

Application of a novel constrained wavelet threshold denoising method in ensemble-based background-error variance

HUANG QunBo, LIU BaiNian, ZHANG WeiMin, ZHU MengBin, SUN JingZhe, CAO XiaoQun, XING Xiang, LENG HongZe and ZHAO YanLai

Citation: [SCIENCE CHINA Technological Sciences](#), 1 (2017); doi: 10.1007/s11431-016-9098-3

View online: <http://engine.scichina.com/doi/10.1007/s11431-016-9098-3>

Published by the [Science China Press](#)

Articles you may be interested in

[Tunable Q-factor wavelet transform denoising with neighboring coefficients and its application to rotating machinery fault diagnosis](#)

SCIENCE CHINA Technological Sciences **56**, 1956 (2013);

[Local thresholding with adaptive window shrinkage in the contourlet domain for image denoising](#)

SCIENCE CHINA Information Sciences **56**, 92107 (2013);

[Gearbox fault diagnosis of rolling mills using multiwavelet sliding window neighboring coefficient denoising and optimal blind deconvolution](#)

Science in China Series E-Technological Sciences **52**, 2801 (2009);

[A data-driven threshold for wavelet sliding window denoising in mechanical fault detection](#)

SCIENCE CHINA Technological Sciences **57**, 589 (2014);

[Wavelet denoising via sparse representation](#)

Science in China Series F-Information Sciences **52**, 1371 (2009);

Application of a novel constrained wavelet threshold denoising method in ensemble-based background-error variance

HUANG QunBo^{1,23*}, LIU BaiNian^{1,2}, ZHANG WeiMin^{1,2}, ZHU MengBin^{1,2}, SUN JingZhe^{1,2}, CAO XiaoQun^{1,2}, XING Xiang^{1,2}, LENG HongZe^{1,2} & ZHAO YanLai^{1,2}

¹ Academy of Ocean Science and Engineering, National University of Defense Technology, Changsha 410073, China;

² College of Computer, National University of Defense Technology, Changsha 410073, China;

³ Weather Center of Chinese People's Liberation Army Air Force, Beijing 100843, China

Received December 14, 2016; accepted July 12, 2017; published online August 18, 2017

A more efficient noise filtering technique is needed in ensemble data assimilation, to improve traditional spectral filtering methods that cannot reflect the local characteristics of spatial scales. In this paper, we present the design of a novel constrained wavelet threshold denoising method (CWTNDM) by introducing an improved threshold value and a new constraining parameter. The proposed method aims to filter noise swamped over different scales. We prepared an ideal experiment object based on the two-dimensional barotropic vorticity equation. A suitable wavelet basis function (i.e., Db11) and the optimal number of decomposition levels (i.e., five) were first selected. The results show that, given the wavelet coefficients are constrained by the parameter, the CWTNDM can produce better filtering results with the smallest root mean square error (RMSE) compared to similar methods. In addition, the filtering accuracy of 10 ensemble sample variances using the CWTNDM is equivalent to that estimated directly from 80 ensemble samples, but with the runtime reduced to approximately one-seventh. Furthermore, a large peak signal-to-noise ratio, which implies a low RMSE, suggests that the proposed method suitably preserves most of the information after denoising.

two-dimensional wavelet, threshold denoising, background-error variance, ensemble data assimilation (EDA)

Citation: Huang Q B, Liu B N, Zhang W M, et al. Application of a novel constrained wavelet threshold denoising method in ensemble-based background-error variance. Sci China Tech Sci, 2017, 60: doi: 10.1007/s11431-016-9098-3

1 Introduction

The background-error covariance plays an important role in information transmission, information smoothing, equilibrium relation, and flow structure construction [1]. An ensemble data assimilation (EDA) system integrates information from multiple samples that can reflect the distribution characteristics of the background-error covariance to determine errors more accurately [2,3]. This system provides an effective way to improve the existing assimilation meth-

ods and has been widely used for future data assimilation. However, using a static, homogeneous, and uniform background-error variance model has become the bottleneck in operational systems [4,5]. Thus, it is necessary to improve the current data assimilation framework and enhance its capabilities.

The main objective is to obtain a suitable tradeoff between the ensemble member size and the validity period of an operational EDA system [6]. First, the number of samples should be sufficient to accurately describe the structure and characteristics of the background-error covariance introduced during the assimilation process and short-term prediction. In addition, a sufficiently large number of ensemble member re-

*Corresponding author (email: hqb09@163.com)

sults in reduced random sampling noise. Then, to meet operational timeliness, the number of samples must be limited to the order of 10^2 . Next, it is necessary to optimize the perturbation techniques that represent error sources in a data assimilation system, such as background-field perturbation, observation data disturbance, and boundary perturbation.

Extending computational resources and using powerful computational techniques (e.g., FPGA or GPU) might lead to an improved system performance. However, a more convenient solution is to develop novel techniques for data processing (e.g., noise filtering) that reduce the computational cost.

Methods of small-sample noise filtering can be divided into Wiener, spectral spatial, and wavelet spatial filtering according to the filtering space. Spectral filtering uses statistical or objective methods for determining a spectrum and its inverse transform, with the noise scales approximately obtained either statistically or empirically. However, these methods cannot describe in detail the local variation of noise, and the intensity of filtering at a given scale is the same for a global area, neglecting the fact that noise and signal vary among regions.

Houtekamer and Mitchell [7,8], Buehner and Chiron [9], and Berre et al. [10] studied the pseudo-non-zero value over a long distance. Raynaud et al. [11] investigated background-error correlations, and their results show that an optimal local spatial filtering can improve the accuracy of the background-error covariance. However, the wavenumber truncation of low-pass filters is difficult to adjust because it relates to the variable, vertical level, and characteristic scale of background error. Raynaud et al. [12] proposed an improved method to automatically calculate the optimal wavenumber truncation by determining the characteristic scale length of the signal and noise. Liu et al. [13] successfully applied this method into the YH4DVAR operational system, and 10-member filter variances exhibited better performance than 30-member estimations. Nevertheless, some limitations remain on spectral filtering: 1) the background-error variance is still estimated from statistics and approximations; 2) handling the small-variance energy spectrum on small scales results arbitrary; and 3) for the same wavenumber, the filter coefficients are also the same, which cannot represent the local spatial characteristics of the signal.

Therefore, spectral filtering has limitations for certain type of applications such as weather prediction. For instance, a global numerical weather prediction (NWP) model with ultrahigh grid resolution would have the capability of multi-scale simulation and forecasting. This model would certainly include various types of weather systems with different atmospheric scales, such as planetary scale, large scale, mesoscale, and small scale. There are different spatial and temporal scales in weather systems composed of large amounts of multi-scale information. For this type of

application, the multi-scale wavelet analysis technique has the intrinsic attribute to represent the complex atmospheric motion information. For instance, a currently operational EDA system designs and implements the background-error covariance based on the spherical wavelet and can be used for multi-scale analysis [14]. Therefore, it would be convenient to apply the wavelet method for denoising in background-error variance.

The wavelet threshold denoising method is usually applied in signal processing [15–18]. Given that the noise and signal components change with the increase in scales, large and small wavelet coefficients are respectively caused by signal singularity and noise [19,20]. Therefore, threshold denoising aims to determine a suitable threshold and select an appropriate threshold function [21]. For instance, Donoho and Johnstone [22] considered the threshold as a function of the signal size and noise variance. In addition, without prior signal information, wavelet threshold denoising can reduce the filtering error defined in both the Hölder and Besov spaces. The median absolute deviation (MAD) method uses the MAD of the minimum-scale coefficient norm as the noise energy level [23]. For small-scale weather phenomena, Bonavita et al. [24] confirmed that the spectral filter overestimates the statistical significance of the sample variances in the tropics and underestimates their significance in the extra-tropics. In addition, although the wavelet filtering method can use its multi-scale and multi-resolution features to realize an inhomogeneous filter, the filter coefficients are regarded as a function of the scales without clearly defining the variability among locations and scales [24]. Therefore, the application of multi-scale wavelet threshold de-noising for NWP needs to be further investigated.

The filtering in NWP can be considered analogous to that in image and signal processing, with some developments given as follows. Chen et al. [25] proposed an improved rank-reduction method by slightly modifying the traditional truncated singular value decomposition formula, and they obtained an almost perfectly reconstruction of 3D seismic data even in the case of low signal-to-noise ratio (SNR). To solve problems caused by interference, it is possible to perform direct imaging and inversion of the blended data, thus attenuating the interference during the inversion process [26,27]. In addition, Chen and Ma [28] used the f-x predictive filtering for dithered slip-sweep data, and Chen et al. [29] proposed a general iterative deblending framework using sparsity or coherency constraints. Several other types of filtering methods have been proposed for image and signal processing [30–34].

In real-world systems such as NWP, the system can be complex, and the noise can present non-Gaussian distributions. For instance, we can focus on the combination of complex networked systems and filtering in EDA. A networked system is a collection of agents that can interact with each other, similar to EDA. Every ensemble member can be seen as a network

distributed over grid points on earth and evolving with space and time. Thus, an ensemble is a collection of many such networks. A theoretical framework to analyze the evolutionary dynamics on complex networks would be beneficial [35]. For instance, Liu et al. [36] developed a finite-time consensus criterion for a class of multi-agent systems with nonlinear dynamics. Han et al. [37] presented three kinds of optimal filters, namely optimal Kalman filter, Markov jump linear filter, and constant-gain filter, for networks with random transmission delays. Moreover, the systematic framework based on the micro-macro-intervention mechanism is useful to understand evolutionary system dynamics [38].

If we consider wavelet analysis for complex systems, the selection of optimal wavelet basis functions has been studied by many researchers. In particular, Daubechies (Db) basis functions have been found optimal for several applications such as: denoising electrocardiogram signals [39] and other types of biological signals [40], denoising electrochemical data [41], and even multi-scale analysis of gravity data [42].

Aiming to improve denoising while preserving most of the original information, we designed a novel denoising method, called the constrained wavelet threshold denoising method (CWTNDM), which can be employed to filter noise from the background-error variance in an EDA system.

2 Filtering principles

2.1 Spectral filtering

The first spectral filtering technique for denoising in EDA is objective filtering. It assumes that the background error follows a Gaussian distribution and establishes the following relation:

$$E[\mathbf{G}^e \mathbf{G}^{eT}] = \frac{1}{N-1} \tilde{\mathbf{B}}^* \otimes \tilde{\mathbf{B}}^*, \quad (1)$$

where $E[\mathbf{G}^e \mathbf{G}^{eT}]$ is the covariance matrix of sample noise \mathbf{G}^e , N is the size of the ensemble sample, $\tilde{\mathbf{B}}^*$ is the true background-error covariance, and \otimes denotes the Hadamard product. The average correlation length of the sample noise is less than that of the background error, which means that the background-error covariance signal has a larger large-scale feature than that of the noise in the original signal. Their energy spectrums start to separate from wavenumber N_{trunc} in the spectral space, which is the theoretical basis for spectral filtering in EDA. The following low-pass filter is adopted [12]:

$$\rho(n_p) = \cos^2 \left(\frac{\pi}{2} \min(n_p, N_{\text{trunc}}) / N_{\text{trunc}} \right), \quad (2)$$

where $\rho(n_p)$ represents the filter coefficient at spectral wavenumber n_p , and N_{trunc} is the truncation wavenumber of the filter. It can be seen that $\rho(n_p)$ gradually degenerates from 1 to 0 with increasing wavenumber n_p . When $n_p > N_{\text{trunc}}$,

the coefficients are set to zero to filter high-frequency components, where the noise is concentrated. In contrast, when $n_p < N_{\text{trunc}}$, $\rho(n_p)$ determines the amplitude of low-frequency noise. This low-pass filter in the spectral space is equivalent to add a weighted average to the error variance in the grid space, so that the large-scale signal is preserved and the small-scale sample noise is filtered.

2.2 Two-dimensional wavelet threshold filtering

The two-dimensional wavelet transform can be easily derived from its one-dimensional counterpart. The two-dimensional scale and wavelet functions are obtained from the tensor product of the corresponding one-dimensional functions. For the multi-scale wavelet transform, one level consists of high-frequency information having three components in the horizontal, vertical, and diagonal directions, whereas low-frequency information can be further and sequentially decomposed at a lower-scale level. Each decomposition makes the scale of the information become one-half of the upper level. The original information is separated into subparts with different scales, where the low-frequency part retains most of the original information, whereas the high-frequency part contains detailed information on boundary, regions, and other detail aspects. The basic principle of the transform can be expressed as

$$\begin{aligned} f(x, y) &= A_j f(x, y) + \sum_{k=1}^3 D_{j,k}^k f(x, y) \\ &= A_{j+1} f(x, y) + \sum_{k=1}^3 D_{j+1,k}^k f(x, y), \end{aligned} \quad (3)$$

where A is the low-frequency coefficient that represents the large-scale approximate information, D is the high-frequency coefficient that represents the small-scale detailed information, j is the scale, and k denotes the horizontal, vertical, and diagonal directions.

Taking the fifth-order wavelet transform as an example, eq. (3) can be written as

$$f(x, y) = A_5 f(x, y) + \sum_{m=1}^5 \sum_{k=1}^3 D_{m,k}^k f(x, y), \quad (4)$$

and its diagram is shown in Figure 1.

The wavelet method has the attributes of multi-scale and multi-resolution. Thus, it is able to represent the global-scale information as well as the local information; hence, inhomogeneous filtering can be realized using the wavelet transform. By transforming the signal and length-scale information into the wavelet space, the expression for inhomogeneous wavelet filtering is

$$S^{\text{filt}} = \sum_{j=1}^K \rho_j \otimes S_j^{\text{raw}}, \quad (5)$$

where S_j^{raw} is the original sampling error in EDA at the j -th wavelet component and ρ_j is the corresponding filter coefficient.

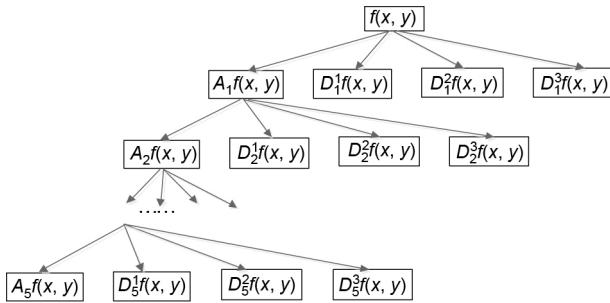


Figure 1 Diagram of a fifth-order wavelet decomposition transform.

The steps for a general wavelet threshold denoising of two-dimensional signals are shown in [Figure 2](#) and described as follows:

Step 1. Multilevel wavelet decomposition. A suitable wavelet basis function and decomposition level L are selected, and the two-dimensional signal is decomposed until the L -th level.

Step 2. Threshold quantization for the decomposed high-frequency coefficients. An appropriate threshold is selected for every decomposition level. The threshold determines the components to be filtered.

Step 3. Reconstruct the two-dimensional signal. The modified signals are reconstructed by means of the low-frequency coefficient at the L -th level and the high-frequency coefficients at every level, which are determined by the thresholds from Step 2.

2.3 Performance metrics

In the image compression field, the peak SNR (PSNR) is an objective measurement method to evaluate the quality of an image after compression [\[43\]](#). It describes the relationship between the signal and background noise, given as follows:

$$\text{PSNR} = 10 \cdot \log_{10} \left(\frac{(2^n - 1)^2}{MSE} \right), \quad (6)$$

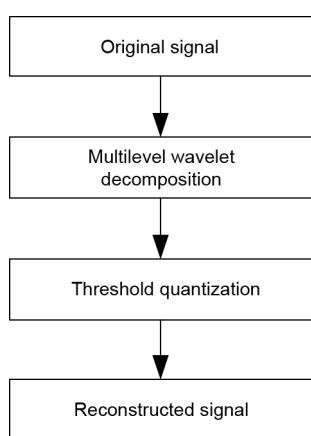


Figure 2 Process for two-dimensional wavelet threshold denoising.

where MSE is the mean square error between the original and compressed images, n is the number of bits per sample, and $(2^n - 1)$ is the number of gray levels, which is usually 255 (i.e., 8-bit images). High PSNR values indicate an increased similarity between the compressed and original images, which represent a smaller distortion after compression.

For our proposed method, we used the PSNR to represent the amount of information preserved after denoising, whereas the RMSE was used to measure the divergence between the original and filtered information.

3 CWTNDM

We use the wavelet threshold method mainly to adapt the filtering to local variations of the input signal and estimate the true signal. For an adaptive filter based on the wavelet threshold denoising method, coefficients $\widetilde{X}_{i,j}$ above threshold T , $|\widetilde{X}_{i,j}| > T$, which indicate sharp changes caused by small scales, are preserved. In contrast, the coefficients $|\widetilde{X}_{i,j}| \leq T$, which represents the relatively flat change of the function, are set to zero and thus filtered.

However, for the background-error variance in EDA, noise usually presents a spatial and scale correlation; hence, it cannot be treated as white Gaussian noise. To filter correlated noise, Nguyen van Yen et al. [\[44\]](#) proposed a method that sets different thresholds on different scales in the wavelet space

$$T_N(j) = \sigma(j) \sqrt{2 \ln(n_j)}, \quad (7)$$

where $\sigma(j)$ and $n_j = 2^j$ are the standard deviation and number of wavelet coefficients on scale j , respectively. However, since the n_j corresponding to the individual scale is too small, the statistical $\sigma(j)$ has a large error. Hence, a global threshold algorithm was developed to manually adjust the thresholds as follows:

$$T_P = \beta \times \sigma_w \sqrt{2 \ln(n)}, \quad \beta \geq 1, \quad (8)$$

where σ_w is the average standard deviation of the correlated noise and n represents the signal size. Thus, by changing parameter β , some large noise coefficients that satisfy $\sigma(j) > \sigma_w$ can also be set to zero to determine the new threshold, T_P .

When $\beta=1$, [eq. \(8\)](#) represents the method proposed by Azzalini et al. [\[45\]](#), and the threshold can be expressed as

$$T_A = \sigma_w \sqrt{2 \ln(n)}. \quad (9)$$

Pannekoucke et al. [\[46\]](#) verified the effectiveness of the relation described in [eq. \(8\)](#), but parameter β is difficult to estimate because noise is non-Gaussian and its theoretical modeling and derivation are complex.

Moreover, two situations should be considered when adjusting threshold T_P . First, it must be large enough to filter

the noise with high energy levels. Second, it cannot be infinite to ensure the useful signal is not affected by the threshold. As shown in Figure 3, the smallest circle (black dashed line) represents white Gaussian noise, whereas the ellipse indicates non-Gaussian noise with spatial and scale correlation. When some noise energy levels are larger than the mean energy level on a given scale, the ellipse exceeds the boundaries of the smallest circle. There could be two extreme cases. If $\sigma = \sigma_w$, the sampling noise in the background-error variance is considered as white Gaussian noise, and the calculated threshold is weak because most of the noise is not filtered. On the other hand, if $\sigma = \max(\sigma(j))$, represented by the black solid line in Figure 3, the filter affects the useful signal, hence resulting in distortion.

To approach these scenarios, noise \tilde{X}^e can be divided into two terms, expressed as

$$\tilde{X}^e = \tilde{X}^G + \tilde{X}^{NG}, \quad (10)$$

where \tilde{X}^G is the Gaussian and \tilde{X}^{NG} is the non-Gaussian noise. In the wavelet space, the number of wavelet coefficients of non-Gaussian noise \tilde{X}^{NG} is much smaller than that of Gaussian noise \tilde{X}^G . Therefore, the statistical characteristics (e.g., standard deviation $\sigma(\tilde{X}^G)$) of \tilde{X}^G can be considered as an approximation for noise \tilde{X}^e . The initial threshold, T_A , is obtained by the Azzalini method [45]. The truncated remainder term, $\tilde{R}_{i,j} \in \|X_{i,j}\| \leq T_A$, contains a small part of the wavelet coefficients, $\|X_{i,j}^*\| \leq T_A$, from the signal, a part of the wavelet coefficients, $\|\tilde{X}^{NG}\| \leq T_A$, from the non-Gaussian noise, and a part of the wavelet coefficients, $\|\tilde{X}^G\| \leq T_A$, from the Gaussian noise. Then, standard deviation σ_w of the truncated remainder, $\tilde{R}_{i,j}$, is used as the initial value of σ_w . Therefore, the part of wavelet coefficients, $\|\tilde{X}^{NG}\| > T_A$, from the non-Gaussian noise, located in the interval $[-2\sigma_w\sqrt{2\ln(n)}, 2\sigma_w\sqrt{2\ln(n)}]$, is considered as an event

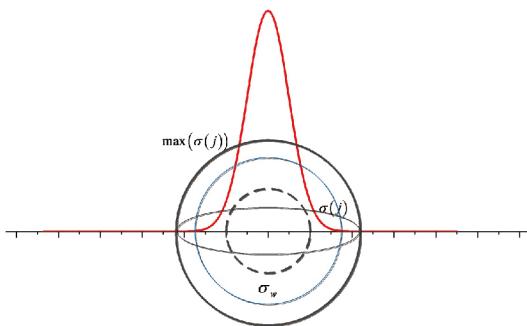


Figure 3 White Gaussian noise (black dashed line), correlated noise (black ellipse), maximum noise (black solid line), filtering noise (blue line), and probability density function of noise (red line).

with small probability of occurrence that follows the truncated remainder distribution.

To preserve the most valuable information, we improve the threshold calculation as follows:

$$T_s = \max(2\sigma_w\sqrt{2\ln(n)}, T_A), \quad (11)$$

$$\sigma_w = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_{i,j} - \bar{X}_{i,j})^2}, \text{ for } \|X_{i,j}\| \leq T_A,$$

where $\bar{X}_{i,j}$ is the expectation of $X_{i,j}$. From eq. (11), 95% of the noise can fall into the interval $[-2\sigma_w\sqrt{2\ln(n)}, 2\sigma_w\sqrt{2\ln(n)}]$. Hence, it considers non-Gaussian noise as much as possible but avoids values of T_s that can cause signal distortion by filtering many wavelet coefficients.

Based on the development described above, we designed a constraining parameter α to reduce noise on some relatively large scales and thus obtain the proposed CWTNDM. The algorithm for our method is described as follows:

Step 1. The original input signal, X , is transformed into the wavelet space, \tilde{X} , and then it is divided into two parts: noise \tilde{X}^e and signal \tilde{X}^* . Initially, the complete signal is regarded as noise (i.e., $\tilde{X}^e = \tilde{X}$).

Step 2. The noise variance and threshold are calculated as $\sigma_{W,k}^2 = \frac{1}{n} \sum_{i,j} |\tilde{X}_{i,j}^e|^2$ and $T_k = \sigma_{W,k}\sqrt{2\ln(n)}$, respectively, at the k -th iteration and $\tilde{X}_{i,j}^e$ denotes the i -th wavelet coefficient of noise on scale j .

Step 3. The wavelet threshold truncation is performed by

$$\tilde{X}_{i,j}^* = \rho_T(\tilde{X}_{i,j}^e) = \begin{cases} \tilde{X}_{i,j}^e, & \text{if } |\tilde{X}_{i,j}^e| > T \\ 0, & \text{if } |\tilde{X}_{i,j}^e| \leq T \end{cases}, \text{ and the new noise signal}$$

is calculated as $\tilde{X}_{i,j}^e = (1 - \rho_T)(\tilde{X}_{i,j}^e)$.

Step 4. Steps 2 and 3 are repeated until the following condition is satisfied: $|\sigma_{W,k}^2 - \sigma_{W,k+1}^2| \leq \Delta$, where Δ is the predefined threshold precision (e.g., $\Delta = 10^{-6}$).

Step 5. The updated noise variance and threshold are obtained as $\sigma_w^2 = \sum_{i,j} |\tilde{X}_{i,j}^e|^2$ and $T = \sigma_w\sqrt{2\ln(n)}$, respectively.

Step 6. The new wavelet coefficients, calculated from the above steps, are constrained by parameter α , and the new threshold function is

$$\tilde{X}_{i,j}^* = \begin{cases} \tilde{X}_{i,j} - \alpha \frac{T}{\tilde{X}_{i,j}}, & \text{if } |\tilde{X}_{i,j}| \geq T, \\ 0, & \text{if } |\tilde{X}_{i,j}| < T. \end{cases}$$

The optimal α can be iteratively obtained, and it is used for constraining the deviation of the wavelet coefficients before and after treatment. Finally, the filtered signal is represented by $X_{\text{wave}} = \sum_{i,j} \tilde{X}_{i,j}^* \psi_{i,j}$.

Considering step 3 in the algorithm, note that the wavelet coefficients with small amplitude are also set to zero, and the

wavelet coefficients with large amplitude are either not preserved or reduced according to the threshold, and gradually approach $\widetilde{X}_{i,j}$. The deviation between $\widetilde{X}_{i,j}$ and $\widetilde{X}_{i,j}^*$ reduces with the increase of $|\widetilde{X}_{i,j}|$. In addition, this dynamically reduces the amplitude attenuation of the wavelet coefficients, and thus retains more detail information obtaining an improved denoised signal.

Note that the CWTNDM degrades to a hard-threshold wavelet denoising method (denoted as Method-A below) when $\alpha=0$.

4 Experiments and discussion

An ideal experiment object was implemented based on the two-dimensional barotropic vorticity equation that represents some physical phenomena of the atmosphere.

4.1 Experimental setup

If the flow is purely horizontal, the relationship between velocity v and horizontal stream function ψ is

$$v = k \times \nabla \psi. \quad (12)$$

The stream function solution is given by

$$\nabla^2 \psi = \zeta = k \cdot \nabla \times v, \quad (13)$$

where ζ is the absolute horizontal vorticity. Then, we obtain potential vorticity η as follows:

$$\frac{d\eta}{dt} = \frac{d(\zeta + f)}{dt} = 0, \quad (14)$$

which indicates that absolute vorticity is conserved following the horizontal motion [47], where $f = 2Q\sin\theta$ and represents the Coriolis term. Eq. (14) can be written in the advective form

$$\frac{d\eta}{dt} = \frac{\partial \eta}{\partial t} + v \cdot \nabla \eta. \quad (15)$$

By integrating eq. (15) using the semi-Lagrangian transport time integration and the particle tracking algorithm we can determine the departure points. This relation follows the earth system modeling framework specification of a gridded component [48,49].

We selected a two-dimensional region with a resolution of 128×64 pixels, and a Rossby-Haurwitz fourth wave was used as the initial state. The integration length was 24 h with 2-h time steps. The unperturbed integral state is considered as the true state, as shown in Figure 4(a). Then, by adding random non-Gaussian perturbation on 10 ensemble samples, their ensemble average among the background-noise variance was taken as the filtering object. Clearly, the ensemble sample variance contains substantial small-scale noise with nonnegligible geographical variations, as shown in Figure 4(b), seriously affecting its quality. A signal like this will show several

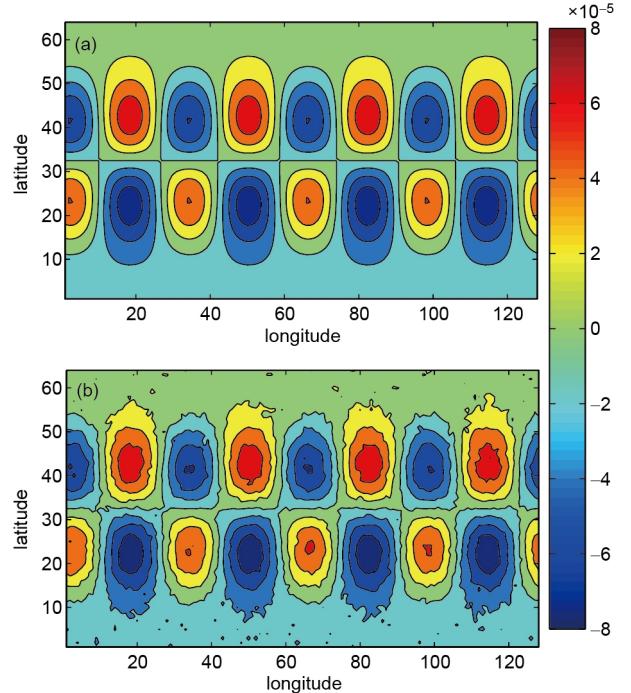


Figure 4 (Color online) True (a) and ensemble (b) average of 10 samples used for the filtering experiment.

negative effects if it is applied directly to a deterministic prediction model.

The results generated from the spectral filtering method, general two-dimensional wavelet threshold filtering method (denoted as Method-0), Method-A, and CWTNDM are denoted by Exp_spec, Exp_0, Exp_A, and Exp_CW, respectively.

4.2 Decomposition levels and wavelet basis function

For the wavelet analysis, it is necessary to first determine a suitable wavelet basis function and the number of decomposition levels. Different wavelet basis functions have different time-frequency attributes to describe the characteristics of a signal. For the analysis, the following factors should be considered: 1) a wavelet filter should have linear phase and preferably generalized linear phase to prevent distortion; 2) the filter should provide localized scales and reduce the computational complexity; 3) orthogonality among the subspaces is propitious for multi-scale decomposition and decreases the dependency among the sub-band coefficients; and 4) the number of decomposition levels is usually between 2 and 5. For specific applications, several tests are usually required and the tradeoff between the computational cost and final denoising effect should be considered.

Based on the Db11 wavelet basis function and 3 to 7 decomposition levels, we compared the RMSE values of the background-error variances obtained by the three wavelet methods described above. Table 1 shows that the optimal decomposition level is 5, because the RMSE of the CWTNDM is the

Table 1 RMSE obtained by Method-0, Method-A, and CWTDNM for different decomposition levels (DL)

DL	3	4	5	6	7
Exp_0	1.02×10^{-6}	8.93×10^{-7}	9.26×10^{-7}	1.01×10^{-6}	9.25×10^{-7}
Exp_A	8.61×10^{-7}	8.70×10^{-7}	8.55×10^{-7}	9.00×10^{-7}	9.27×10^{-7}
Exp_CW	8.55×10^{-7}	8.69×10^{-7}	8.42×10^{-7}	8.91×10^{-7}	9.22×10^{-7}

smallest, thus achieving a better estimation among the other levels. Other basis functions, such as Coif5, Sym8, and Bior6.8, showed similar trends, but they were not used for the analysis as explained next.

The results for different wavelet basis function are listed in **Table 2**. It is clear that by using Db11, the smallest RMSE is obtained compared to the other basis functions.

Figure 5 shows the grid distribution of the relative errors for the background-error variance obtained from the tested methods. It is clear that the filtering effects of the three wavelet methods are better than those of spectral filtering, with the CWTDNM being the best, as shown in **Figure 5(c)**. Even if the same large-scale features are present after spectral filtering, there is also a large amount of small-scale noise. Thus, most of the smaller-scale features that are indistinguishable from background error have been preserved, as shown in **Figure 5(d)**. The small-scale noise is usually concentrated in meteorologically active areas, which disturbs the forecasting operation and leads to unpredictability.

It is worth noting that the decomposition and reconstruction of the initial signal using the wavelet transform inevitably results in reconstruction errors. This is measured by the relative error that represents the deviation between the signal after denoising and the true values. The results show that the order of the reconstruction error is 10^{-18} , which is much smaller than

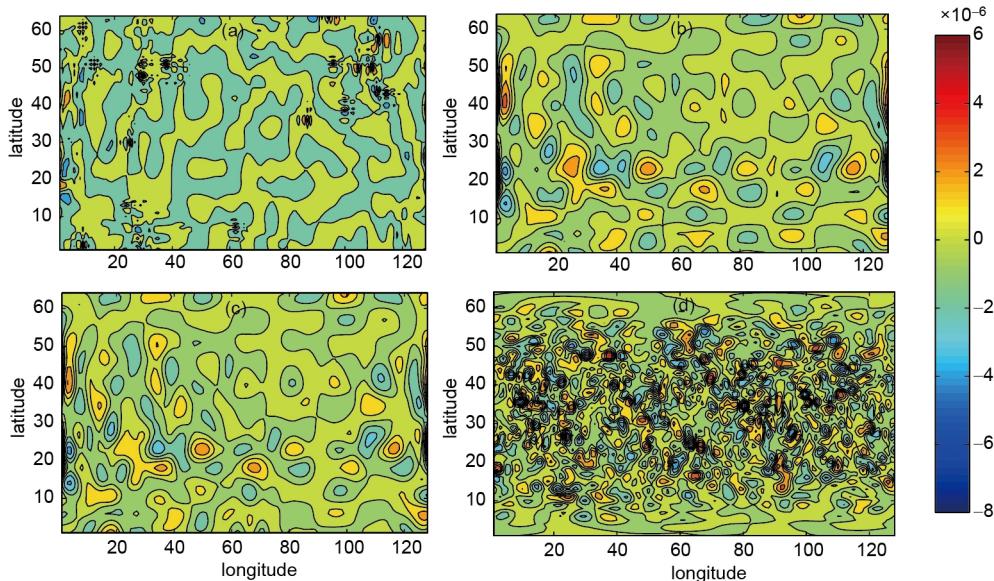
Table 2 RMSE obtained by Method-0, Method-A, and CWTDNM for different wavelet basis functions (BF)

BF	Bior6.8	Coif5	Sym8	Db11
Exp_0	1.03×10^{-6}	9.49×10^{-7}	1.01×10^{-6}	9.26×10^{-7}
Exp_A	1.14×10^{-6}	8.96×10^{-7}	1.07×10^{-6}	8.55×10^{-7}
Exp_CW	1.14×10^{-6}	8.77×10^{-7}	1.07×10^{-6}	8.42×10^{-7}

that of the relative error in 10^{-6} ; hence, it can be ignored.

4.3 Results

Figure 6 shows the filtering results for the vorticity background-error variance. In contrast to the true values and the filtering object in **Figure 4**, it can be seen that, although the spectral filtering is able to filter random noise, the overall structure of variance is influenced after denoising (**Figure 6(d)**). Compared to the other tested methods, the CWTDNM with the optimally constrained parameter $\alpha = 0.21$ shows the best outcome, exhibiting an improved smoothing and accurate structure of background-error variance by retaining more small-scale detail information, being similar to the truth, as shown in **Figure 6(c)**. This verifies that the CWTDNM is an effective tool to filter a large amount of small-scale random noise without affecting the useful variance signal.

**Figure 5** (Color online) Grid distribution of relative errors for background-error variance obtained using Method-0 (a), Method-A (b), CWTDNM (c), and spectral filtering method (d).

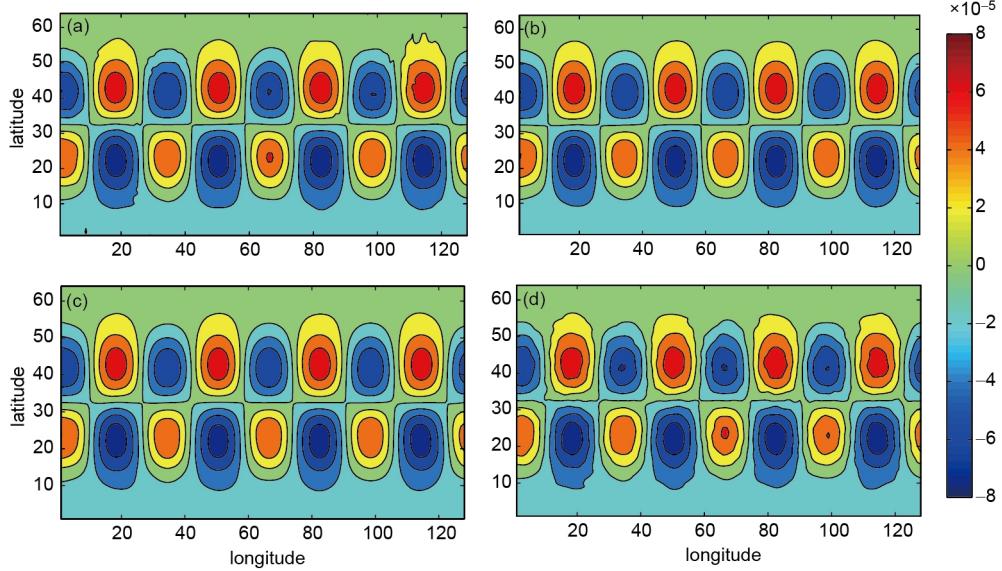


Figure 6 (Color online) Filtering results of the vorticity background-error variance in experiments Exp_0 (a), Exp_A (b), Exp_CW (c), and Exp_spec (d).

Figure 7 shows the RMSE and PSNR trends for different values of α in Exp_CW. At the eighth iteration, the smallest RMSE and corresponding largest PSNR (i.e., PSNR=108.88 dB) are obtained. This PSNR represents that most of the useful signal is preserved after denoising, indicating the suitable denoising ability for the original filtering object.

Moreover, we also investigated the efficiency of the CWTNDM by comparing its outcome with the results of

10, 30, 50, 80, and 100 ensemble samples. We found that the accuracy of the background-error variance for the CWTNDM is comparable to that obtained from 80 ensemble samples, as shown in Table 3. However, the CWTNDM runtime is approximately one-seventh of that taken for the ensemble-sample computation, as shown in Figure 8.

Our proposed filtering method is essentially a postprocessing technique, which is different from the direct error statistics obtained by the ensemble average with a large sample

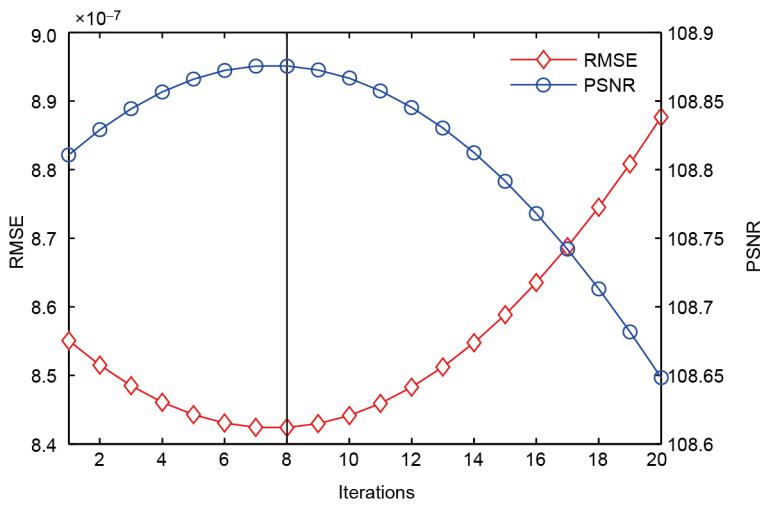


Figure 7 (Color online) RMSE and PSNR for different α values in Exp_CW.

Table 3 RMSE of the background-error variance obtained by experiments with different numbers of ensemble samples (NS)

NS	10	30	50	80	100	CWTNDM
RMSE	3.06×10^{-5}	1.37×10^{-6}	1.07×10^{-6}	8.64×10^{-7}	7.71×10^{-7}	8.42×10^{-7}

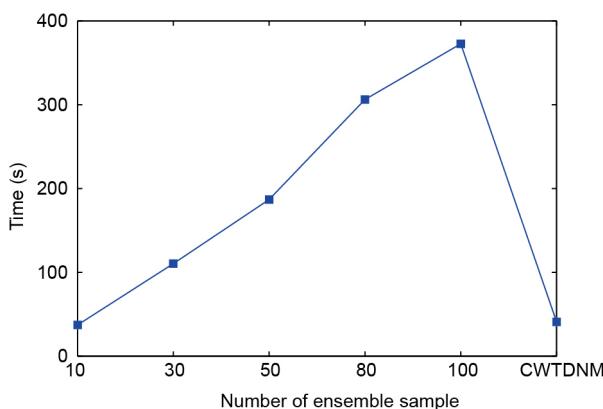


Figure 8 (Color online) Runtime of the experimental system with different numbers of ensemble sample and that of the CWTNDM.

size. In practice, an operational EDA system generally requires from a few to several hours to run. Hence, it is unfeasible to implement real-time statistics to the background error by using a large sample size because of limitations in computational sources. On the other hand, the CWTNDM only requires a few minutes to perform denoising, which greatly improves the operational efficiency.

Finally, we compare the CWTNDM with the traditional spectral filtering method. The latter can be optimized by adjusting the truncation wavenumbers, and the corresponding results are shown in Figure 9. They indicate that spectral filtering achieves its best result when the truncation wavenumber is 50. The corresponding RMSE is 8.50×10^{-7} , which is still larger than that of the CWTNDM (8.42×10^{-7}). Overall, our results suggest that the CWTNDM is better than the spectral filtering method in denoising background error-variance by preserving more useful multi-scale information.

In summary, the correctness and effectiveness of the CWTNDM were verified in an ideal two-dimensional scenario. The spectral filter has a large distortion on the structure of the background-error variance. After denoising by CWT-

DNM, a large amount of small-scale information is filtered. However, most information is preserved by the proposed enhanced filtering. The CWTNDM can be helpful to improve operational timeliness and effectiveness. Moreover, this novel method can help in the analysis of multi-scale noise information of the background-error variance in the field of NWP.

5 Conclusions

The results of this study suggest that the proposed CWTNDM successfully filters non-Gaussian noise. In addition, it can quickly and accurately filter the noise of the background-error variance in an EDA system.

We tested the method on a two-dimensional ideal experiment object, aiming at filtering the noise in the background-error variance produced by an ensemble model with 10 samples. The results show that the filtering effect of the CWTNDM is better than that of spectral filtering and other two wavelet filtering methods. With the constraint effect of the introduced optimal parameter, our method can filter the noise on larger scales while preserving useful original information. The accuracy of the filtering results is similar to that using the average of 80 ensemble samples, but with a smoother spatial structure and less time cost. In addition, a large PSNR suggests that the proposed method preserves almost all the original information.

Considering the influence of multi-scale noise on the background-error variance information, the CWTNDM shows potential to be applied on an operational data assimilation system. Thus, we aim to investigate the influence of our method in a real operational system, as well as the denoising effect based on the spherical wavelet framework and the application of the method in image processing. We expect that this method will provide useful insights to improve operational predictions and the quality of the analyses. Moreover, we aim to adapt this denoising approach for other types of networked systems.

This work was supported by the National Natural Science Foundation of China (Grant Nos. 41375113, 41475094, 41305101 & 41605070). We would like to thank Editage [www.editage.cn] for English language editing.

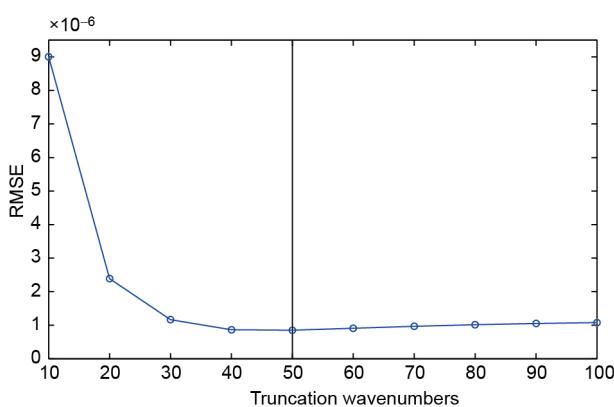


Figure 9 (Color online) RMSE for different truncation wavenumbers using the spectral filtering method.

- 1 Bannister R N. A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances. *QJR Meteorol Soc*, 2008, 134: 1951–1970
- 2 Bonavita M, Raynaud L, Isaksen L. Estimating background-error variances with the ECMWF ensemble of data assimilations system: Some effects of ensemble size and day-to-day variability. *QJR Meteorol Soc*, 2011, 137: 423–434
- 3 Huang Q B, Cao X Q, Zhu M B, et al. New data assimilation system DNDAS for high-dimensional models. *Chin Phys B*, 2016, 25: 0505021
- 4 Laroche S, Gauthier P. A validation of the incremental formulation of

- 4D variational data assimilation in a nonlinear barotropic flow. *Tellus A-Dynamic Meteor Oceanography*, 1998, 50: 557–572
- 5 Derber J, Bouttier F. A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus A-Dynamic Meteor Oceanography*, 1999, 51: 195–221
- 6 Pereira M B, Berre L. The use of an ensemble approach to study the background error covariances in a global NWP model. *Mon Wea Rev*, 2006, 134: 2466–2489
- 7 Houtekamer P L, Mitchell H L. A sequential ensemble kalman filter for atmospheric data assimilation. *Mon Wea Rev*, 2001, 129: 123–137
- 8 Houtekamer P L, Mitchell H L. Data assimilation using an ensemble Kalman filter technique. *Mon Wea Rev*, 1998, 126: 796–811
- 9 Buehner M, Charbonneau M. Spectral and spatial localization of background-error correlations for data assimilation. *QJR Meteorol Soc*, 2007, 133: 615–630
- 10 Berre L, Pannekoucke O, Desroziers G, et al. A variational assimilation ensemble and the spatial filtering of its error covariances: Increase of sample size by local spatial averaging. In: Proceedings of ECMWF Workshop on Flow-dependent Aspects of Data Assimilation. Reading, England, 2007. 151–168
- 11 Raynaud L, Berre L, Desroziers G. Spatial averaging of ensemble-based background-error variances. *QJR Meteorol Soc*, 2008, 134: 1003–1014
- 12 Raynaud L, Berre L, Desroziers G. Objective filtering of ensemble-based background-error variances. *QJR Meteorol Soc*, 2009, 135: 1177–1199
- 13 Liu B N, Zhang W M, Cao X Q, et al. Investigations and experiments of variances filtering technology in the ensemble data assimilation. *Chin J Geophys*, 2016, 59: 33–42
- 14 Cao X Q, Zhang W M, Song J Q, et al. Design of background error covariance's model with spherical wavelet. *Comput Eng Appl*, 2014, 50: 49–55
- 15 Chen Y M, Zi Y Y, Cao H R, et al. A data-driven threshold for wavelet sliding window denoising in mechanical fault detection. *Sci China Tech Sci*, 2014, 57: 589–597
- 16 Ma H W, Fan H W, Mao Q H, et al. Noise reduction of steel cord conveyor belt defect electromagnetic signal by combined use of improved wavelet and EMD. *Algorithms*, 2016, 9: 62
- 17 Xu J, Wang Z, Tan C, et al. Adaptive wavelet threshold denoising method for machinery sound based on improved fruit fly optimization algorithm. *Appl Sci*, 2016, 6: 199
- 18 Chou C M. Application of set pair analysis-based similarity forecast model and wavelet denoising for runoff forecasting. *Water*, 2014, 6: 912–928
- 19 Vidakovic B, Lozoya C B. On time-dependent wavelet denoising. *IEEE Trans Signal Process*, 1998, 46: 2549–2554
- 20 Krim H, Tucker D, Mallat S, et al. On denoising and best signal representation. *IEEE Trans Inform Theor*, 1999, 45: 2225–2238
- 21 Yang T. The application research of signal and image processing using wavelet multi-scaling analysis. Dissertation for the Doctoral Degree. Guangzhou: Sun Yat-sen University, 2005
- 22 Donoho D L, Johnstone J M. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 1994, 81: 425–455
- 23 Mallat S, Papanicolaou G, Zhang Z. Adaptive covariance estimation of locally stationary processes. *Ann Stat*, 1998, 26: 1–47
- 24 Bonavita M, Isaksen L, Hólm E. On the use of EDA background error variances in the ECMWF 4D-Var. *QJR Meteorol Soc*, 2012, 138: 1540–1559
- 25 Chen Y, Huang W, Zhang D, et al. An open-source Matlab code package for improved rank-reduction 3D seismic data denoising and reconstruction. *Comp Geosci*, 2016, 95: 59–66
- 26 Chen W, Yuan J, Chen Y K, et al. Preparing the initial model for iterative deblending by median filtering. *J Seism Explor*, 2017, 26: 25–47
- 27 Chen Y, Yuan J, Zu S, et al. Seismic imaging of simultaneous-source data using constrained least-squares reverse time migration. *J Appl Geophys*, 2015, 114: 32–35
- 28 Chen Y, Ma J. Random noise attenuation by f-x empirical-mode decomposition predictive filtering. *Geophysics*, 2014, 79: V81–V91
- 29 Chen Y, Fomel S, Hu J. Iterative deblending of simultaneous-source seismic data using seislet-domain shaping regularization. *Geophysics*, 2014, 79: V179–V189
- 30 Chen W, Chen Y K, Liu W. Ground roll attenuation using improved complete ensemble empirical mode decomposition. *J Seism Explor*, 2016, 25: 485–495
- 31 Chen Y, Fomel S. Random noise attenuation using local signal-and-noise orthogonalization. *Geophysics*, 2015, 80: WD1–WD9
- 32 Chen Y. Dip-separated structural filtering using seislet transform and adaptive empirical mode decomposition based dip filter. *Geophys J Int*, 2016, 206: 457–469
- 33 Chen Y. Deblending using a space-varying median filter. *Explor Geophys*, 2015, 46: 332–341
- 34 Chen W, Xie J, Zu S, et al. Multiple-reflection noise attenuation using adaptive randomized-order empirical mode decomposition. *IEEE Geosci Remote Sens Lett*, 2017, 14: 18–22
- 35 Tan S, Lü J, Lin Z. Emerging behavioral consensus of evolutionary dynamics on complex networks. *SIAM J Control Optim*, 2016, 54: 3258–3272
- 36 Liu K X, Wu L L, Lü J H, et al. Finite-time adaptive consensus of a class of multi-agent systems. *Sci China Tech Sci*, 2016, 59: 22–32
- 37 Han C, Zhang H, Fu M. Optimal filtering for networked systems with Markovian communication delays. *Automatica*, 2013, 49: 3097–3104
- 38 Tan S, Wang Y, Lu J. Analysis and control of networked game dynamics via a microscopic deterministic approach. *IEEE Trans Automat Contr*, 2016, 61: 4118–4124
- 39 Singh B N, Tiwari A K. Optimal selection of wavelet basis function applied to ECG signal denoising. *Digital Signal Process*, 2006, 16: 275–287
- 40 Rafiee J, Rafiee M A, Prause N, et al. Wavelet basis functions in biomedical signal processing. *Expert Syst Appl*, 2011, 38: 6190–6201
- 41 Darowicki K, Zieliński A. Optimal wavelet choice in electrochemical experiments. *Fluct Noise Lett*, 2006, 06: L215–L225
- 42 Mou L, Chen Z X. The optimal choice of wavelet bases in gravity data multi-scale analysis. *Geophys Geochem Explor*, 2015, 39: 1013–1019
- 43 Huynh-Thu Q, Ghanbari M. Scope of validity of PSNR in image/video quality assessment. *Electron Lett*, 2008, 44: 800–801
- 44 Nguyen van Yen R, Farge M, Schneider K. Scale-wise coherent vorticity extraction for conditional statistical modeling of homogeneous isotropic two-dimensional turbulence. *Phys D-Nonlinear Phenom*, 2012, 241: 186–201
- 45 Azzalini A, Farge M, Schneider K. Nonlinear wavelet thresholding: A recursive method to determine the optimal denoising threshold. *Appl Comp Harmonic Anal*, 2005, 18: 177–185
- 46 Pannekoucke O, Raynaud L, Farge M. A wavelet-based filtering of ensemble background-error variances. *QJR Meteorol Soc*, 2014, 140: 316–327
- 47 Holton J R. An Introduction to Dynamic Meteorology. San Diego: Elsevier Academic Press, 2004. 108
- 48 Collins N, Theurich G, DeLuca C, et al. Design and implementation of components in the earth system modeling framework. *Int J High Perform Comput Appl*, 2005, 19: 341–350
- 49 Drake J B, Worley P, D'Azevedo E. Spherical harmonic transform algorithms. *ACM Trans Math Softw*, 2010, 35: 995–999