This is a PDF of a manuscript that has been peer-reviewed and accepted for publication. As the article has not yet been formatted, copy edited or proofread, the final published version may be different from the early online release.

This pre-publication manuscript may be downloaded, distributed and used under the provisions of the Creative Commons Attribution 4.0 International (CC BY 4.0) license. It may be cited using the DOI below.

The DOI for this manuscript is DOI:10.2151/jmsj.2018-048

J-STAGE Advance published date: June 22nd, 2018

The final manuscript after publication will replace the preliminary version at the above DOI once it is available.
Refinement of the Use of Inhomogeneous Background Error Covariance Estimated from Historical Forecast Error Samples and its Impact on Short-term Regional Numerical Weather prediction

Yaodeng Chen*, Jia Wang, Yufang Gao

Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters/Key Laboratory of Meteorological Disaster of Ministry of Education, NUIST, Nanjing, China

Xiaomeng Chen
Fujian Meteorological Service Center, Fuzhou, China

Hongli Wang*
National Center for Atmospheric Research, Boulder, CO, USA

Xiang-Yu Huang
Centre for Climate Research Singapore, Meteorological Service, Singapore

*Corresponding author:
Dr. Yaodeng Chen, keyu@nuist.edu.cn;
Dr. Hongli Wang, hungli.wong@gmail.com;
ABSTRACT

Background error covariance (BEC) is one of the key components in the data assimilation systems for numerical weather prediction. Recently, a scheme of using an inhomogeneous and anisotropic BEC estimated from historical forecast error samples has been tested by employing the extended alpha control variable approach (BEC-CVA) in the framework of the Variational Data Assimilation system for the Weather Research and Forecasting model (WRFDA). In this paper, the BEC-CVA approach is further examined by conducting single observation assimilation experiments and continuously cycling data assimilation and forecasting experiments covering a 3-weeks period. Moreover, additional benefits of using a blending approach (BEC-BLD), which combines a static, homogeneous BEC with an inhomogeneous and anisotropic BEC, are also assessed.

Single observation experiments indicate that the noises in the increments in BEC-CVA can be somehow reduced by using BEC-BLD, while the inhomogeneous and multivariable correlations from the BEC-CVA are still taken into account. The impact of BEC-CVA and BEC-BLD on short-term weather forecasts is compared with three-dimensional variational data assimilation scheme (3DVar), and compared with the hybrid ensemble transform Kalman filter and 3DVar (ETKF-3DVar) in WRFDA also. Results show that the BEC-CVA and BEC-BLD outperform the use of 3DVar. It is shown that BEC-CVA and BEC-BLD underperform ETKF-3DVar as expected, however the computational cost of BEC-CVA and BEC-BLD is considerably less expensive since no ensemble forecasts are required.

Keywords: Numerical Weather Prediction; Data Assimilation; Background Error Covariance; Inhomogeneous; Computational Cost
1. Introduction

Accurate numerical weather forecasts depend on the accuracy of the initial conditions used by numerical weather prediction (NWP) models, which are usually estimated and optimized by using data assimilation techniques. Among many factors, the background error covariance (BEC) plays a key role in data assimilation systems such as 3/4 dimensional variational data assimilation system (3/4DVar), ensemble-based Kalman filters and hybrid ensemble variational data assimilation systems. This is because BEC spatially spreads the observation information around it and defines the correlations among control variables. When BEC was introduced into 3/4DVar, usually a control variable transform along with the assumption of spatial homogeneity and isotropy in static error covariance estimation are used that simplify the error modeling process and increase computation efficiency (Wu et al. 2002; Barker et al. 2004). However, 4DVar uses a forecast model as a dynamic constraint of the analysis and thus the flow dependent BEC is implicitly included (e.g. Sun and Crook, 1997; Rabier et al. 2000; Honda et al. 2005; Rawlins et al. 2007; Huang et al. 2009; Wang et al. 2013; Lorenc et al. 2015). The establishment and maintenance of the adjoint model, however, is a tremendous effort, and the computational cost of 4DVar remains expensive for mesoscale and convective scale data assimilation.

In addition to 4DVar, ensemble-based Kalman filters are other implementations that estimate flow-dependent BEC in an ensemble subspace (e.g. Evensen 1994; Anderson 2001; Bishop et al. 2001; Whitaker and Hamill 2002; Hunt et al. 2007 and many others). However, BEC calculated from an ensemble with limited number of members might suffer
from sampling errors and require localizations to suppress long distance spurious correlations and inflation techniques to enlarge ensemble spread (Houtekamer and Mitchell, 1998; Hamill and Snyder, 2000; Bannister, 2007).

Hybrid data assimilation approach has been developed to take advantage of an existing variational data assimilation system and a flow-dependent estimation of BEC provided by an ensemble-based Kalman filter or other ensemble generation methods (Hamill and Snyder 2000; Lorenc 2003; Buehner 2005; Wang et al. 2007b; Clayton et al. 2013; Lorenc et al. 2015; Bowler et al. 2017). Hamill and Snyder (Hamill and Snyder 2000) constructed a hybrid scheme that directly replacing the BEC term in the cost function by a linear combination of a static BEC with an ensemble-based flow-dependent covariance. Lorenc (2003) proposed another form of the hybrid variational scheme where the control variables in the cost function were augmented by another set of control variables, preconditioned upon the square root of the ensemble covariance (Wang et al. 2007b). In general, the hybrid data assimilation method combines a flow-dependent BEC estimated from an ensemble with a static BEC in the variational framework.

Many studies have suggested that the hybrid method yields better forecasts than the 3DVar method that does not incorporate the flow-dependent BEC, and is also more robust than ensemble-based Kalman filters (e.g. Wang et al. 2008a, b; Wang 2011; Kuhl et al. 2013; Wang et al. 2013; Zhang et al., 2013; Kleist and Ide. 2015a,b). The hybrid method requires a set of ensemble forecasts to provide a flow-dependent BEC during each data assimilation cycle. The computational cost is significantly higher than that of 3DVar that does not involve ensemble forecasts. The computational cost is still a burden for some operational centers and research communities with limited computational resources. In
recent studies, some weight to static climatological BEC improved forecasts using operational global models (Clayton et al. 2013; Kleist and Ide 2015a, b; Lorenc et al. 2015; Bowler et al. 2017). Moreover, the value of climatological BEC information in ensemble-based Kalman filters has also shown by Kretschmer et al. (2015). Their research indicates that the use of climatological error perturbations has significant potential for improving analyses and forecasts.

The static BEC estimation used in the hybrid method requires a set of historical forecasting error samples. They can be obtained from either historical ensemble perturbations or forecast differences obtained via the NMC method or a combination of both. Though some significant works have been done to improve the anisotropic and inhomogeneous BEC of 3D/4DVar using recursive filters and wavelets (Deckmyn and Berre, 2005; Bannister 2007; Oliveira 2009), in most of studies, the static BEC usually takes assumption of spatial homogeneity and isotropy in error covariance estimation and uses various control variable transforms (Bannister 2008). This assumption ignores the most valuable inhomogeneous and anisotropic information and somehow weakens multivariate correlations (Wang et al. 2014) and even introduces questionable solutions in wind analysis through control variable transform (Xie and MacDonald, 2012; Wang et al. 2014; Sun et al. 2016).

To take advantage of BEC in the historical forecast error samples, Wang et al. (2014) explored a scheme that uses the extended alpha control variable approach (Lorenc 2003) to introduce an inhomogeneous and anisotropic historical BEC (BEC-CVA) for WRFDA (Data Assimilation system for the Weather Research and Forecasting model). The final goal of Wang et al.’s (2014) work is to best use of both historical forecast error samples
and error samples of the day (e.g. from a short-term ensemble forecast at an analysis time) to produce a best estimation of BEC for data assimilation. However, as a first step, they introduced the BEC-CVA approach for better use of historical forecast error samples. Their study shows that the BEC-CVA approach is capable of extracting inhomogeneous and anisotropic climatological information from historical forecast differences obtained via the NMC method (Parrish and Derber 1992).

However, similar to ensemble-based Kalman Filters, the BEC-CVA approach suffers from sampling errors (Wang et al. 2014). In addition, Wang et al. (2014) only conducted a case study to examine the impact of the BEC-CVA method on short term weather forecast where the radar data was one of main data resource. More studies will help to understand the implementations of the method and evaluate its performance. This research puts a light on ideas on better use of static BEC, and its potential role in a hybrid BEC data assimilation system.

This paper further examines the BEC-CVA approach by conducting single observation assimilation experiments and various real data assimilation and forecasting experiments. Moreover, additional benefits of considering the contribution from a static homogeneous BEC, which helps to reduce some noises in analysis increments, are also assessed. Based on the fact that the noises in the analysis increments in the BEC-CVA approach can be somehow reduced by combining contribution of a static, homogeneous BEC, the use of the blended BEC (BEC-BLD) that combines a static, homogeneous BEC and an anisotropic and inhomogeneous BEC on short-term regional weather forecasts is tested. Continuously cycling data assimilation and forecasting experiments over a 3-weeks period are performed to examine the impact of a few variations of use the BEC such as BEC-CVA and BEC-
BLD on short-term weather forecasts for heavy rainfall events occurred in east China. We also compare the BEC-CVA and BEC-BLD approaches to the traditional hybrid data assimilation scheme in WRFDA.

The rest of this paper is structured as follows. In section 2, the method used to refine use of BEC is described. Section 3 introduces the experimental design and the rainfall event examined for the study. The single observation experiments are discussed in section 4. In section 5, the computational costs and verification scores for the continuously cycling data assimilation and forecasting experiments over 3-weeks period are compared. The diagnosis for a rainfall event using different methods are presented in sections 6. Finally, conclusions are presented in section 7.

2. Method

WRFDA is used to explore the impact of variations of BEC formulations on NWP. The cost function that combines two different BECs in this paper can be written as a traditional hybrid method (Wang et al. 2008a; Lorenc 2003b):

\[ J(\delta x, \alpha) = \beta_1 \frac{1}{2} \delta x^T B^{-1} \delta x + \beta_2 \frac{1}{2} \alpha^T A^{-1} \alpha + \frac{1}{2} (d - H\delta x)^T R^{-1} (d - H\delta x) \]  

(1)

In Eq.1, \(d = y - H(x^b)\) is the innovation vector with \(y\) denoting the observation and \(H\) is the nonlinear observation operator whose linearized observation operator is denoted by \(H\). The \(R\) matrix represents the observation error covariance.

The first term on the right side in Eq.1 is the traditional background term and \(\delta x_1\) is the increment associated with the static BEC. In WRFDA, \(B\) is a homogenous and isotropic BEC typically estimated via the NMC method by taking the difference between forecast samples of different lead times valid at the same time (Parrish and Derber 1992). The
second term on the right side is the background term associated with the inhomogeneous BEC in which $a$ is the extended control variable and $A$ is the correlation matrix of the extended control variable. $\delta x = \delta x_1 + \sum_{k=1}^{N}(a_k \cdot x_k^e)$ is the final analysis increment, the vectors $a_k$ ($k = 1, \ldots, K$) denote the extended control variables for each ensemble member. $\alpha$ is a vector formed by concatenating by $K$ vectors $a_k$. In other words, $\alpha^T = (a^T_1, a^T_2, \ldots, a^T_K)$. The coefficients $\beta_1$ and $\beta_2$ represent the weights applied for the static homogeneous BEC and the inhomogeneous BEC respectively, and $1/\beta_1 + 1/\beta_2 = 1$.

Recently, Wang et al. (2014) used the second term on the right side in Eq. 1 to incorporate the inhomogeneous and anisotropic BEC by using historical forecast error samples, which can be obtained from a time series of ensemble perturbations and/or NMC-style forecast differences that are defined as the difference between the forecasts with different forecast time lengths but valid at the same time.

In this paper, the historical NMC-type forecast differences are used. $x_k^e$ is redefined as:

$$x_k^e = (x_k^{diff} - \bar{x})/\sqrt{2M}$$  \hspace{1cm} (2)

$$x_k^{diff} = x_k^{T1} - x_k^{T2}$$  \hspace{1cm} (3)

Where, $M$ is the total number of the forecast difference; $x_k^{diff}$ is the forecast difference defined by Eq.3; $T1$ and $T2$ are the forecast lead times; $\bar{x}$ is the time-averaged bias of forecast differences. It is noted that the $x_k^e$ here needs to be scaled for a specific application because $\left(x^e \right)^T x^e$ is combinations of true forecast error covariances at the forecast lead times $T1$ and $T2$, and their covariances (Bannister 2008; Wang et al. 2014).

Compared to Wang et al.’s (2014) formation, the contribution from a static BEC is considered. As shown in Wang et al. (2014) and single observation assimilation
experiments that will be depicted in Section 3, the analysis increments from BEC-CVA include small scale noises. Addition of the static BEC term will reduce the weight of the inhomogeneous and anisotropic BEC and add smooth analysis increment components into the final analysis at the same time. Another benefit is that it expands the analysis increment in solution subspace. In practice, the BEC in Eq.1 can combine a static B from a longer period of NMC-type forecast differences with an anisotropic and inhomogeneous BEC from another shorter period of forecast error samples. If the latter replaced by ensemble forecast perturbations at an analysis time, then it is the standard hybrid scheme defined by Wang et al. (2008a).

3. The Experiment Setup and Rainfall Event

3.1 Experiment Setup

The Advanced Research Weather Research and Forecasting (ARW-WRF) Model (V3.6.1) and WRFDA are used to investigate impacts of BEC. All experiments are conducted over a single domain that covers the Yangtze-Huaihe area and its surrounding areas with a 181×151 horizontal mesh grid employing 12-km spacing and 41 vertical levels up to 50hPa. The initial and boundary fields are interpolated from the NCEP GFS 0.5°×0.5° analyses and forecasts. The WRF physics components are the WRF single-moment five-class (WSM) microphysics scheme (Hong et al. 2004), the Yonsei University (YSU) boundary layer scheme (Hong et al. 2006), the Kain–Fritsch cumulus parameterization scheme (Kain and Fritsch 1990), the Rapid Radiative Transfer Model (RRTM) long wave radiation scheme (Mlawer et al. 1997), and the Dudhia shortwave radiation scheme (Dudhia 1989).
A variety of conventional and satellite observations can be assimilated by the WRFDA (Barker et al, 2012). The data assimilated include surface and upper-air observations of temperature, wind, surface pressure, and specific humidity in addition to aircraft reports of temperature and wind. Furthermore, satellite-tracked wind (SATOP) and aircraft reports (AIREP) observations are assimilated. Surface observations from surface synoptic observation (SYNOP), and aviation routine weather report (METAR) platforms are also assimilated. Observations taken within 6h window of each analysis are assimilated. All observations are assumed to be valid at the analysis time for each experiment.

The NMC-type forecast differences are used to derive BECs in all the experiments. For the isotropic and homogeneous BEC as denoted in $B$ in Eq.1, WRF 12h and 24h forecast differences valid at the same time during a two-month period (1 June-31 July 2010) are utilized as inputs for generating an isotropic and homogeneous BEC using the “GEN-BE” code in WRFDA system with Control Variables Transform option CV5 (Chen et al. 2013). The forecasts were initiated at 00 and 12 UTC each day.

The anisotropic and inhomogeneous BEC presented by the alpha control variable in Eq. 1 is estimated by using the NMC-type forecast differences as well, but with different sampling period that starts from 00 UTC 15 June and ends at 18 UTC 14 July 2011. The forecast differences are generated every 6 hours, so there are 118 forecast error samples in total (Fig. 1). Here the two BECs from different sample periods are blended in BEC-BLD to better capture the errors statistics and increase the freedom of analysis solutions.
In order to examine the differences BEC-CVA and BEC-BLD methods to the traditional hybrid approach, a hybrid experiment based on the hybrid ensemble transform Kalman filter (ETKF) and 3DVar (ETKF-3DVar) in WRFDA is also carried out. The hybrid experiment used ETKF scheme to initiate ensemble forecasts that provide flow-dependent BEC, a 32-member ensemble was used. After 6 hours of spin-up, the initial ensemble at the very beginning of the data assimilation cycles and the lateral boundary condition ensembles during the cycles were generated by adding 32 perturbations to GFS analyses. These perturbations were sampled from a static background error covariance (named as “RandomCV” in WRFDA) (Wang et al.2008a,b). Then the ensembles in following cycle were updated by ETKF. The weighting coefficient of ensemble-based BEC is 75%; the horizontal covariance localization length scale is 200km. The static BEC is used as describe before. More detailed descriptions on the ETKF-3DVar method are referred to papers by Wang et al. (2008a, b).
3.2 Rainfall Event

Heavy rainfall events took place in the Yangtze-Huaihe River Basin in China during summer season usually caused economical and life losses in east China. The experimental period starting on 17 July and ending on 9 August 2011 covered a few precipitation events.

Fig. 2a displays the 24h accumulated precipitation reported by the China Hourly Merged Precipitation Analysis (CHMPA) (Shen et al. 2014), starting from 12UTC 18 to 12UTC 19 July. The 850hPa large-scale horizontal winds and precipitable water at 12 UTC of 18 July 2011 are shown in Fig. 2b and Fig. 2c, which are derived from Final global tropospheric analyses produced by NCEP's Global Forecast System (GFS) (Rutledge et al. 2006). The 24h accumulated precipitation distribution shows a South-North (S-N) rainfall belt, and the heaviest rainfall center (exceeding 100mm) is located at 32.5°N-34.5°N along 117.5°E (Fig. 2a). The southeasterly flow associated with the cyclonic vortex occurs around (27°N, 119°E) (Fig. 2b) and strong water vapor flux convergence (Fig. 2c) provide a favorable environment for the heavy rainfall event.
Fig. 2 (a) The 24h accumulated precipitation (shaded; mm) from the China Hourly Merged Precipitation Analysis initial at 12 UTC 18 July 2011, red box area is the main rainfall area. (b) the horizontal wind (vector; m/s) at 850hPa and precipitable water (shaded; mm) derived from GFS analyses at 12 UTC 18 July 2011; (c) water vapor flux convergence at 850hPa (contour; $10^{-6}$g*cm$^{-2}$*(hPa*s)$^{-1}$; dotted contours with negative value represent convergence) at 12 UTC of 18 July 2011 is shown in Fig. 2c, which is derived from MICAPS (Meteorological Information Comprehensive Analysis and Processing System).

4. Single Observation Experiments Results

Before performing real data assimilation experiments, 7 single observation assimilation experiments (Table 1) are conducted to demonstrate analysis increment difference among 3DVar, ETKF-3DVar, BEC-CVA and BEC-BLD when different weights given to inhomogeneous BEC and different horizontal localization length scales are used. The experiment BEC-BLD_J0.75_L200 (BEC-BLD_J0.50_L200) means the weighting coefficient of the inhomogeneous ensemble-based BEC is 75% (50%) and the horizontal localization length scale is 200 km. The experiment BEC-BLD_J0.75_L100 (BEC-BLD_J0.75_L300) means the weighting coefficient of the inhomogeneous ensemble-based BEC is 75% and the horizontal localization length scale is 100 km (300km). The BEC-CVA (3DVar) is a specific application of the blended method giving the weight with value of 100% (0) to the inhomogeneous BEC. The location of the single observation is at (32°N, 111°E) and 21$^{st}$ vertical level (approximate pressure level 650 hPa).
Structures of Specific Humidity Increments

Fig. 3 shows the specific humidity increments at the 21st level as a result of assimilating a single specific humidity observation (the innovation of specific humidity is 0.001 g kg$^{-1}$, and the observation error is 0.001 g kg$^{-1}$). The pattern of the specific humidity increment in 3DVar shows characteristics of isotropy and homogeneity (Fig. 3a). In contrast, the increments from the BEC-BLD, BEC-CVA and ETKF-3DVar experiments are clearly characterized by the anisotropic and inhomogeneous features (Figs. 3b-g). The increments of specific humidity in the BEC-BLD, BEC-CVA and ETKF-3DVar experiments are more physically reasonable than 3DVar since the increments are extended along the background specific humidity with large gradients. It is also seen that the amplitude of the increments in ETKF-3DVar is less than the increments of BEC-BLD and BEC-CVA, which indicates that the variance of the moisture background errors of BEC-BLD and BEC-CVA are larger than ETKF-3DVar.

Comparing the three BEC-BLD experiments with different horizontal localization length scales (100km, 200km and 300km) (Figs. 3c, 3e and 3f), it is obvious that the smaller

<table>
<thead>
<tr>
<th>Experiment name</th>
<th>Weighting coefficient ($1/\beta^2$)</th>
<th>Horizontal localization scaling</th>
</tr>
</thead>
<tbody>
<tr>
<td>3DVar</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>BEC-BLD_J0.50_L200</td>
<td>0.50</td>
<td>200</td>
</tr>
<tr>
<td>BEC-BLD_J0.75_L200</td>
<td>0.75</td>
<td>200</td>
</tr>
<tr>
<td>BEC-CVA</td>
<td>1.00</td>
<td>200</td>
</tr>
<tr>
<td>BEC-BLD_J0.75_L100</td>
<td>0.75</td>
<td>100</td>
</tr>
<tr>
<td>BEC-BLD_J0.75_L300</td>
<td>0.75</td>
<td>300</td>
</tr>
<tr>
<td>ETKF-3DVar</td>
<td>0.75</td>
<td>200</td>
</tr>
</tbody>
</table>

Table 1. List of the experiments of single observation tests
horizontal localization length scale restricts the increments to a local area around the position of single observation.

4.2 Multivariate Analyses

Temperature and wind increments, resulting from the assimilation of a single specific humidity observation, are shown in Figs. 4 and 5.

The temperature and wind increments in 3DVar are zero (Figs. 4a, 5a), because the control variables adopted in 3DVar for this study are the default control variable options (CV option=5) in WRFDA that does not take into account the correlations between the moisture and other control variables (Chen et al. 2013).

Figs. 4b-4g show that the patterns of temperature increment by assimilating a single specific humidity observation in the BEC-BLD experiments, BEC-CVA and ETKF-3DVar distribute roughly along the region where the background temperature has large gradients. And in Figs. 5b-5f, there are noticeable cyclonic wind increments near the location of the low-pressure system by assimilating a single specific humidity observation. The results show that the multivariate correlations between moisture, temperature and wind can be modeled by the blended BEC. From Figs. 4e-f and Figs. 5e-f, one can also conclude that the smaller horizontal localization length scale limits the increments to a local area around the single observation.
Fig. 3. The specific humidity increments (shaded; g/Kg) at the 21st level as a result of assimilating a single specific humidity observation (a) 3DVar, (b) BEC-BLD_J0.50_L200, (c) BEC-BLD_J0.75_L200, (d) BEC-CVA, (e) BEC-BLD_J0.75_L100, (f) BEC-BLD_J0.75_L300, (g) ETKF-3DVar. The black contours in (a)-(g) are the background specific humidity (g/Kg) at the 21st level, which is derived from GFS analyses at 00 UTC 17 July 2011.
Fig. 4. The same as Fig. 3 but for temperature increments (shaded; K). The black contours in (a)-(g) are the background temperature (K) at the 21st level, which is derived from GFS analyses at 00 UTC 17 July 2011.
Fig. 5. The same as Fig.3 but for wind increments (vector; m/s). The color shades in (a)-(g) are the pressure field (hPa) at the 21st level, which is derived from GFS analyses at 00 UTC 17 July 2011.
4.3 The smooth effect

To show the noise problem caused by sampling error in BEC-CVA more clearly, the temperature (T) increment values along the South-North (Fig. 6a) and West-East (Fig. 6b) which go through the single observation point (32°N, 111°E) by assimilating a single temperature observation (the innovation of temperature is 1.0 K, and the observation error is 1.0 K) with 3DVar, BEC-CVA and BEC-BLD are compared. It should be noted that the weighting coefficient in BEC-BLD is 50% and the length scales is 200km.

It can be seen that the increment of 3DVAR is the unimodal shape with smooth curve; and the general increment structure of BEC-CVA and BEC-BLD are similar to 3DVAR, while parts of the non-smooth feature in BEC-CVA may be caused by sampling error. There still have some noisy features in BEC-BLD. However, due to the blending static BEC and inhomogeneous BEC, BEC-BLD is smoother than BEC-CVA and closer to 3DVAR. Particularly, the smooth effect is more obvious in the distance from the single observation point, which indicate that the noise in BEC-CVA can be reduced by BEC-BLD.
Fig. 6. The temperature increments (K) along two lines that go through the single observation point (32°N, 111°E) at the 21st level as a result of assimilating a single temperature observation. (a) South-North (b) West-East. The weighting coefficient in BEC-BLD is 50% and the length scales is 200km.
5. Continuously cycling data assimilation and forecasting

A series of 6-hourly 3-weeks period (00UTC 17 July–18UTC 9 August, 2011) the continuously cycling data assimilation and forecasting experiments (Table 2) are carried out using 3DVar, BEC-BLD, BEC-CVA and ETKF-3DVar. The weighting coefficient of ensemble-based BEC in ETKF-3DVar, BEC-BLD and BEC-CVA are 75%, 75% and 100% respectively. The background for each following assimilation cycle is the 6-h forecast of the previous cycle. 24-h forecasts are made every 6h. In this section, the computational costs, the root-mean-square errors (RMSE) and accumulated precipitation Fractions Skill Score (FSS) (Roberts et al. 2008) for the 3-weeks cycling are discussed.

Table 2. List of the 3-weeks cycling experiments

<table>
<thead>
<tr>
<th>Experiment name</th>
<th>Weighting coefficient ($\frac{1}{\beta^2}$)</th>
<th>Horizontal localization scaling</th>
</tr>
</thead>
<tbody>
<tr>
<td>3DVar</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>BEC-BLD</td>
<td>0.75</td>
<td>200</td>
</tr>
<tr>
<td>BEC-CVA</td>
<td>1.00</td>
<td>200</td>
</tr>
<tr>
<td>ETKF-3DVar</td>
<td>0.75</td>
<td>200</td>
</tr>
</tbody>
</table>

5.1 Computational Cost

Compared to 3DVar, the additional computation cost of ETKF-3DVar mainly comes from: 1) the ensemble Kalman filter analysis; 2) the ensemble forecasts; 3) the computation of extended control variables in variational cost function. The first two steps take up most of the additional cost in ETKF-3DVar, while BEC-BLD avoids the two steps.

Table 3 lists the total computation cost (wall clock time), on a Linux workstation with 32 CPU processors, used by the experiments 3DVar, BEC-BLD, and ETKF-3DVar for the 3-weeks cycling. The weight to ensemble-based BEC in BEC-BLD and ETKF-3DVar is 75%. It can be seen that 3DVar uses 756 minutes. Compared to 3DVar, the BEC-BLD
experiment adds 195 minutes because of the use of extended control variables in analysis step. However, ETKF-3DVar needs the ensemble forecasts and the corresponding EnKF analysis, thus ETKF-3DVar with 32 members uses 6792 minutes of wall clock time, which is about 9 times of 3DVar and 7 times of BEC-BLD. Such computational cost might be attractive for real-time implementations for some operational centers and research communities with limited computational resources.

Table 3. Total computational cost of the 3-weeks cycling data assimilation and forecasting

<table>
<thead>
<tr>
<th>Experiment name</th>
<th>Weighting coefficient (1/\beta^2)</th>
<th>Total Computational Cost (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3DVar</td>
<td>-</td>
<td>756</td>
</tr>
<tr>
<td>BEC-BLD</td>
<td>0.75</td>
<td>951</td>
</tr>
<tr>
<td>ETKF-3DVar</td>
<td>0.75</td>
<td>6792</td>
</tr>
</tbody>
</table>

5.2 Verification Score

In this sub-section, the RMSE from the continuously cycling data assimilation and forecasting experiments over 3-weeks period are calculated against the GFS analyses. Fig. 7 displays the vertical profiles of mean analysis RMSE against GFS analyses for U, V, T, Q. The RMSE in ETKF-3DVar, BEC-CVA and BEC-BLD are smaller than 3DVar. The ETKF-3DVar performs best among all the experiments. The BEC-BLD slightly reduces the RMSE in wind fields compared to BEC-CVA.
Fig. 7. Vertical RMSE profiles of averaged analysis against GFS analyses. Red line denotes 3DVar, black line denotes BEC-CVA, orange line denotes BEC-BLD, blue line denotes ETKF-3DVar. Error bars show the confidence interval of the mean RMSE for that level (95% confidence limit).

Fig. 8 shows time series of averaged RMSE against GFS analyses. It is seen that BEC-BLD and BEC-CVA are better than 3DVar, and worse than ETKF-3DVar for almost all forecast time. The RMSE difference of the four experiments become smaller with longer forecasting time, since the effect of the initial field decreased gradually.

Fig. 9 shows time series of averaged FSS of 6h accumulated precipitation over the 3-weeks (Fig.9a) and its statistically significance at 90% level (Fig.9b). It can be seen that BEC-BLD and BEC-CVA is better than 3DVar but worse than ETKF-3DVar for almost all forecast time. Furthermore, it is seen that comparing to 3DVar and BEC-CVA, the refinement of the use of BEC-BLD can improve precipitation forecasts slightly but not significant.
Fig. 8. Time series of averaged RMSE following forecast time (00, 06, 12, 24) against GFS analyses. Red line denotes 3DVar, black line denotes BEC-CVA, orange line denotes BEC-BLD, blue line denotes ETKF-3DVar. Error bars show the confidence interval of the averaged RMSE for that forecast time (95% confidence limit).
6. More details for a rainfall event

The first rainfall event during 17–19 July in the 3-weeks period was examined in details in the section.

6.1 Hourly accumulated precipitation FSS

The hourly accumulated precipitation FSS is shown in Fig. 10. The horizontal scale of FSS in this study is 24km. And the FSS is averaged over the 9 forecasts, which are initialized every 6 hours (00, 06, 12, 18) during the two-day assimilation cycles (00 UTC 17 to 00 UTC 19 July 2011).

The hourly accumulated precipitation FSS with a threshold of 1 mm h$^{-1}$ in the BEC-BLD experiments with different weighting coefficients compared to 3DVar, BEC-CVA and ETKF-3DVar, are shown in Fig. 10a. The scores of the BEC-BLD and BEC-CVA experiments are better than that of the 3DVar, worse than that of ETKF-3DVar, for almost
all forecast times. For the two BEC-BLD experiments, the averaged FSS values in BEC-BLD_J0.75_L200 are clearly greater than BEC-BLD_J0.50_L200 when the forecast time is less than 18h, and after 18h BEC-BLD_J0.50_L200 is the best one.

Fig. 10b presents the FSS of BEC-BLD experiments with different horizontal localized scale. It can be seen that the FSS of the BEC-BLD experiments are greater than that of the 3DVar. The FSS of BEC-BLD_J0.75_L200 is the best one before 21h, and BEC-BLD_J0.75_L200 is worse than BEC-BLD_J0.75_L100 after 21h. Therefore, horizontal localization 200km and weighting coefficients 75%, are appropriate for the simulation of this rainfall case.
Fig. 10. The hourly accumulated precipitation FSS (averaged over the 9 forecasts which are initialized every 6 hours during two-day assimilation cycles) with thresholds of 1 mm h$^{-1}$. (a) the 3DVar, BEC-CVA, ETKF-3DVar and the BEC-BLD experiments with different weighted coefficients, (b) the 3DVar and the BEC-BLD experiments with different localized length scales. The horizontal axis is the forecast time (h), the vertical coordinates is the average value of FSS.

6.2 Diagnosis for the rainfall case

To better understand the performance of the different experiments more detail, precipitation distribution, vertical velocity and vapor flux divergence for the rainfall case, are diagnosed in this section. For the sake of brevity, the results of BEC-BLD_J0.75_L200 are selected to represent BEC-BLD. The results of BEC-BLD_J0.75_L100, BEC-BLD_J0.75_L300, and BEC-BLD_J0.50_L200 will not be shown.

1) Precipitation distribution

The 24h accumulated precipitation (shaded; mm) initialized at 12 UTC 18 July 2011 are shown in Fig 11. Fig. 11a shows the real 24h accumulated precipitation from CHMPA, and Figs. 11b-e are the simulated 24h accumulated precipitation initiated at 12 UTC 18 July 2011. The precipitation from CHMPA shows a rainfall band that extends roughly in the N-S direction, and the rainfall center is located at approximately 32.5°N-34.5°N along 117.5°E. It can be seen that the accumulated precipitation amount in 3DVar (Fig. 11b) is
much lower than CHMPA, and BEC-CVA (Fig. 11d) is closer to that of CHMPA than 3DVar, but the precipitation distribution of BEC-CVA is still different from CHMPA. The location of the rainfall center and precipitation distribution of BEC-BLD (Fig. 11c) is closer to CHMPA than 3DVar and BEC-CVA. For the location of the rainfall center in ETKF-3DVar (Fig. 11e) is closest to that of CHMPA, but the rainfall area, especially where the precipitation is over 50mm, of ETKF-3DVar is smaller than that of CHMPA.
Fig. 11. The 24h accumulated precipitation (shaded; mm) initialized at 12 UTC 18 July 2011. (a) CHMPA, (b) 3DVar, (c) BEC-BLD, (d) BEC-CVA, (e) ETKF-3DVar.
(2) Vapor Flux Convergence

The vapor flux convergence is useful to diagnose the potential intensity of the precipitation. The 6h forecast for vapor flux divergence in the main rainfall region at 850hPa initialized at 12 UTC 18 July 2011 are presented in Fig. 12. The negative vapor flux divergence means water vapor convergence. It is seen that 3DVar has weak vapor flux convergence in the rainfall center (Fig. 12a), which also explains the weaker model-simulated precipitation in 3DVar (Fig. 11b). Comparing 3DVar which is more southerly than rainfall center located at 32.5°N-34.5°N along 117.5°E, the distributions of water vapor convergence in BEC-BLD and BEC-CVA (Figs. 12b, 12c) are consist with the rainfall distribution (Fig. 11c, Fig. 11d). And the convergence intensity in BEC-BLD (Fig. 12b) is larger than BEC-CVA (Fig. 12c), which can be one of the reasons that the area of 24h accumulated precipitation bigger than 50mm in BEC-BLD is larger and closer to CHMPA than BEC-CVA. And for the ETKF-3DVar (Fig. 12d), the location of water vapor convergence is more similar to the rainfall center and the intensity is greater than that of other experiments.
Fig. 12. The 6h forecasts vapor flux divergence (shaded; gcm-2hPa-1s-1) at 850hPa initialized at 12 UTC 18 July 2011. (a) 3DVar, (b) BEC-BLD, (c) BEC-CVA, (d) ETKF-3DVar.

7 Conclusions and Discussions

In this paper, the inhomogeneous and anisotropic BEC from BEC-CVA approach and the homogenous and isotropic BEC from WRF-Var BEC modeling are blended within the hybrid framework of WRFDA system. The performance of the blended BEC (BEC-BLD) was assessed by conducting single observation experiments and 3-weeks continuously cycling data assimilation and forecasting experiments, and the details were discussed of a heavy rainfall case in the 3-weeks period that occurred over the Yangtze-Huaihe River Basin China.
Single observation assimilation experiments indicate that by using BEC-BLD with the blended BEC, the noises produced by the BEC-CVA approach caused by the sample errors are reduced, and the multivariate correlations between moisture and other control variables are introduced. Furthermore, the increments of the BEC-BLD experiments using the blended BEC are anisotropic and inhomogeneous.

The 3-weeks cycling data assimilation and forecasting experiments show that the BEC-BLD experiments perform better than the 3DVar and BEC-CVA experiments both in analyses and precipitation forecasts. The diagnostic study on the rainfall case occurred during 17-19 July 2011 shows that, compared with 3DVar and BEC-CVA, the BEC-BLD experiments provide more physically favorable dynamical and water vapor environments for the heavy precipitation event and thus lead to an improvement in the location and intensity of the precipitation forecast.

It is noted that ETKF-3DVar produced the best forecasts among all the experiments, because ETKF-3DVar uses the flow-dependent BEC. However, the ETKF-3DVar requires an ensemble of real-time short-term forecasts to provide the flow-dependent BEC. Whereas, the BEC-CVA or BEC-BLD method only requires an ensemble of historical forecast errors samples and its computational cost is similar to 3DVar and notably less than ETKF-3DVar (hybrid) method. Such computational cost might be important for real-time implementations for some operational centers and research communities with very limited computational resources. The experiment results are sensitive to horizontal localization scales. However, it is not easy to choose one optimal localization scale for all the experiments since forecast errors exist multiple scale features. A better solution is to develop multi-scale localization schemes which is a subject of future works.
Acknowledgements

This work is jointly sponsored by the National Natural Science Foundation of China (41675102), the National Key Research and Development Program of China (2017YFC1502102), the Special Fund for Meteorological Scientific Research in Public Interest (GYHY201506002) and the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD). The authors also thank Dr. Juanzhen Sun and Dr. Zhiqian Liu at NCAR, and Dr. Tanya Peevey at ERSL/NOAA for improving the presentation of this manuscript.

References


Barker, D. M., W. Huang, Y.-R. Guo, A. Bourgeois, and Q. Xiao, 2004: A three-


Dudhia, J., 1989: Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *Journal of the Atmospheric


Kain, J. S., and J. M. Fritsch, 1990: A one-dimensional entraining/detraining plume model


