A Modified Nonlinear Wavelet thresholding filter in ensemble data Assimilation

Bainian Liu, Weimin Zhang, Kaijun Ren, Hongze Leng, Mengbin Zhu

Academy of Ocean Science and Engineering, NUDT, Changsha, 410075, China.

Introduction

Ensemble-based estimates of background-error in Ensemble data assimilation often contain sampling noise due to the limited ensemble size. Objective filter technology (Raynaud, et al., 2009) has been successfully applied in server operational ensemble data assimilation systems, such as in ECMWF and in Mete-France. But one of the main shortages is that this homogeneous filter cannot be adaptable to the local structure of the signal. Thus, heterogeneous filtering methods such as nonlinear wavelet thresholding technology is employed. As the noise level varies in different scales, the threshold determined by iterative algorithms (Azzalini, 2005) is no longer suitable for noises. To address this problem, Pannekoucke et al. (2014) uses a multiplicative factor to adjust the filtering strength based on the optimization of the trade-off between the removal of the noise and the averaging of the useful signal. However, tuning α is not easy, especially in real operational context. In our work, the threshold of wavelet is modified accounting to the distribution of the wavelet coefficients, whose modulus is smaller than the threshold value. Its validity and performance are examined in a one-dimensional model. Results show that our method outperforms previous filters. The filtering performance has been improved by 13.28%.

Methods

I. Wavelet-based filtering algorithm

As wavelet transform can provide full spatial resolution, denoising by wavelet thresholding is regarded as an effective method. The main ideal of wavelet thresholding is to keep only the wavelet coefficients whose modulus is above a assuming threshold value T.

$$\tilde{X}^{*}_{i,j} = \rho_{T_0}\left(\tilde{X}_{i,j}\right) = \begin{cases} \tilde{X}_{i,j}, & if \left|\tilde{X}_{i,j}\right| > T \\ 0, & if \left|\tilde{X}_{i,j}\right| \le T \end{cases} \qquad (1) \qquad\qquad \mathbf{X}_{wave} = \sum_{i,j} \tilde{X}^{*}_{i,j}\psi_{i,j} \qquad (2)$$

The rest of the coefficients are used to reconstruct the denoised signal. Thus, just like objective filtering, the key issue of wavelet thresholding is how to get a suitable threshold. Azzalini (2005) proposed an recursive algorithm to determine the threshold. Its basic process are listed as follows.

- ◆The noise signal is firstly decomposed into an orthogonal wavelet space $\tilde{X} = \sum_{j=0}^{J-1}\sum_{i=0}^{2^J-1}\tilde{X}_{i,j}\psi_{i,j}$,
- ◆$T_0 = \sigma_W\sqrt{2ln(n)}$ is used as the primary threshold to split coefficients. Only smaller ones remain to calculate the next variance σ_W and threshold T_1.
- ◆Repeat this algorithm until the difference between T_i and T_{i+1} small enough.

III. Modified wavelet threshold

Under the assumption of a Gaussian white noise, wavelet thresholding is a very efficient method, but in practical ensemble data assimilation system, the error in estimates is often correlated and scaled depended. It means in some scales noise level is too high, which need a more strength filter, while others filtering is over-strength. To dress this problem, Pannekoucke et al. (2014) uses a multiplicative factor to adjust the filtering strength based on the optimization of the trade-off between the removal of the noise and the averaging of the useful signal.

$$T_D' \approx \alpha \times \sigma_w \sqrt{2\ln(n)} \qquad (3)$$

However, the tuning of choose α is not so obvious, especially in real operational context. We propose an alternative method to automatically determine a suitable α. Instead of considering the full noise, we just focus on the denoised part of noisy signal which is used to modified the threshold.

$$T_s = \max\left(2\sigma_s\sqrt{2\ln(n)}, T_D\right), \quad \sigma_s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}_i\right)^2} \qquad (4)$$

Where σ_S is standard deviation of the coefficients whose modulus is below the recursive threshold value T_D. The theoretical basis is that the denoised part of noisy signal X is still a quasi-Gaussian error even X is correlated and scale depended error. On the one hand, $\pm 2\sigma_S$ range can include about 95% noise. On the other hand, threshold can be restricted in a reasonable range, so there remains enough useful signal wavelet coefficient for signal restructuring. The process of modified wavelet threshold is similar with Azzalini's method except an additional step to calculate T_D.

Experiments

I. Configuration

We choose the Earth's great circle as the whole domain with radius r=6370km. This circle is divided into n=401 equally spaced grid points. variance varies with grid points, and a very sample correlation function is applied to construct background error covariance model .

$$c(i,i+r) = \exp\left(-\frac{r^2}{2L_{\varepsilon^b}^2}\right) \qquad \begin{cases} \alpha_k \sim N(0,1) \\ \varepsilon_k^b = B^{1/2}\alpha_k, k = 1,\cdots,n \end{cases} \qquad \begin{cases} V_i = \frac{1}{N-1}\sum_{k=1}^{N}\left(\varepsilon_{k,i}^b - \overline{\varepsilon_i^b}\right)^2 \\ \overline{\varepsilon_i^b} = \frac{1}{N}\sum_{k=1}^{N}\varepsilon_{k,i}^b \end{cases}$$

The length-scale of correlation L_{ε^b} is set to 300km. N random error realizations are generated from B using randomization technique (Fisher and Courtier, 1995)

II. Filtering results

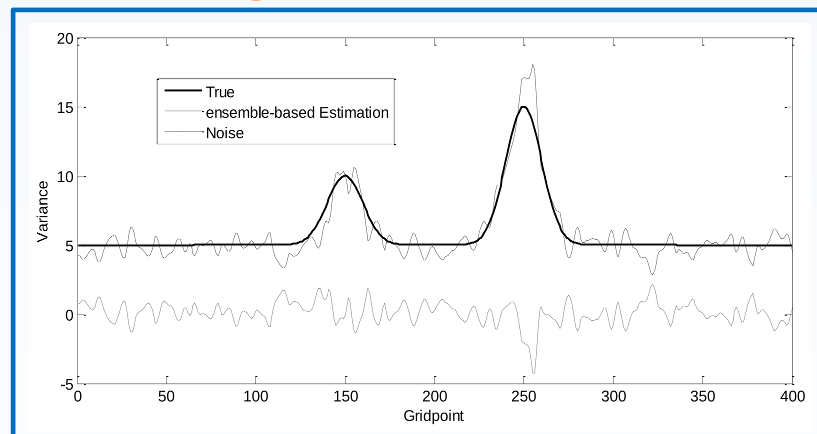


Fig. 1. True (bold solid) and ensemble-based estimated (thin solid) variances and corresponding sampling noise (dashed)

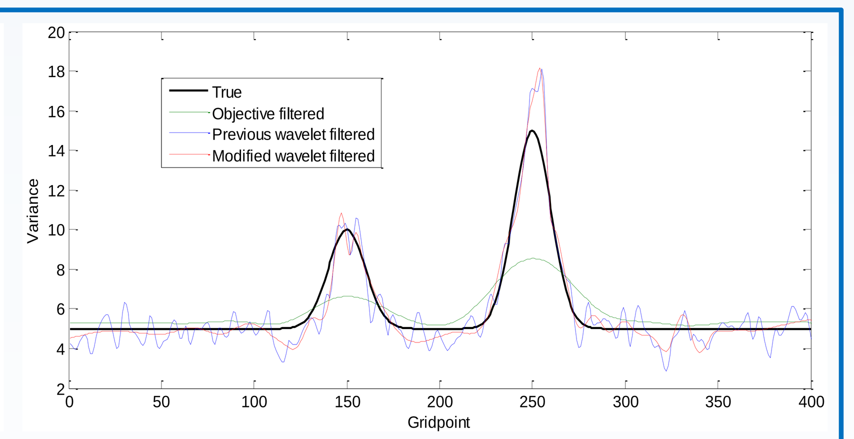


Fig. 2. Comparison of filtered variances through spectral filtered (green), previous wavelet filtered (blue) and modified wavelet filtered (red), the true variance is denoted by bold solid line

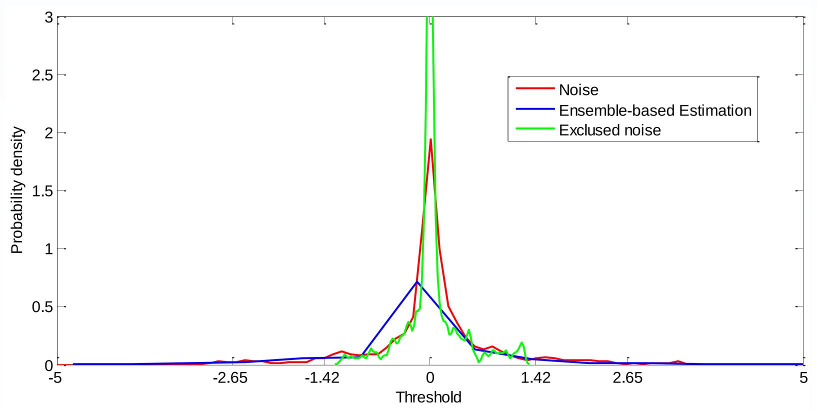


Fig.3 Probability density distribution of the wavelet coefficients, the red, blue and green line corresponding to noise, estimate variance and filtered noise respectively whose coefficients magnitude are small than threshold

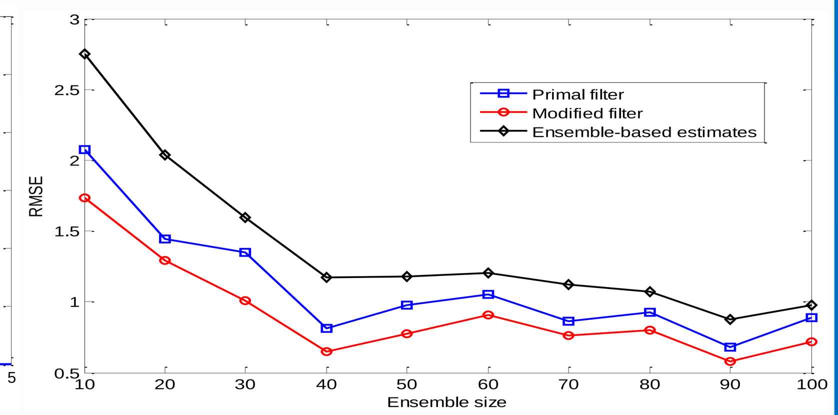
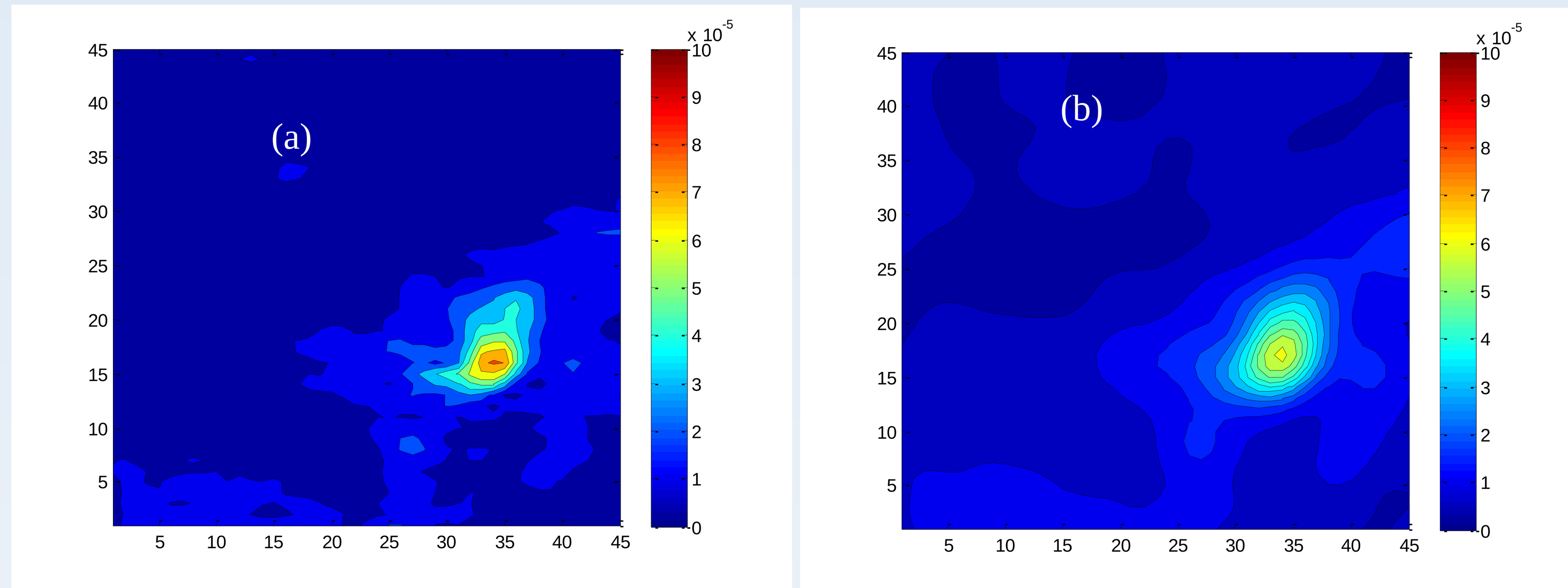


Fig. 4 The RMSE of ensemble-based variance estimates (black) and filtered results processed by previous wavelet filter (blue) and modified wavelet filter (red)

Fig.1 shows the assumed background error variance (bold solid) and ensemble-based estimates (thin solid) . Sampling noise (dashed line) drifts around the true value. The maximum values and sharp change appeared at grid 150th and 250th are used to represent the error distribution correlation to rapid development system, such as storm. Fig. 2 indicates that the modified wavelet filtered result owns the lowest RMSE error. Probability density distribution of the wavelet coefficients for noise, estimate variance and filtered noise are showed in Fig.3. threshold value T_D=1.42 is modified to T_S=2.65 using modified method. It helps to filter higher energy level noise in some scales. Fig.5 intend to illustrate the versatility and reliability of the modified filter with different ensemble size from 10 to 100. With increasing the member size, the error is gradually reduced. Comparing the RMSE of filtered result, we can see that the wavelet filtering method outperforms spectral filtering. Comparing to previous filter, the RMSE of filtered variances is reduced 13.28% by using the modified and revised threshold filter.



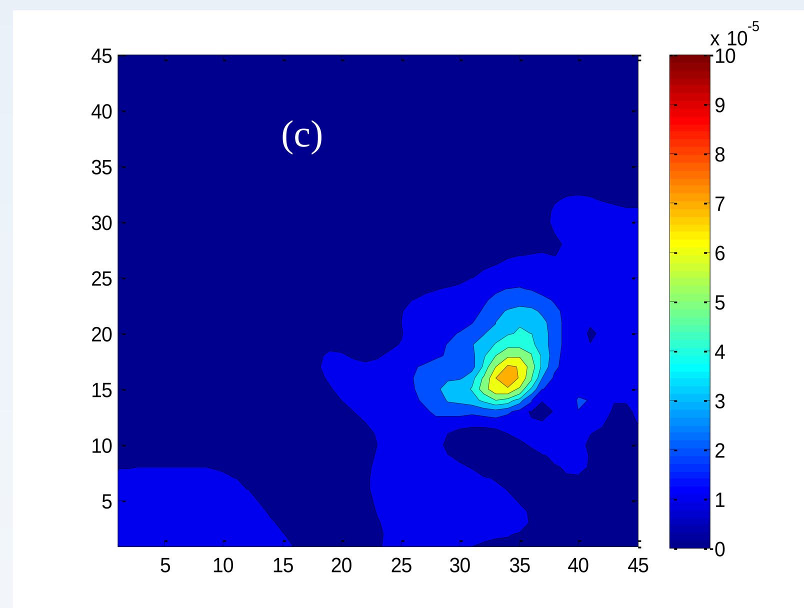


Fig. 5. Standard deviations of vorticity at ML=91, corresponding to 2 August 2013 at 09 UTC (a). Unit: 10^{-5} s^{-1}. Filtered result based on 10 members raw with: (b) objective filter, (c) modified wavelet filter.

We also apply it to a experimental ensemble data assimilation system. The 10-members estimate of vorticity at model level 91, on 2 August 2013 at 2100 UTC (Fig.5 Unit: 0.7×10^{-5} s^{-1}.) is used for testing. It corresponding to the 9th typhoon "Jebi" passing by the southeast part of Wenchang City, Hainan Province of China. Although the objective filtering result (b) appears more smooth, its maximum is 7.11 $\times10^{-5}$, while the wavelet filter's (c) is 8.21 $\times10^{-5}$ which more closer to presumed value 8.31$\times10^{-5}$.

Conclusion

We introduced a modified wavelet filtering method to deal with correlated and scale dependent error in ensemble data assimilation. Instead of tuning the multiplicative factor manually, we used the standard deviation of the wavelet coefficients, whose modulus is smaller than threshold value T_D, to adjust the threshold automatically. This process made the wavelet filter suitable for reducing the residual noise arising from scales with higher noise level, keeping the useful signal as while.

We tested the proposed algorithm in 1D analytical framework with correlated and heterogeneous noise. Comparing with the wavelet filtering algorithm, the RMSE of filtered variances was reduced by 13.28% using our modified threshold. 2D system experiments also shows the effectiveness of this method.

References

[1] Raynaud L, Berre L, Desroziers G. Objective filtering of ensemble-based background-error variances[J]. Quarterly Journal of the Royal Meteorological Society, 2009, 135(642):1177-1199.

[2] Azzalini A, Farge M, Kai S. Nonlinear wavelet thresholding: A recursive method to determine the optimal denoising threshold[J]. Applied & Computational Harmonic Analysis, 2005, 18(18):177-185.

[3] Pannekoucke O, Raynaud L, Farge M. A wavelet-based filtering of ensemble background-error variances[J]. Quarterly Journal of the Royal Meteorological Society, 2014, 140(140):316–327.

[4] Fisher, Courtier M, P. Estimating the covariance matrices of analysis and forecast error in variational dataassimilation[J]. Ecmwf, 1995.