term and long-term predictions. For example, the RMS for 30 d and 60 d predictions can be maintained under 36 ns and 90 ns, respectively. The results show that the proposed method possesses a better synthetic performance.

In addition, the prediction results in Table 1 for IIR Rb of PRN02 perform the best, and that for IIA Cs of PRN10 perform the worst. When the prediction length is no more than 15 days, the prediction results for rubidium are better than caesium. When prediction lengths are 30 d and 60 d, the prediction results for the modern atomic clock IIF caesium are better than for IIR-M rubidium. The results and analysis show that the prediction results of the proposed method perform better for rubidium and modern caesium atomic clocks, but perform worse for caesium atomic clocks which have been working for a comparatively long period. The reason may be that there is a larger level of uncertainty about the performance of aging atomic clocks, which indicate that the atomic clock should be replaced or upgraded. The facts are also consistent with these inferences; the caesium atomic clock IIA of PRN10 terminated its service on 16 July 2015, and was replaced by a modern caesium atomic clock in a IIF satellite on 31 October 2015.

4.3. *Determination of parameters during prediction.* The parameters involved in the fusion-based prediction method were recorded and analysed. The parameters include the quantity of components obtained by EMD, the selected models for each component, the primary fitted error and the length of observed data needed for optimum prediction. The time consumed from data decomposition to prediction was also recorded, as well as that by the combination prediction methods LS+AR and ComT. The experiment was conducted on an ordinary laptop (Windows 7 64bit, Intel(R) Core(TM) i5-4210U @1.70 GHz 2.40 GHz, RAM: 4 GB). The results are listed in Table 2, where f_i is the prediction model determined for the i -th component.

It can be seen from the results in Table 2 that, for the prediction of various lengths of future SCB, the length of observed data and quantity of components obtained by EMD are different for different satellites. This indicates that there are differences between characteristics of SCB generated by different types of on board atomic clocks. For example, for the prediction of PRN27 1 d, 3 d, 15 d, 30 d and 60 d into the future, all the lengths of

Table 1. Clock bias prediction results comparison (ns).

Prediction length	Satellite number	Criterion	QP	SA	GM	ARIMA	LS+AR	ComT	MergM	
1d	PRN02	RMS	5.07	0.16	1.46	0.99	0.31	1.29	0.32	
		Range	9.34	0.85	2.77	2.16	0.91	2.45	1.15	
	PRN05	RMS	12.05	2.28	1.98	1.63	1.99	1.11	0.56	
		Range	23.20	3.92	3.44	2.15	3.35	3.18	1.99	
	PRN10	RMS	3.65	8.89	14.98	8.66	9.03	8.96	2.35	
		Range	11.90	19.92	31.75	19.83	20.24	20.83	9.10	
	PRN27	RMS	2.48	0.25	0.66	1.94	0.40	0.36	0.20	
		Range	4.71	0.82	1.36	2.87	0.92	0.91	0.81	
	3d	PRN02	RMS	17.06	0.83	0.40	7.43	1.05	6.07	0.42
			Range	31.65	2.31	2.00	15.38	3.71	11.45	1.78
PRN05		RMS	26.63	6.34	1.97	3.22	6.08	3.26	0.49	
		Range	51.09	11.38	3.83	4.28	11.35	7.89	2.01	
PRN10		RMS	25.22	24.68	19.52	73.39	23.25	23.10	7.93	
		Range	64.41	39.96	34.90	138.27	38.67	39.51	19.17	
PRN27		RMS	0.43	1.11	0.66	3.50	0.43	1.06	1.09	
		Range	1.66	3.00	1.55	4.65	1.48	2.08	2.77	
15d		PRN02	RMS	33.50	5.09	4.66	14.49	14.74	6.70	0.81
			Range	65.95	12.06	11.12	29.72	34.32	14.08	3.52
	PRN05	RMS	29.32	13.50	13.13	11.66	63.03	4.25	4.36	
		Range	56.41	25.63	25.46	21.87	137.84	8.68	9.36	
	PRN10	RMS	18.28	86.49	76.54	325.53	218.42	117.07	36.83	
		Range	70.98	149.75	130.21	674.03	452.82	226.42	118.48	
	PRN27	RMS	1.69	10.40	10.35	20.09	20.04	11.48	11.32	
		Range	5.27	21.16	20.98	42.10	43.04	23.91	22.82	
	30d	PRN02	RMS	86.51	29.48	29.33	31.30	4.86	43.11	2.94
			Range	162.94	53.73	54.17	66.56	22.91	82.63	14.24
PRN05		RMS	33.30	40.34	50.25	77.84	48.08	53.59	9.26	
		Range	83.96	97.22	121.31	178.34	114.73	126.92	33.28	
PRN10		RMS	407.72	188.39	188.61	252.03	67.72	258.27	35.34	
		Range	717.48	301.08	300.86	439.59	126.38	438.14	133.66	
PRN27		RMS	40.55	30.36	30.25	17.30	23.35	28.17	25.28	
		Range	70.51	51.49	51.35	26.19	35.63	46.66	76.98	
60d		PRN02	RMS	273.30	11.04	10.44	79.92	170.02	89.82	10.35
			Range	543.44	43.39	40.42	194.24	365.40	198.66	39.39
	PRN05	RMS	250.12	365.48	473.34	244.19	448.11	348.71	89.54	
		Range	788.72	996.04	1247.05	714.82	1187.40	960.04	358.95	
	PRN10	RMS	691.57	502.78	526.90	218.62	374.07	481.79	46.37	
		Range	1425.79	1067.69	1123.11	434.34	771.39	1008.16	108.92	
	PRN27	RMS	374.13	48.82	42.64	22.97	164.40	49.67	26.49	
		Range	769.97	147.75	134.28	68.75	334.98	145.52	54.54	

observed data needed for optimum prediction are a constant 7 d, and the quantities of EMD components are a constant 7 d. This indicates that the characteristics of SCB of PRN27 can be described adequately with 7 d observed data. In comparison, for PRN05, the length of observed data needed for predicting future 1 d, 3 d, 15 d and 30 d are 8 d, while for 60 d it is 59 d. These results could provide a better reference to determine the length of observed data required when designing a prediction method.

It can also be seen from the selected models that, although the quantity of components for various satellites are different, only SA and ARIMA are selected for these components.

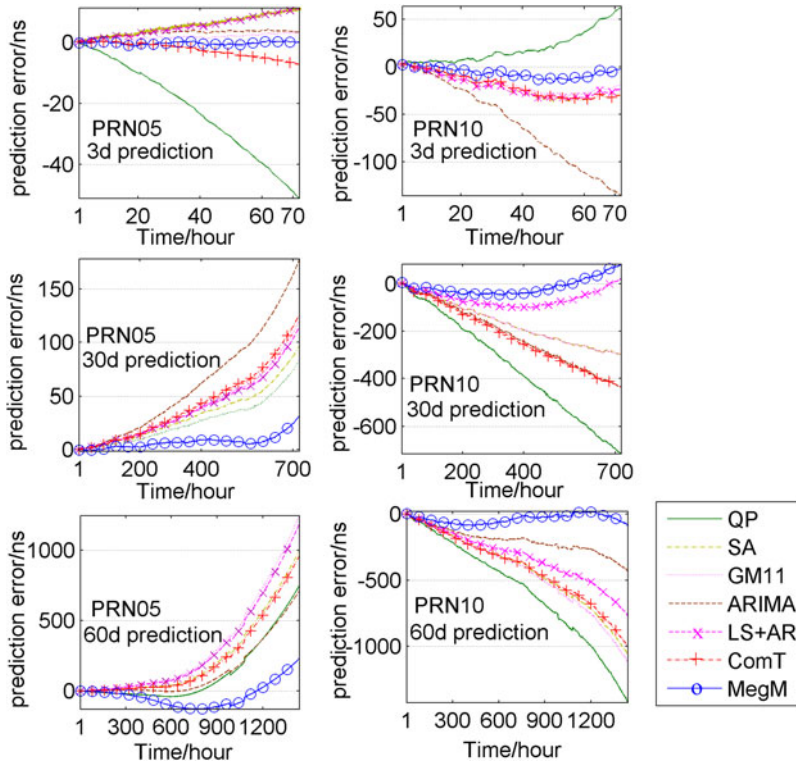


Figure 4. Comparison of SCB prediction errors for PRN05 in the left panels and PRN10 in the right panels.

What is more, SA is mostly selected for the last component. This is because the last component is the main one, which generally shows periodic characteristics and is less affected by stochastic factors. Therefore, SA is more appropriate than ARIMA, which coincides with the results pointed out in Section 2.1. For the primary fitted error, GM11 is generally selected for short-term and mid-term prediction, while SA is generally selected for long-term prediction. These results show that it is important to decompose and predict SCB according to the periodic or stochastic characteristics.

In addition, for future prediction of SCB no longer than 30 d, apart from the time consumed for the future 30 d prediction of PRN10, which is 66.95 seconds, all the other calculations take less than 60 seconds (1 minute). For the future 60 d prediction for all the satellites, the times consumed are all within 90 seconds (1.5 minutes). Compared with the two existing combination methods, the time consumed by the proposed method is longer. The main reason is that most of the IMFs are determined by using the ARIMA model, which is more complex and time consuming. This indicates that the complexity is influenced by associated single models. Therefore, one way to improve the proposed method and reduce the complexity is to select more time saving models with higher performance. For practical SCB prediction applications using observed data sampled at 15 minutes' interval, the proposed fusion-based prediction method is a more appropriate selection for higher precision and stability.

Table 2. Parameters used by the proposed method and the consumed time.

Prediction length	Satellite number	Length of observed data (day)	Quantity of component	Selected models for each component	Selected model for primary fitted error	Consumed time (s)		
						LS+AR	ComT	MergM
1d	PRN02	4	7	$f_i = \text{ARIMA}, i \in [1,5];$ $f_i = \text{SA}, i \in [6,7]$	GM11	1.15	2.12	20.37
	PRN05	8	8	$f_i = \text{ARIMA}, i \in [1,7];$ $f_i = \text{SA}, i = 8$	GM11	1.20	2.32	27.46
	PRN10	7	8	$f_i = \text{ARIMA}, i \in [1,7];$ $f_i = \text{SA}, i = 8$	GM11	1.19	2.74	27.23
	PRN27	7	7	$f_i = \text{ARIMA}, i \in [2,6];$ $f_i = \text{SA}, i \in \{1,7\}$	GM11	1.28	1.74	26.48
3d	PRN02	9	8	$f_i = \text{ARIMA}, i \in [1,8]$	GM11	1.24	2.47	36.47
	PRN05	8	8	$f_i = \text{ARIMA}, i \in [1,7];$ $f_i = \text{SA}, i = 8$	GM11	1.23	2.53	28.74
	PRN10	3	7	$f_i = \text{ARIMA}, i \in [1,6];$ $f_i = \text{SA}, i = 7$	GM11	1.23	3.09	18.37
	PRN27	7	7	$f_i = \text{ARIMA}, i \in [2,6];$ $f_i = \text{SA}, i \in \{1,7\}$	GM11	1.30	1.68	28.38
15d	PRN02	12	9	$f_i = \text{ARIMA}, i \in [1,8];$ $f_i = \text{SA}, i = 9$	GM11	1.46	4.02	44.54
	PRN05	8	8	$f_i = \text{ARIMA}, i \in [1,7];$ $f_i = \text{SA}, i = 8$	GM11	1.46	6.05	33.63
	PRN10	3	7	$f_i = \text{ARIMA}, i \in [1,6];$ $f_i = \text{SA}, i = 7$	SA	1.47	4.40	23.13
	PRN27	7	7	$f_i = \text{ARIMA}, i \in [1,6];$ $f_i = \text{SA}, i = 7$	SA	1.48	5.25	27.55
30d	PRN02	18	9	$f_i = \text{ARIMA}, i \in [1,9]$	SA	2.05	4.55	54.48
	PRN05	8	8	$f_i = \text{ARIMA}, i \in [1,7];$ $f_i = \text{SA}, i = 8$	GM11	2.01	5.05	33.79
	PRN10	28	11	$f_i = \text{ARIMA}, i \in [1,11]$	SA	2.02	6.07	66.95
	PRN27	7	7	$f_i = \text{ARIMA}, i \in [2,6];$ $f_i = \text{SA}, i \in \{1,7\}$	GM11	2.07	6.19	30.24
60d	PRN02	39	10	$f_i = \text{ARIMA}, i \in [1,10]$	SA	3.85	8.04	89.56
	PRN05	59	9	$f_i = \text{ARIMA}, i \in [1,9]$	SA	3.88	8.81	86.08
	PRN10	29	11	$f_i = \text{ARIMA}, i \in [1,11]$	SA	3.80	10.22	84.79
	PRN27	7	7	$f_i = \text{ARIMA}, i \in [2,6];$ $f_i = \text{SA}, i \in \{1,7\}$	GM11	3.58	10.35	46.58

In summary, the experiments covered almost all types of GPS satellites and current operational on board atomic clocks, such as BLOCK IIR Rb, BLOCK IIR-M Rb, BLOCK IIA Cs and BLOCK IIF Cs. The prediction results of the proposed method perform well for all of them, and are better for modern types of satellites and atomic clocks. The above experimental results and analysis indicate that the proposed method has good practical usability and outstanding performance.

5. CONCLUSIONS. For SCB prediction, a fusion-based framework is proposed through the analysis of data characteristics and fitted residue. The proposed framework also considers issues such as length and pre-processing of the observed data, and fusion of multiple models. We take four typical models as examples, and a prediction algorithm

following the proposed framework is presented. Experimental results show that the prediction accuracy and stability of the proposed approach performs better compared with single models and existing combination models, which indicates the validity and usability of the proposed method.

Performance of the prediction model by fusing multiple models will be influenced by the single models involved to a certain extent. As the quantity and performance of the involved models increase, the performance of the proposed framework could also be improved further. The proposed method is validated by experimental analysis, and the models for components are selected by simple variance analysis. In future work, the principle of each single model will be theoretically analysed, and then, the fusion-based method with better performance will be presented by optimising the procedure of model selection and parameter setting based on rigorous theoretical analysis.

ACKNOWLEDGMENTS

This work was supported in part by grants from the National Natural Science Foundation of China (nos. 61602508 and 61379151).

REFERENCES

- Ai, Q.S., Xu, T.H., Li, J.J. and Xiong, H.W. (2016). The short-term forecast of Beidou satellite clock bias based on wavelet neural network. *Proceedings of China Satellite Navigation Conference*, 145–154, doi: 10.1007/978-981-10-0934-1_14.
- Batista, P. (2015). Long baseline navigation with clock offset estimation and discrete-time measurements. *Control Engineering Practice*, **35**, 43–53, doi: 10.1016/j.conengprac.2014.10.009.
- Cacciapuoti, L., Dimarcq, N., Santarelli, G., Laurent, P., LEMONDE, P., Clairon, A., Berthoud, P., Jornod, A., Reina, F., Feltham, S. and Salomon, C. (2007). Atomic clock ensemble in space: Scientific objectives and mission status. *Nuclear Physics B (Proceedings Supplements)*, **166**, 303–306, doi: 10.1016/j.nuclphysbps.2006.12.033.
- Davis, J., Bhattarai, S. and Ziebart, M. (2012). Development of a Kalman filter based GPS satellite clock time-offset prediction algorithm. *Proceedings of European Frequency and Time Forum*, 152–156, doi: 10.1109/EFTF.2012.6502355.
- Galleani, L. and Tavella, P. (2010). Time and the Kalman filter. *IEEE Control Systems Magazine*, **30**(2), 44–65, doi: 10.1109/MCS.2009.935568.
- Galleani, L., Sacerdote, L., Tavella, P. and Zucca, C. (2003). A mathematical model for the atomic clock error. *Metrologia*, **40**(3), S257–S264, doi: 10.1088/0026-1394/40/3/305.
- Griggs, E., Kursinski, E.R. and Akos, D. (2014). An investigation of GNSS atomic clock behavior at short time intervals. *GPS Solutions*, **18**(3), 443–452, doi: 10.1007/s10291-013-0343-7.
- Guo, H.R. (2006). Study on the analysis theories and algorithms of the time and frequency characterization for atomic clocks of navigation satellites. *Dissertation, Zhengzhou Information Science and Technology Institute*.
- Heo, Y.J., Cho, J. and Heo, M.B. (2010). Improving prediction accuracy of GPS satellite clocks with periodic variation behavior. *Measurement Science and Technology*, **21**(7), 073001–1–8, doi: 10.1088/0957-0233/21/7/073001.
- Huang, G.W., Zhang, Q. and Xu, G.C. (2014). Real-time clock offset prediction with an improved model. *GPS Solutions*, **18**(1), 95–104, doi: 10.1007/s10291-013-0313-0.
- Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q.A., Yen, N., Tung, C.C. and Liu, H.H. (1998). The empirical mode decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society A Mathematical Physical and Engineering Sciences*, **454**(1971), 903–995, doi: 10.1098/rspa.1998.0193.
- Lei, Y., Hu, Z.P. and Zhao, D.N. (2014). A novel method for navigation satellite clock bias prediction considering stochastic variation behavior. *Proceedings of China Satellite Navigation Conference*, 369–379, doi: 10.1007/978-3-642-54740-9_32.
- Li, X.X. and Zhao, X.H. (2012). Improving the estimation of uncalibrated fractional phase offsets for PPP ambiguity resolution. *Journal of Navigation*, **65**, 513–529, doi: 10.1017/S0373463312000112.

- Liu, J.Y., Chen, X.H., Liu, Q. and Sun, J.Z. (2013). Prediction of satellite clock errors using LS-SVM optimized by improved artificial Fish Swarm algorithm. *Proceedings of IEEE International Conference on Signal Processing*, 1–5, doi: 10.1109/ICSPCC.2013.6664130.
- Pratt, J., Axelrad, P., Larson, K.M., Lesage, B., Gerren, R. and DiOrto, N. (2013). Satellite clock bias estimation for iGPS. *GPS Solutions*, **17**(3), 381–389, doi: 10.1007/s10291-012-0286-4.
- Senior, K.L., Ray, J.R. and Beard, R.L. (2008). Characterization of periodic variations in the GPS satellite clocks. *GPS Solutions*, **12**(3), 221–225, doi: 10.1007/s10291-008-0089-9.
- Shi, J.B., Xu, C.Q., Li, Y.H. and Gao, Y. (2015). Impacts of real-time satellite clock errors on GPS precise point positioning-based troposphere zenith delay estimation. *Journal of Geodesy*, **89**(8), 747–756, doi: 10.1007/s00190-015-0811-7.
- Stein, S.R. and Evans, J. (1990). The application of Kalman filters and ARIMA models to the study of time prediction errors of clocks for use in the Defense Communication System (DCS). *Proceedings of the 44th Annual Symposium on Frequency Control*, 630–635, doi: 10.1109/FREQ.1990.177553.
- Tang, G.F., Xu, X.Q., Cao, J.D., Liu, X.P. and Wang, Q.Y. (2015). Precision analysis for Compass satellite clock prediction based on a universal clock offset model. *Science China: Physica, Mechanica & Astronomica*, **45**(7), 079502–1–6, doi: 10.1360/SSPMA2015-00121.
- Wang, J.G. (2010). Research on time comparison based on GPS precise point positioning and atomic clock prediction. *Dissertation, University of Chinese Academy of Sciences*.
- Wang, J.G., Hu, Y.H., He, Z.M., Wu, J.F., Ma, H.J. and Wang, K. (2011). Prediction of clock errors of atomic clocks based on modified linear combination model. *Chinese Astronomy and Astrophysics*, **35**(3), 318–326, doi: 10.1016/j.chinastron.2011.07.009.
- Wang, Y.P., Lv, Z.P., Gong, X.C., Zhou H.T., Wang, N. (2015) Analysing and comparing the prediction performance of several satellite clock bias prediction models. *Journal of Geodesy and Geodynamics*, **35**(3): 373–378, doi: 10.14075/j.jgg.2015.03.003
- Wang, Y.P., Lv, Z.P., Qu, Y.Y., Li, L.Y. and Wang, N. (2017). Improving prediction performance of GPS satellite clock bias based on wavelet neural network. *GPS Solutions*, **21**(2), 523–534, doi: 10.1007/s10291-016-0543-z.
- Wang, Y.P., Lv, Z.P., Wang, N., Li, L.Y. and Gong, X.C. (2016). Prediction of navigation satellite clock bias considering clock's stochastic variation behavior with robust least square collocation. *Acta Geodaetica et Cartographica Sinica*, **45**(6), 646–655, doi: 10.11947/j.AGCS.2016.20150569.
- Xi, C., Cai, C.L., Li, S.M., Li, X.H., Li, Z.B. and Deng, K.Q. (2014). Long-term clock bias prediction based on an ARMA model. *Chinese Astronomy and Astrophysics*, **38**(3), 342–354, doi: 10.1016/j.chinastron.2014.07.010.
- Xu, B., Wang, Y. and Yang, X.H. (2013a). A hybrid model for navigation satellite clock error prediction. *Computational Intelligence*, **465**, 307–316, doi: 10.1007/978-3-642-35638-4_20
- Xu, B., Wang, Y. and Yang, X.H. (2013b). Navigation satellite clock error prediction based on functional network. *Neural Processing Letters*, **38**(2), 305–320, doi: 10.1007/s11063-012-9247-8.
- Xu, X.Q., Zhou, S.S., Shi, S., Hu, X.G. and Zhou, Y.H. (2016). Performance evaluation of the Beidou satellite clock and prediction analysis of satellite clock bias. *Proceedings of China Satellite Navigation Conference*, 27–35, doi: 10.1007/978-981-10-0940-2_3.
- Yuan, H.B., Wang, Z.M., Dong, S.W., Wu, H.T. and Qu, L.L. (2008). Dynamic grey-autoregressive model of an atomic clock. *Metrologia*, **45**(6), S1–S5, doi: 10.1088/0026-1394/45/6/S01.
- Zheng, Z.Y., Lu, X.S. and Chen, Y.Q. (2008). Improved grey model and application in real-time GPS satellite clock bias prediction. *Proceedings of the 4th International Conference on Natural Computation*, 419–423, doi: 10.1109/ICNC.2008.630.
- Zucca, C. and Tavella, P. (2015). A mathematical model for the atomic clock error in case of jumps. *Metrologia*, **52**(4), 514–521, doi: 10.1088/0026-1394/52/4/514.