Acta Geodyn. Geomater., Vol. 21, No. 1 (213), 37–50, 2024
DOI: 10.13168/AGG.2024.0004
journal homepage: https://www.irsm.cas.cz/acta

ORIGINAL PAPER

STATISTICALLY OPTIMAL SLEPIAN METHOD FOR PROCESSING GRACE LEVEL 2 DATA

Yu CAO [1], Guobin CHANG[1] *, Yong FENG [1], Zhengqiang WEI [2] and Yujia WANG [1]

[1] *School of Environment Science and Spatial Informatics, China University of Mining and Technology, Xuzhou, China*
[2] *Institute of Surveying, Mapping and Geoinformation, China Railway Design Corporation, Tianjin, China*

**Corresponding author's e-mail:* guobinchang@hotmail.com

ARTICLE INFO

Article history:
Received 24 October 2023
Accepted 20 January2024
Available online 5 February 2024

Keywords:
GRACE
Slepian
Tikhonov regularization
Power law model
DDK
GCV

ABSTRACT

The global surface mass variations obtained by the Gravity Recovery and Climate Experiment (GRACE) satellite Level-2 products show significant north-south strip noise, which seriously affects the estimation of regional surface mass variations. Most existing filtering methods are based on the processing of spherical harmonic basis functions to remove stripes. However, because spherical harmonic basis functions are only orthogonal globally, they may not be orthogonal in regions, which hinders the effective constraint of quality changes in specific regions. Therefore, this paper converts the spherical harmonic base of GRACE Level-2 to the Slepian base of the interested region and truncates it based on the Shannon number of the region. Even though the signals are concentrated in the region, there are still many stripes present. In order to remove the stripes, it is necessary to introduce the regularization and consider the statistical information of the spherical harmonic coefficient. The Tikhonov regularization matrix of the Slepian coefficient is obtained, combined with a power-law model to construct the prior covariance matrix of the signal and the optimal regularization coefficient is selected by using the generalized cross-validation (GCV) method, which is represented as the statistically optimal Slepian method (SO-Slepian) in this study. The results show that the ability of SO-Slepian and decorrelation and denoising Kernel (DDK) filtering with the same regularization parameters as SO-Slepian to remove stripes and retain signals in the selected region is comparable. The line chart comparing the differences between SO-Slepian and DDK not only demonstrates the close similarity in results between SO-Slepian and traditional DDK filtering in the regional domain but also emphasizes the logical application of regularization during the Slepian modeling stage. It further supports the rationality of developing a regularization scheme for GRACE data processing based on Slepian.

1. INTRODUCTION

GRACE is a gravity satellite launched by the German Aerospace Center and NASA in March 2002, its main purpose is to acquire monthly time-varying Earth gravity fields to explore hydrosphere, ice, and ocean mass changes (Tapley et al., 2004). From its launch in March 2002 to its end in July 2017, GRACE satellites have observed the global gravity field during their orbit and accumulated a large amount of gravity data, which is widely used in the study of changes in land water storage, changes in the mass of the Antarctic and Greenland ice sheets, and global sea level changes (Huan et al., 2022). The north-south flight trajectory of the GRACE satellite leads to an uneven distribution of subsatellite points. This leads to a denser data distribution at higher latitudes, causing the presence of distinct north-south stripes in the GRACE satellite Level-2 product. This phenomenon is unrealistic and referred to as strip noise. In addition to strip noise, the time-variable gravity field model of GRACE also exhibits significant high-frequency noise, particularly in the spherical harmonic coefficients beyond the 40th degree. The noise level surpasses the signal. This is mainly attributed to the attenuation of gravity field signals with height, design flaws in the GRACE satellite orbit, errors in the background models, and observation errors resulting from the inaccuracies in on-board load measurements (Chambers et al., 2004; Velicogna and Wahr, 2006). In order to reduce the impact of errors on the true signal, some constraints need to be added to the Level-2 products released by various centers. Currently, the existing constraint methods can be broadly classified into three categories. The first category is spatial domain constraints, which utilize the spatial correlations in the data to impose constraints. Examples include isotropic or anisotropic smoothing (Wahr et al., 1998), DDK filtering based on inversion or regularization (Klees et al., 2008; Kusche, 2007; Kusche et al., 2009), and so on. The second category is time domain constraints, which use a suitable method to characterize the characteristics of the observations in the time domain by forming a time series from a continuous set of observations. Examples

include wavelet analysis (Andrew et al., 2017), time differencing regularization (Ditmar et al., 2018), and so on. The third category is spatial-temporal joint constraints, such as principal component analysis (PCA) (Chao and Liau, 2019), singular spectrum analysis (SSA) (Guo et al., 2018), and so on. There are some other filters such as: Gaussian filtering (Wahr et al., 1998), Swenson filtering (Swenson and Wahr, 2006), P3M6 filtering (Chen et al., 2007), Low-pass filtering (Yang et al., 2022) and so on. The filtering methods mentioned above are primarily designed for global-scale processing, aiming to attenuate global striping errors while retaining the true geophysical signals from all regions as much as possible. However, these methods may overlook the signal preservation in special areas to some extent. The focus of this study is to use the Slepian method to transform GRACE spherical harmonic coefficients to a region of interest while considering the statistical properties of the spherical harmonic coefficients. The study aims to gradually search for optimal Slepian coefficients to obtain "clean" geophysical signals in the region and ultimately invert the regional surface mass changes.

The Slepian functions were first introduced by American applied mathematician Slepian in the 1960s. They are a group of orthogonal functions that can be constructed by maximizing signal energy in a certain frequency band. They can be used in various fields such as spectrum estimation, signal filtering, noise analysis and so on, to improve the accuracy and reliability of signal processing (Simons et al., 2006). The main purpose of introducing the Slepian functions is to solve problems such as finite signal truncation and improve the estimation of signal power spectral density (Slepian, 1983). In general, the problems that can be addressed using the Slepian method can be divided into two categories: spatial concentration of finite spectral signals and spectral concentration of finite spatial signals. The GRACE Level 2 data provided by various analysis centers are finite spectral or band-limited. Therefore, when applying the Slepian method to process this data, it addresses the spatial concentration problem of finite spectral signals.

After Simons applied the Slepian functions to the sphere, more and more scholars began to use the Slepian functions for data processing. Some of them have also applied Slepian functions in processing GRACE satellite data products, such as using Slepian functions for mass estimation of the Antarctic ice sheet (Gao et al., 2019) and using Slepian functions to invert for local-scale hydrological information in the Congo and Nile river sub-basins (Ramillien et al, 2021); Using Slepian functions to monitor water storage in the Aral Sea and its adjacent basins (Tao et al., 2020) and so on. Some of these applications are in regions with weaker strip errors in high-latitude areas, some involve pre-processing of GRACE data, and others involve post-processing after using Slepian functions. Therefore, in this paper, the spherical harmonic basis of GRACE Level-2 data is converted to the Slepian basis of the region of interest. Regularization is applied during the Slepian modeling stage, and a reasonable regularization scheme based on Slepian for processing GRACE data is derived.

2. METHOD

2.1. PRINCIPLES OF THE SLEPIAN METHOD

Any finite spectral signal on the sphere Ω can be represented by the following spherical harmonic model (Simons and Dahlen, 2006),

$$\boldsymbol{f(x)} = \sum_{l=0}^{L}\sum_{m=-l}^{l} \boldsymbol{f_{lm}Y_{lm}(x)} \quad (1)$$

Here, x represents the location (θ, λ) of any point on the sphere Ω, where θ is the latitude and λ is the longitude. Y_{lm} represents the normalized spherical harmonics, and the f_{lm} represents the spherical harmonic coefficients. For a given local region R on the sphere Ω, we aim to find a signal within that region that maximizes the proportion of energy (minimizes the energy outside of that region). This can be expressed as follows:

$$\frac{\int_{R}[\boldsymbol{f(x)}]^2 \boldsymbol{d\sigma(x)}}{\int_{\Omega}[\boldsymbol{f(x)}]^2 \boldsymbol{d\sigma(x)}} = \boldsymbol{max} \quad (2)$$

Where $\sigma(\boldsymbol{x})$ represents the area element at $\boldsymbol{x}$. Substituting equation (1) into the above equation, we can derive the corresponding relationship in the spectral domain. The relationship that the spherical harmonic coefficients of the signal (with maximized energy within region R) must satisfy. Introducing the following variable:

$$\boldsymbol{d_{lml'm'}} = \int_{R} \boldsymbol{Y_{lm}(x)Y_{l'm'}(x)d\sigma} \quad (3)$$

Considering $\int_{\Omega} Y_{lm}(x)Y_{l'm'}(x)d\sigma = 1$, the equation (2) is equivalent to the following expression.

$$\frac{\sum_{l=1}^{L}\sum_{m=-l}^{l}\sum_{l=1}^{L}\sum_{m=-l}^{l} \boldsymbol{d_{lml'm'}f_{lm}f_{l'm'}}}{\sum_{l=1}^{L}\sum_{m=-l}^{l}\sum_{l=1}^{L}\sum_{m=-l}^{l} \boldsymbol{f_{lm}f_{l'm'}}} = \boldsymbol{max} \quad (4)$$

By adopting the vector-matrix form, the equation (4) can be equivalently expressed as follows.

$$\frac{\boldsymbol{f^T Df}}{\boldsymbol{f^T f}} = \boldsymbol{max} \quad (5)$$

The left side of equation (5) represents the Rayleigh quotient. Considering that the scale of the Rayleigh quotient is invariant, for any given constant c, we can obtain the following. $\frac{(cf^T)D(cf)}{(cf^T)(cf)} = \frac{c^2 f^T Df}{c^2 f^T f} = \frac{f^T Df}{f^T f}$ Therefore, to solve equation (5), it is sufficient to solve the following equation.

$$\boldsymbol{f^T Df = max,\ subject\ to\ f^T f = 1} \quad (4)$$

Solving equation (2) is thus equivalent to solving equation (6), and the solution of equation (6) is given by the saddle point of the following Lagrangian.

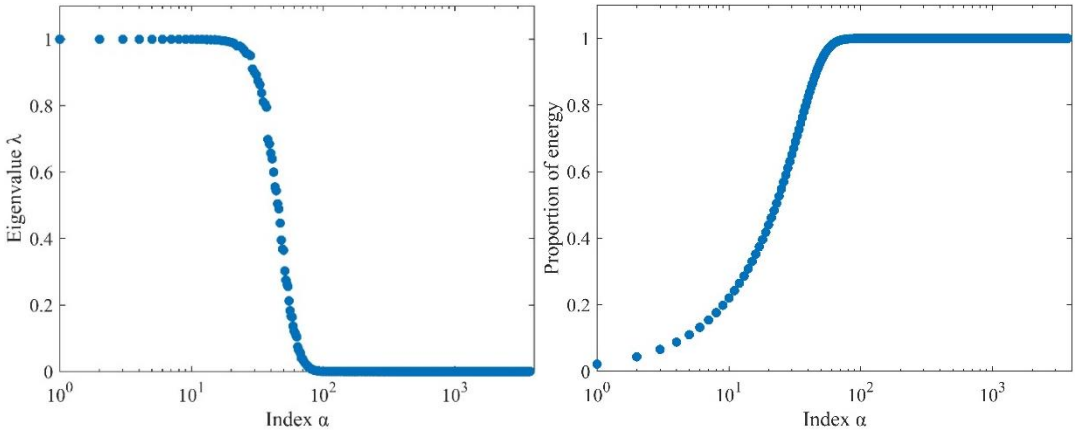

Fig. 1 The eigenvalue λ corresponding to the 3721 Slepian basis function expanded up to degree-60 (left) and the corresponding ratio of the sum of the energy concentration of the first items to the total energy (right), (Amazon as an example).

$$\boldsymbol{\varphi} = \boldsymbol{f}^T\boldsymbol{D}\boldsymbol{f} - \boldsymbol{\lambda}\,(\boldsymbol{f}^T\boldsymbol{f} - \boldsymbol{1}) \tag{5}$$

Thus, we have the necessary conditions for the saddle point, as shown below (setting the first derivative of the above equation to zero):

$$\boldsymbol{D}\boldsymbol{f} = \boldsymbol{\lambda}\,\boldsymbol{f} \tag{6}$$

It can be observed that the solution $\boldsymbol{f}$ of equation (6) must be an eigenvector of $\boldsymbol{D}$. Let the i-th eigenvector of $\boldsymbol{D}$ be represented by $\boldsymbol{g}_i$, where i=1, 2, ..., n=(L+1)^2. Then, the above equation is equivalent to $\boldsymbol{D}\boldsymbol{g}_i=\lambda\boldsymbol{g}_i$, where λ_i is the i-th eigenvalue. Assuming the eigenvalues λ_i are arranged in descending order, let us define the following eigen-signals, also known as Slepian bases.

$$f_i(x) = \sum_{l=0}^{L}\sum_{m=-l}^{l} g_{ilm}Y_{lm}(x) \tag{9}$$

The element of the feature vector $\boldsymbol{g}_i$ corresponding to the l-th degree m-th order is denoted as $\boldsymbol{g}_{ilm}$. For any $i \neq j$, we can get the following equation.

$$\begin{cases}\int_R f_i(\boldsymbol{x})\, f_j(\boldsymbol{x})\mathrm{d}\sigma = \boldsymbol{g}_i^{\mathrm{T}}\boldsymbol{D}\boldsymbol{g}_j = \lambda_j\ \boldsymbol{g}_i^{\mathrm{T}}\boldsymbol{g}_j = 0\\ \int_\Omega f_i(\boldsymbol{x})\, f_j(\boldsymbol{x})\mathrm{d}\sigma = \boldsymbol{g}_i^{\mathrm{T}}\boldsymbol{g}_j = 0\end{cases} \tag{10}$$

It is easy to observe from equation (10) that the aforementioned feature signal $\boldsymbol{f}_i$ is orthogonal on both the spherical region Ω and the domain R. Then substituting equation (9) into the $\frac{\int_R[f(x)]^2d\sigma}{\int_\Omega[f(x)]^2d\sigma}$, get the following equation.

$$\frac{\int_R[f_i(x)]^2d\sigma}{\int_\Omega[f_i(x)]^2d\sigma} = g_i^T D g_i = \lambda_i g_i^T g_i = \lambda_i \tag{11}$$

The above equation indicates that the energy of the feature signal $\boldsymbol{f}_i$ in the region R decreases progressively.

Finally, any spectrum-limited signal of the form given in equation (1) can be represented as a linear combination of the feature signals $\boldsymbol{f}_i$. Thus, we have the following expression:

$$f(x) = \sum_{l=0}^{L}\sum_{m=-l}^{l} f_{lm}Y_{lm}(x) = \sum_{i=1}^{(L+1)^2}\alpha_i f_i(x) = \sum_{i=1}^{(L+1)^2}\alpha_i\sum_{l=0}^{L}\sum_{m=-l}^{l} g_{ilm}Y_{lm}(x) \tag{12}$$

Where the a_i is the Slepian coefficient. By substituting f_i(x) into equation (12), we can obtain:

$$\sum_{l=0}^{L}\sum_{m=-l}^{l} f_{lm}Y_{lm}(x) = \sum_{i=1}^{(L+1)^2}\sum_{l=0}^{L}\sum_{m=-l}^{l}\alpha_i g_{ilm}Y_{lm}(x) \tag{13}$$

Considering the arbitrariness of the position $\boldsymbol{x}$, the above equation is equivalent to the following expression:

$$f_{lm} = \sum_{i=1}^{(L+1)^2} g_{ilm}\alpha_i \tag{14}$$

Written in vector-matrix form, it is as follows:

$$\boldsymbol{f} = \boldsymbol{G}\boldsymbol{\alpha} \tag{15}$$

It is easy to understand from equation (15) that the spherical harmonic coefficients $\boldsymbol{f}$ and Slepian coefficients $\boldsymbol{\alpha}$ can be converted to each other, indicating that $\boldsymbol{f}$ and $\boldsymbol{\alpha}$ are just coordinates of the gravitational signal in different coordinate systems and are one-to-one corresponding to each other. Figure 1 shows the 3721 energy concentrations (means eigenvalues λ) and the corresponding sum of the first energy concentrations as a ratio of the total energy, obtained for the 60th-degree Slepian basis functions in the Amazon. Figure 2 shows the spatial distribution of the first 16 Slepian basis functions in the Amazon. Due to the smooth spreading from the center, Slepian basis functions located near the boundary and far away from the center of the integration area exhibit a reduced sensitivity (Cheng et al., 2021; Harig and Simons, 2012). Therefore, in order to concentrate the signal in the target region as much as possible in the study area, a buffer zone analysis was performed on the study area to establish an appropriately sized buffer zone. In this paper, the buffer selection for the test area is three degrees. The outer blue border in Figure 2 represents the buffer zone

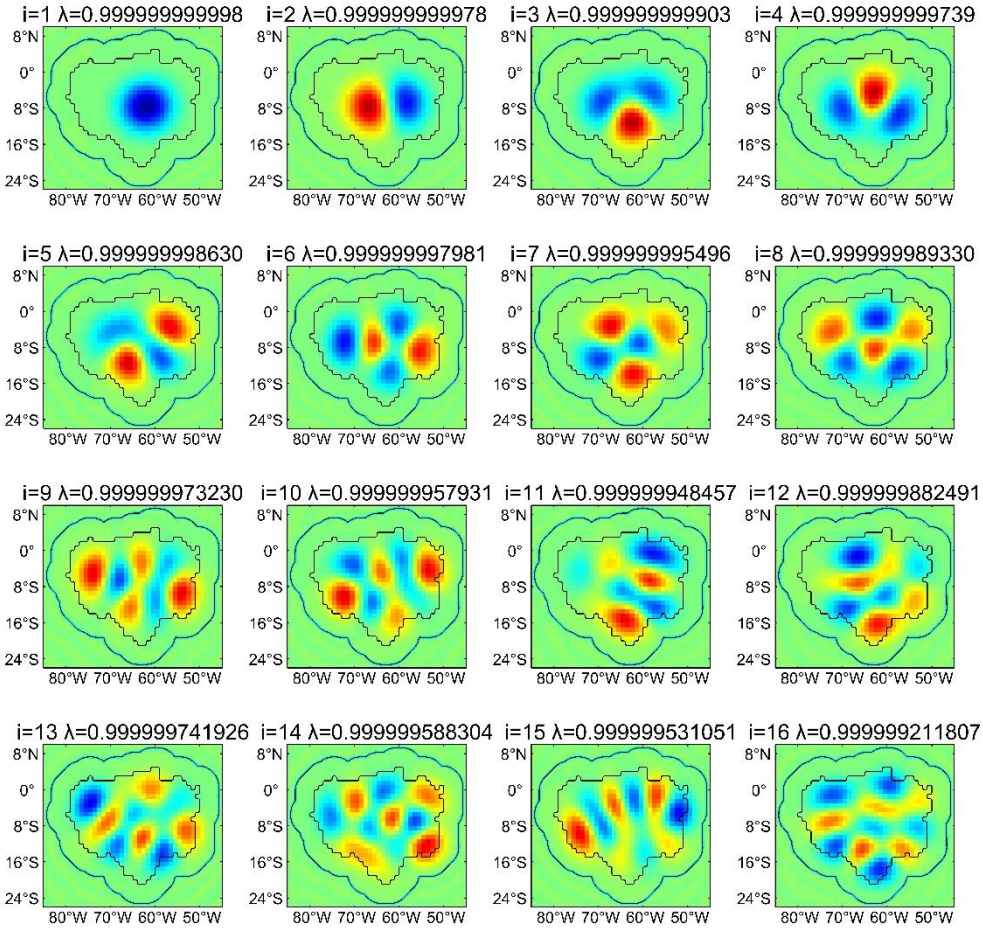

Fig. 2 Signal distribution of the first 16 Slepian basis function (Amazon as an example) i is the number of basis function items, λ is the energy concentration, the inner black contour is the boundary of Amazon and the outer blue contour is the extended boundary for integral calculation.

boundary, which is the new boundary of the integration area. From Figure 1, it can be observed that the concentration of energy for Slepian basis functions becomes closer to 1 as they appear earlier. Combining this with the signal distribution of the Slepian basis functions shown in Figure 2, it is not difficult to find that Slepian basis functions appearing earlier have a larger presence in the specified region, thus allowing for better concentration of the signal in the region of interest. According to equations (11) and (12), the energy contained in the region R for each term on the right side of these equations is different. As the index 'i' becomes sufficiently large, the corresponding λ_i becomes smaller, and as a result, the energy of f_i becomes small enough to be negligible. Based on this, it is possible to truncate the right side of equation (12) to obtain a simpler model with fewer unknowns. For instance, equation (16) computes the Shannon number N, which only includes a finite number of significant non-zero feature signals.

$$N = (L+1)^2 \frac{\int_R d\sigma}{\int_\Omega d\sigma} = (L+1)^2 \frac{A}{4\pi} = \sum_{i=1}^{(L+1)^2} \lambda_i = \text{trace}[\boldsymbol{D}] \quad (16)$$

Therefore, equation (12) is approximated as follows:

$$f(x) = \sum_{i=1}^{(L+1)^2} \alpha_i f_i(x) \approx \sum_{i=1}^{N} \alpha_i f_i(x) \quad (17)$$

Similar to the reasoning process from equation (13) to equation (14), we can get the following observation equation:

$$f_{lm} = \sum_{i=1}^{N} \alpha_i g_{ilm} \quad (18)$$

Now the observation equation is an overdetermined equation containing $(L+1)^2$ observations and N unknowns. It can be expressed in the following vector-matrix form as:

$$\boldsymbol{f} = \bar{\boldsymbol{G}}\bar{\boldsymbol{\alpha}} \quad (19)$$

Where $\bar{\boldsymbol{G}}$ is a submatrix of the orthogonal matrix $\boldsymbol{G}$, and the elements of $\bar{\boldsymbol{\alpha}}$ are the first N elements of α. When we don't consider the statistical properties of the observation vector f and the priori variance of $\bar{\boldsymbol{\alpha}}$, the parameter estimation is shown as follows.

$$\hat{\bar{\boldsymbol{\alpha}}} = (\bar{\boldsymbol{G}}^{\boldsymbol{T}}\bar{\boldsymbol{G}})^{-1}\bar{\boldsymbol{G}}^{\boldsymbol{T}}\boldsymbol{f} = \bar{\boldsymbol{G}}^{\boldsymbol{T}}\boldsymbol{f} \quad (20)$$

However, without considering the statistical information of the observation *f* in equation (20), (its covariance matrix **Q**), the statistical information of measurement errors (including striping error and high-frequency noise) can't be reflected, and the effect of stripe removal can't be achieved. Moreover, it is also important to consider the prior information of the variable $\bar{\boldsymbol{\alpha}}$ (its prior covariance matrix $\bar{\boldsymbol{P}}$). This allows for the contraction of high-degree signals and is also the key to reducing high-frequency noise. Therefore, the optimal estimation of the parameter will be elaborated in detail in the next section.

2.2. OPTIMAL ESTIMATION OF PARAMETER $\bar{\alpha}$

Compared to the ordinary estimation in equation (20), it is difficult to achieve effective error removal. Therefore, a regularized estimation that weights the data and constrains the parameter $\bar{\boldsymbol{\alpha}}$ should be adopted. It is not difficult to observe that using the inverse of the covariance matrix of spherical harmonic coefficients as the weighting matrix for the data is the optimal choice. Similar to DDK filtering, this paper constructs the prior covariance matrix of parameters on the well-known power law model (generalized Kaula criterion) and uses it as the Tikhonov regularization matrix. According to the generalized Kaula criterion, a suitable prior variance for f_{lm} is chosen as shown below (Sasgen et al., 2006).

$$var[f_{lm}] = l^{-\mu} \tag{21}$$

The μ is a hyperparameter that needs to be adjusted. The prior covariance matrix *S* of vector *f* is a diagonal matrix with the above variances as diagonal elements. The equation (15) is reversible and one-to-one correspondence, the prior covariance matrix of parameter *a* is as follows:

$$\boldsymbol{P} = cov[\boldsymbol{\alpha}] = cov[\boldsymbol{G}^T\boldsymbol{f}] = \boldsymbol{G}^T\boldsymbol{S}\boldsymbol{G} \tag{22}$$

Considering that $\bar{\boldsymbol{\alpha}}$ is the truncation of parameter *a* ($\bar{\alpha}$=$\boldsymbol{F}$ *a*) where $\boldsymbol{F}$ is the first N rows of the $(L+1)^2\times(L+1)^2$ identity matrix, the following equation can be obtained:

$$\bar{\boldsymbol{P}} = cov[\bar{\boldsymbol{\alpha}}] = cov[\boldsymbol{F}\boldsymbol{\alpha}] = \boldsymbol{F}\boldsymbol{P}\boldsymbol{F}^T = \boldsymbol{F}\boldsymbol{G}^T\boldsymbol{S}\boldsymbol{G}\boldsymbol{F}^T = \bar{\boldsymbol{G}}^T\boldsymbol{S}\bar{\boldsymbol{G}} \tag{23}$$

The definition and calculation equation of the Tikhonov regularization estimation with equation (19) as the observation equation is as follows, respectively:

$$\hat{\bar{\boldsymbol{\alpha}}} = \underset{\bar{\alpha}}{argmin}[(\boldsymbol{f}-\bar{\boldsymbol{G}}\bar{\boldsymbol{\alpha}})^T\boldsymbol{Q}^{-1}(\boldsymbol{f}-\bar{\boldsymbol{G}}\bar{\boldsymbol{\alpha}}) + \sigma\bar{\alpha}^T\boldsymbol{P}^{-1}\bar{\boldsymbol{\alpha}}] = (\bar{\boldsymbol{G}}^T\boldsymbol{Q}^{-1}\bar{\boldsymbol{G}} + \sigma\bar{\boldsymbol{P}}^{-1})^{-1}\bar{\boldsymbol{G}}^T\boldsymbol{Q}^{-1}\boldsymbol{f} \tag{24}$$

According to the defined equation, as shown below, the prediction matrix (also known as the hat matrix $\boldsymbol{H}$) is obtained:

$$\boldsymbol{H} = \bar{\boldsymbol{G}}(\bar{\boldsymbol{G}}^T\boldsymbol{Q}^{-1}\bar{\boldsymbol{G}} + \sigma\bar{\boldsymbol{P}}^{-1})^{-1}\bar{\boldsymbol{G}}^T\boldsymbol{Q}^{-1} \tag{25}$$

The σ is also a hyperparameter that needs to be adjusted. The commonly used criteria for determining the regularization parameter are the L- curve method (Hansen and O'Leary, 1993), Generalized Cross-Validation (GCV) (Golub et al., 1979; Qian et al., 2021), and minimum Mean Squared Error (MSE) criteria (Hoerl and Kennard, 2000; Ji et al., 2022; Shen et al., 2012). Each of these criteria has its own merits and limitations (Xu, 1992). In this work we choose the GCV.

$$GCV = \frac{(\boldsymbol{f}-\bar{\boldsymbol{G}}\hat{\bar{\boldsymbol{\alpha}}})^T\boldsymbol{Q}^{-1}(\boldsymbol{f}-\bar{\boldsymbol{G}}\hat{\bar{\boldsymbol{\alpha}}})}{((L+1)^2-tr[\boldsymbol{H}])^2} \tag{26}$$

The selection of the two hyperparameters μ and σ in the equation (21) and (24) is performed based on the GCV method as shown in equation (26). This paper takes into account the statistical information of the spherical harmonic coefficients, namely the covariance matrix $\boldsymbol{Q}$, and the prior variance $\boldsymbol{P}$ of parameter *a*. Among the models corresponding to different parameters, the parameter selected by minimizing the GCV is optimal within the chosen parameters. This is how the name "SO- Slepian" method used in this paper is derived.

3. EXPERIMENTAL AND RESULT DISCUSSION

3.1. ANALYSIS OF FILTERING EFFECT

Due to the availability of the covariance matrix of 136 months of GRACE Level-2 RL05 (Release-05) spherical harmonic coefficients up to degree 60, only provided by the Center for Space Research (CSR), University of Texas at Austin, for the period from April 2002 to June 2014. Therefore, in this paper, the 60-degree unconstrained spherical harmonic coefficients of GRACE Level-2 data for the period from April 2002 to June 2014 are used (excluding the missing months), and the data is processed by removing the effects of non-tidal atmospheric signals, high-frequency ocean signals, various tidal components, solid tides, and polar motion. The spherical harmonic coefficients and their corresponding covariance matrix used in the experiments of this paper can be downloaded from the website: http://download.csr.utexas.edu/outgoing/grace/. Before conducting the experiments, the following pre-processing steps are applied to the spherical harmonic coefficients:

1: The first four terms of the spherical harmonic coefficients are omitted.

2: The mean value of the coefficients during the period from 2004 to 2009 is subtracted.

3: The coefficients are adjusted for glacier isostatic adjustment (GIA) using the ICE-6G_D model.

Considering the similarity between the Decorrelation and Denoising Kernel (DDK) filtering method and the SO-Slepian method proposed in this paper, the same parameters are often used for comparison analysis. Therefore, the DDK filtering and SO-Slepian method adopt the same hyperparameters μ and σ. The arbitrariness of variable x in Equation (1) allows for the arbitrary selection of experimental

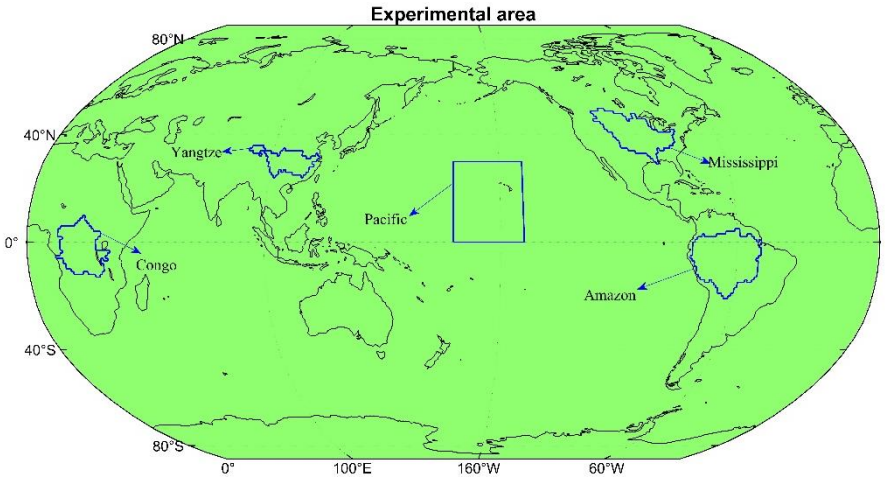

Fig. 3 Experimental area.

Table 1 The optimal parameters for the five regions obtained from equation (24).

Region	Parameter	
	μ	σ
Amazon	3.5	1.1×10^{19}
Congo	4.35	1.0×10^{19}
Pacific	6.8	5.6×10^{19}
Mississippi	4.13	6.36×10^{17}
Yangtze	4.35	5.24×10^{18}

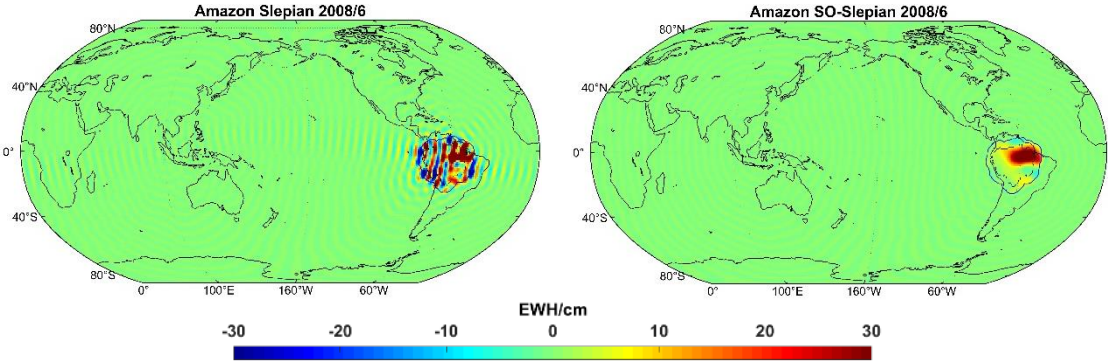

Fig. 4 The global grid map in June 2008 after Slepian (left) and SO-Slepian (right) processing (Amazon as an example).

regions. Therefore, in this paper, five regions were randomly selected for the experiments: a region in the Pacific, the Amazon, the Congo, the Mississippi, and the Yangtze River. Figure 3 shows the selected regions for the experiments.

Changes in quality are usually represented by Equivalent Water Height (EWH), and the following equation can be used to obtain changes in EWH on a global scale.

$$\Delta H = \frac{a\rho ave}{3\rho_w}\sum_{l=0}^{L}\sum_{m=o}^{l}\frac{2l+1}{1+kl}\bar{P}_{lm}(cos\,\theta)(\Delta C_{lm}\,cos(m\lambda) + \\ +\Delta S_{lm}\,sin(m\lambda)) \quad (27)$$

In the equation, $\boldsymbol{\Delta H}$ is the change in Equivalent Water Height, a is the radius of the Earth, ρ_{ave} is the average density of the Earth, ρ_{w} is the density of water, k_l is the load Love number, and θ and λ is the geocentric latitude and longitude, respectively. By substituting the result obtained from equation (9) into equation (17) and considering equation (21) and (24), the change in Equivalent Water Height in the region R can be obtained as follows:

$$\Delta H_{region} = \frac{a\rho ave}{3\rho_w}\sum_{l=0}^{L}\sum_{m=-l}^{l}\frac{2l+1}{1+kl}\sum_{i=1}^{N}\hat{\tilde{\alpha}}_{A_i}g_{ilm}Y_{lm}(x) \quad (28)$$

The optimal parameters obtained from equation (26) in the five regions are shown in Table 1. It can be seen that both μ and σ are highest in oceanic regions, as there is very little signal and almost all of it is noise. As latitude increases, the regularizing parameter σ decreases because the influence of strip errors weakens with increasing latitude and the limitations naturally decrease.

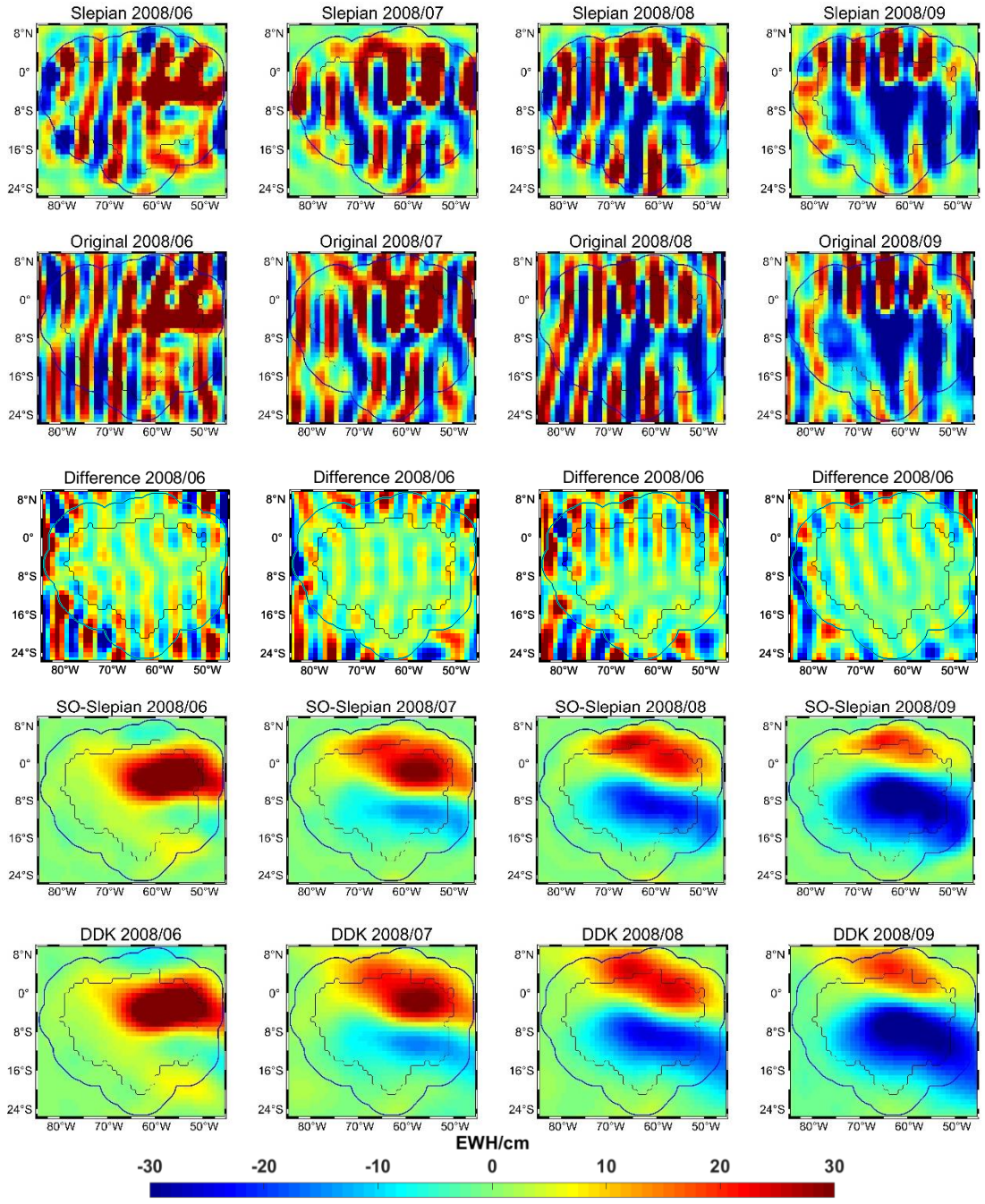

Fig. 5 The processed CSR data results of the Amazon in June, July, August and September 2008 using DDK, Slepian and SO-Slepian methods, as well as the corresponding original results and the difference between Slepian and original (the inner black contour is the boundary of Amazon and the outer blue contour is the extended boundary for integral calculation).

Figure 4 shows that the global grid maps after Slepian and SO-Slepian processing have significantly concentrated signals in the regions of interest, while there are minimal signals outside those regions. The results after Slepian processing still exhibit a considerable number of strip errors, whereas the results after SO-Slepian processing practically eliminate strip errors. Figures 5, 6, and 7 demonstrate the Equivalent Water Height variations in the Amazon, Mississippi, and a specific region of the Pacific in June, July, and August 2008. The figures display the results of Slepian, DDK, and SO-Slepian processing and the corresponding raw data. From the first three lines of these three figures, which represent the results after Slepian processing, the corresponding raw data and the difference between Slepian and original, it can be clearly observed that the Slepian results closely resemble the raw data. This further indicates that the results obtained from Slepian basis functions and spherical harmonic basis functions are just different coordinate representations of the gravity signal, and they are one-to-one correspondences with

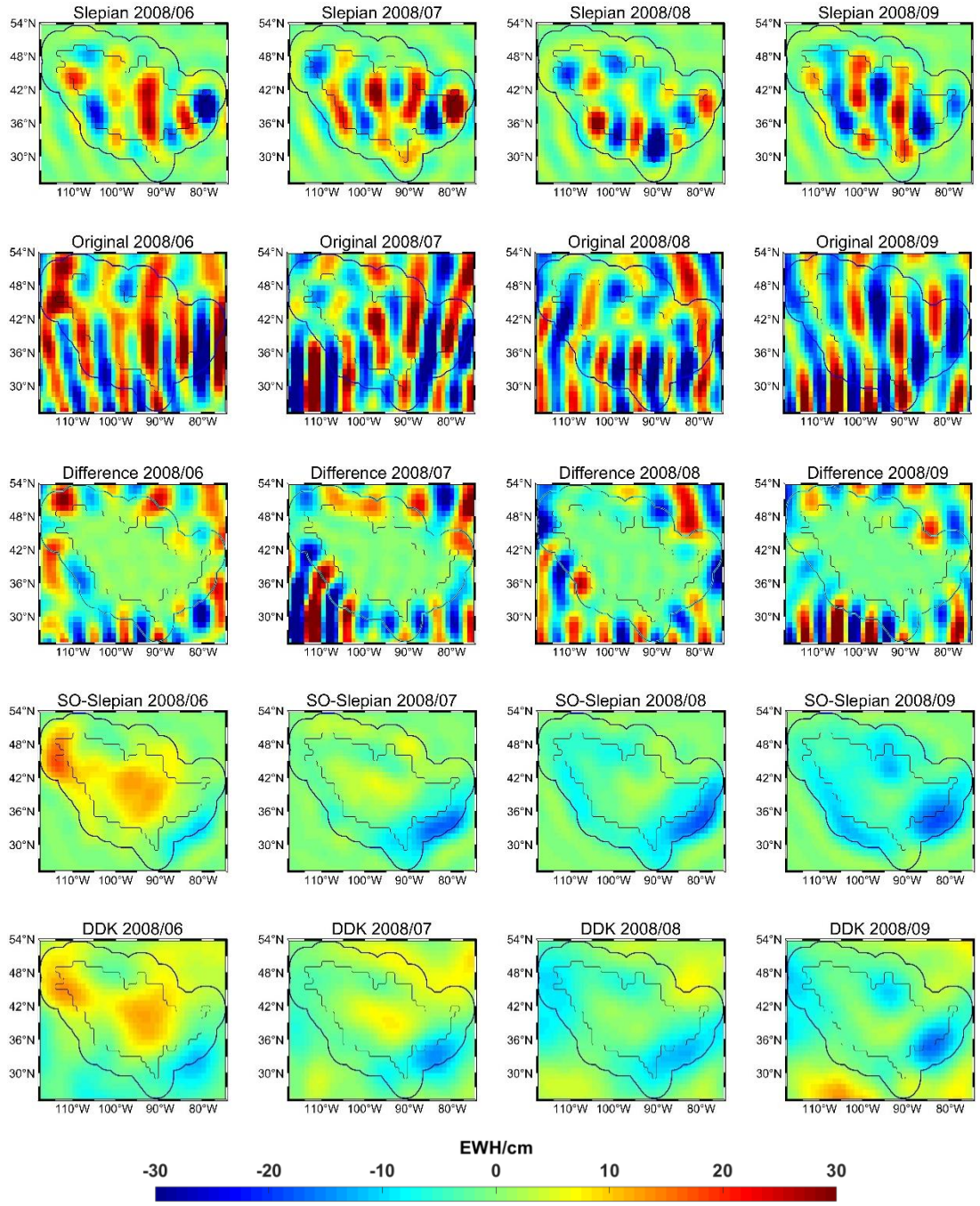

Fig. 6 The processed CSR data results of the Mississippi in June, July, August and September 2008 using DDK, Slepian, and SO-Slepian methods, as well as the corresponding original results and the difference between Slepian and original (the inner black contour is the boundary of Mississippi and the outer blue contour is the extended boundary for integral calculation).

each other. There are some residual signals in the regional difference (the third line) due to the error caused by the introduction of Shannon numbers for simplified calculation. The Slepian functions do not have the capability to remove strip errors. Then looking at the remaining two lines which represent the results after SO-Slepian and DDK processing, it is evident that strip errors have been effectively removed. Furthermore, the filtering effects of these two methods have highly similar results. It emphasizes the logical application of regularization during the Slepian modeling stage and further supports the rationality of developing a regularization scheme for GRACE data processing based on Slepian. From the results in Figure 7, it can be observed that both SO-Slepian and DDK show weak signals in the oceanic region. This is due to the use of the GSM model data in this study, which has already subtracted the atmospheric and oceanic signals. Therefore, only a residual oceanic mass signal that the model did not

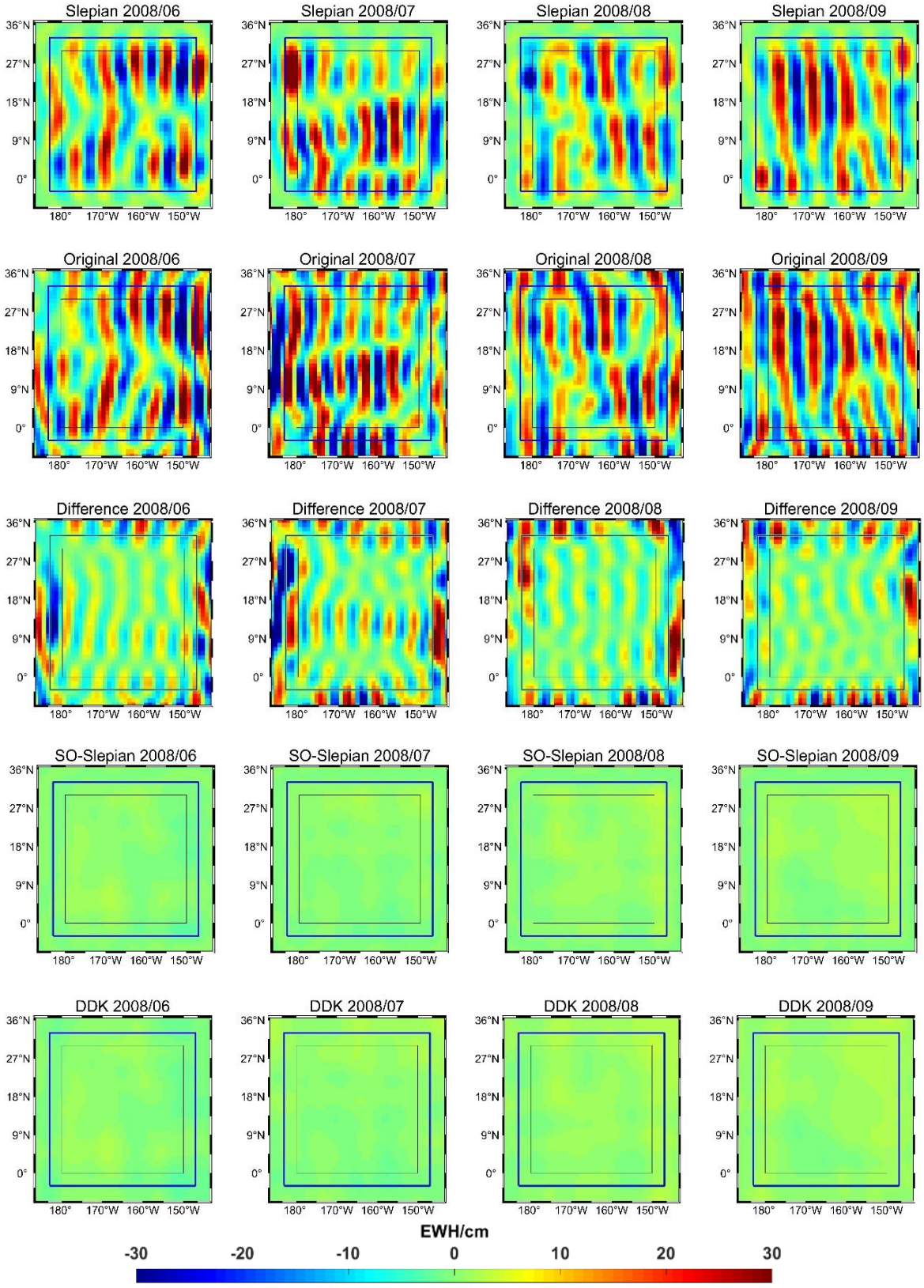

Fig. 7 The processed CSR data results of an area in the Pacific in June, July, August, and September 2008 using DDK, Slepian, and SO-Slepian methods, as well as the corresponding original results and the difference between Slepian and original (the inner black contour is the boundary of an area in the Pacific and the outer blue contour is the extended boundary for integral calculation).

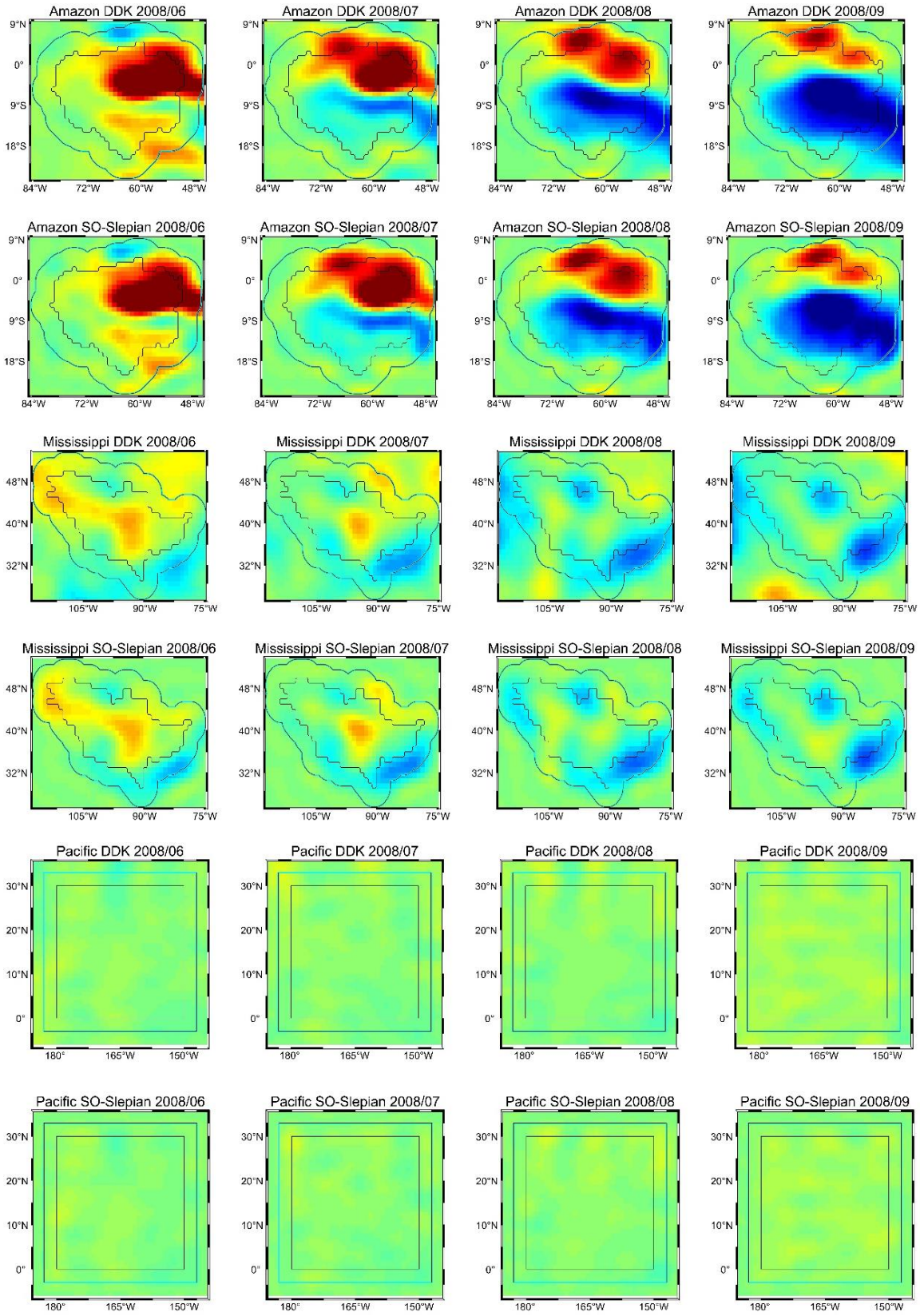

Fig. 8 The processed ITSG data results of Amazon, Mississippi and an area in the Pacific in June, July, August, and September 2008 using DDK and SO-Slepian methods (the inner black contour is the boundary of an area in the Pacific and the outer blue contour is the extended boundary for integral calculation).

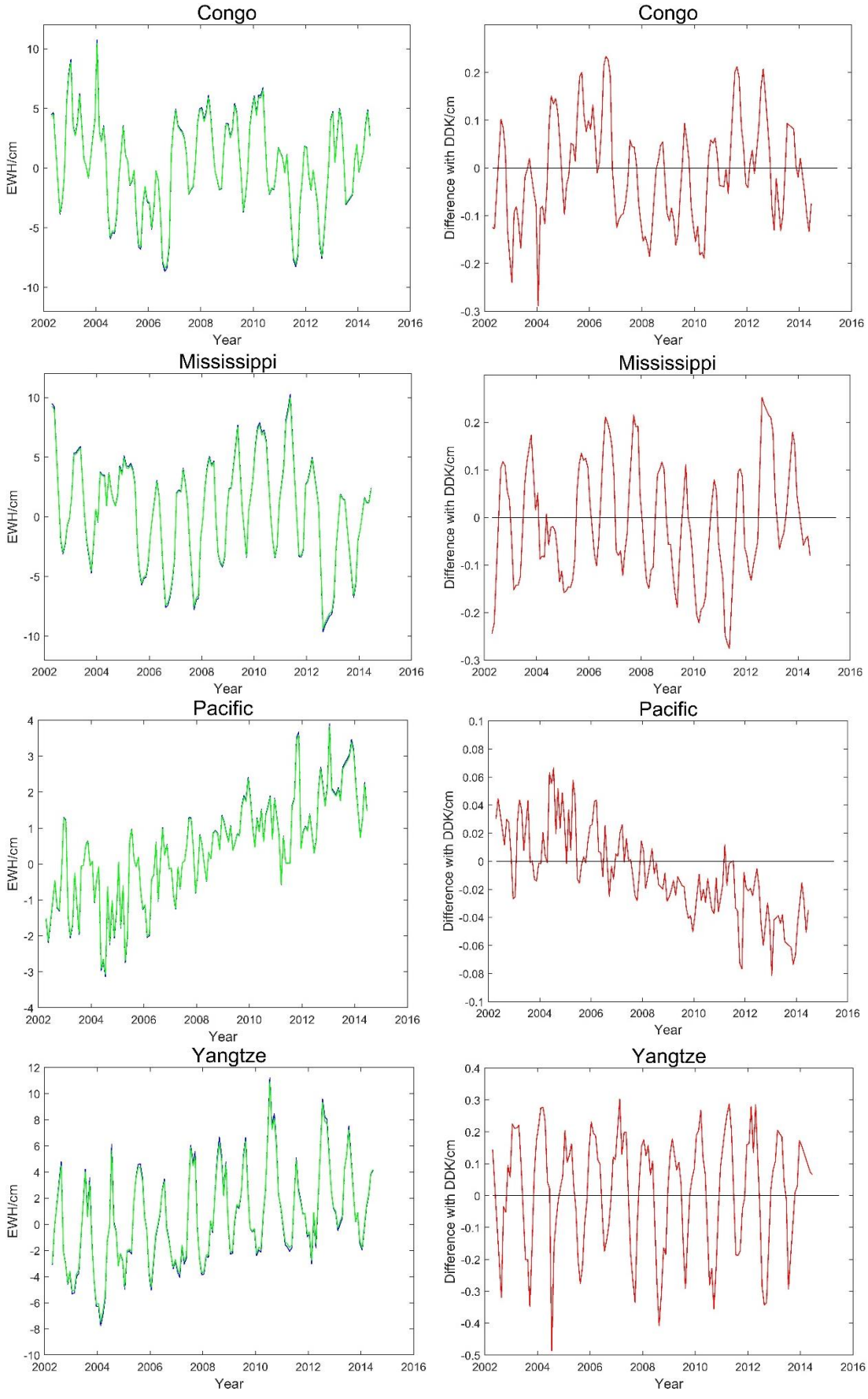

Fig. 9 Changes in EWH over time in five regions filtered by SO-Slepian and DDK (left) and differences between the two filtering (right).

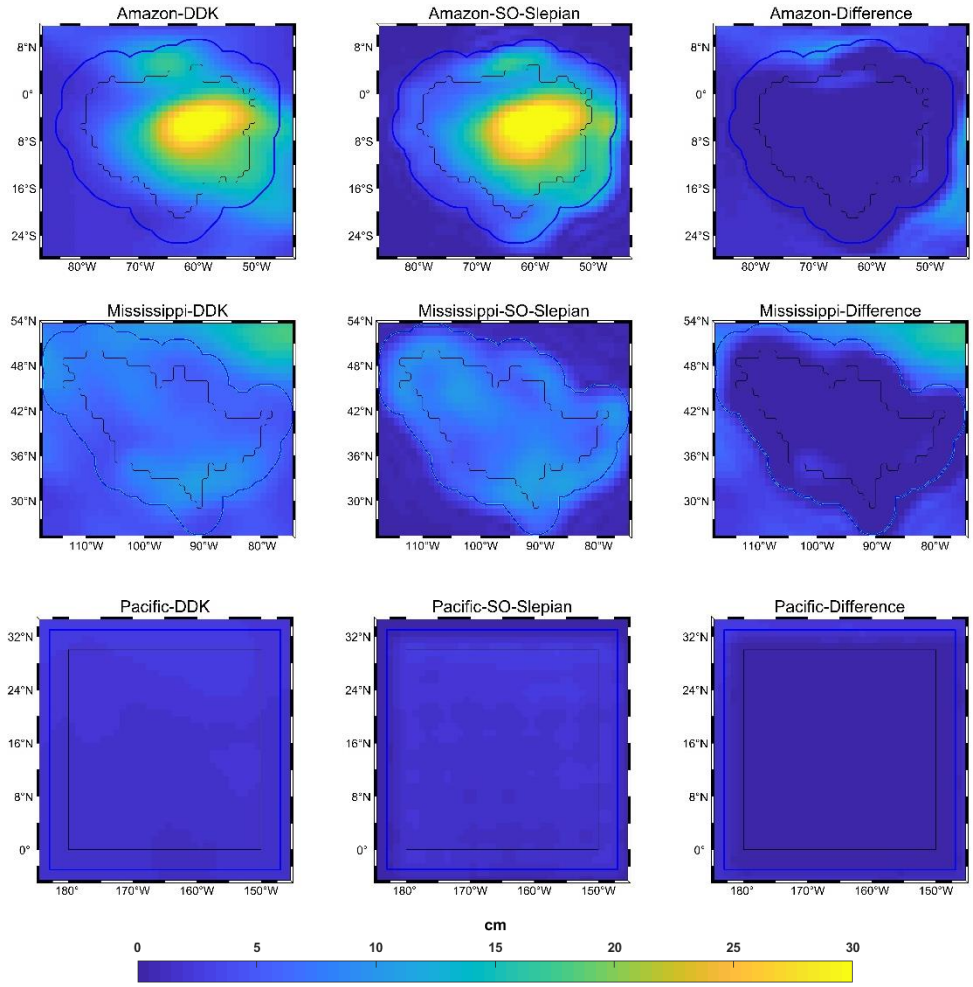

Fig. 10 RMS of EWH for DDK and SO-Slepian filtering methods and difference in the Amazon, Mississippi, and Pacific, April 2002 to June 2014.

completely remove may be present. As a result, this leads to weak signals in the oceanic region. The purpose of selecting the oceanic region in this paper is mainly to examine the effect of stripe errors removal. For the Amazon and Mississippi, the grid maps obtained after SO-Slepian and DDK processing clearly show the variations in EWH within these regions. For instance, in the Mississippi, there is a noticeable positive EWH change around June 2008, followed by a significant negative EWH change around August 2008. Interestingly, these patterns of change are consistent between the results obtained from both SO-Slepian and DDK filtering methods.

To verify the universality of the filtering method proposed in this paper, we added the ITSG-Grace2018 data for experiment. On the website (http://ftp.tugraz.at/outgoing/ITSG/GRACE/ITSG-Grace2018/monthly), it only provides the ITSG-Grace2018 covariance matrix of degree 96th. We selected the same months (June, July, August, and September of 2008) as the CSR data mentioned above and truncated them to 60th degree for experiments.

From the processing results of ITSG data, it can be seen that the SO-Slepian method used in this paper is universal and reflects the effectiveness of the Slepian-based regularization scheme derived in this paper.

3.2. TIME SERIES ANALYSIS

This paper uses a total of 136 months of GRACE spherical harmonic coefficients from April 2002 to June 2014 (including missing months) to obtain EWH through different filtering methods, including Slepian, SO-Slepian, and DDK constraints. Figure 9 shows the EWH time series obtained by SO-Slepian and DDK methods in the five selected regions (left), as well as the differences between the two filtering methods (right) (missing months are ignored in the plot). From the EWH time series (left column), it can be seen that SO-Slepian results are essentially in agreement with DDK for all five regions. From the differences relative to DDK (right column), the filtering results show very little difference, and this combined with the grid maps presented in Figures 5, 6, and 7 suggests that these two filtering methods produced nearly identical results.

Figure 10 illustrates the root mean square (RMS) of the grid in the Amazon, Mississippi, and a specific area of the Pacific region. In terms of error level, both DDK and SO-Slepian exhibit close values and minimal errors. Judging by the differences between DDK and SO-Slepian, the filtering effects of them are very close. Combining the above information with the results shown in Figure 8, it can be concluded that SO-Slepian and DDK filtering methods have nearly identical effects in the specified regions.

4. CONCLUSION

This paper applies the Slepian function with regularization to the GRACE Level-2 data. The data used in this study are the CSR RL05 GRACE Level-2 products from April 2002 to June 2014. The regional mass changes are obtained by applying constraints and compared with the DDK filtering method using the same parameters as SO-Slepian. The following results are obtained.

1. By comparing the grid maps globally and regionally, it is evident that the simple Slepian method only represents the gravity signal in another coordinate system. It can concentrate the signal in a specific region of interest. However, it fails to effectively remove the stripe errors in the region.
2. From the time series in the five selected regions and the differences relative to the DDK solution, it can be observed that the results obtained from SO-Slepian are almost identical to those of DDK. The slight discrepancies can be attributed to the truncation introduced using the Shannon number to simplify the computation in SO-Slepian.
3. Considering the RMS of the regional EWH, as well as the time series and corresponding grid maps in the five selected regions, this not only confirms the consistency between SO-Slepian and DDK results but also highlights the rationality of applying regularization during the Slepian modeling stage in this study, which leads to a regularized scheme for GRACE data processing based on Slepian.

FUNDING

The National Natural Science Foundation of China (42074001), The Open Fund of the State Key Laboratory of Geographic Information Engineering (SKLGIE2020-Z-1-1)

REFERENCES

Andrew, R., Guan, H. and Batelaan, O.: 2017, Estimation of GRACE water storage components by temporal decomposition. J. Hydrol., 552, 341–350. DOI: 10.1016/j.jhydrol.2017.06.016

Chambers, D.P., Wahr, J. and Nerem, R.S.: 2004, Preliminary observations of global ocean mass variations with GRACE. Geophys. Res. Lett., 31, 4. DOI: 10.1029/2004gl020461

Chao, B.F. and Liau, J.R.: 2019, Gravity changes due to large earthquakes detected in GRACE satellite data via empirical orthogonal function analysis. J. Geophys. Res., Solid Earth, 124, 3024–3035. DOI: 10.1029/2018jb016862

Chen, J.L., Wilson, C.R., Famiglietti, J.S. and Rodell, M.: 2007, Attenuation effect on seasonal basin-scale water storage changes from GRACE time-variable gravity. J. Geod., 81, 237–245. DOI: 10.1007/s00190-006-0104-2

Cheng, S., Yuan, L.G., Jiang, Z.S., Liu, Z.G., Zhang, D. and Xu, X.F.: 2021, Investigating terrestrial water storage change in Sichuan, Yunnan, and Chongqing using Slepian basis functions. Chin. J. Geophys.-Ch, 64, 1167–1180, (in Chinese). DOI: 10.6038/cjg202100194

Ditmar, P., Tangdamrongsub, N., Ran, J.J. and Klees, R.: 2018, Estimation and reduction of random noise in mass anomaly time-series from satellite gravity data by minimization of month-to-month year-to-year double differences. J. Geodyn., 119, 9–22. DOI: 10.1016/j.jog.2018.05.003

Gao, C.C., Lu, Y., Shi, H.L., Zhang, Z.Z., Xu, C.Y. and Tan, B.: 2019, Detection and analysis of ice sheet mass changes over 27 Antarctic drainage systems from GRACE RLO6 data. Chin. J. Geophys.-Ch, 62, 864–882, (in Chinese). DOI: 10.6038/cjg2019M0586

Golub, G.H., Heath, M. and Wahba, G.: 1979, Generalized cross-validation as a method for choosing a good ridge parameter. Technometrics, 21, 215–223. DOI: 10.1080/00401706.1979.10489751

Guo, J.Y., Li, W.D., Chang, X.T., Zhu, G.B., Liu, X. and Guo, B.: 2018, Terrestrial water storage changes over Xinjiang extracted by combining Gaussian filter and multichannel singular spectrum analysis from GRACE. Geophys. J. Int., 213, 397–407. DOI: 10.1093/gji/ggy006

Hansen, P.C. and O'Leary, D.P.: 1993, The use of the L-Curve in the regularization of discrete Ill-Posed Problems. SIAM J. Sci. Comput., 14, 1487–1503. DOI: 10.1137/0914086

Harig, C. and Simons, F.J.: 2012, Mapping Greenland's mass loss in space and time. Proc. Natl. Acad. Sci. U S A, 109, 19934–19937. DOI: 10.1073/pnas.1206785109

Hoerl, A.E. and Kennard, R.W.: 2000, Ridge regression: Biased estimation for nonorthogonal problems. Technometrics, 42, 80–86. DOI: 10.1080/00401706.2000.10485983

Huan, C.M., Wang, F.W. and Zhou, S.J.: 2022, Empirical mode decomposition for post-processing the GRACE monthly gravity fielf models. Acta Geodyn. Geomater., 19, 281–290. DOI: 10.13168/agg.2022.0013

Ji, K.P., Shen, Y.Z., Chen, Q.J., Li, B.F. and Wang, W.: 2022, An adaptive regularized solution to inverse Ill-Posed Models. IEEE Trans. Geosci. Remote Sens., 60, 15, DOI: 10.1109/tgrs.2022.3205572

Klees, R., Revtova, E.A., Gunter, B.C., Ditmar, P., Oudman, E., Winsemius, H.C. and Savenije, H.H.G.: 2008, The design of an optimal filter for monthly GRACE gravity models. Geophys. J. Int., 175, 417–432. DOI: 10.1111/j.1365-246X.2008.03922.x

Kusche, J.: 2007, Approximate decorrelation and non-isotropic smoothing of time-variable GRACE-type gravity field models. J. Geod., 81, 733–749. DOI: 10.1007/s00190-007-0143-3

Kusche, J., Schmidt, R., Petrovic, S. and Rietbroek, R.: 2009, Decorrelated GRACE time-variable gravity solutions by GFZ, and their validation using a hydrological model. J. Geod., 83, 903–913. DOI: 10.1007/s00190-009-0308-3

Qian, N.J., Chang, G.B., Gao, J.X., Pan, C., Yang, L., Li, F.C., Yu, H.P. and Bu, J.W.: 2021, Vehicle's instantaneous velocity reconstruction by combining GNSS Doppler and Carrier Phase Measurements through Tikhonov Regularized Kernel Learning. IEEE Trans. Veh. Technol., 70, 4190–4202. DOI: 10.1109/tvt.2021.3076056

Ramillien, G., Seoane, L. and Darrozes, J.: 2021, An innovative Slepian approach to invert GRACE KBRR for localized hydrological information at the sub-basin scale. Remote Sens., 13, 16. DOI : 10.3390/rs13091824

Sasgen, I., Martinec, Z. and Fleming, K.: 2006, Wiener optimal filtering of GRACE data. Stud. Geophys. Geod., 50, 499–508. DOI: 10.1007/s11200-006-0031-y

Shen, Y., Xu, P. and Li, B.: 2012, Bias-corrected regularized solution to inverse ill-posed models. J. Geod., 86, 597–608. DOI: 10.1007/s00190-012-0542-y

Simons, F.J. and Dahlen, F.A.: 2006, Spherical Slepian functions and the polar gap in geodesy. Geophys. J. Int., 166, 1039–1061. DOI: 10.1111/j.1365-246X.2006.03065.x

Simons, F.J., Dahlen, F.A. and Wieczorek, M.A.: 2006, Spatiospectral concentration on a sphere. SIAM Rev., 48, 504–536. DOI: 10.1137/s0036144504445765

Slepian, D.: 1983, Some comments on Fourier analysis, uncertainty and modeling. SIAM Rev., 25, 379–393. DOI: 10.1137/1025078

Swenson, S. and Wahr, J.: 2006, Post-processing removal of correlated errors in GRACE data. Geophys. Res. Lett., 33, 4. DOI: 10.1029/2005gl025285

Tao, D.L., Shi, H.L., Gao, C.C., Zhan, J.G. and Ke, X.P.: 2020, Water storage monitoring in the Aral Sea and its Endorheic Basin from multisatellite data and a hydrological model. Remote Sens., 12, 23. DOI: 10.3390/rs12152408

Tapley, B.D., Bettadpur, S., Watkins, M. and Reigber, C.: 2004, The gravity recovery and climate experiment: Mission overview and early results. Geophys. Res. Lett., 31, 4. DOI: 10.1029/2004gl019920

Velicogna, I. and Wahr, J.: 2006, Acceleration of Greenland ice mass loss in spring 2004. Nature, 443, 329–331. DOI: 10.1038/nature05168

Wahr, J., Molenaar, M. and Bryan, F.: 1998, Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE. J. Geophys. Res., Solid Earth, 103, 30205–30229. DOI: 10.1029/98jb02844

Xu, P.: 1992, Determination of surface gravity anomalies using gradiometric observables. Geophys. J. Int., 110, 321–332.

Yang, T.L., Yu, H.W. and Wang, Y.: 2022, An efficient low-pass-filtering algorithm to de-noise global GRACE data. Remote Sens. Environ., 283, 15. DOI: 10.1016/j.rse.2022.113303