On the ~ 7 year periodic signal in length of day from a frequency domain stepwise regression method

Can-Can Hsu1,3 · Peng-Shuo Duan1,2 · Xue-Qing Xu1,2 · Yong-Hong Zhou1,2,3 · Cheng-Li Huang1,2,3

Received: 18 August 2020 / Accepted: 15 March 2021 / Published online: 22 April 2021
© The Author(s) 2021

Abstract
The intradecadal variations in length-of-day (LOD) and their time-varying characteristics still need to be further studied. Given that the corresponding signal periods on the intradecadal scales are quite close to each other and the span of currently observed ΔLOD data (i.e., 1962–2019, only 57 years) is not long enough, accurate detection of these signals depends on using effective mathematical method. On the basis of the traditional harmonic retrieval model, this work proposes a frequency domain stepwise regression method, which can well identify the periodic components with close periods and recognize the weaker signals to estimate the relevant harmonic parameters (i.e., amplitude, frequency and phase). Furthermore, we apply this method to detection of the actual LOD intradecadal variations, the result of which shows that there are three components (i.e., the ~5.9 years, ~8.3 years and ~7.3 years) existing in LOD intradecadal variations. Here, we firstly give the time-domain harmonic expression of the ~7.3 years signal, i.e., $y(t) = A \cos \left( \frac{2\pi}{T} (t - 1962 + \psi) \right)$, where the period $T = 7.33$ years, amplitude $A = 0.0413(\pm 0.0099)$ ms, phase $\psi = -1.58(\pm 0.30)$ years. The mechanism of this ~7.3 years signal needs to be further studied.

Keywords LOD · Intradecadal variations · AR-z spectrum · Frequency domain stepwise regression · Frequency domain least squares

1 Introduction
The intradecadal variations (i.e., 5–10 year scales) in length-of-day (LOD) is an interesting topic, which may closely correlate with the physics within the deep Earth interior, e.g., Earth’s core motions (e.g., Gillet et al. 2010, 2015), the fast changes of geomagnetic field (e.g., Mandea, et al. 2010; Silva et al. 2012; Holme and de Viron 2013). Therefore, accurately studying the periodic harmonic components in ΔLOD on the intradecadal scales and their physical origins are significant to understand the Earth’s core motions, geomagnetic field intensity, etc. (e.g., Mound and Buffett 2003; Gillet et al. 2010; Holme and de Viron 2013; Duan et al 2018). As Duan and Huang (2020) indicated, the intradecadal variations in LOD during 1962–2019 are more complex than current findings, i.e., besides two primary harmonic components (i.e., ~5.9 years and ~8.6 years signals with average amplitudes of ~0.124 ms and ~0.08 ms, respectively), a relatively weaker periodic component with period between 5.9 and 8.6 years also exists. It should be emphasized that Duan and Huang (2020) explicitly referred to the existence of a ~7.2 years signal in LOD using the normal Morlet wavelet transformation method (see Page 11 in the Peer Review File of Duan and Huang (2020), which can be downloaded from https://www.nature.com/articles/s41467-020-16109-8#Sec13), though Duan and Huang (2020) did not give the specific harmonic expression of this ~7.2 years signal.
Some previous works seemed to show the existence of this ~7.0 years signal as well: For example, Vondrák (1977) indicated that two prominent signals (i.e., ~8.67 years and ~6.75 years) are present on the intradecadal scales in ΔLOD during 1955.5–1976.5 based on a filtering approach, while using the traditional Fourier analysis, Silva et al. (2012) showed that there are three frequency peaks (i.e., ~6.0 years, ~7.0 years and ~8.5 years) existing in the first-order derivative of LOD (i.e., dLOD/dt) on 5–10 year scales (see Fig. 1), nevertheless, they did not refer to any discussion of the ~7.0 years and ~8.5 years signals. Furthermore, the existence of this ~7.0 years signal is still worthy of further studying in light of the following two points:

1. As mentioned above, Vondrák (1977) only exhibited two signals (i.e., ~7.0 and ~8.67 years), however, recently many works (e.g., Holme and de Viron 2013; Chao et al. 2014; Duan et al. 2015) all referred to the existence of a ~5.9 years oscillation signal in ΔLOD, that is to say, Vondrák (1977) might not well clarify these two signals (i.e., ~5.9 years and ~7.0 years) in that time. Besides, this ~7.0 years signal is very close to the ~5.9 years signal in frequency domain, so that other methods (such as the band-pass filtering) may easily confuse these two signals due to their poor frequency resolution.

2. The currently observed ΔLOD data (i.e., 1962–2019, only ~57 years) are not long enough and the target periods of these intradecadal signals are close to each other, as well as the traditional Fourier analysis owns the so-called spectral-leakage phenomenon, which means that the ~7.0 years peak shown in Fig. 1 cited from Silva et al. (2012) may be a false signal.

Consequently, in order to further confirm the existence of this ~7.0 years signal mentioned by Duan and Huang (2020) and estimate its related harmonic parameters in a new way (which is different from the methods mentioned above), we need to adopt an effective mathematical harmonic analysis method to accurately analyze the signals on the intradecadal scales.

The traditional harmonic retrieval model (see Formula (1) in this work) is a parameterized method, which has been applied to many different research fields, for instance, the communications, radar and earthquake. This harmonic retrieval model is just to extract the harmonic signals from the original observed data with white or colored noises (Stoica and Moses 2005). According to the nonlinear least-squares theory (Stoica and Nehorai 1989), if we can obtain frequency estimation of each harmonic signal, then the following problem is just to deal with the issues of simple linear-least-squares estimation. However, here the difficult issues are as following: How to determine the number of harmonic components existing in the original data and how to solve the nonlinear issue about the frequency estimation (Chao 1990).

In order to solve the above questions, a series of super-resolution methods has been developed, which can be divided into three categories: nonlinear regression model, auto-regressive and moving average (ARMA) model and co-variance matrix model (Stoica and Moses 2005). On the basis of the auto-regressive (AR) method developed by Chao and Gilbert (1980), which belongs to the category of the ARMA model methods, Ding and Chao (2018) adopted the z-domain auto-regressive (AR-z) spectrum. This method can significantly simplify the above nonlinear frequency estimation issues through the low-order AR form in frequency domain, which has been applied to some data analysis fields (e.g., Ding and Chao 2018; Ding 2019). Despite this, the frequency resolution of AR-z is still limited, e.g., AR-z cannot effectively separate the ~7.0 years component in ΔLOD data (Ding 2019) and this work will further show this point.

On the basis of the current harmonic retrieval model, we propose a new scheme on the nonlinear regression model in this work, i.e., frequency domain stepwise regression (FDSR), which not only can overcome the frequency-resolution limitation when the original data length is not long enough, but also can detect the relative weaker harmonic components from the original series with close periods. Nevertheless, this method depends on the frequency initialization of harmonic signals, which is just the limitation of this type of nonlinear regression model (Stoica and Moses 2005; Stoica et al. 1989). Of course, if some priori frequencies information (which may be determined by previous works or AR-z spectrum) can be

![Fig. 1 Fourier spectrum of the dLOD/dt time series (cited from Fig. 1b in Silva et al. (2012), in which all red annotations are added by us to indicate the main periods explicitly)](image-url)
introduced into the harmonic retrieval model in advance, the stability of FDSR method will be further improved, and the frequency estimation will be closer to the globally optimal solution. We apply this method to the detection of the intra-decadal variations in LOD, and we confirm the existence of a ≈7.0 years signal in ΔLOD; meanwhile, we firstly give the relevant parameters (i.e., amplitude, period and initial phase) of the target ≈7.0 years signal.

2 Methodology

The traditional harmonic retrieval model (Formula (1)) is to retrieve \( M \) complex sinusoids components from a white or colored stochastic noise, where the frequencies, phase and amplitudes of these \( M \) sinusoids are all unknown. Considering the series used here are real, the related expression can be described by

\[
y[n] = \sum_{m=1}^{M} C_m \cos(\omega_m n) + \sum_{m=1}^{M} S_m \sin(\omega_m n) + \epsilon[n]
\]

where \( y[n] \) is the uniform sampled observed data (the observed data can be non-uniform in a more general case, but which is out of the scope of this work), the index number \( n \in \{0, 1, 2, \ldots, N-1\} \), where \( N \) represents the data length; \( \epsilon[n] \) is the white or colored stochastic noise; \( \omega_m \) is the frequency of the \( m \)-th component; we can take the sum \( C_m \cos(\omega_m n) + S_m \sin(\omega_m n) \) as a harmonic signal, while \( C_m \) and \( S_m \) are the amplitudes of sine and cosine components, respectively.

2.1 Frequency domain least squares

If the frequencies of \( M \) harmonic components are known, the problem of harmonic retrieval can be transformed into a simple least squares problem. Here, we consider the least squares problem in the frequency domain and introduce the main principle of this method.

Step 1 Turning the formula (1) into the matrix form, where \( X_{(N \times 2M)} \) is composed of \( \cos(\omega_m n) \) and \( \sin(\omega_m n) \); \( \hat{\beta}_{(2M \times 1)} \) is made up of \( C_m \) and \( S_m \); \( \epsilon_{(N \times 1)} \) is the random noise.

\[
Y = X \hat{\beta} + \epsilon
\]

Step 2 Applying the windowed Fourier transformation to formula (2), which equals to multiplying transformational matrix on both sides, that is:

\[
F \Lambda Y = FAX \beta + FA \epsilon
\]

where \( \Lambda_{(N \times N)} \) is a diagonal matrix, the diagonal element being the value of window function (such as rectangular window, Hanning window, Blackman window), \( F_{(2N \times N)} \) is a discrete Fourier transformation matrix; in order to separate the real and imaginary parts, we set the first \( N \) lines as the cosine components, and the following \( N \) lines as the sine components. The expression of \( F \) is written as

\[
F = \begin{bmatrix} \text{Re}(Y) \\ \text{Im}(Y) \end{bmatrix}, Y_{nk} = e^{-j \frac{2\pi}{N} nk}, k = 0, 1, \ldots, N - 1
\]

It can be seen that windowed Fourier transform is actually a linear transformation of the vectors, which can concentrate the information that we need into a certain subspace. If we do the least squares in this subspace, the interference of irrelevant signals will be greatly reduced. We call this subspace as the inspection frequency band.

Step 3 In order to filter out the important frequency bins, we only need to multiply a row filtering matrix \( J_{(2N \times N)} \). The expression is as follows.

\[
J = \begin{bmatrix} \tilde{J}^T \end{bmatrix}, \tilde{J}_{(K \times N)} = \begin{bmatrix} 0_k \cdots 0_k E_k \cdots 0_k \end{bmatrix}
\]

where \( 0_k \) is a zero vector with the length \( k \); \( E_k \) is a \( k \times k \) unit matrix, where \( k \) reflects the number of frequency bins within the target frequency band, and \( k << N \).

\[
JFAX \beta = JFAX \beta + JFAX \epsilon
\]

Thus, we can establish a regression equation in the target frequency band. Here, we write \( F \triangleq \text{JFAX} \) and assuming that the co-variance matrix of \( \epsilon \) is \( \Sigma_{(N \times N)} \), then the co-variance matrix of \( F \epsilon \) is \( F \Sigma F^T \), which is expressed as \( \Sigma_{(2N \times 2N)} \).

Step 4 Multiplying by \( \Sigma^{-1/2} \) both sides of formula (4) and further denoting \( \Sigma^{-1/2} Y_{(2N \times 1)}, \Sigma^{-1/2} X_{(2N \times M)}, \Sigma^{-1/2} \Sigma F X_{(2N \times 1)}, \) \( \Sigma^{-1/2} \Sigma F X_{(2N \times M)} \), respectively. Then, we can obtain

\[
\bar{Y} = \bar{X} \beta + \bar{\epsilon}
\]

Here, \( \bar{\epsilon} \) respects the Gauss-Markov assumptions, then \( \beta \) can be solved by ordinary least squares method. The co-variance matrix of \( \bar{\epsilon} \) is an identity matrix, so we can further assume that \( \bar{\epsilon} \) follows the multivariate normal distribution, and our frequency domain stepwise regression method is established on this basis.

2.2 Frequency domain stepwise regression

Stepwise regression is a mature method, which has been widely used in many areas (Myers 1986; Montgomery et al. 2012). When the harmonic frequencies are unknown, it can be used to solve the problem: In the target frequency band, we can set up \( M \) candidate harmonics with equal frequency interval. As long as \( M \) is large enough, a certain combination of some candidate harmonics can achieve the globally optimal fitting of Eq. (1). In essence, stepwise regression is an efficient search algorithm, and we can obtain the result that is closer to the globally optimal combination.
Here, each time we can include a harmonic component, which means to include two variables, i.e., the cosine and sine components. Therefore, the null hypothesis and the statistics F-test have changed; for example, in order to determine whether the \( q \)-th candidate harmonic can be introduced into the model, the null hypothesis is expressed as follows.

\[
H_0 : C_m = 0, S_m = 0
\]

According to the general linear hypothesis theory (see Sect. 3.3.4 of Montgomery et al. 2012 for detail), our F-test statistics is as follow:

\[
F = \frac{\text{RSS}_q - \text{RSS}_{q+1}}{\text{RSS}_{q+1}/(k - q - 1)} \sim F(2, 2(k - q - 1))
\]  

(6)

where \( q \) expresses the number of the included harmonic components, \( \text{RSS}_q \) expresses the Residual Sum of Squares (RSS) of the \( q \) harmonic components, while \( \text{RSS}_{q+1} \) expresses the RSS after including the \( q \)-th candidate harmonic.

Based on Eq. (5), the \( \text{RSS}_q \) can be calculated and expressed as follow:

\[
\text{RSS}_q = \left\| \overline{Y} - \overline{X}_q \hat{\beta}_q \right\|^2
\]  

(7)

where \( \| \cdot \| \) is the 2-norm of vector, \( \overline{X_q} \) equals \( \Sigma^{-1/2} \Gamma X_q \) is the estimate value of \( \beta_q \). Here, \( X_q \) and \( \beta_q \) have the same meaning as \( X \) and \( \beta \) in Eq. (2), but both \( X_q \) and \( \beta_q \) are composed of \( q \) harmonic components, so \( X_q \) and \( \beta_q \) have the size of \( N \times 2q \) and \( 2q \times 1 \), respectively.

\[
Y_i(t) = 0.12 \sin \left( \frac{2\pi}{5.9} (t - 1962) + \phi_1 \right) + 0.05 \sin \left( \frac{2\pi}{7.3} (t - 1962) + \phi_2 \right)
+ 0.08 \sin \left( \frac{2\pi}{8.5} (t - 1962) + \phi_3 \right) + \varepsilon(t); \quad \varepsilon(t) \sim N(0, 0.025^2)
\]

If there are many harmonic components existing in the target frequency band (or the data length is not long enough to distinguish them), there will be multi-collinearity between them. In order to measure the collinearity of the two adjacent harmonics, we can use Pearson correlation coefficient, which can be used for a general linear model (Kraha et al. 2012) and expressed by \( C_H \). The value range of \( C_H \) is \([0, 1]\); when \( C_H > 0.8 \), it means that the frequencies of the two adjacent harmonics are very close, and the estimation of their parameters will produce large bias; when the \( C_H \) value is smaller, the separation of the introduced harmonics will be better, the disturbances of them are weaker and the result will be more accurate.

As for the harmonic signals located at both ends of the target frequency band, there will be a large systematic bias in their estimation. This is because only half or even less of the effective data for these harmonics will be involved in the regression process (whose data in the target frequency band are almost zeros). Therefore, we provide two solutions: (1) Putting the two ends of the target frequency band at the minimum point of the Fourier spectrum; (2) If there are known harmonics (for example, the frequencies determined by AR-z spectrum), we can artificially include them (or part of) into the model in advance, which is helpful to improve the stability of the results.

3 Comparison of tests between FDSR and AR-z spectrum

Considering the relatively short \( \Delta \text{LOD} \) data (i.e., 1962–2020, only \( \sim \) 58 year), while the periods of intradecadal signals in \( \Delta \text{LOD} \) are just quite close, we attempt to apply FDSR method to obtain the periodic components and the related harmonic parameters (i.e., period, phase and amplitude) of the intradecadal LOD variations. Before this, we would like to carry out a simulation analysis to check the reliability of FDSR in estimating the target signals.

3.1 Test A

Here, a simulation is applied to a synthesized signal \( Y_1 \), which is composed of three steady signals, the periods of which are, respectively, 5.9 years, 7.3 years and 8.5 years. Meanwhile, the time span of the simulation signal is 1962/01/01–2018/12/31, and their respective amplitudes are 0.12, 0.05 and 0.08. The expression of the synthetic signal is written in the form.
On the ~7 year periodic signal in length of day from a frequency domain stepwise regression method

If we adopt the FDSR method, then the results of the three signals are listed in Table 1 indeed, we find that this ~7.0 years signal is automatically introduced into the final results, and the other two estimated periods generally coincide with the periods of the original simulated signals, though there are some small deviations. Importantly, above test shows that FDSR method can detect this relatively weaker ~7.0 years signal.

Table 1 FDSR results of the simulated signals $Y_1$, the introduced harmonic signals and their parameter estimation

| Period (years) | Amplitude          | C     | S     | $C_{H}$  |
|---------------|--------------------|-------|-------|----------|
| 8.254         | +8.354e−02         | −8.308e−02 | +3.229e−03 | 0.09     |
|               | ±8.876e−03         | ±8.942e−03 | ±8.825e−03 |          |
| 7.031         | +3.170e−02         | +2.434e−02 | +1.809e−02 | 0.04     |
|               | ±8.794e−03         | ±8.975e−03 | ±9.036e−03 |          |
| 5.890         | +1.228e−01         | −3.179e−02 | +1.183e−01 |          |
|               | ±9.166e−03         | ±9.113e−03 | ±9.160e−03 |          |

The first line values of each signal are estimated parameters and the second line are STD (standard deviations) of those values. The $C_{H}$ value in the middle of two adjacent signals is the Pearson correlation coefficient between them two.

Fig. 2 Synthetic signal of Test A. a The three simulated harmonic signals, whose amplitudes are stable and their periods are 5.9 years (blue line), 7.3 years (orange line) and 8.5 years (yellow line), respectively; b shows the synthesized signal with white noise added (standard deviation = 0.025).

Fig. 3 The stabilized AR-z spectrum analysis of the simulated signal $Y_1$.

Moses 2005), but this spectrum cannot represent the harmonic amplitude, similar phenomena can be seen in Burg’s maximum entropy spectrum (see Figure 4 of Malkin and Terentev 2007).

If we adopt the FDSR method, then the results of the three signals are listed in Table 1 indeed, we find that this ~7.0 years signal is automatically introduced into the final results, and the other two estimated periods generally coincide with the periods of the original simulated signals, though there are some small deviations. Importantly, above test shows that FDSR method can detect this relatively weaker ~7.0 years signal.
The variables introduced by stepwise regression algorithm may be not globally optimal, but they are close to the globally optimal solution. If the frequency of each harmonic needs to be estimated more accurately, the harmonic frequency extracted by FDSR can be taken as the initial value, and a certain algorithm can be used to search for the globally optimal solution which meets a certain criterion. Taking result of Table 1 as an example, we use the following scheme to improve the period estimation:

**Step 1** Taking the three periods estimated by FDSR as initial values, which is denoted as \( \{T_1, T_2, T_3\} \), and then adding random numbers with zero means into these initial values, thus three new periods \( \{\bar{T}_1, \bar{T}_2, \bar{T}_3\} \) are obtained, here, these new periods satisfy the following relationship \( U > \bar{T}_1 > \bar{T}_2 > \bar{T}_3 > L \), where \( U \) and \( L \) are the designated to be lower and upper boundaries, here, \( U = 9, L = 5.4 \);

**Step 2** Taking these three new periods \( \{\bar{T}_1, \bar{T}_2, \bar{T}_3\} \) from Step 1 as the known values, which are further used to fit the simulation signal \( Y_1 \) through frequency domain least-squares method, where the rectangular window is used. Then, the RSS\((k)\) is subsequently calculated, where \( k \) is the loop index and \( k > 1 \).

Repeating the above two steps for 10,000 times in this work. In each loop, if RSS\((k) < \text{RSS}(k-1)\), then, \( \{T_1, T_2, T_3\} \) (in next loop) is further replaced by \( \{\bar{T}_1, \bar{T}_2, \bar{T}_3\} \) (in current loop); otherwise, if RSS\((k) \geq \text{RSS}(k-1)\), then, \( \{T_1, T_2, T_3\} \) (in next loop) is kept to be unchanged, meanwhile, RSS\((k)\) is set to be equal to RSS\((k-1)\). Finally, we can get the results of \( \{T_1, T_2, T_3\} \) and the corresponding harmonic parameters with the smallest RSS.

This is just the FDSR correction scheme. Although the amount of calculation of this scheme is large, it is easy to obtain the globally optimal solution since this scheme is simple and practical. We call the above scheme as the random correction (RC). The outputs of RC of the results shown in Table 1 are listed in Table 2. On the whole, the frequency estimation of FDSR plus RC (i.e., FDSR + RC) is closer to the real frequency, since it is the globally optimal solution under the criterion of minimum RSS.

As shown in Fig. 4, we further compare the noise-free simulated signals with the results estimated by FDSR + RC. It can be seen from Fig. 4a that there is a deviation between the signal estimated by FDSR and real signal, which is due to some errors from the estimation of the frequencies. Figure 4b shows that the signal estimated by FDSR + RC is more consistent with the real signal, which demonstrates that if the real signal itself is composed of periodic signals.

### Table 2

| Period (years) | Amplitude | C         | S         | \( C_H \) |
|---------------|-----------|-----------|-----------|----------|
| 8.466         | +8.245e−02| −6.845e−02| +4.598e−02| 0.06     |
| 7.238         | +3.758e−02| +3.741e−02| −3.161e−03| 0.20     |
| 5.912         | +1.210e−01| −1.703e−02| +1.198e−01|          |
| 11.010        | ±1.275e−03| ±1.264e−03| ±1.264e−03|          |

More metadata can be found in the header of Table 1.
with stable amplitude, the FDSR + RC results can obtain more accurate results.

### 3.2 Test B

Next, we will further test the FDSR method using signals with amplitude changes. The mathematical expression of the new simulation signal is as follows:

\[
Y_2(t) = 0.15e^{-0.0101(t - 1962)} \sin \left( \frac{2\pi}{5.9} (t - 1962) + \phi_1 \right) + 0.04 \sin \left( \frac{2\pi}{7.3} (t - 1962) + \phi_2 \right) \\
+ 0.05e^{0.0157(t - 1962)} \sin \left( \frac{2\pi}{8.5} (t - 1962) + \phi_3 \right) + \epsilon(t); \quad \epsilon(t) \sim N(0, 0.025^2)
\]

The time span of this simulated signal is set to be 1962/01/01–2018/12/31 with a sampling rate 1 day. The amplitudes of 5.9 years and 8.5 years are from Duan and Huang (2020). Their average amplitudes are, respectively, 0.1140, 0.0400 and 0.0809. Figure 5 shows the total synthetic signal and its harmonic components.

We apply the stabilized AR-z spectrum to analyze the simulated \(Y_2\) signal, which is similar to the procedures shown
in Test A. The result of which is presented in Fig. 6, where the ~ 5.9 years and ~ 8.5 years signals are clearly identified, while the 7.3 years signal is still absent, which is similar to the case mentioned in Test A. The results of FDSR and FDSR + RC are listed in Tables 3 and 4. These two tables show the three periodic signals and the amplitudes estimated by FDSR + RC are consistent with those of the original simulated inputs.

According to the nonlinear least squares theory of the harmonic retrieval model, it is important to estimate the period of each harmonic. Moreover, if the estimation of the period is more accurate, the estimated single harmonic is more representative of the real signal. From the above two tests, the estimated periods of FDSR + RC are closer to the real periods, regardless of the periodic signals with stable or unstable amplitudes. Therefore, the result of FDSR + RC is more reasonable and reliable.

We compare the noise-free simulation signal \( Y_2 \) with the signal estimated by FDSR and FDSR + RC, as shown in Fig. 7. Comparing Fig. 4 of Test A with Fig. 7 of Test B, it can be seen that FDSR + RC can greatly improve the fitting effect for the periodic signal with stable amplitude, but there is little improvement for the fitting effect for the periodic signal with time-varying amplitude. That is to say, the estimation of the signal by FDSR + RC is different from that of the noise-free simulation signal \( Y_2 \), which is determined by the nature of FDSR method, that is, using the harmonics with stable amplitudes to fit the target harmonics with variable amplitudes. Even so, compared with test A and B, it can be seen that FDSR + RC scheme is helpful to obtain the most accurate period estimations.

### 3.3 Demo: the necessity of applying FDSR method

From Test A and Test B, a question may arise: When the original data are not long enough (e.g., ~ 57 years for the length of the currently real LOD data), whether we can accurately detect the target intradecadal signals via combining...
the least squares fitting (LSF) method and fast Fourier transform (FFT) or so-called AR-z spectrum. If we can do it, then it will be unnecessary to adopt the FDSR method. However, the fact is that we cannot do it. In this section, we will further show this point and emphasize the necessity of applying FDSR method.

Here, we denote the noise-free simulate signal $Y_1$ (i.e., $0.12\sin(2\pi(t - 1962)/5.9 + \phi_1) + 0.05\sin(2\pi(t - 1962)/7.3 + \phi_2) + 0.08\sin(2\pi(t - 1962)/8.5 + \phi_3)$) as $Y$. The Taylor windowed Fourier transform of $\tilde{Y}$ is shown in Fig. 8a, from which only the period information of 5.9 years and 8.5 years signals is displayed, while the 7.3 years signal cannot be identified. At this time, using LSF to fit and remove these two signals from $Y_1$, a ~7.3 years signal can be shown in the residual time series (i.e., $R_1$, see the black curve in Fig. 8b), however, further fitting and removing this 7.3 years signal by LSF, the relevant residual results (i.e., $R_2$, see the bold red curve shown in Fig. 8b) are not zero, which demonstrates that using FFT + LSF cannot obtain the ideal target intradecadal signals, which should be due to the mutual interference among these three intradecadal signals with close periods under the case of a relatively short original data. However, the proposed FDSR + RC method can obtain an ideal result (e.g., the $R_3$ result, see the magenta dashed curve in Fig. 8b). In addition, this demo assumes that the periods of the harmonics are exactly known, conversely, if the periods are unknown in advance, the results of LSM + FFT will be worse.

### 4 Application of the FDSR method to detection of the ~7 years signal in real LOD

The ΔLOD datasets with sampling rate of 1 day used in this work are from COMB 2018 (Ratcliff and Gross 2019). Furthermore, according to the IERS Conventions (2010) (Petit and Luzum 2010), we have removed the tidal terms in ΔLOD. The atmospheric angular momentum (AAM) datasets with sampling rate of 6 h are from IERS Special Bureau for Atmosphere, which are computed on the basis of the wind field and surface pressure datasets from NCEP/NCAR reanalysis project (Barnes et al. 1983; Salstein et al. 1993; Kalnay et al. 1996; Zhou et al. 2006). The oceanic angular momentum (OAM) datasets are from the IERS Special Bureau for Ocean, and the ECCO 50 years datasets with sampling rate of 10 days and ECCO kf080 datasets with sampling rate of 1 day are used here. After the related data preprocessing, all the sampling rates of these three datasets are set to be daily, within the time span 1962/01–2018/12. The residual time series in which the tidal, AAM and OAM effects have been removed are shown in Fig. 9a, and we refer it also as ΔLOD.

Ding (2019) detected signals of ~149.0 years, ~68.0 years, ~33.0 years, ~22.3 years, ~18.6 years, ~13.5 years, ~1.0 years, ~8.5 years and ~5.9 years in records using AR-z spectrum. Assuming that all these signals exist, of course, the frequencies of these signals are not necessarily estimated accurately due to the Cramer-Rao lower bound, which is the lower bound of all the minimum variance unbiased estimators. As the frequency gets close to zero, the Cramer-Rao lower bound goes to infinity (see Example-3.5 in Kay 1998), the result of which indicates that the variance of unbiased estimators of any method in analyzing for the long-periods signals is large when the data are not long enough. Here, we use the Vondrák smoothing (Vondrák 1969) filter to weaken the long-term trends, which cannot affect the signals on the intradecadal scales, see Fig. 9b, which also shows the comparison of the ΔLOD series (i.e., the blue line) with the residual series (i.e., ΔLOD$R$, the magenta line, which refers to the output after removing the longer period term with periods > 10 years filtered by the Vondrák filter in the frequency domain). Last but not least, although this filter can weaken the > 10 years scale components, as demonstrated...
in Fig. 9b, it still includes the remaining signals with periods > 10 years, since these signals cannot be fully eliminated by this filtering method.

Next, applying FDSR without RC to ΔLOD_R. Considering the band with period > 5.0 years. We regard the above four periodic signals 22.3 years, 18.6 years, 13.5 years, 11.0 years as the known initiation periods and introduce them into the harmonic retrieval model (i.e., formula (1)) in advance. After the automatic including and excluding procedures of FDSR, we further introduce seven periodic signals with periods of 22.30 years, 17.94 years, 13.50 years, 11.20 years, 8.319 years, 6.957 years, 5.903 years, respectively. It can be seen that the 18.6 years and 11.0 years signals are excluded and then included, but the periods become 17.94 years and 11.20 years signals are 0.51. Therefore, the above phenomenon should originate from the influence of the adjacent harmonics (22.30 years and 13.50 years). It is not difficult to find that without any artificial intervention, three target signals (5.903 years, 6.957 years and 8.319 years) are automatically introduced into the model through the hypothesis test. The FDSR results of ΔLOD_R are listed in Table 5. Here, it should be noted that in Table 5, these signals with periods > 10 years in ΔLOD have been interference by the Vondrák smoothing process (see Fig. 9b), so the amplitudes of harmonic signals with periods > 10 years estimated by FDSR method cannot express the real ΔLOD record.

Comparison of the seven harmonics combination with the original ΔLOD_R in frequency domain shows that the seven harmonics combination can well characterize the ΔLOD_R, as Fig. 10 shows, the frequency-spectral within the band of 5–10 year reduces to 10^{-3} ms from 10^{-1} ms. Figure 10 further shows that if we do not consider the ~ 7.0 years signal, then the combination of the other six harmonics signals cannot well fit the frequency spectrum on the intradecadal scales, since the spectral of ΔLOD_R after removing the six harmonics combination (black curve) within the 5–10 year band is at the same order as the original ΔLOD_R which means that a significant ~ 7.0 years signal exists.

Here, we would like to further determine the statistical significance of the ~ 7.0 years periodic peak. The amplitude estimations of the ~ 7.0 years signal in Table 5 are

---

**Table 5** FDSR results of ΔLOD_R, the estimated harmonic signals and their parameter estimation

| Period (years) | Amplitude (ms) | C (ms)       | S (ms)       | CH  |
|---------------|----------------|--------------|--------------|-----|
| 22.30         | +1.621e−01     | −1.598e−01   | +2.607e−02   |     |
|               | ±7.403e−03     | ±7.458e−03   | ±8.269e−03   | 0.71|
| 17.94         | +1.354e−01     | −3.683e−02   | −1.300e−01   |     |
|               | ±8.417e−03     | ±1.022e−02   | ±8.306e−03   | 0.04|
| 13.50         | +1.857e−01     | −1.546e−01   | +1.025e−01   |     |
|               | ±7.712e−03     | ±8.111e−03   | ±7.331e−03   | 0.51|
| 11.20         | +6.324e−02     | +6.066e−02   | −1.696e−02   |     |
|               | ±6.553e−03     | ±6.705e−03   | ±5.927e−03   | 0.14|
| 8.319         | +7.629e−02     | −7.246e−02   | −2.335e−02   |     |
|               | ±4.731e−03     | ±4.701e−03   | ±4.858e−03   | 0.36|
| 6.957         | +3.993e−02     | −3.716e−02   | +1.380e−02   |     |
|               | ±4.772e−03     | ±4.902e−03   | ±4.577e−03   | 0.31|
| 5.903         | +1.273e−01     | −4.187e−02   | −1.201e−01   |     |
|               | ±4.404e−03     | ±4.398e−03   | ±4.428e−03   |     |
four times larger than their STDs, which means that we have enough confidence to reject the hypothesis that indicates the amplitude of ~7.0 years signal is zero. To be specific and strict, the F-test is used to test the hypothesis:

\[ H_0 : C_7 = 0, S_7 = 0 \]

where \( C_7 \) and \( S_7 \) are the amplitudes of cosine and sine components of the introduced ~7.0 years signal. Based on the general linear hypothesis theory (Montgomery et al. 2012), the F-test statistic is as follow

\[
F = \frac{RSS_6 - RSS_7}{RSS_7/(K - 7)} \sim F(2, 2(K - 7))
\]

Here, \( k = 11 \). Providing the RSS values derived from the process for the results shown in Table 5, the test statistic for \( H_0 : C_7 = 0, S_7 = 0 \) is that \( F = 66.42 \) with the \( P \) value \( P = 1.04 \times 10^{-5} \). This is strong evidence to reject the hypothesis \( H_0 \) (Montgomery et al. 2012), and one would feel very confident in claiming the existence of the ~7.0 years signal.

Furthermore, in order to use FDSR + RC method to obtain a more accurate periods estimation of \( \Delta \text{LOD}_R \), we take the above four periods (i.e., 22.3 years, 18.6 years, 13.5 years, 11.0 years) as known and fixed values, and only adjust the other three intradecadal periods (i.e., 8.319 years, 6.957 years, 5.903 years in Table 5) according to the RC scheme. The results are listed in the following Table 6.

| Period (years) | Amplitude (ms) | \( C_7 \) (ms) | \( S_7 \) (ms) | \( C_H \) |
|----------------|----------------|----------------|----------------|---------|
| 8.284          | +6.988e-02     | -6.689e-02     | -1.808e-02     | 0.45    |
| ±1.016e-02     | ±1.014e-02     | ±1.052e-02     |               |         |
| 7.333          | +4.131e-02     | +8.578e-03     | +3.903e-02     | 0.19    |
| ±9.912e-03     | ±9.761e-03     | ±1.000e-02     |               |         |
| 5.867          | +1.238e-01     | -1.890e-02     | -1.222e-01     |         |
| ±7.549e-03     | ±7.482e-03     | ±7.506e-03     |               |         |

5 Discussion

It should be noted that under the current case of a relatively short observed LOD data, any analysis method should own some errors, hence, this ~8.3 years period estimated by FDSR + RC is also not completely accurate, of course, the ~8.67 years, ~8.5 years or ~8.6 years periods from previous works (i.e., Vondrák 1977; Ding 2019; Duan and Huang 2020) are also the approximately estimated values. Based on the FDSR + RC, a ~8.3 years result can be given, in this work, we call it as ~8.3 years signal. Nevertheless, here, we must stress again that although this ~8.3 years period may be not accurate completely, which is in general well consistent with the estimated values (i.e., ~8.5 years or ~8.6 years) from the different data processing approaches and the data analysis methods used in the previous works.

Here, we further use the normal Morlet wavelet transformation (NMWT from Liu et al. 2007) to compare \( \Delta \text{LOD}_R \) with the signals estimated by FDSR + RC in both time and frequency domain, the results of which are shown in Fig. 11. In order to improve the frequency resolution, the window width parameter \( \sigma \) is set to be 5, and the ~7.0 years signal existing in actual \( \Delta \text{LOD} \) is clearly shown in Fig. 11-left. Moreover, Comparing with Fig. 11-left and 11-right, the \( \Delta \text{LOD}_R \) is in general well consistent with the three target harmonics on the intradecadal scales, where Fig. 11-right shows the synthesized seven harmonic signals with stable amplitudes from FDSR + RC method. However, there are still some subtle but interesting differences:
the absolute value of wavelet coefficients in the magenta dash line box on the left is smaller than that in the same position in the right figure;
the absolute value of the wavelet coefficients in the magenta solid line box in the left figure is larger than that in the same position in the right figure.

That is to say, under the action of a same mathematical rule (i.e., NMWT), the results displayed in Fig. 11-left and 11-right are different. This distinction actually implies that the amplitudes of ~5.9 years and ~8.5 years signals in ΔLOD should be unstable, otherwise, Fig. 11-left and 11-right should be in good accordance with each other. Here, we attribute this distinction to the amplitude attenuation of the ~5.9 years signal (Duan et al. 2015, 2017) and the amplitude amplifying of the ~8.5 years signal in ΔLOD (Duan and Huang 2020).

Besides, we further confirm that the ~7.0 years signal in ΔLOD cannot be attributed to AAM/OAM or the hydrologic angular momentum (HAM). The HAM dataset from 1948 to 2009 provided by Chen et al. (2019) can be found at http://ftp.csr.utexas.edu/pub/ggfc/ham/. The Blackman windowed Fourier spectrum of HAM and ΔLOD is shown in the following Fig. 12, from which we can find that the HAM effect is very weak (~3 × 10^{-3}ms) in the 5–10 years scale. Hence, the ~7.0 years signal origin is unlikely related to HAM effect.

Furthermore, we consider the original LOD record, i.e., the COMB 2018 of Ratcliff and Gross (2019), from which the effects of AAM/OAM/HAM are not removed. After
related pre-processing (see Fig. 12), the NMWT spectrum of the origin LOD record is shown in Fig. 13. The result indicates that this ~7.0 years signal is obviously present on the intradecadal scales. Considering the HAM effect is quite weak on the intradecadal scales (see Fig. 12), and combining Fig. 13 (i.e., the original LOD record) with the above Fig. 11 (which shows the results after removing the AAM/OAM effects), the ~7.0 years signal in ΔLOD is still present, which means that the ~7.0 years signal cannot be attributed to AAM/OAM/HAM effects.

As mentioned above, any series analysis method may have its own advantages and disadvantages, the proposed FDSR method should have the corresponding applicable scope as well, so we are very cautious about the results from the only one data process method. In this work, we study and confirm these target signals not only from one proposed method (i.e., FDSR) but also from the other different methods (e.g., FFT, AR-\(z\) spectrum and NMWT), meanwhile, many simulation cases are also shown to confirm the effectiveness of the relevant methods to verify the existence of the target ~7 years signal. From the currently observed datasets and the methods adopted in this work, the results confirm that this ~7 years signal in LOD is a robust signal with a relatively steady period.

In the end, the question that what is the physical mechanism responsible for this ~7.0 years signal in LOD is interesting, which needs to be further studied. Nevertheless, a most recent work (Chi-Durán et al. 2020) detected a periodic mode signal with 7.08(±0.58) years period from the geomagnetic field changes and the fluid core motion. Besides, the results from Gillet et al (2010, 2015) indicated that the predicted LOD variations from the ensemble average torsional oscillation flow model are consistent with the result of the observed LOD changes on 4–9.5 years, which shows that the fast torsional waves within the FOC may also have the corresponding ~7.0 years periodic component, as the torsional waves can transfer angular momentum from the FOC to the mantle, and then cause corresponding LOD variations. Hence, we tend to consider this ~7.0 years signal as the Earth core origin. Further in-depth mechanism research may resort to the relevant numerical simulation method, we will strive for further this research in the future.

6 Conclusions

In this work, a new method, i.e., FDSR with random correction scheme, is proposed to further detect the intradecadal variations in LOD, and some tests performed with synthetic data show the good performance of the new method. The result shows that there are three components (i.e., the ~5.9 years, ~8.3 years and ~7.3 years) existing in the intradecadal variations of LOD. The mean amplitudes of the ~5.9 years and ~8.3 years signals are estimated to be 0.1238 (±0.0075) ms and 0.0699 (±0.0102) ms, respectively. Furthermore, we estimate the average amplitude, period and phase parameters for the ~7.0 years signal expressed in the form.

\[
y(t) = A \cos \left( \frac{2\pi}{T} (t - 1962 + \psi) \right), t \in [1962, 2019]
\]

And obtain the following values: the period \(T = 7.33\) years, amplitude \(A = 0.0413(±0.0099)\) ms, phase \(\psi = -1.58(±0.30)\) year.

However, limited by the nature of the proposed FDSR scheme, one can only give a steady harmonic model about this ~7.3 years signal. Of course, we do not exclude the alternative possibility that this ~7.3 years signal could perform a time-varying behavior. In addition, finding out the physical mechanism of this ~7.3 years signal, which is shown to be unrelated to the surface factors (e.g., AAM/OAM/HAM.
effects) but may be related to the FOC motions (Gillet et al. 2015; Chi-Durán et al. 2020), is an interesting topic and needs to be further studied in future.

Acknowledgements The authors thank the anonymous reviewers and the associate editor for critical and constructive comments; the authors also thank J.T. Ratchiff and R.S. Gross for providing ΔLOD time series, thank NCEP/NCAR for the wind field and surface pressure datasets, thank R.S. Gross for providing the ECCO OAM datasets, thank J.L. Chen and C.R. Wilson for providing the HAM datasets; the authors also thank X.H. Liao for the suggestions of the FSDR method, and M. Fang for discussion on the mechanism of the intradecadal variations of LOD. This work is supported by the B-type Strategic Priority Program of the Chinese Academy of Sciences (Grant No. XDB41000000), Youth Innovation Promotion Association CAS and National Natural Science Foundation of China (Grant Nos. 11371037, 11673049, 11803064, 11773058, 11773057).

Authors contribution CCH, YZ and PD performed the primary analysis of this manuscript, CCH and MX made the data processing, CCH and PD led the primary writing of this manuscript, CLH also contributed to the writing of this manuscript.

Data availability The ΔLOD time series is got from JPL’s Geodynamics and Space Geodesy Group at https://keof.jpl.nasa.gov/combinatios/2018. The AAM datasets are got from IERS Special Bureau for Atmosphere at http://files.aer.com/ aerweb/AAM. The OAM datasets are got from IERS Special Bureau for Ocean at https://euler.jpl.nasa.gov/sbo/sbo_data.html.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Barnes RTH, Hide R, White AA, Wilson CA (1983) Atmospheric angular-momentum fluctuations, changes in the length of the day and polar motion. Geophys J R Astron Soc 73:280–280

Chao BF (1990) Comment on “A new method of spectral analysis and its application to the earth’s free oscillations—the ’Somp’ method” by S. Hori et al. J Geophys Res 95:19789–19790. https://doi.org/10.1029/JB095iB12p19789

Chao BF, Gilbert F (1980) Autoregressive estimation of complex eigen frequencies in low frequency seismic spectra. Geophys J Int 63:641–657

Chao BF, Chung WY, Shih ZR, Hsieh YK (2014) Earth’s rotation variations: a wavelet analysis. Terra Nova 26:260–264. https://doi.org/10.1111/ter.12094

Chen J, Wilson CR, Kuang W, Chao BF (2019) Interannual oscillations in earth rotation. J Geophys Res Solid Earth 124:13404–13414. https://doi.org/10.1029/2019JB018541

Chi-Durán R, Avery MS, Kneezek N, Buffett BA (2020) Decomposition of geomagnetic secular acceleration into traveling waves using complex empirical orthogonal functions. Geophys Res Lett. https://doi.org/10.1029/2020GL087940

Ding H (2019) Attenuation and excitation of the similar to 6 year oscillation in the length-of-day variation. Earth Planet Sci Lett 507:131–139. https://doi.org/10.1016/j.epsl.2018.12.003

Ding H, Chao BF (2018) Application of stabilized AR-x spectrum in harmonic analysis for geophysics. J Geophys Res Solid Earth 123:8249–8259. https://doi.org/10.1029/2018JB015890

Duan PS, Huang CL (2020) Intradecadal variations in length of day and their correspondence with geomagnetic jerks. Nat Commun 11:2273. https://doi.org/10.1038/s41467-020-16109-8

Duan PS, Liu GY, Liu LT, Hu XG, Hao XG, Huang Y, Zhang ZM, Wang BB (2015) Recovery of the 6-year signal in length of day and its long-term decreasing trend. Earth Planets Space. https://doi.org/10.1186/s40623-015-0328-6

Duan PS, Liu GY, Hu XG, Sun YF, Li HL (2017) Possible damping model of the 6 year oscillation signal in length of day. Phys Earth Planet Inter 265:35–42. https://doi.org/10.1016/j.pepi.2017.02.002

Duan PS, Liu GY, Hu XG, Zhao J, Huang CL (2018) Mechanism of the interannual oscillation in length of day and its constraint on the electromagnetic coupling at the core-mantle boundary. Earth Planet Sci Lett 482:245–252. https://doi.org/10.1016/j.epsl.2017.11.007

Gillet N, Jault D, Canet E, Fournier A (2010) Fast torsional waves and strong magnetic field within the Earth’s core. Nature 456:74–77. https://doi.org/10.1038/nature09010

Gillet N, Jault D, Finlay CC (2015) Planetary gyre, time-dependent eddies, torsional waves, and equatorial jets at the Earth’s core surface. J Geophys Res Solid Earth 120:3991–4013. https://doi.org/10.1002/2014JB017786

Holme R, de Viron O (2013) Characterization and implications of intradecadal variations in length of day. Nature 499:202–205. https://doi.org/10.1038/nature12282

Kalnay E et al (1996) The NCEP/NCAR 40-year reanalysis project. Bull Am Meteor Soc 77:437–471. https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2

Kay SM (1998) Fundamentals of statistical signal processing. Prentice-Hall International, Englewood Cliffs.

Kraha A, Turner H, Nimon K, Zientek LR, Henson RK (2012) Tools to support interpreting multiple regression in the face of multicollinearity. Front Psychol 3:44. https://doi.org/10.3389/fpsyg.2012.00044

Liu L, Hsu H, Grafarend EW (2007) Normal Morlet wavelet transform and its application to the Earth’s polar motion. J Geophys Res. https://doi.org/10.1029/2006JB004895

Malkin Z, Terentev D (2007) Parameters of the free core nutation from VLBI data. https://arxiv.org/abs/physics/0702152v1

Mandea M, Holme R, Pais A, Pinheiro K, Jackson A, Verbanac G (2007) Parameters of the free core nutation from VLBI data. https://arxiv.org/abs/physics/0702152v1

Montgomery DC, Peck EA, Vining GG (2012) Introduction to linear regression analysis. Wiley, Hoboken.

Mound JE, Buffet BA (2003) Interannual oscillations in length of day: implications for the structure of the mantle and core. J Geophys Res Solid Earth. https://doi.org/10.1029/2002JB002054

Myers RH (1986) Classical and modern regression with applications. 2nd edn. PWS-KENT, Boston

Petit G, Luzum B (2010) IERS Conventions (2010). Bureau International Des Poids Et Mesures Sevres (France)
Ratcliff JT, Gross RS (2019) Combinations of earth orientation measurements: SPACE2018, COMB2018, and POLE2018, Pasadena, CA: Jet Propulsion Laboratory, National Aeronautics and Space Administration. http://hdl.handle.net/2014/46964

Salstein DA, Kann DM, Miller AJ, Rosen RD (1993) The sub-bureau for atmospheric angular-momentum of the international earth rotation service—a meteorological data center with geodetic applications. Bull Am Meteorol Soc 74:67–80. https://doi.org/10.1175/1520-0477(1993)074%3c0067:TSBFAA%3e2.0.CO;2

Silva L, Jackson L, Mound J (2012) Assessing the importance and expression of the 6 year geomagnetic oscillation. J Geophys Res Solid Earth. https://doi.org/10.1029/2012jb009405

Stoica P, Moses RL (2005) Spectral analysis of signals. Pearson/Prentice Hall, Upper Saddle River.

Stoica P, Nehorai A (1989) Statistical-analysis of two nonlinear least-squares estimators of sine-wave parameters in the colored-noise case. Circ Syst Signal Process 8:3–15. https://doi.org/10.1007/BF01598742

Stoica P, Moses RL, Friedlander B, Soderstrom T (1989) Maximum-likelihood estimation of the parameters of multiple sinusoids from noisy measurements. IEEE Trans Acoust Speech 37:378–392. https://doi.org/10.1109/29.21705

Vondrák J (1969) A contribution to the problem of smoothing observational data. Bull Astron Inst Czechoslov 20:349

Vondrák J (1977) The rotation of the earth between 1955.5 and 1976.5. Stud Geophys Geod 21:107–117. https://doi.org/10.1007/bf01634821

Zhou YH, Salstein DA, Chen JL (2006) Revised atmospheric excitation function series related to Earth’s variable rotation under consideration of surface topography. J Geophys Res. https://doi.org/10.1029/2005jd006608