

INVESTIGATIONS AND EXPERIMENTS OF VARIANCES FILTERING TECHNOLOGY IN THE ENSEMBLE DATA ASSIMILATION

LIU Bai-Nian, ZHANG Wei-Min, CAO Xiao-Qun, ZHAO Yan-Lai, HUANG Qun-Bo, LUO Yu

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

Abstract An experimental Ensemble Data Assimilation (YH-EDA) system has been built with 10 ensemble members based on the operational YH4DVAR system. The system can provide flow-dependent background-error variances, which are superior to the operational ones both in structure and magnitude. However, the finite ensemble size implies a detrimental sampling noise for the variances estimation. To solve this problem, a spectral filtering technique is implemented to formulate a low-passing filter. Taking into account the typical horizontal length scales of noise and signal, the filter can eliminate the sampling noise while extracting the signal of interest. In the ensemble variance filtering experimentations of the 9th typhoon “Jebi” in 2013, our results show that the 10-members’ filtered variances exhibit better performance than 30-members’s estimation. The successful implementation of the spectral filtering reduces the requirement of large ensemble size of Ensemble Data Assimilation system, which indicates that spectral filtering has become an important and necessary technology in EDA operational implementation.

Key words Ensemble Data Assimilation; Spectral filtering; Truncated wavenumber; Random sampling noise

1 INTRODUCTION

The aim of data assimilation (DA) is to combine the observations with background information to provide the best estimate of the atmosphere states for numeric weather prediction model at the initial time (Lin, 1999; Yue et al., 2010; Qian, 2012). In this process, observation error covariance matrix \mathbf{R} and background error covariance matrix \mathbf{B} should be well specified as they determine the weights of these two sources of information in the analysis. However, limited information and the huge dimensions of \mathbf{B} leave the accurate definition of \mathbf{B} one of the main challenges in data assimilation. In order to solve this problem, several reasonable approximations have been made, such as stationarity, homogeneity and isotropy (Mitchell et al., 1990; Courtier et al., 1998; Berre, 2000; Weaver and Courtier, 2001). Such simple \mathbf{B} makes the implementation of data assimilation more convenient at the expense of its universality and flow dependence, especially for some high impact weather events (Laroche et al., 1998; Derber and Bouttier, 1999). For this reason, many researches attempt to get a flow dependent \mathbf{B} model (Evensen, 1994; Houtekamer and Mitchell, 1998; Buizza et al., 2005) through different methods, e.g. NMC (Parrish and Derber et al., 1992), Monte Carlo (Fisher and Andersson, 2001), ensemble data assimilation (EDA) (Houtekamer et al., 1996; Fisher., 2003; Kalnay., 2003; Pereira and Berre., 2006; Isaksen et al., 2006; Raynaud et al., 2009; Bonavita et al., 2011), etc. EDA follows the idea of ensemble Kalman filter. Based on an ensemble of perturbed data assimilations, EDA can provide estimates of day-to-day background error statistics for operational determined DA system. By now, EDA scheme has been adopted by ECMWF (Bonavita et al., 2011) and Mete-France (Berre and Desroziers, 2010) to update the quasi-climatological \mathbf{B} model.

A critical aspect of the EDA system is the choice of a reasonable ensemble size. EDA ensemble is obtained through random perturbation of the observations, the boundary conditions and the background fields. Since EDA is a pure stochastic method, the accuracy of error estimates sampled from the ensemble is proportional to the square root of the ensemble size (Pereira and Berre, 2006). According to this relationship, the size of ensemble could be scaled-up to be large enough to get desired accuracy, but it also means an increasing

computational cost at the same time. Although some acceleration technologies such as FPGA, GPU can be utilized to reduce the elapsed time and can introduce a larger ensemble size, the low rate of convergence may causes poor efficiency. A suitable approach need to be proposed to design a nearly cost-free ad-hoc filtering technique which can preserve interested signals of the raw ensemble estimates and filter out useless sampling noise.

For this purpose, Raynaud (2008) proposed a local spatial averaging filter according to the characteristic length-scales of noise and signal respectively. Such low-pass filter can get rid of the small length-scale noise by local spatial averaging. Some experiment results show that 10-members' filtered background-error variances can outperform the 50-members' estimation. But an obvious limitation of this filtering scheme is that the optimal smooth regions have to be set for each variable, each level and each analysis time manually. Thus, it is not suitable for automatic operational implementation. On the basis of this pioneer work, an objective spectral filter of ensemble variances based on an evaluation of the noise-to-signal ratio was invented by Raynaud in 2009 (Raynaud et al., 2009), and the spatial correlation length-scale of the sampling noise is calculated automatically following the relationship between the length-scale of the noise and \mathbf{B} . As the noise energy mainly locates at the large wavenumber or small length-scale space, the truncation wavenumber is determined according to the energy separation of noise and signal in spectral space. Then an experiential low-pass filter, whose coefficients are simple cosine function of the wavenumber, is applied to raw estimate to enable the large-scale signal of interest to be extracted as Wiener's method (Wiener, 1949), and to filter out the small-scale sampling noise effectively. Several researches showed its attractive excellences such as low-cost, reduced requirement of the ensemble size, and improved accuracy of the error statistical estimates (Bonavita et al., 2011, 2012; Raynaud et al., 2008, 2009).

YH4DVAR is a four-dimensional variational data assimilation system and still relies on a static climatological background error covariance matrix, whose goal in the near future is to implement Ensemble Data Assimilation (EDA). It motivates the paper to look deeply into the related theories and efficiency of spectral filtering method in real operational context. The paper is organized as follows. Section 2 reviews the technique of the spectral filtering method. An experimental Ensemble Data Assimilation system, containing 10 members, is then built through perturbing three main uncertain sources of YH4DVAR's background error in Section 3. In Section 4, ensemble-based background-error variances for 9th severe tropical typhoon in 2013 is examined. The flow dependence of the ensemble-based background-error variances and its impact on the filtering algorithm are also presented in this part. Conclusions are given in Section 5.

2 THEORETICAL ASPECT OF FILTER

The uncertainties of data assimilation system are represented by perturbing observations, sea surface temperature and model physical tendencies. The differences between several independent perturbed forecasts can be used to estimate the background error covariance (Parrish and Derber, 1992; Fisher, 2003; Isaksen et al., 2006). The differences between estimates and its true values may be caused by the finite number of statistical samples, perturbation methods and imperfect numerical weather prediction (NWP) model. It can be theoretically divided into two parts as system error and random error (sampling noise) (Raynaud et al., 2009):

$$\tilde{B}_{ij} - B_{ij}^* = [\tilde{B}_{ij} - E(\tilde{B}_{ij})] + (E[\tilde{B}_{ij}] - B_{ij}^*), \quad (1)$$

where \tilde{B}_{ij} is the ensemble estimate of background error covariance between grid i and j , and B_{ij}^* is the true value. The first term $\tilde{B}_{ij} - E(\tilde{B}_{ij})$ on the right side of the equation stands for system error which indicates how much EDA error substitution deviates from the truth. This kind of error, caused by imperfect disturbing method and numerical model, can be calibrated via spread-error relationship (Palmer et al., 2009). Detailed procedure can be found in Isaksen et al., (2006); Fisher, (2007). The second term $E[\tilde{B}_{ij}] - B_{ij}^*$ is random error (sampling noise). It will get close to zero as the ensemble size increases. However, it cannot be ignored that in operational system the ensemble size is limited. It can be expressed as a simple function of the expectation of

the ensemble-based error covariance matrix (Mallat et al., 1998; Raynaud et al., 2009):

$$\begin{cases} E[G^e(i)G^e(j)] = \frac{2}{N-1}(E[B_{ij}])^2, \\ G^e(i) = [\tilde{B}_{ii} - E(\tilde{B}_{ii})], \end{cases} \quad (2)$$

where N is the ensemble size. The Daley length-scale relationship between sampling noise and background error covariance deduced from this equation is formed as

$$\begin{cases} L_{G^e}^2(i) = \left(-\frac{d^2 c^2}{dr^2} \right)^{-1} \Big|_{i,r=0} = \frac{L_{\varepsilon^b}^2(i)}{2}, \\ L_{G^e}(i) = \frac{L_{\varepsilon^b}(i)}{\sqrt{2}}, \end{cases} \quad (3)$$

where $L_{G^e}(i)$ and $L_{\varepsilon^b}(i)$ stand for the length-scale of noise and signal at grid i respectively. This equation indicates that the length-scale of sampling noise is shorter than that of background error covariance, and the latter seems to be focused on large scale. The ensemble estimate of error variance \tilde{V} and its true value V^* are transformed into spectral space denoted by \tilde{S} and \tilde{S}^* respectively. In spectral space, the length-scale is represented by the wavenumber, the longer length has a smaller wavenumber. According to Eq.(3), spectral energy of sampling noise is located at large wavenumbers, while variance energy is at small ones. Thus, a simple low-pass filter can be applied to the raw signal to filter out sampling noise (Berre et al., 2007).

$$\begin{cases} \tilde{S}^* \sim \rho \tilde{S}, \\ \rho = \frac{\text{cov}(\tilde{S}^*, \tilde{S})}{V(\tilde{S})} = \frac{1}{1 + \mathbf{P}(S^e)/\mathbf{P}(\tilde{S}^*)}, \end{cases} \quad (4)$$

where $\mathbf{P}(S^e)$ is sampling noise power spectrum, the noise-free signal energy spectrum can be expressed as $\mathbf{P}(\tilde{S}^*) = \mathbf{P}(\tilde{S}) - \mathbf{P}(S^e)$. Filter coefficient ρ varies between 0 and 1 representing the percentage of noise in raw variances. For example, there is no noise left when $\rho = 1$, while $\rho = 0$ means raw variances at this wavenumber are considered as noise. Such filter in Eq.(4) is named raw filter whose coefficients are determined by signal-to-noise ratio.

Another kind of filter is given by Raynaud (Raynaud et al., 2009).

$$\rho'(n) = \left[\cos \left\{ 0.5\pi \frac{\min(n, N_{\text{trunc}})}{N_{\text{trunc}}} \right\} \right]^2, \quad (5)$$

where n means wavenumber, and N_{trunc} is the special wavenumber at which $\mathbf{P}(S^e)$ is very close to $\mathbf{P}(\tilde{S}^*)$. This kind of filter is named smooth filter. In the following part, the large length-scale signal refers to those signals whose wavenumber is larger than N_{trunc} , otherwise they are called small length-scale signals.

Both of the two filters share same process, except the definition of filter coefficients. The differences are further discussed in Section 4.

3 APPLICATION TO YH4DVAR-EDA SYSTEM

YH4DVAR provides high-quality and accurate initial fields for global numerical weather predication spectral model (YHGSM) twice a day (Cao et al., 2008; Zhang et al., 2010; Zhang et al., 2012).

An experimental ensemble data assimilation system has been built to estimate flow-dependent background-error variances. In order to simulate the inherent uncertainty source of the YH4DVAR system, the experimental EDA system adopts the configuration as follows:

(1) The EDA is running at T399L91 resolution (i.e. spectral triangular truncation is T399 with linear grid, and model has 91 levels) with one control (unperturbed) analysis and 30 perturbed analyses.

(2) The analysis time window of EDA is 12-hour, with two minimization steps, first at T95, then at T159 with advanced linearized physics. The first minimization is running with 50 iterations, and the second is 30, corresponding to 50 of the operational YH4DVAR.

(3) For perturbed members, observations from sounding, surface, airplane, radiance and GPS radio occultation are perturbed randomly by adding a sampling from a Gaussian distribution with zero mean and standard deviation equal to the estimate of the observation error standard deviation (square root of \mathbf{R}). Model error is represented by spectral stochastic backscatter (SPBS) scheme, in which a fraction of the dissipated energy is backscattered to be up-scaled and acts as stream function forcing to the resolved-scale flow (Shutts and Palmer, 2004, Shutts, 2005, Berner et al., 2009).

(4) Such EDA system is running twice a day, with the midnight analysis using observations from 2100 UTC to 0900 UTC (incl.), and the midday analyses using observations from 0900 UTC to 2100 UTC (incl.).

The process of the EDA scheme is shown in Fig. 1. Firstly, we suitably perturb three inputs of EDA and the background field perturbation is implemented implicitly by adding perturbation of physical processes in numerical prediction models. Secondly, N sets of disturbing data are input to EDA to obtain the N perturbed analysis fields. At last, with YHGSM, N background fields (forecasting fields) are produced with integrations forward to 15 hours.

Using these EDA members, the flow-dependent background-error variances can be estimated as follows

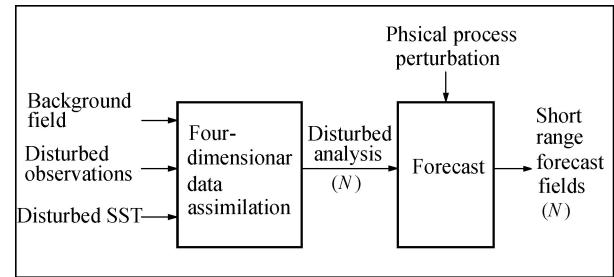


Fig. 1 The process for ensemble of flow-dependent background fields generated by the experimental Ensemble Data Assimilation system

$$\tilde{\mathbf{B}}_{ii} \approx \frac{1}{N-1} \sum_{i=1}^N \langle (\mathbf{x}_b^i - \bar{\mathbf{x}}_b)(\mathbf{x}_b^i - \bar{\mathbf{x}}_b)^T \rangle,$$

where \mathbf{x}_b^i represents the i -th ensemble member, $\bar{\mathbf{x}}_b$ is the mathematical expectation of ensemble members. The process of filtering $\tilde{\mathbf{B}}_{ii}$ is described as follows:

(1) According to Eq.(2), the spectral energy of sampling noise can be calculated by the equation $\mathbf{P}(\mathbf{S}^e) = \frac{2}{N-1} \text{Trans}(\overline{\mathbf{B}}^2)$, where $\text{Trans}(\cdot)$ represents transform operator from grid to the spectral space. Since the approximation of \mathbf{B} is known, the spectral energy of sampling noise can be pre-calculated before running EDA.

(2) Transform the variance $\tilde{\mathbf{B}}_{ii}$ into the spectral space and calculate the raw signal's energy spectra $\mathbf{P}(\tilde{\mathbf{S}}) = \text{Trans}(\tilde{\mathbf{B}}_{ii})$.

(3) Calculate raw filter $\rho(n)$, truncated wavenumber N_{trunc} and smoothing filter $\rho'(n)$ according to the distribution of $\mathbf{P}(\tilde{\mathbf{S}})$ and $\mathbf{P}(\mathbf{S}^e)$ in the spectral space.

(4) Apply the coefficients of raw filter or smoothing filter to $\mathbf{P}(\tilde{\mathbf{S}})$.

(5) Complete filtering by transforming filtered $\mathbf{P}_f(\tilde{\mathbf{S}})$ to grid space.

4 EXPERIMENT ANALYSIS

4.1 Improvements of Background Error in EDA

In order to verify the improvements on background error of EDA, ensemble-estimated standard deviations from 10, 20, 30 samples are compared with the results of YH4DVAR operational “stochastic method”, which refers to generate the standard deviation by randomly sampling from the static \mathbf{B} matrix of YH4DVAR. The flow-dependent features here are obtained by the non-linear balance equation and the linearization quasi-geostrophic Omega (vertical speed) equation around the background fields (Fisher et al., 1995; Kalnay, 2003).

Raw standard deviation maps derived from 10, 20 and 30 members are displayed in Fig. 2(a-c) for vorticity at model level 91 at 0900 UTC on 2nd Aug. 2009, while Fig. 2d shows the distribution of background error produced by random sampling statistics in East Asia region. What can be seen from the figures is that the maximums of background errors obtained by EDA (Figs. 2a, 2b, 2c) are all located near the southeast part of Wenchang City, Hainan province, corresponding to the low-pressure center of the 9th typhoon “Jebi” in 2013. While the operational background error has a relatively smooth distribution in the entire region, which is unable to reflect the structure and location of the typhoon.

As shown in Table 1, all maximums of EDA variance exceed 12×10^{-5} , which is six times larger than that of the random method, those locations are in the neighborhood of (112°E, 18°N), close to the low pressure center of typhoon (111.2°E, 19.2°N). And its accuracy is significantly higher than the maximum position of the random method (120.93°E, 15.40°N). Additionally, we also calculate the background errors of the following time sequences such as at 2100 UTC on August 2nd, and 0900 UTC on August 3rd. All the structures and the maximums correspond with changes in the typhoon while the background errors adopted by operational ones stay unchanged during this period.

The analysis above shows that the background error estimated by EDA approach has “flow-dependent” features with the change of weather situations, which can help to improve the quality control decision-making capacity of a variety of available observations in the affected area of stronger convective situation. In this way, more “legitimate” observations data can be assimilated into the system, which is ultimately improving the forecast skill of extreme weather.

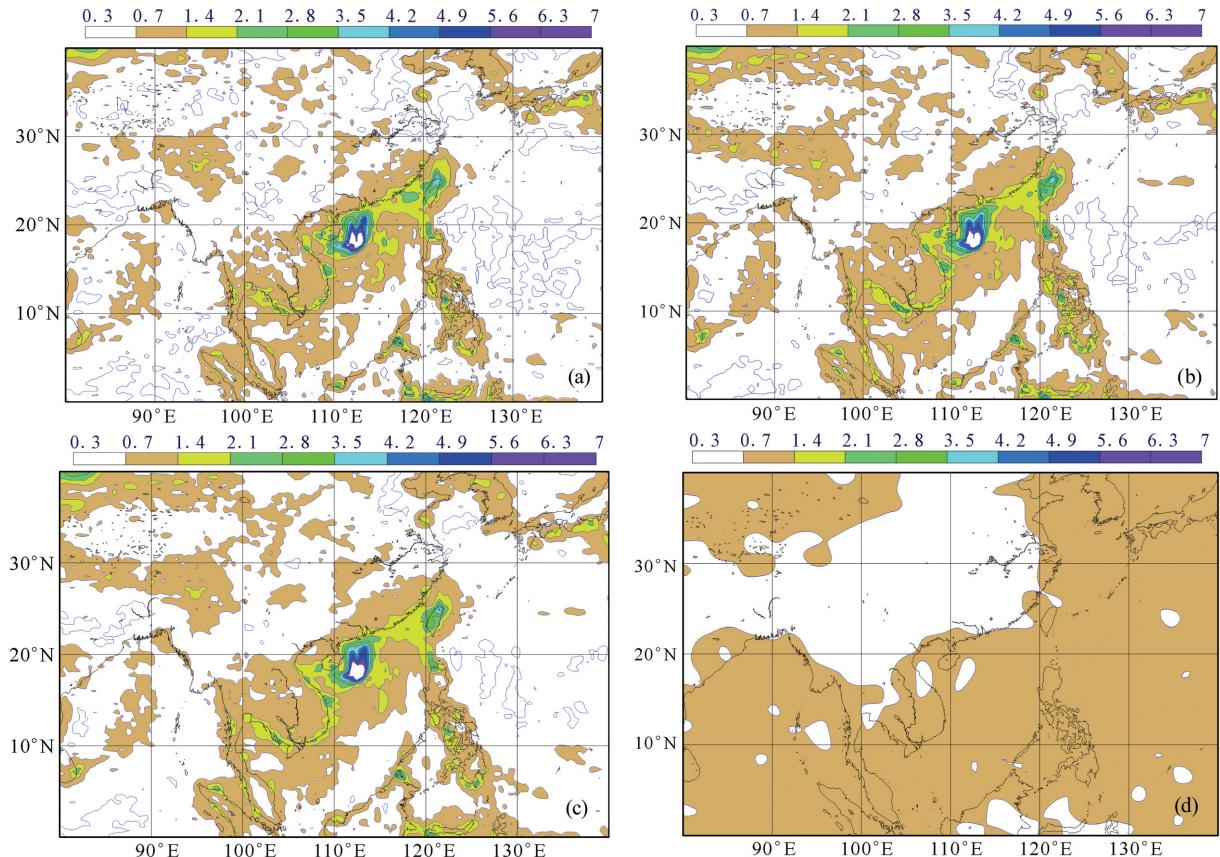


Fig. 2 Standard deviations of vorticity at model level 91, corresponding to 2 August 2013 at 0900 UTC. Unit: $0.7 \times 10^{-5} \text{ s}^{-1}$. Raw estimates from 10, 20, 30 member ensemble separately for (a), (b) and (c); background error estimate from the “randomization” technique valid at the same time(d), contour: $0.7 \times 10^{-5} \text{ s}^{-1}$

Table 1 The maximums and corresponding positions of background error according to Fig. 1 and the center position of the typhoon from observation

Methods	EDA-based			Random method	Observation
	10 members	20 members	30 members		
Max. (10^{-5} s^{-1})	12.559	12.521	12.253	2.135	—
Lon./Lat.	(112.91°E, 18.17°N)	(112.95°E, 17.93°N)	(112.21°E, 18.19°N)	(120.93°E, 15.40°N)	(111.2°E, 19.2°N)

4.2 Effect of Ensemble Size on Background Error Structure

The ensemble size will affect the structure of background error variance in EDA. Comparing the horizontal distribution of the background-error variances in Figs. 2a, 2b, 2c, we can see that spatial structures represented by 10, 20 and 30-members are very similar, except for different levels of associated noise. The relatively small-scale signals, appearing in Fig. 2a, are markedly reduced in Fig. 2c, which proves the correctness of the Eq.(2). Noise reduces the precision of background-error variance estimates, making the variance estimates not suitable to be directly used in assimilation system. Therefore, filtering is needed in EDA system.

In this experiment, the background-error variances are directly estimated with 10-members as the filtering object, and the reason behind is that the estimation results of 10-members are very similar to that with 30-members. Except for the decrease of noise, there are no essential changes of background error, the ensemble size has been close to or has reached convergence.

Figure 3 shows the statistic divergence criteria obtained with 10, 20 and 30-members. The maximum ratio (Fig. 3b) of 34.02% appears at model level 11 (about 44.4 km altitude). It may be related to the enough observations and rapid spread of the initial perturbation in the first 15 hours of model integration in high-altitude atmospheric circulation.

These results above indicate that 10-members can accurately capture the structure of background errors, which is very consistent with the conclusions of the research (Bonavita et al., 2012). Hereby, it is reasonable to apply 10-members for spectrum filtering process.

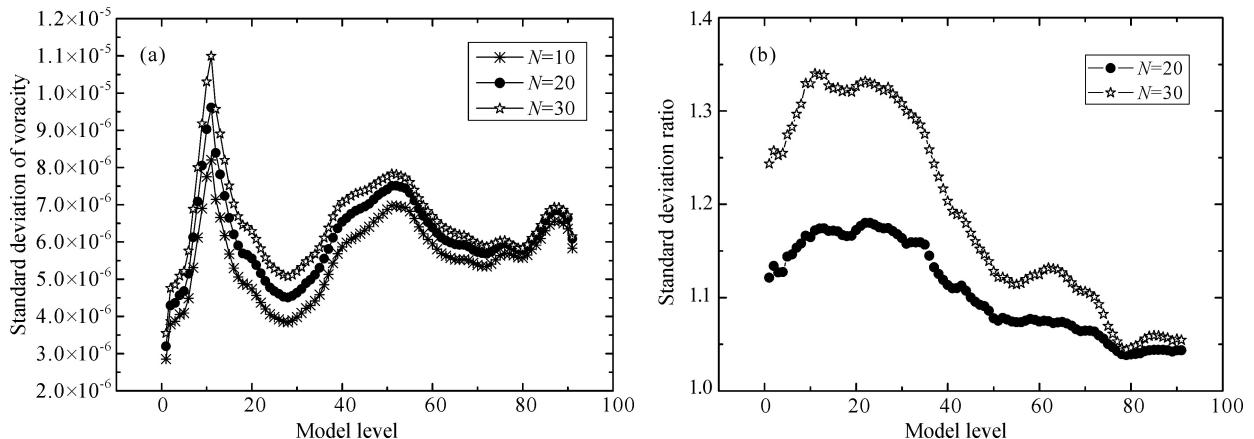


Fig. 3 Horizontal averages of standard deviations of vorticity estimated with 10, 20, 30 members (a), and their ratios to 10 members (b), as a function of model level, valid at 0900 UTC on 2 August 2013

4.3 Filtering Results

Figure 4a is the spectral energy of 10-members variance $\mathbf{P}(\tilde{\mathbf{S}})$ and associated noise $\mathbf{P}(S^e)$. When wavenumber N is less than 100, the noise energy accounts for less than 30% of the raw signal energy and in this case variance signals with larger scale are dominating. When the wavenumber N is no less than 195, noises will dominate at the scale corresponding to those wavenumbers, there is barely any useful signal variance.

According to the Raynaud's theory (Raynaud et al., 2009), the filtering residual is $\mathbf{S}_r = \tilde{\rho} \tilde{\mathbf{S}} - \tilde{\mathbf{S}}^*$, where $\tilde{\rho}$ is the filter coefficient, and $\tilde{\rho} \in (\rho, \rho')$. The strength of filtering is controlled by the value of $\tilde{\rho}$. So the

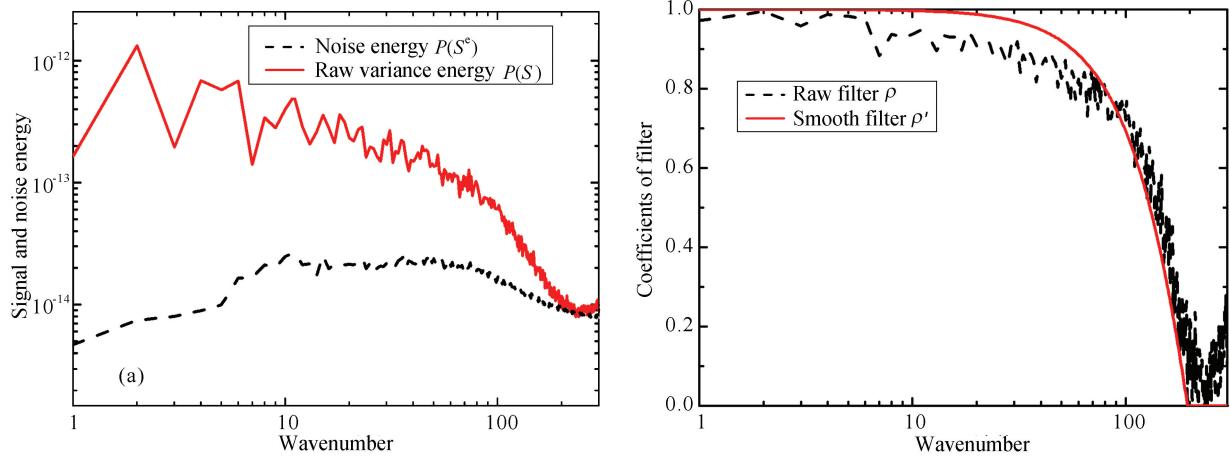


Fig. 4 (a) Power spectra of ensemble standard deviations (full red line) and of the sampling noise (dashed black line) associated to the estimation of vorticity variances at model level 91 with a 10-member ensemble, on 2th August 2013 at 0900 UTC. Unit: s^{-2} . (b) The resulting raw filter (dashed black line), the smoothed filter (full red line) obtained by Eq.(4) and (5) respectively

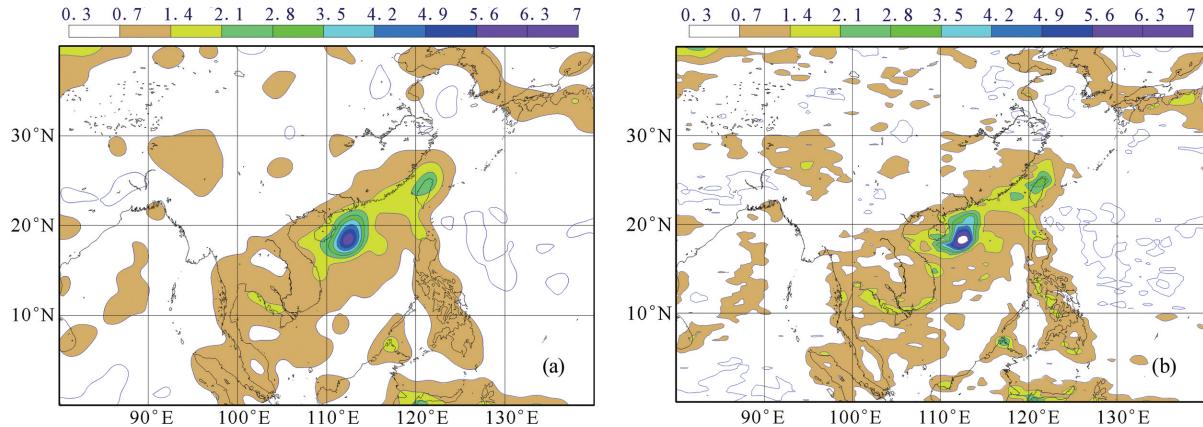


Fig. 5 Standard deviations of vorticity near the surface, corresponding to 2 August 2013 at 09 UTC. Unit: $10^{-5} s^{-1}$. Filtered result based on 10 members raw estimates same as Fig. 3 with (a) smooth filter and (b) raw filter

smoothing filter defined by Eq.(5) can weaken the noises at the larger scale ($N > 195$), and meanwhile completely filter noises at small scale when N is no less than 195.

Figure 4b shows the coefficients change of smoothing filter ρ' and raw filter ρ with wavenumber, where a similar trend can be seen. The difference is that when wavenumber N is larger than 195, the coefficient of raw filter is not zero, suggesting the filter strength of the raw filter is less than that of smoothing filter.

In fact, the essence of filtering is equivalent to a weighted average in grid space. The filtering coefficient is playing the role of weighting factor, which determines the average range of the space, namely length scale. After converting grid space to spectral space, length scale corresponds to wavenumber. Then filtering will be achieved by applying different filter coefficients to the spectral coefficients.

Figure 5 shows two filtering results. Fig. 5a is the results of the smoothing filtering while Fig. 5b is the raw filter. Compared with the raw estimated from 10-members (Fig. 2a), the smoothing filter, along with the raw filter, filtered small-scale noises and meanwhile retained the useful signal in the raw estimate of background error. And the definition of error structure out-performs the estimation of 30-ensembles. The difference is that the smoothing filter sets all coefficients to zero when $N > N_{\text{trunc}}$, thus treated all the raw estimates as noise to be processed together when $N > N_{\text{trunc}}$.

The advantage of this approach is that most of the noises are filtered at the cost of the loss of a small portion of the signal, and the efficiency of filtering is improved. On one hand, as the filter coefficient of the raw filter in $N > N_{\text{trunc}}$ is not equal to zero, a lot of noises are brought in when we are retaining these small-scale signals, which deteriorates the filtering efficiency. On the other hand, although smoothing filtering is an empirical method, the smoothness of background error obtained in this way is better than that of raw filtering for the reason that the coefficients of the raw filter change intensively with wavenumber.

Furthermore, random noises of background error estimation for variables like vorticity, divergence, temperature on other model levels are processed in the same way. For the sake of simplicity, only the truncated wavenumber change of vorticity with the number of layers is presented in Fig. 6, where the truncated wavenumber gradually decreases as the altitude increases. The reason behind is that the upper atmosphere is not affected by geography and turbulence, making the overall length of fluid scale larger than the lower level and the corresponding signal and noise also larger than the lower level.

5 CONCLUSION

Ensemble data assimilation is the future direction of YH4DVAR system, and exploratory researches on related technologies are still ongoing.

Based on the YH4DVAR analysis, we built an experimental Ensemble Data Assimilation system. With 10 ensemble members, the background-error variances are calculated when the 9th typhoon “Jebi” in 2013 passed nearby the area of Wenchang City, Hainan Province, and successfully applied spectral filtering technique to the processes of background error.

The results showed that flow-dependent background error obtained by EDA can accurately reflect the error characteristics of typhoon-struck region, and structural features and magnitude were both significantly better than the statistical results of operational stochastic method. However, because there are obvious small-scale random sampling noises in the estimation, the accuracy of background-error variances estimation is reduced.

In the end, we employ a spectral filtering technique to improve the estimation of background-error variances in EDA. The spectral filtering technique can calculate reasonable truncated wavenumber according to the energy spectrum relationship of signal and noise, and formulate a low-passing filter to further eliminate the noise in the raw estimations. Results show that the 10-members’ filtered variances exhibit better performance than 30-members’s raw estimation.

Nevertheless, the spectrum filtering technique also has some limitations. Firstly, noise energy spectrum is estimated with the use of climate-state \mathbf{B} . Therefore, the accuracy and the correctness are determined entirely by \mathbf{B} and ensemble members. Due to the use of static approximation in \mathbf{B} , the noise energy spectrum of statistical random sampling is a static constant, unable to reflect weather systems with larger local error variance. Secondly, noises are filtered in spectral filtering by multiplying the spectrum coefficients with filtering coefficients. Meanwhile the signals are also weakened and some background error variances with very tiny scale properties may be smoothed out. Finally, filter coefficients acting on the same wave number are equal, suggesting smoothing treatment on the same scale of grid space are globally identical, which fails to reflect local characteristics. All limitations above need further research, such as to develop the heterogeneous wavelet filtering techniques.

In addition, we merely described the random sampling noise entry of background error variance term.

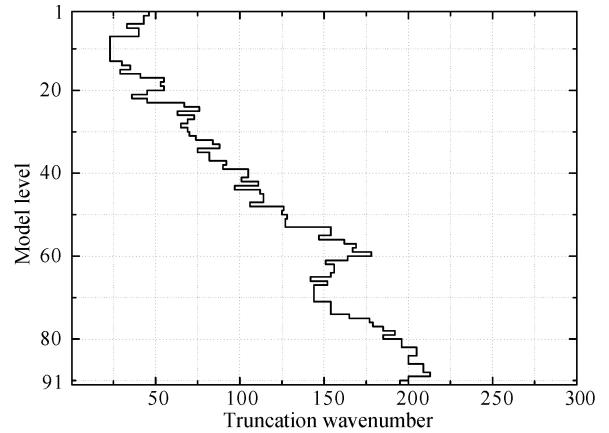


Fig. 6 Vertical profile of objective truncated wave number for the filtering of raw ensemble-based standard deviation of vorticity from 10 members, on 2th August 2013 at 09 UTC

However, more EDA samples are needed to compute the statistical background error covariance, and its noise processing is also different and requires further study.

The filtering technique and related experiments in this paper are conducive to strengthen the awareness of Ensemble Data Assimilation variance filtering and to provide theoretical guidance for our EDA operational design and implementation, which could be of certain significance.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (41375113, 41105063, 41475094). We would like to thank all reviewers for their constructive comments and suggestions.

References

- Berre L. 2000. Estimation of synoptic and mesoscale forecast error covariances in a limited-area model. *Mon. Wea. Rev.*, 128(3): 644-667.
- Berre L, Pannekoucke O, Desroziers G, et al. 2007. A variational assimilation ensemble and the spatial filtering of its error covariances: Increase of sample size by local spatial averaging. // Proc. ECMWF Workshop on Flow-Dependent Aspects of Data Assimilation, 151-168. (Available from: <http://www.ecmwf.int/publications/>).
- Berre L, Desroziers G. 2010. Filtering of background error variances and correlations by local spatial averaging: A review. *Mon. Wea. Rev.*, 138(10): 3693-3720.
- Bonavita M, Raynaud L, Isaksen L. 2011. Estimating background-error variances with the ECMWF Ensemble of Data Assimilations system: some effects of ensemble size and day-to-day variability. *Quart. J. Roy. Meteor. Soc.*, 137(655): 423-434.
- Bonavita M, Isaksen L, Hólm E. 2012. On the use of EDA background error variances in the ECMWF 4D-Var. *Quart. J. Roy. Meteor. Soc.*, 138(667): 1540-1559.
- Buizza R, Houtekamer P L, Pellerin G, et al. 2005. A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems. *Mon. Wea. Rev.*, 133(5): 1076-1097.
- Cao X Q, Huang S X, Du H D. 2008. The new method of modeling horizontal error functions in variational assimilation with orthogonal wavelet. *Acta Physica Sinica* (in Chinese), 57(3): 1984-1989.
- Courtier P, Andersson E, Heckley W, et al. 1998. The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation. *Quart. J. Roy. Meteor. Soc.*, 124(550): 1783-1807.
- Daley R. 1993. Atmospheric Data Analysis. Cambridge: Cambridge University Press.
- Derber J, Bouttier F. 1999. A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus A*, 51(2): 195-221.
- Evensen G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, 99(C5): 10143-10162.
- Fisher M, Courtier P. 1995. Estimating the covariance matrices of analysis and forecast error in variational data assimilation. ECMWF Tech. Memo. No. 220. 1-29. (Available from: <http://www.ecmwf.int/publications/>).
- Fisher M, Andersson E. 2001. Developments in 4D-Var and Kalman filtering: ECMWF Tech. Memo. No. 347. (available from European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, Berkshire RG2 9AX, UK).
- Fisher M. 2003. Background error covariance modelling. // Proc. ECMWF Seminar on Recent Development in Data Assimilation for Atmosphere and Ocean, 45-63. (Available from: <http://www.ecmwf.int/publications/>).
- Fisher M. 2007. The sensitivity of analysis errors to the specification of background error covariances. // Proc. ECMWF Workshop on Flow-Dependent Aspects of Data Assimilation, 27-36. (Available from: <http://www.ecmwf.int/publications/>).
- Houtekamer P L, Lefavre L, Derome J, et al. 1996. A system simulation approach to ensemble prediction. *Mon. Wea. Rev.*, 124(6): 1225-1242.
- Houtekamer P L, Mitchell H L. 1998. Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, 126(3): 796-811.

- Isaksen L, Fisher M, Berner J. 2006. Use of analysis ensembles in estimating flow-dependent background error variances. // Proc. ECMWF Workshop on Flow-Dependent Aspects of Data Assimilation, 65-86. (Available from: <http://www.ecmwf.int/publications/>).
- Kalnay E. 2003. Atmospheric Modeling, Data Assimilation and Predictability. Cambridge: Cambridge University Press.
- Laroche S, Gauthier P. 1998. A validation of the incremental formulation of 4D variational data assimilation in a nonlinear barotropic flow. *Tellus A*, 50(5): 557-572.
- Lin Z S. 1999. The discuss of the predictability of climatic or atmospheric system. *Chinese J. Geophys.* (in Chinese), 42(S1): 13-17.
- Mallat S, Papanicolaou G, Zhang Z F. 1998. Adaptive covariance estimation of locally stationary processes. *Ann. Statist.*, 26(1): 1-47.
- Mitchell H L, Charette C, Chouinard C, et al. 1990. Revised interpolation statistics for the Canadian data assimilation procedure: Their derivation and application. *Mon. Wea. Rev.*, 118(8): 1591-1614.
- Palmer T N, Buizza R, Doblas-Reyes F, et al. 2009. Stochastic Parametrization and Model Uncertainty. ECMWF Tech. Memo. No. 598. (available from European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, Berkshire RG2 9AX, UK).
- Parrish D F, Derber J C. 1992. The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.*, 120(8): 1747-1763.
- Pereira M B, Berre L. 2006. The use of an ensemble approach to study the background error covariances in a global NWP model. *Mon. Wea. Rev.*, 134(9): 2466-2489.
- Qian W H. 2012. How to improve the skills of weather and climate predictions. *Chinese J. Geophys.* (in Chinese), 55(5): 1532-1540, doi: 10.6038/j.issn.0001-5733.2012.05.010.
- Raynaud L, Berre L, Desroziers G. 2008. Spatial averaging of ensemble-based background-error variances. *Quart. J. Roy. Meteor. Soc.*, 134(633): 1003-1014.
- Raynaud L, Berre L, Desroziers G. 2009. Objective filtering of ensemble-based background-error variances. *Quart. J. Roy. Meteor. Soc.*, 135(642): 1177-1199.
- Weaver A, Courtier P. 2001. Correlation modelling on the sphere using a generalized diffusion equation. *Quart. J. Roy. Meteor. Soc.*, 127(575): 1815-1846.
- Wiener N. 1949. Extrapolation, Interpolation, and Smoothing of Stationary Time Series. Cambridge, MA: MIT Press.
- Yue X A, Wan W X, Liu L B, et al. 2010. Development of an ionospheric numerical assimilation nowcast and forecast system based on Gauss-Markov Kalman filter: An observation system simulation experiment taking example for China and its surrounding area. *Chinese J. Geophys.* (in Chinese), 53(4): 787-795, doi: 10.3969/j.issn.0001-5733.2010.04.003.
- Zhang W M, Cao X Q, Xiao Q N, et al. 2010. Variational Data Assimilation Using Wavelet Background error covariance: Initialization of typhoon KAEMI (2006). *J. Trop Meteorol.*, 16(4): 333-340.
- Zhang W M, Cao X Q, Song J Q. 2012. Design and implementation of four-dimensional variational data assimilation system constrained by the global spectral model. *Acta Physica Sinica* (in Chinese), 61(24): 249202-249214.