A DRP-4DVar-based Data Assimilation System for Global NWPs: System Description and Observing System Simulation Experiment

S. J. Zhu[1], B. Wang[2,1,4], L. Zhang[3], J. J. Liu[2,4], Y. Z. Liu[3], J. D. Gong[3], S. M. Xu[1], Y. Wang[1], W. Y. Huang[1], L. Liu[1], Y. J. He[2] and X. J. Wu[3]

[1]Department of Earth System Science, Tsinghua University, Beijing 100084, China.

[2]State Key Laboratory of Numerical Modelling for Atmospheric Sciences and Geophysical Fluid Dynamics, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, 100029, China.

[3]Research and Development Division, Numerical Weather Prediction Center of China Meteorological Administration, Beijing, 100081, China

[4]School of Ocean, University of Chinese Academy of Sciences, Qingdao, 266400, China.

Corresponding author: Bin Wang (wab@lasg.iap.ac.cn)

Key Points:

- A DRP-4DVar-based 4DEnVar system with flow-dependent BEC was developed for global numerical weather predictions
- The DRP-4DVar system significantly outperforms the 4DVar system in terms of improving analyses and forecasts in data-sparse areas
- Higher quality of analyses and forecasts can be produced by the DRP-4DVar system relative to the 4DVar system

Abstract

A four-dimensional ensemble-variational hybrid data assimilation (DA) system based on the dimension-reduced projection (DRP) technique was developed for global numerical weather predictions (NWPs). Instead of the adjoint technique, an ensemble approach is utilized in this technique to calculate the gradient of the cost function of the standard four-dimensional variational (4DVar) DA. The flow-dependence of the background error covariance (BEC) is realized in the variational configuration by dynamically updating the initial perturbation samples during the assimilation cycle. A limited number of leading eigenvectors of the correlation function of localization are selected to filter out the spurious correlations in the BEC matrix (B-matrix). A linear combination of the ensemble analysis increment sample with the random perturbation sample satisfying balance constraints is used as the inflation technique to prevent the BEC from underestimation and to achieve the hybrid of the flow-dependent and static B-matrices when updating the initial perturbations. In order to evaluate the new system, single-point observation experiments (SOEs) and observing system simulation experiments (OSSEs) were conducted with sounding and cloud-derived wind data. The flow-dependent characteristic was verified by the SOEs that utilized localized ensemble covariance. In the OSSEs, DRP-4DVar produced better analysis than 4DVar. Moreover, the ensemble mean forecast with DRP-4DVar analyses as the initial conditions reduced errors in geopotential height, 24-h accumulated precipitation and other variables relative to the 4DVar-based forecast. Significant improvements of analysis and forecast achieved in the data-sparse Southern Hemisphere by DRP-4DVar indicate a remarkable advantage of the ensemble covariance in sufficient use of sparse observations.

Plain Language Summary

Medium-range numerical weather prediction (NWP) is of great significance to disaster mitigation and to the improvement of human living standards. It is typically calculated by obtaining a future 1-10 days weather forecast using an initial state with a global NWP model based on a set of partial differential equations. Therefore, a high quality initial state is crucial to the medium-range NWP, which is produced using a data assimilation (DA) system combining observations and model constraints. The four-dimensional variational (4DVar) method is recognized as one of the most advanced DA methods, but it still has its disadvantages. The adjoint

technique for minimizing its cost function limits its application in operational NWP centers worldwide. Also, static estimations of background error covariance (BEC) that determines its analysis quality limits further improvements to its analysis accuracy. To overcome these disadvantages, this study aims to develop a DA system for global NWPs with an adjoint-free technique, i.e., the dimension-reduced projection (DRP) technique, to minimize the cost function and to estimate BEC dynamically. Compared with the 4DVar system, the DRP-4DVar system can further reduce analysis errors of basic state variables and also improve the forecasts of these variables, especially in the data-sparse areas.

1 Introduction

Accurately predicting future weather and climate states is of great significance to disaster mitigation and to the improvement of human living standards. The accuracy of global numerical weather prediction (NWP) can be significantly improved through the use of new types of data such as from satellites (Simmons & Hollingsworth, 2002). Therefore, it is necessary to develop an effective data assimilation (DA) system to make good use of observations to provide more accurate initial conditions (ICs) for NWPs.

The four-dimensional variational (4DVar) DA is recognized as one of the most advanced DA methods. This method produces the analysis field constrained dynamically and physically (Rabier et al., 2000; Wang et al., 2010a), and implicitly implements the flow-dependent background error covariance (BEC) matrix (simply B-matrix), which propagates information within the assimilation window by the tangent linear model (TLM) and the adjoint model (ADM; Lorenc, 2003). The uses of the ADM (Lewis & Derber, 1985; Le Dimet & Talagrand, 1986) and the incremental 4DVar scheme (Courtier et al. 1994) make the operational application of 4DVar possible (Rabier et al., 2000; Gauthier & Thépaut, 2001; Koizumi et al., 2005; Rawlins et al., 2007; Gauthier et al., 2007; Zhang et al., 2019). However, this advanced DA method has not been applied in most NWP centers in the world except very few major advanced centers, e.g., the European Centre for Medium-range Weather Forecasts (ECMWF), due to the use of the adjoint model of which the development and maintenance are costly and difficult to keep abreast of the development and updating of the corresponding forecast model. In addition, the 4DVar approach fails to dynamically update the B-matrix during an assimilation cycle currently, given that the same static covariance model is used at the beginning of each assimilation window (Buehner et al., 2010a).

Ensemble Kalman Filter (EnKF) is also a commonly used ensemble DA method. It uses a finite number of ensembles to estimate the globally flow-dependent B-matrix using the Monte Carlo method (Evensen, 1994), while avoiding the modeling of the B-matrix and the use of the ADM. Moreover, EnKF has the advantage of saving time, implicitly through concurrently generating the ensembles on a parallel computer system due to the mutual independence of ensemble members. Due to these advantages, EnKF has many applications (Houtekamer & Mitchell, 1998, 2001; Houtekamer et al., 2005; Whitaker et al., 2008, 2009; Buehner et al., 2010a, 2010b). There have been some studies comparing the performance of the variational and EnKF systems. Whitaker et al. (2008) compared 3DVar and EnKF using low-resolution operational model and observations, except satellite radiation, and found that the ensemble system outperforms the 3DVar system, especially in data-sparse areas. Whitaker et al. (2009) further compared 3DVar, 4DVar and EnKF using sparse surface pressure observations, and discovered that 4DVar and EnKF have comparable performance. Buehner et al. (2010b) found a slight degradation (improvement) in the short-term (medium-range) forecasts based on the EnKF ensemble mean analysis over the 4DVar-based forecasts in the extratropics. There is not enough evidence to prove that the analysis provided by EnKF is better than that provided by 4DVar systems for the NWP models. Also, it is noted that the limited size of ensembles can result in sampling errors in the ensemble B-matrix.

However, the ensemble method may provide the estimation of the B-matrix that evolves with flows for the variational method. Likewise, the variational method can supply the ensemble method with proven modules, e.g., quality control and minimization iteration modules. Therefore, several hybrid DA methods combining the variational and ensemble ideas have continuously been developed (Hamill & Snyder, 2000; Lorenc, 2003; Qiu et al., 2007; Liu et al., 2008, 2009; Tian et al., 2008, 2011; Wang et al., 2010a).

The main idea of the hybrid DA system is to incorporate the ensemble covariance into the variational framework, which can be achieved in different ways. What are the effects of introducing the ensemble covariance into a variational system? Can the ADM be avoided in the minimization iterative procedure? Can the covariance localization be implemented in subspaces spanned by the ensemble members?

Different methods of incorporating the ensemble covariance make the classification of hybrid methods different. The hybrid ensemble-4DVar methods are mainly divided into En4DVar methods that include the ADM and 4DEnVar methods that avoid the ADM. En4DVar methods typically incorporate the ensemble covariance into the variational framework by a weighted sum of the static and ensemble covariances (Hamill & Snyder, 2000) or extending the original control variables by the control variables preconditioned by the square root of the ensemble covariance (Lorenc, 2003). In addition, the effects of the hybrid BEC on forecast skills have been investigated in simple models (Hamill & Snyder, 2000), regional models (Wang et al., 2008a, 2008b; Zhang & Zhang, 2012) and global models (Raynaud et al., 2011; Bonavita et al., 2012; Buehner et al., 2010a, 2010b, 2013, 2015; Clayton et al., 2013; Lorenc, 2015; Wang et al., 2013; Wang & Lei, 2014; Kleist & Ide, 2015a, 2015b).

Avoiding the use of the ADM plays a crucial role in hybrid DA methods, due to the difficulty in producing the ADM as well as updating it following the rapid development of the corresponding forecast model. Thus, 4DEnVar method, which applies the variational framework and the idea of using ensembles valid at multiple time slots to avoid the ADM to obtain the optimal analysis, is a promising DA method. Several ensemble-based methods, which can reduce the dimension from the model space to a subspace composed of a limited number of base vectors in optimization and avoid the use of the ADM, have been proposed in recent decade (Qiu et al., 2007; Tian et al., 2008; Wang et al., 2010a). The dimesion-reduced projection 4DVar (DRP-4DVar) is one of the 4DEnVar methods that has been successfully applied in regional meso-scale weather forecasts (Wang et al., 2010a; Zhao & Wang, 2010; Liu & Wang, 2011; Zhao et al., 2012) and global decadal climate predictions (He et al., 2017, 2020a, 2020b; Li et al., 2021a, 2021b; Shi et al., 2021). In global medium-range NWPs, the DRP-4DVar approach has not been widely applied and systematically evaluated, although a DRP-4DVar system (Shen et al., 2015) was preliminarily established using an old version of the global forecast system of the Global/Regional Assimilation and Prediction System (GRAPES-GFS) based on the 3DVar system of this version (Chen et al., 2008; Xue et al., 2008). This method uses a limited number of base vectors of initial perturbation to project the incremental analysis in model space onto a low-dimensional subspace spanned by these base vectors, and directly obtains an optimal analysis solution to the minimization of the 4DVar cost function in the subspace. Furthermore, this method calculates the gradient of the cost function based on the statistical relationship between the model space and observation space, thereby avoiding the use of the ADM (Wang et al., 2010a).

The limited ensemble size is easy to introduce sampling errors, which can result in spurious correlations in the B-matrix (Evensen, 2003), and localization techniques (Liu et al., 2009; Hamill et al., 2001; Wang et al., 2010b, 2018) can effectively alleviate the aforementioned problem and ameliorate analyses and forecasts. In comparison to the observation space localization, implementing localization in model space is confirmed to be more beneficial to analyses and forecasts when assimilating the data of which the locations are not well defined, such as satellite radiances (Campbell et al., 2010). However, the especially troubling thing is that, conducting localization in model space is quite inconvenient and computationally expensive. Adopting ensemble-sample-based subspace localization schemes is thought to be an economical choice (Wang et al, 2018). Localization is typically conducted as a Schür product between the ensemble-based B-matrix and the correlation matrix of which the elements are decided by the distances among the corresponding variables, so how to decompose the correlation matrix to avoid the expensive multiplication between high-dimensional matrices caused by the Schür product is the key to reduce computational cost. Liu et al. (2009) used the empirical orthogonal function (EOF) to decompose the correlation function and selected a limited number of eigenvectors on a lower-resolution grid, but this lower-resolution decomposition and its interpolation back to the original resolution grid may lead to a degraded accuracy. To avoid this problem, a limited number of leading eigenvectors expressed by orthogonal functions (e.g., sine function and spherical harmonic function) were used to expand the correlation function so that the high-dimensional correlation matrix is decomposed into the sum of a set of products between an eigenvector and its transpose (Buehner et al., 2010a, 2010b; Bishop et al., 2011; Kuhl et al., 2013; Wang et al., 2010b, 2018). This approach not only alleviates the spurious correlations and rank deficiency of the B-matrix, but also efficiently produces the extended ensemble samples, which converts a very costly schür product between two high-dimension matrices to much more economical schür products between ensemble samples and eigenvectors.

Motivated by these studies, many research and operational centers have not only established their standalone variational systems, but also have been developing hybrid DA systems for their global NWPs. These centers explicitly realized the flow-dependence of the B-matrix during the assimilation cycle based on the original 4DVar framework, so that the forecast skills were further improved. The ECMWF (Bonavita et al., 2012) and Météo-France (Raynaud et al., 2011) have developed hybrid DA systems, which include ensemble BECs estimated by the EDA system based on 4DVar analyses. The Met Office incorporated the flow-dependent BEC estimated by EnKF into the 4DVar system to develop a hybrid system (Clayton et al., 2013). Unlike these systems relying on the ADM, some centers have developed 4DEnVar systems avoiding the use of the ADM. Environment Canada combined the static BEC with the 4D ensemble BEC obtained from EnKF to develop a 4DEnVar system, which is considered to be a potential alternative to 4DVar considering the simplicity, computational efficiency and forecast quality (Buehner et al., 2010a, 2010b, 2013, 2015). The Met Office developed a hybrid 4DEnVar system (Lorenc et al., 2015; Bowler et al., 2017a) and used an ensemble of 4DEnVars instead of the ETKF system to generate ensembles for the hybrid system (Bowler et al., 2017b). Wang et al. (2013) and Kleist & Ide (2015a) proved the benefits of including ensemble BECs into 3DVar. Then, perturbations valid at multiple time slots during the DA window were used to estimate the 4D ensemble BEC to develop a 4DEnVar system in NCEP (Wang et al., 2014; Kleist & Ide., 2015b).

The aim of the paper is to develop a 4DEnVar system based on the DRP-4DVar approach for the hybrid system with low costs in development and computation, and to evaluate it by comparing with the corresponding 4DVar system. The successful applications of the DRP algorithm and economical localization technique provide a good foundation to develop the 4DEnVar system for the new version of the GRAPES-GFS model, after version 2.4 (Su et al., 2020). As the first and necessary step to evaluate the impact of the 4DEnVar system on analyses and forecasts, single-point observation experiments (SOEs) and observing system simulation experiments (OSSEs) were conducted. SOEs are easy to study the flow-dependent characteristic of the BEC. OSSEs can help us evaluate the realistic analysis error because the "truth" state is known. The remainder of the paper is organized as follows. Section 2 introduces the formulation and implementation of the DRP-4DVar system, Section 3 explains the localization technique used for the ensemble covariance, Section 4 follows with the experimental description and the inflation technique used to alleviate the problem of filtering divergence during the assimilation cycle, and Section 5 evaluates the performance of the DRP-4DVar system on analyses and forecasts relative to the 4DVar system. A summary and prospect for future work is presented in the last section.

2 Description of method

2.1 Incremental 4DVar algorithm

The variational system used in this paper (Zhang et al, 2019) adopts the incremental 4DVar scheme (Courtier et al., 1994), which usually obtains the optimal analysis of IC by minimizing a cost function on a low-resolution grid:

$$J(x') = \frac{1}{2}(x')^T B^{-1}(x') + \frac{1}{2}\sum_{k=1}^{K}(y_k' - d_k)^T R^{-1}(y_k' - d_k), \qquad (1)$$

where x' is the perturbation of the IC, B is the static B-matrix, and R is the observation error covariance matrix. $d_k = y_k - H_k M_k x_b$ contains the observation innovations at different time (with subscript k) in the assimilation window, where x_b is the background state vector, y_k

contains the observations at different times (with subscript k), involving the observation operator H_k at k time level and the nonlinear forecast model integration M_k from the analysis time to k time level. $y'_k = \mathbf{H}_k\mathbf{M}_k x'$ contains the projection of the IC perturbation onto the observation variables at different times (with subscript k), $\mathbf{H}_k$ is the tangent linear observation operator corresponding to H_k, and $\mathbf{M}_k$ is the TLM of M_k.

The convergence rate of the gradient for the optimization problem is dependent on the condition number of the Hessian matrix (Zupanski, 1996). Operational DA systems generally reduce the condition number of the Hessian matrix of Eq. (1) through the physical and preconditioning transformations, that is $x' = Uw$. Thus, the static BEC can be estimated by

$$B_c = UU^T, \tag{2}$$

where w is a column vector of the N-dimensional control variables before the physical and preconditioning transformations. U contains the physical transformation operator that transforms independent variables to model variables, the diagonal matrix composed of the background error variance square root of the independent variables, and the background error correlation transformation matrix that can be decomposed into the horizontal correlation matrix by the spectral method and the vertical correlation matrix by the EOF decomposition method (Zhang et al., 2019). After the aforementioned physical and preconditioning transformations, Eq. (1) becomes

$$J(w) = \frac{1}{2}w^T w + \frac{1}{2}\sum_{k=1}^{K}(\mathbf{H}_k\mathbf{M}_k Uw - d_k)^T R^{-1}(\mathbf{H}_k\mathbf{M}_k Uw - d_k). \tag{3}$$

The analysis increment that minimizes Eq. (3) satisfies the necessary condition that the corresponding gradient equals to zero, to which the solution needs to use the TLM and ADM. Moreover, the calculations of the TLM and ADM respectively require the forward and backward model trajectories typically provided by the nonlinear forecast model, which are expensive in development, calculation and storage. For practical difficulties in the use of the TLM and ADM, finding adjoint-free algorithms to implement 4DVar is an important research point for operational applications.

2.2 DRP-4DVar algorithm

The DRP-4DVar approach projects the initial increment x' in model space onto the subspace expanded by a limited number of IC perturbation samples as its basis vectors, and obtains the optimal solution directly in the subspace (Wang et al, 2010a).

For the convenience of implementing DRP-4DVar in the standard 4DVar framework, the IC perturbation samples are obtained by the method of Baker (2005) $X = [x'_1, x'_2, \cdots, x'_s]$, which contains the IC perturbation samples, where s is the ensemble size. The corresponding observational perturbation samples $Y = [y'_1, y'_2, \cdots, y'_s]$ are calculated using the TLM. Thus, an ensemble of IC perturbation samples and observational perturbation samples are chosen to define the following projection matrices:

$$\begin{cases} p_x = \frac{1}{\sqrt{s-1}}[x'_1 - \bar{x}', x'_2 - \bar{x}', \cdots, x'_s - \bar{x}'] \\ p_y = \frac{1}{\sqrt{s-1}}[y'_1 - \bar{y}', y'_2 - \bar{y}', \cdots, y'_s - \bar{y}'] \end{cases}, \tag{4}$$

where $\begin{cases} \bar{x}' = \frac{1}{s}[x'_1 + x'_2 + \cdots + x'_s] \\ \bar{y}' = \frac{1}{s}[y'_1 + y'_2 + \cdots + y'_s] \end{cases}$. Defining $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_s)^T$ as a s-dimensional vector composed of the weight coefficients of the basis vectors, x' and $\mathbf{H}_k\mathbf{M}_k x'$ can be projected onto the subspace spanned by the ensemble samples via the following transformation:

$$\begin{cases} x' = p_x\alpha \\ \mathbf{H}_k\mathbf{M}_k x' = p_{y,k}\alpha \end{cases}, \tag{5}$$

where $p_{y,k}$ is the observational projection matrixes at k time level. Thus, the ensemble BEC can be represented by

$$B_e = p_x p_x^T \tag{6}$$

and the new cost function can be written as:

$$J(\alpha) = \frac{1}{2}\alpha^T\alpha + \frac{1}{2}\sum_{k=1}^{K}\left(p_{y,k}\alpha - d_k\right)^T R^{-1}\left(p_{y,k}\alpha - d_k\right). \tag{7}$$

To minimize Eq. (7), α must satisfy $\nabla_\alpha J = 0$. Here, no ADM is needed. It is noted that a degraded analysis may result from the approximation in Eq. (5) due to the much smaller ensemble size than the dimension of the original IC perturbation, which can be alleviated by localization techniques.

3 Localization

The major drawback to the ensemble-based method is its spurious correlations and very small rank in the BEC due to the limited number of the IC perturbation samples, which excessively constrains the solving subspace of the optimal analysis increment. Localization is considered to be an effective technique to alleviate the aforementioned problems (Hamill et al., 2001).

The localized B-matrix can be typically expressed as the Schür product between the ensemble BEC B_e and the correlation matrix of the covariance localization C. Because the direct use of the localized B-matrix may lead to much computational cost according to Wang et al. (2018), this matrix should be expressed as a form that can be used easily and economically. The correlation matrix can be approximately decomposed into a limited number of leading eigenvectors and extended IC perturbation samples can be obtained:

$$Ep_x = \left[\left(p_{x,1} \circ \boldsymbol{\rho}_{x,1}, \cdots, p_{x,1} \circ \boldsymbol{\rho}_{x,L}\right), \cdots, \left(p_{x,s} \circ \boldsymbol{\rho}_{x,1}, \cdots, p_{x,s} \circ \boldsymbol{\rho}_{x,L}\right)\right], \qquad (8)$$

where $\boldsymbol{\rho}_{x,j}$ $(j = 1, 2, \cdots, L)$ is a leading eigenvector in model space, and L is the number of the selected leading eigenvectors according to the cumulative contribution of variance. In implementation, each leading eigenvector can be decomposed into zonal, meridional and vertical components: $\boldsymbol{\rho}_{x,j} = \boldsymbol{\rho}^{z}_{x,j_z} \circ \boldsymbol{\rho}^{m}_{x,j_m} \circ \boldsymbol{\rho}^{v}_{x,j_v}$. The EOF decomposition method is used to obtain the zonal and vertical components:

$$\begin{cases} \boldsymbol{\rho}^{z}_{x,j_z} = \boldsymbol{E}^{z}_{x,j_z}\left(\lambda^{z}_{x,j_z}\right)^{1/2} \\ \boldsymbol{\rho}^{v}_{x,j_v} = \boldsymbol{E}^{v}_{x,j_v}\left(\lambda^{v}_{x,j_v}\right)^{1/2} \end{cases}, \qquad (9)$$

where $\boldsymbol{E}^{z}_{x,j_z}$ and $\boldsymbol{E}^{v}_{x,j_v}$ are eigenvectors for the zonal and vertical components, respectively, obtained using the empirical orthogonal decomposition. λ^{z}_{x,j_z} and λ^{v}_{x,j_v} are their corresponding eigenvalues. Then, the sine expansion method is utilized (Wang et al., 2018) to obtain the meridional component:

$$\boldsymbol{\rho}^{m}_{x,j_m} = \boldsymbol{E}^{m}_{x,j_m}\beta^{1/2}_{x,j_m}\,. \qquad (10)$$

Here, $\boldsymbol{E}^{m}_{x,j_m}$ is the eigenvector for the meridional component, and β_{x,j_m} is its eigenvalue. When defining the correlation function model, we used the GC correlation function (Gaspari & Cohn, 1999) for the horizontal components:

$$c(r) = \begin{cases} -\frac{1}{4}r^5 + \frac{1}{2}r^4 + \frac{5}{8}r^3 - \frac{5}{3}r^2 + 1, 0 \le r \le 1 \\ \frac{1}{12}r^5 - \frac{1}{2}r^4 + \frac{5}{8}r^3 + \frac{5}{3}r^2 - 5r + 4 - \frac{2}{3}r^{-1}, 1 < r \le 2 \\ 0, 2 < r \end{cases} \tag{11}$$

and the following correlation function for the vertical component:

$$c(r) = \sqrt{ln\{1.0 + K_p[\Delta log(r)]^2\}}. \tag{12}$$

According to Eq. (8) and ignoring the time-variation of the localization leading eigenvectors, the extended observational perturbation samples can be then represented as

$$Ep_y = \left[\left(p_{y,1} \circ \boldsymbol{\rho}_{y,1}, \cdots, p_{y,1} \circ \boldsymbol{\rho}_{y,L}\right), \cdots, \left(p_{y,s} \circ \boldsymbol{\rho}_{y,1}, \cdots, p_{y,s} \circ \boldsymbol{\rho}_{y,L}\right)\right]. \tag{13}$$

Redefining the control variables as an $s \times L$-dimensional vector β, the analysis increment and observational increment can be modified as

$$\begin{cases} x' = Ep_x\beta \\ \mathbf{H}_k\mathbf{M}_k x' = Ep_y\beta \end{cases}. \tag{14}$$

Finally, the localized cost function is formulated on the extended sample space. In the generation of the extended observational perturbation samples, the TLM is called for only s times. On one hand, the ensemble size can be greatly increased from the original samples to the extended samples without any additional computational cost for TLM calling, and the leading eigenvectors in both the model and observation space can be pre-calculated according to the coordinates of the model grid and observation locations. On the other hand, spurious correlations among the original samples can be significantly eliminated, and the calculation accuracy of the cost function and its gradient can be improved given that the extended samples have better independence from each other than the original samples.

4 Experimental design

4.1 Configuration of DRP-4DVar system

In this study, the model used in the DRP-4DVar system is the GRAPES-GFS model version 3.0, which is the new version after version 2.4 (Su et al., 2020) and contains 87 vertical levels. The horizontal resolution of the system is 0.5° × 0.5° for the outer loop and 1.0° × 1.0° for the inner loop. The DRP-4DVar system combines the ensemble BEC estimated by 60 samples and the original variational framework to solve the assimilation problem, and is evaluated in comparison to the available 4DVar system (Zhang et al, 2019) with the same model and same resolutions. The schematic flowchart in Figure 1 describes the operation process of the DRP-4DVar system. In order to mitigate the sampling errors and spurious correlations in the BEC due to the limited ensemble size (Hamill et al., 2001; Lorenc et al., 2003; Wang et al., 2010b, 2018), the localization scheme is designed according to the implementation introduced in Section 3, with 7° for the filtering radius in the horizontal direction and 3 for the filtering parameter K_p in the vertical direction. The minimization problem of the DRP-4DVar system is solved in the subspace spanned by the extended samples derived from the Schür products between the ensemble members and the leading eigenvectors of the localization correlation function. The DRP-4DVar system not only can implicitly realize the flow-dependence of the BEC from the beginning to the end of a single assimilation window like the standard 4DVar system, but also can explicitly achieve the flow-dependent BEC from one assimilation window to the next. The perturbed observations are assimilated to the ensemble samples, and the flow-dependent samples in different windows are generated during the assimilation cycle at 6-h intervals. 60 random samples with balance constraints are obtained by randomly perturbing the control variable w of the incremental 4DVar system and implementing the physical and preconditioning transformations $(x_i')_r = Uw_i$ $(i = 1, 2, \cdots, 60)$ based on the climate B-matrix of the 4DVar system (Baker, 2005). They are linearly combined with the 60 analysis increment samples $(x_i')_a$ $(i = 1, 2, \cdots, 60)$ with the weighting coefficients $\gamma_1 = 0.2$ for the former and $\gamma_2 = 0.9$ for the latter to achieve the inflation of the BEC for the next assimilation after a model integration from the beginning to the end of the assimilation window with the inflated analysis increments $(x_i')_a^{inf}$ $(i = 1, 2, \cdots, 60)$ as ICs, where

$$\begin{cases} (x_i')_a^{inf} = \gamma_1 (x_i')_r + \gamma_2 (x_i')_a \\ (x_i')_f = M\left(x_b + (x_i')_a^{inf}\right) \end{cases} \quad (i = 1, 2, \cdots, 60). \tag{15}$$

A obvious advantage of this inflation method is that it is convenient and easy to generate the random samples with balance constraints directly through the preconditional process of the 4DVar system. Moreover, the aforementioned inflation technique incorporate the climate BEC into the ensemble BEC to construct the hybrid B-matrix, which can be represent by

$$(B_e)_{inf} = (x_i')_f (x_i')_f^T. \tag{16}$$

Thus, the inflation method not only alleviates the underestimation of the B-matrix, but also implicitly realizes the hybrid BEC for the DRP-4DVar system. Collaborated with the localization and perturbing techniques of observation and SST, the inflation may alleviate the filtering divergence problem during the assimilation cycle. The DRP-4DVar system uses the perturbations of observation and SST to characterize observation errors and boundary analysis errors, respectively. Observational perturbations are obtained by superimposing normal distribution random perturbations with zero as their expectations (or mean values) and the observation errors as their standard deviations onto the observations, and SST perturbations are obtained similarly except that the standard deviations of random perturbations adopts the SST analysis errors.

4.2 Experiment design

In order to evaluate the performance of the DRP-4DVar system efficiently, the OSSE is considered as one of the best choices. Here, two OSSEs are designed using the $0.25° \times 0.25°$ version of GRAPES-GFS for both the DRP-4DVar and standard 4DVar systems. The OSSE for the latter is to provide a reference for comparisons.

A previous study has demonstrated that the 4DVar system using the GRAPES-GFS model significantly outperforms the 3DVar system using the same model on both analyses and medium-range forecasts, especially in the Southern Hemisphere (Zhang et al., 2019). OSSEs can be used to fairly evaluate the performance of the assimilation system (Wang et al., 2008a; Wang et al., 2010a; Kleist et al., 2015a, 2015b). In order to further study the influence of the 4DEnVar system based on the DRP-4DVar algorithm, comparisons between it and the 4DVar system are necessary.

Both experiments utilized a 6-h assimilation window (0900 UTC 13 September 2016 - 1500 UTC 13 September 2016) after a 2-day assimilation cycle covering the period from 0900 UTC 11 September 2016 to 0900 UTC 13 September 2016 to alleviate the influence of the spin-up, and the analysis time was taken at the beginning of the assimilation window. In the OSSEs, the results from an uninterrupted free run for 15 days with the higher-resolution ($0.25° \times 0.25°$) version of GRAPES-GFS were used as the "truth" state, which is crucial to OSSEs. To eliminate the impact of spin-up, the "truth" state was initiated from the time 24 hours prior to the analysis time of the first assimilation window with the ERA-5 reanalysis field as the IC, which was verified to be consistent with the realistic atmospheric state in terms of geopotential height and precipitation. For example, we investigated the rationality of the "truth" state in the Northern and Southern Hemispheres based on a comparison of the 500hPa geopotential height between the ERA-Interim reanalysis and the "truth" state at 1200 UTC on 14, 16 and 18 September 2016 (Figure 2). Figure 2a shows the 500hPa geopotential height from the ERA-Interim reanalysis in the Northern Hemisphere at 1200 UTC on 14 September 2016, with a low-pressure system near the Arctic and 4 troughs extended from the low-pressure system near 60°E, 180°, 120°W and 30°W. The low-pressure system extends along 180° and 30°W , and the locations and intensities of other main systems change slightly as the integration time increases (Figures 2e and 2i). The "truth" state captures these main features and their time-variations (Figures 2b, 2f and 2j). Similarly, Figure 2c shows the results from the ERA-Interim reanalysis in the Southern Hemisphere. A low-pressure system exists near the Antarctic at 180° with 3 troughs near 0°, 90°W and 90°E, and some troughs at low and middle latitudes. As the integration time increases, the intensity of the low-pressure system near the Antarctic weakens and a high value center appears near 60°E, and the locations and intensities of the main systems at low and middle latitudes change slightly (Figures 2g and 2k). The "truth" state simulates these main systems well (Figures 2d, 2h and 2l). In general, the "truth" state reasonably captures the main features and the time-variations of the 500hPa geopotential height from the ERA-Interim reanalysis and gradually degrads following the increase of integration time.

The "observations" were produced by interpolating the "truth" state to the positions at which sounding and cloud-derived wind observations are located, and then superimposing normal distribution random perturbations with zero as their expectations and the observation errors as their standard deviations onto them. Figure 3 shows the spatial distribution of these observations. Sounding observations are typically sampled in the continental areas of the Northern Hemisphere and are valid at 1200 UTC 13 September 2016. Cloud-derived wind observations are sampled every 30 minutes, mainly in the central and eastern North Pacific, the eastern South Pacific, the northern Indian Ocean, the Atlantic Ocean, as well as some continents such as the America and Africa. The observation errors were taken the same as the 4DVar system.

For the first assimilation window of the 4DVar system, the background was obtained from a 15-h forecast by the 0.5° × 0.5° version of GRAPES-GFS with the 6-h forecast of the ERA-

Interim dataset as the IC, so that it is different from the "truth" state. Meanwhile, for the first assimilation window of the DRP-4DVar system, 60 IC samples were generated by superimposing 60 random perturbation samples onto this background. These perturbation samples were generated according to the method of Baker (2005), which has been mentioned in the inflation method introduced in section 4.1. The background for each assimilation window of the DRP-4DVar system is the ensemble mean of the IC samples of this window, which are derived from 60 6-h forecasts by the 0.5° × 0.5° version of GRAPES-GFS with 60 analysis samples produced in the previous assimilation window as their ICs, respectively, except for the first assimilation window. The DRP-4DVar system has the same background as the 4DVar system in the first assimilation window because the ensemble mean of 60 superimposed random perturbation samples is zero.

In addition, based on the OSSE for the DRP-4DVar, two SOEs were also conducted for both the DRP-4DVar and 4DVar systems within a 6-h window covered the period from 0900 UTC 13 September 2016 to 1500 UTC 13 September 2016 after a 2-day assimilation cycle to verify the flow-dependent characteristic of the ensemble-based BEC across the assimilation windows. Both SOEs adopted the same filtering radius that is 15° in the horizontal direction and the same background that is the ensemble mean of the IC samples produced by the DRP-4DVar system. The observation error was set to 0.95, and the observation innovation was -1.543 K. The single-point observation was selected from the "observations" valid at 1200 UTC 13 September in the OSSEs, which is the single-point temperature observation located upstream at the top of the short-wave ridge in the middle troposphere.

5 Results

5.1 Single-point observation experiments

Figure 4 shows the analysis increments from the 4DVar and DRP-4DVar systems at the beginning of the assimilation window (the analysis time) when the same single-point temperature observation is taken at the middle of the assimilation window in two SOEs. Both increments show the minimum values near the observation as a response to the low temperature observation. The 4DVar analysis increment of temperature appears a quasi-Gaussian distribution around the observation location (Figure 4a). Given that there is only 3 hours elapse between the analysis time the resultant analysis increment is obtained and the time the single-point observation is located at, this distribution is reasonable. The minimum value of the increment of both systems shifts towards the northwest of the observation, which shows that both systems have flow-dependent BECs within the assimilation window (Figures 4a and 4b). Moreover, the DRP-4DVar analysis increment of temperature, obtained using the flow-dependent BEC across the assimilation windows, extends along the gradient of geopotential height, which is consistent with the northwestern background flow. Satisfactorily, no spurious correlations are sighted near the analysis increment produced by the DRP-4DVar system when the signal of the observation is preserved in the analysis increment (Figure 4b). Two experiments also show cyclone wind responses around the temperature increment, which suggests that the BECs satisfy some balance constraints. These results are consistent with the SOE introduced in Kleist et al. (2015b).

5.2 Observing system simulation experiments

Considering that the root mean square error (RMSE) mainly measures the random error that is not as correctable as the systematic bias, it is usually applied to statistically analyze the random errors of the background and analysis fields. To exclude the systematic error from the RMSE, we use a metrics called anomaly RMSE (ARMSE) instead of RMSE (He et al. 2020a):

$$ARMSE = \sqrt{\frac{\sum_{n=1}^{N} w_{(n)} \times \left(M_{(n)} - truth_{(n)} - bias\right)^2}{\sum_{n=1}^{N} w_{(n)}}}. \tag{17}$$

Here, $M_{(n)}$ and $truth_{(n)}$ represent the analysis (or background) and the “truth” state at the n-th grid point, respectively. $w_{(n)}$ denotes the weighted coefficients at the n-th grid point, and $bias = \frac{\sum_{n=1}^{N} w_{(n)} \times \left(M_{(n)} - truth_{(n)}\right)}{\sum_{n=1}^{N} w_{(n)}}$ represents the systematic bias.

To facilitate comparisons with reanalysis data that are located at the middle of the assimilation window, all background and analysis fields from both DA systems are transformed from the beginning to the middle of the window through 3-h forecasts using the $0.5° \times 0.5°$ version of GRAPES-GFS, which is similar to Zhang et al. (2019). Figure 5 shows the vertical profiles of the ARMSEs of the background and analysis fields from the 4DVar and DRP-4DVar systems relative to the "truth" state. On one hand, comparing with the background fields, the analysis fields from both assimilation approaches basically improve most variables at most vertical levels. These analyses significantly reduce the ARMSE of zonal wind at almost all vertical levels in the Northern and Southern Hemispheres, Tropics and East Asia (Figures 5a-5d). They improve the temperature at most vertical levels in the Northern and Southern Hemispheres and at the levels between 500hPa and 200hPa in the East Asia, although the temperature from the analyses in the lower troposphere (lower and middle troposphere) in the Northern Hemisphere (East Asia) by the DRP-4DVar system (both the DRP-4DVar and 4DVar systems) is obviously degraded (Figures 5e, 5f and 5h). In the Tropics, no significant differences of temperature between the backgrounds and analyses can be observed (Figure 5g). As for the specific humidity, no distinct changes from the backgrounds to analyses can be found except the degradations near the surface in the Northern Hemisphere and East Asia by both assimilation approaches (Figures 5i-5l). On the other hand, DRP-4DVar fully outperforms 4DVar on the backgrounds and analyses of zonal wind, temperature and specific humidity in the aforementioned four regions. DRP-4DVar makes the biggest improvement in zonal wind (temperature) in the stratosphere (lower troposphere) in the Tropics (Southern Hemisphere) relative to 4DVar (Figures 5c and 5f). Significant improvements in specific humidity by DRP-4DVar are mainly in the lower troposphere comparing with 4DVar. The degradations in the temperature and specific humidity may be caused by model forecast errors introduced by using 3-h forecasts as the analysis and background fields.

The analysis error structures of the DRP-4DVar and 4DVar experiments are very similar (Figure 6 left and middle), which are also consistent with the analysis error structures of the 3DVar experiment and the corresponding 3D hybrid assimilation experiment in Kleist et al. (2015a). As shown in Figures 6a-6b, the zonal wind error maxima are distributed in the middle and upper troposphere at middle latitudes of the Southern Hemisphere, and large zonal wind errors even extend to the lower troposphere near 60 °S. Comparing with 4DVar, DRP-4DVar reduces the analysis errors of zonal wind mainly at the latitudes between 30°S and 30°N, although it increases the analysis errors at high latitudes in the Northern Hemisphere where analysis errors are not too large (Figure 6c). Large temperature errors in the analyses of both assimilation approaches are concentrated in the lower troposphere, especially in the region from the Antarctica to 30°S, which extend to the middle and upper troposphere near 60°S (Figures 6d-6e). DRP-4DVar has smaller ARMSEs of temperature than 4DVar over most latitudes except for high latitudes in the Northern Hemisphere and the latitudes around 60°S (Figure 6f). Specific humidity shows analysis error structures quite different from the zonal wind and temperature, which have semicircular shapes located between 60°S and 60°N in the lower and middle troposphere (Figures 6g-6h). In the regions large humidity errors locate at, DRP-4DVar improves the accuracies of almost all humidity analyses (Figure 6i). In a word, DRP-4DVar reduces most analysis errors of zonal wind, temperature and specific humidity in comparison to 4DVar.

From the above discussions, it can be found that the analysis accuracy of the DRP-4DVar system is basically higher than that of the 4DVar system. Based on these encouraging results, our attention is now drawn to the impact of these more realistic analysis ICs on medium-range forecasts. We want to know whether the improved analysis IC can lead to improved forecasts. For this reason, the analysis fields at 1200 UTC 13 September 2016 produced by the DRP-4DVar and 4DVar systems were used as ICs to conduct 10-day forecasts. Because the DRP-4DVar is an ensemble-based assimilation approach that produced 60 analysis ICs in the OSSE, 60 sets of medium-range forecasts were obtained using these analysis ICs. For convenience to compare with the single set of 10-day forecast with the 4DVar analysis as its IC, the ensemble mean of the 60 sets of DRP-4DVar-based 10-day forecasts was used. The forecasts were evaluated using the “truth” state as the reference and adopting the anomaly correlation coefficient (ACC), ARMSE and the relative change rate of ARMSE (RCRA) as the metrics.

ACC is one of the important metrics to investigate the skill of a forecast, which is used to qualitatively measure the similarity between the anomalies of this forecast and the “truth” state. In terms of this metrics, the DRP-4DVar-based 10-day forecast of 500hPa geopotential height has higher skills than the 4DVar-based forecast on most lead forecast days (Figure 7). In the Northern Hemisphere (Figures 7a and 7e), these two forecasts have comparable skills on the lead days 1-5, and the DRP-4DVar-based forecast has slightly lower skill on the lead days 6-8 and slightly higher skill on the lead days 9-10. On the whole, there are no obvious difference between these two forecasts in the Northern Hemisphere. In contrast, the DRP-4DVar-based forecast outperforms the 4DVar-based forecast on all lead days in the Southern Hemisphere (Figures 7b and 7f), where observations are much sparser than in the Northern Hemisphere (Figure 3). In particular, significant improvements on the lead days 6-10 can be easily sighted in the DRP-4DVar-based forecast comparing with the 4DVar-based forecast. Similarly, as shown in Figures 7c-7d and 7g-7h, DRP-4DVar has better performance on all lead days except the 5^{th}-6^{th} days (7^{th} day) than 4DVar in the Tropics (East Asia). In addition, both forecasts have the highest skills in the East Asia and the lowest skills in the Tropics possibly due to the worse constraint of geostrophic balance in the Tropics. Meanwhile, 4DVar has better performance in the Northern Hemisphere than in the

Southern Hemisphere because of much denser observations in the Northern Hemisphere. In comparison, DPR-4DVar has equivalent skills in the Northern and Southern Hemispheres, suggesting the DRP-4DVar analysis IC has much better capability to improve the prediction skill of 500hPa geopotential height in the regions with sparse observations. In summary, more accurate ICs from DRP-4DVar generally achieve to higher prediction skills in the ensemble mean forecast on almost all lead days than that from 4DVar in the single forecast.

ARMSE is also an indispensable metrics to evaluate the skill of a forecast, which is used to quantitatively measure the difference between the anomalies of this forecast and the "truth" state. To facilitate the comparison between DRP-4DVar and 4DVar on their contributions to prediction skill, a relative change rate of ARMSE (RCRA) from 4DVar to DRP-4DVar is defined as the following:

$$RCRA = \frac{ARMSE(DRP-4DVar)-ARMSE(4DVar)}{ARMSE(4DVar)}. \qquad (18)$$

The RCRA with a negative (positive) value indicates a further improvement (degradation) of the forecast by DRP-4DVar comparing with that by 4DVar. Figure 8 shows the RCRA of the 500hPa geopotential height forecasts. The skill of the DRP-4DVar-based forecast under this metrics basically matches that under the metrics of ACC, which is better than the skill of the 4DVar-based forecast on most lead days. In all four regions except the Northern Hemisphere where the DRP-4DVar-based and 4DVar-based forecasts have comparable performances (Figure 8a), the analysis IC from DRP-4DVar leads to better forecasts than that of 4DVar on almost all lead days (Figures 8b-8d). Large improvements occur on the 8th day in the Tropics and on the 10th day in the East Asia (Figures 8c-8d). The results here support the conclusion made under the metrics of ACC.

Because DRP-4DVar leads to the maximum improvements in the forecast of the 500hPa geopotential height on the 9th day in both the Northern and Southern Hemispheres comparing with 4DVar (Figures 7e-7f and 8a-8b), the horizontal distributions of the forecasts on this lead day were also analyzed and compared. Figure 9 shows the "truth" state and the 216-h forecasts of the 500hPa geopotential height respectively using the 4DVar and DRP-4DVar analyses as their ICs in the Northern Hemisphere (20°N-90°N). In the "truth" state, a low-pressure system is distributed around the Arctic, with three troughs near 75°E, 160°W and 60°W, respectively. In addition, there is a high value center at middle and high latitudes near 0° (Figure 9a). The main circulation situations in the 4DVar-based (DRP-4DVar-based) forecast are basically similar to those in the "truth" state (Figures 9b-c), although the low-pressure trough (high value center) near 75°E (0°) is not correctly presented. As shown in Figure 9d, the 4DVar-based forecast has large errors with a "positive-negative-positive" distribution between 30°W and 0° at middle and high latitudes. It also presents significant negative errors in the regions between 30°E and 60°E at the Arctic and near 90°W at middle latitudes, and positive errors near 160°W at middle latitudes. In comparison, the DRP-4DVar-based forecast mainly reduces the magnitude of the errors in the regions from 30°E to 90°W at middle and high latitudes, but increases the errors between 30°W and 0° at middle and high latitudes (Figure 9e). Similar to Figure 9, Figure 10 shows the results in the Southern Hemisphere. There is a low-pressure system near the Antarctica, extending out troughs near 150°E, 30°E and 90°W in the "truth" state (Figure 10a). In addition, there is a low-pressure trough near the 90°E at middle latitudes. The 4DVar-based and DRP-4DVar-based forecasts basically represent the circulation situations in the "truth" state (Figures 10b-10c), but the former does not capture the troughs at 90°W and 90°E very well. The forecast errors of 4DVar in the Southern Hemisphere are significantly higher than those in the Northern Hemisphere (Figure 10d), while DRP-4DVar reduces almost all the significant forecast errors of 4DVar (Figure 10e). In short, compared with 4DVar, DRP-4DVar improves the 500hPa geopotential height forecasts in some areas in the Northern Hemisphere and almost all areas in the Southern Hemisphere, which basicly coincide with the data-sparse areas shown in Figure 3. This suggests that the analysis IC of DRP-4DVar has a stronger ability to improve the 500hPa geopotential height forecast in data-sparse areas, which is consistent with the conclusions obtained by the ACC and RCRA metrics.

The DRP-4DVar-based forecast of geopotential height also has similar performances at most other vertical levels (Figure 11). In the Northern Hemisphere, the 4DVar-based forecast has significant errors at the vertical levels between 400hPa and 200hPa, which keeps increasing following the lead time and reaches the maximum on the 10th day (Figure 11a). These errors gradually extend to lower and upper levels following the lead time. As shown in Figure 11b, The DRP-4DVar-based forecast has a similar error structure with the 4DVar-based forecast except for larger errors in the stratosphere, which extends to troposphere on the 1th and 6th-7th days, and smaller errors in the troposphere on the 2nd-4th days and since the 8th day. The maximum improvements (degradations) by DRP-4DVar comparing with 4DVar are located below 800hPa (near 100hPa) on the 8th-9th days (7th-9th days). In contrast, DPR-4DVar has significant improvements in the Southern Hemisphere (Figure 11d) when large forecast errors of 4DVar appear in this region with sparse observations (Figure 11c). The forecast with the IC from DRP-4DVar has smaller errors than that with the IC from 4DVar since the 2nd day at almost all vertical levels in this region. Large improvements by DRP-4DVar relative to 4DVar are achieved since the 7th day and the maximum improvements appear at the levels from 500hPa to 200hPa on the 7th-9th days and in the stratosphere on the 10th day where and when the 4DVar-based forecast has large errors (Figure 11c). In the Tropics, the forecast errors based on 4DVar are much smaller than in the Northern and Southern Hemispheres (Figures 11e), but the relative improvements by DRP-4DVra are still significant (Figures 11f). DRP-4DVar reduces the forecast error at most levels on most lead days relative to 4DVar, and the maximum improvements are located near the surface on the 7th-8th days, which correspond to the large errors in the 4DVar-based forecast (Figure 11f). In the East Asia, the forecast errors by 4DVar are larger than in the Tropic but smaller than in the Northern and Southern Hemispheres (Figure 11g). DRP-4DVar improves the forecast at all vertical levels on all lead days except in the stratosphere before the 6th day and in the middle and upper troposphere on Day 7 (Figure 11h). The maximum improvements are sighted in the stratosphere on the 8th-9th days and in the lower troposphere since the 9th day.

The 4DVar-based zonal wind forecast has an error structure similar to the geopotential height forecast at most vertical levels with the largest errors in the Southern Hemisphere and smallest errors in the Tropics (Figures 12a, 12c, 12e and 12g). However, quite different from the geopotential height forecast, the DRP-4DVar-based forecast of zonal wind reduces the errors at almost all vertical levels on almost all lead days in the Northern and Southern Hemispheres, Tropics and East Asia, relative to the 4DVar-based forecast (Figures 12 b, 12d, 12f and 12h). In the Northern Hemisphere, DRP-4DVar reduces the forecast errors of zonal wind at all levels on all lead days except in the upper troposphere and stratosphere after Day 2 and near the surface on Days 6-7, and the largest improvements of DRP-4DVar over 4DVar are located in the lower troposphere on Days 8-9 (Fig. 12b). In the Southern Hemisphere, DPR-4DVar improves the zonal wind forecasts at almost all vertical levels on all lead days compared with 4DVar (Figure 12d). In particular, similar to the geopotential height forecast, the improvements in the zonal wind forecast are much more significant in the sparsely observed Southern Hemisphere than in the Northern Hemisphere. Large improvements can be found in the upper troposphere (near the surface) on the 7th-10th days (7th day). In the Tropics, the improvements in the DRP-4DVar-based forecast are not as obvious as in the Southern Hemisphere, but more distinct than in the Northern Hemisphere. The significant improvements appear in the middle troposphere on the 9th-10th days (Figure 12f). The locations of significant improvements appeared in the Northern and Southern Hemispheres and Tropics correspond to relatively larger but not the largest errors in the 4DVar-based forecast. In the East Asia, comparing with 4DVar, DRP-4DVar improves the forecast at all vertical levels on all lead days except the levels between 200hPa and 100hPa on Days 1-5 and near the surface around Day 4. The largest improvements are seen in the lower troposphere on Days 9-10 and the upper troposphere on Days 6-10, and extend to the lower troposphere (middle troposphere) around Day 6 (Days 7 and 9), where and when the errors of the 4DVar-based forecast are large (Figure 12h).

The error distribution of the 4DVar-based temperature forecast is not quite the same as those of the geopotential height and zonal wind forecasts, with the error size sorting as same as in geopotential height and zonal wind forecasts (Figures 13a, 13c, 13e and 13g). In the Northern Hemisphere, DRP-4DVar reduces the errors of the 4DVar-based forecast at all vertical levels on all lead days except the levels near 200hPa on Days 1-5 and 7-8 and near 100hPa on Days 3-9 with increased or unchanged errors (Figure 13b). The improvements in the temperature prediction by DPR-4DVar are more significant in the Southern Hemisphere than in the Northern Hemisphere (Figure 13d), where the forecast errors of 4DVar are also much larger than those in the Northern Hemisphere (Figure 13c). The largest improvements occur between the surface and 300hPa on the 7th-9th days. In the Tropics, DRP-4DVar improves the temperature forecast at all vertical levels on all lead days. The largest improvements are located near the surface on Days 5-6 and 8-9 and in the middle troposphere after Day 8 (Figure 13f). In the East Asia, DRP-4DVar reduces the forecast errors at almost all vertical levels on all lead days, although it increases the errors in the upper troposphere on Days 1-4. The most significant error reductions appear across the levels from 700hPa to 400hPa on Day 6, from 600hPa to 300hPa on Day 8 and from 500hPa to 200hPa on Day 10, which do not correspond to the locations of large errors in the 4DVar-based forecast (Figure 13h).

The 4DVar-based specific humidity forecast has an error structure different from other variables, with large errors between 900hPa and 700hPa in each of four regions, increasing with the lead time and reaching a maximum since the 9th day (Figures 14a, 14c, 14e and 14g). These errors gradually extend to lower and higher levels with the lead time. Similar to the zonal wind and temperature forecasts, the DRP-4DVar-based forecast of specific humidity reduces errors at almost all vertical levels on almost all lead days in all regions compared with the 4DVar-based forecast (Figures 14 b, 14d, 14f and 14h). In the Northern Hemisphere, DRP-4DVar reduces the errors at all levels on all lead days except the levels near 100hPa on Days 4-6 relative to the 4DVar-based forecast (Figure 14b). The largest improvements from 4DVar to DRP-4DVar are located between 800hPa and 200hPa on Days 6-8. Although the 4DVar-based forecast errors in the Southern Hemisphere are not significantly different from those in the Northern Hemisphere (Figure 14c), the improvements of DPR-4DVar are more significant in this region than in the Northern Hemisphere (Figure 14d). The most significant improvements in the DRP-4DVar-based forecast relative to the 4DVar-based forecast can be found near 200hPa on all lead days except Day 3. They also occur near 700hPa on the 6th day or between 925hPa and 800hPa after the 8th day, where and when the 4DVar-based forecast has large errors (Figure 14c). In the Tropics, the 4DVar-based forecast errors are comparable to those in the Northern and Southern Hemispheres (Figure 14e), while the improvements in the DRP-4DVar-based forecast are comparable to those in the Northern Hemisphere but less significant than in the Southern Hemisphere (Figure 14f). The largest improvements appear near 200hPa on all lead days except Days 7-8. In the East Asia, the forecast errors of 4DVar are smaller than in the other three regions (Figure 14g), but the improvements in the DRP-4DVar-based forecast after Day 5 are basically comparable to those in the Southern Hemisphere. DRP-4DVar improves the forecast at almost all the vertical levels on all lead days except the levels above 100hPa before the 7th Day. The greatest improvements can be found at the levels near the surface and in the upper troposphere on Days 6-7, and in the lower troposphere since the 8th day where and when large forecast errors of 4DVar are located (Figure 14h).

Figure 15 shows the RCRAs of 24-h accumulated precipitation forecast. The DRP-4DVar-based forecast skill of precipitation under this metric outperforms the 4DVar-based forecast skill in all the four regions on all lead days except in the East Asia on the 2nd day. The largest improvements occur on the 9th day in the Northern Hemisphere, 8th day in the Southern Hemisphere and Tropics, and 10th day in the East Asia (Figures 15), respectively.

The computational efficiency of the DRP-4DVar system is also a key concern. Taking the 6-h assimilation window (0900 UTC 11 September 2016 - 1500 UTC 11 September 2016) as an example, the computational time was about 25 minutes for a 4DVar DA using 480 cores on the high-performance computer PI-SUGON of the China Meteorological Administration. In comparison, the DRP-4DVar system took only 13 minutes since the ensemble members of the DRP-4DVar system are independent and all members can be analyzed concurrently using a total of 60×480 cores. The aforementioned results may be slightly impacted by several factors, such as the high-performance computer state, but overall, the DRP-4DVar system has the advantage of saving time.

6 Summary and discussion

In this study, a new 4DEnVar hybrid DA system was developed based on the DRP-4DVar approach. This system is novel in some aspects. It introduces the idea of ensemble into the variational framework to achieve the time-variant BEC and minimizes the cost function without using the ADM. It can easily take full advantage of the mature variational framework and construct the hybrid BEC. A new inflation method based on random perturbations with balance constraints produced using the static B-matrix is applied to alleviate the filter divergence during the assimilation cycle, which can be conveniently and efficiently implemented. Moreover, a limited number of leading eigenvectors of the localization correlation function are used to perform the localization of the B-matrix and rapidly increase the ensemble size without any extra model integrations.

Preliminary tests including SOEs and OSSEs were conducted to evaluate the performance of the DRP-4DVar system, using the 4DVar system as a reference for comparison. The first assimilation window for the SOEs and OSSEs was set at the time after a 2-day assimilation cycle for spin-up. The SOEs show that both DRP-4DVar and 4DVar assimilated the single-point observation effectively and satisfied certain balance constraints. Moreover, DRP-4DVar using the ensemble BEC exhibits obvious flow-dependent features.

In the OSSEs, on one hand, the DRP-4DVar and 4DVar analyses are basically more accurate than the corresponding background for most variables at most vertical levels, especially for the zonal wind in all the four regions including the Northern and Southern Hemispheres, Tropics and East Asia and the temperature in the first two regions, while they have performances on temperature in Tropics (in the lower and middle troposphere in the East Asian) and the specific humidity in the middle troposphere (near the surface in the East Asian) comparable to (worse than) the background. On the other hand, DRP-4DVar outperforms 4DVar in terms of ARMSE on the background and analysis fields of the zonal wind, temperature and specific humidity in the aforementioned four areas. In addition, the largest improvements in the analyses of zonal wind (temperature) of DRP-4DVar compared to those of 4DVar are mainly located in the stratosphere in the Tropics (the lower troposphere in the Southern Hemisphere), and the largest improvements in the specific humidity analysis are mainly in the lower troposphere. According to the latitudinal-vertical distribution of the analysis errors, DRP-4DVar reduces analysis errors of the zonal wind between 30°S and 30°N, the temperature in all regions except the high latitudes of the Northern Hemisphere and near 60°S, and the specific humidity in almost all regions, relative to 4DVar. In conclusion, the analysis accuracy of the DRP-4DVar system is basically higher than that of the 4DVar system.

In this study, the ACC, ARMSE and RCRA metrics are used to evaluate the 10-day forecasts using the DRP-4DVar and 4DVar analyses as ICs. For the forecast of 500hPa potential height, there are no significant differences between the DRP-4DVar-based and 4DVar-based forecasts in the Northern Hemisphere in terms of ACC and RCRA indicators, but the forecast skills of DRP-4DVar in the Southern Hemisphere, Tropics, and East Asia are significantly improved, especially on the 8^{th}-10^{th} days, in comparison to those of 4DVar. Furthermore, DRP-4DVar reduces the forecast errors of 500hPa geopotential height in some areas in the Northern Hemisphere, and almost all areas in the Southern Hemisphere, especially in the data-sparse areas. This indicates that the analysis of DRP-4DVar is more capable of sufficiently using sparse observations to improve the forecast, which is consistent with the conclusions obtained by the ACC and RCRA metrics.

For the forecast of the geopotential height at all vertical levels in the Northern Hemisphere, the largest improvements (degradations) by DRP-4DVar comparing with 4DVar are located below 800hPa (near 100hPa) on the 8^{th}-9^{th} days (7^{th}-9^{th} days). The DRP-4DVar-based forecast has more significant improvements in the sparsely observed Southern Hemisphere than in the Northern Hemisphere, improving the geopotential height at almost all vertical levels after the 2^{nd} day. DRP-4DVar reduces the forecast errors at most levels on most lead days in the Tropics compared with 4DVar. DRP-4DVar improves the forecast in the East Asian at almost all vertical levels on all lead days, although there is significant degradation in the stratosphere before the 6^{th} day and in the middle and upper troposphere on Day 7. Compared to the 4DVar-based forecast, the DRP-4DVar-based forecast improves the forecasts of the zonal wind in almost all regions on all lead days, and the most significant improvements are seen in the upper troposphere on the 7th-10th days and near the surface on the 7th day (in the lower troposphere on Days 9-10 and the upper troposphere on Days 6-10) in the Southern Hemisphere (East Asia). In addition, there are obvious degradation in the stratosphere after Day 2 and near the surface on Days 6-7 in the Northern Hemisphere, and between 200hPa and 100hPa on Days 1-5 in the East Asia. For the forecast of the temperature, DRP-4DVar reduces errors in almost all regions on all lead days compared to 4DVar, although there is obvious degradation in the stratosphere on Days 3-9 in the Northern Hemisphere, and in the upper troposphere on the lead days 1-4 in the East Asia. The improvements in temperature are most significant from the surface to 300hPa (middle and upper troposphere) on Day 7- 9 (the 6th, 8th and 10th days) in the Southern Hemisphere (East Asia). A closer look at where the DRP-4DVar-based forecasts degrades in zonal wind and temperature compared to the 4DVar-based forecasts reveals a good match with where the geopotential height degrades, most likely due to the localization scheme that conducted upon the basic state variables and the use of low-resolution TLM to generate observational variable samples, which degrades the balance constraints between model variables and thus negatively affects the geopotential height forecast. For the forecast of the specific humidity, DRP-4DVar steadily improve it in all regions at almost all vertical levels compared to the 4DVar-based forecast. The DRP-4DVar-based forecast also steadily improves the 24-h accumulated precipitation forecasts in all regions at all lead days except in the East Asia on Day 2, with the largest improvements occurring on the 9th day in the Northern Hemisphere, 8th day in the Southern Hemisphere and Tropics, and 10th day in the East Asia, which is consistent with the above specific humidity forecasts.

Overall, the DRP-4DVar system shows great promise in terms of both high quality of analyses and ensemble forecasts. The significantly improved forecasts from DRP-4DVar suggest that the analysis ensembles produced by the DRP-4DVar system have the potential to provide high quality flow-dependent ensemble BECs for hybrid DA systems. The apparent improvements in data-sparse areas imply that the DRP-4DVar system using the flow-dependent ensemble BEC can more sufficiently incorporate observational information in these areas. Moreover, DRP-4DVar saves more time than 4DVar given that the ensemble members of DRP-4DVar can be analyzed concurrently. There is still much room for further improving the performance of the DRP-4DVar system. For example, the current localization is performed upon the basic state variables, which may negatively affect the balance constraints. In the future, this could be done upon the unbalanced part of the basic state variables. In addition, the current DRP-4DVar system relies on the low-resolution TLM to produce the observational variable samples. In the future, high-resolution NLM can be used, which not only avoids the development of the TLM, but also improves the quality of the ensemble samples. In addition, longer OSSEs and assimilation experiments using real observations, especially satellite radiance observations, should be carried out to further evaluate the performance of the DRP-4DVar system.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (2018YFC1506700). The assimilation and forecast experiments were performed on the high-performance computer PI-SUGON of the China Meteorological Administration. The sounding and cloud-derived wind observations were provided by the Global Telecommunications System (https://public.wmo.int/en/programmes/global-telecommunication-system). The ERA-5 reanalyses were downloaded from https://apps.ecmwf.int/data-catalogues/era5/?class=ea&stream=oper&expver=1&type=an. The 6-h forecasts of ERA-Interim dataset and the ERA-Interim reanalyses were downloaded from https://apps.ecmwf.int/datasets/data/interim-full-daily/levtype=sfc/.

References

Barker, D. M. (2005). Southern high-latitude ensemble data assimilation in the Antarctic mesoscale prediction system. *Monthly Weather Review, 133*(12), 3431-3449.

Bishop, C. H., Hodyss, D., Steinle, P., Sims, H., Clayton, A. M., Lorenc, A. C., Barker, D. M., & Buehner, M. (2011). Efficient ensemble covariance localization in variational data assimilation. *Monthly Weather Review, 139*, 573-580.

Bonavita, M., Isaksen, L., & Hólm, E. (2012). On the use of EDA background error variances in the ECMWF 4D-Var. *Quarterly Journal of the Royal Meteorological Society, 138*(667), 1540-1559.

Bowler, N. E., Clayton, A. M., Jardak, M., Jermey, P. M., Lorenc, A. C., Wlasak, M. A., Barker, D. M., Inverarity, G. W., & Swinbank, R. (2017a). The effect of improved ensemble covariances on hybrid variational data assimilation. *Quarterly Journal of the Royal Meteorological Society, 143*: 785-797.

Bowler, N. E., Clayton, A. M., Jardak, M., Lee, E., Lorenc, A. C. and Piccolo, C., Pring, S. R., Wlasak, M. A., Barker, D. M., Inverarity, G. W., & Swinbank, R. (2017b). Inflation and localization tests in the development of an ensemble of 4D-ensemble variational assimilations. *Quarterly Journal of the Royal Meteorological Society, 143*(704): 1280-1302.

Buehner, M., Houtekamer, P. L., Charette, C., Mitchell, H. L., & He, B. (2010a). Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part I: Description and single-observation experiments. *Monthly Weather Review, 138*(5), 1550-1566.

Buehner, M., Houtekamer, P. L., Charette, C., Mitchell, H. L., & He, B. (2010b). Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part II: One-month experiments with real observations. *Monthly Weather Review, 138*(5), 1567-1586.

Buehner, M., Morneau, J., & Charette, C. (2013). Four-dimensional ensemble-variational data assimilation for global deterministic weather prediction. *Nonlinear Processes in Geophysics, 20*(5), 669-682.

Buehner, M., Mctaggart-Cowan, R., Beaulne, A., Charette, C., Garand, L., Heilliette, S., Lapalme, E., Laroche, S., Macpherson, S. R., Morneau, J., & Zadra, A. (2015). Intercomparison of deterministic weather forecasting systems based on ensemble–variational data assimilation at Environment Canada. Part I: The global system. *Monthly Weather Review, 143*, 2532-2559.

Campbell, W. F., Bishop, C. H., & Hodyss, D. (2010). Vertical covariance localization for satellite radiances in ensemble kalman filters. *Monthly Weather Review, 138*, 282-290.

Chen, D. H., Xue, J. S., Yang, X. S., Zhang, H. L., Shen, X. S., Hu, J. L., Wang, Y., Ji, L. R., & Chen, J. B. (2008). New generation of multi-scale NWP system (GRAPES): general scientific design. *Chinese Science Bulletin, 53*(22), 3433–3445.

Clayton, A. M., Lorenc, A. C., & Barker, D. M. (2013). Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office. *Quarterly Journal of the Royal Meteorological Society, 139*(675), 1445-1461.

Courtier, P., Thépaut, J. N., & Hollingsworth, A. (1994). A strategy for operational implementation of 4D-Var, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society, 120*(519), 1367-1387.

Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research Oceans, 99*(C5), 10143–10162.

Evensen, G. (2003). The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dynamics, 53*(4), 343-367.

Gaspari, G., & Cohn, S. E. (1999). Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society, 125*(554), 723-757.

Gauthier, P., & Thépaut, J. N. (2001). Impact of the digital filter as a weak constraint in the preoperational 4dvar assimilation system of Météo-France. *Monthly Weather Review, 129*(8), 2089-2102.

Gauthier, P., Tanguay, M., Laroche, S., & Pellerin, S. (2007). Extension of 3dvar to 4dvar: Implementation of 4dvar at the meteorological service of Canada. *Monthly Weather Review, 135*(6), 2339-2354.

Hamill, T. M., & Snyder, C. (2000). A hybrid ensemble Kalman filter–3d variational analysis scheme. *Monthly Weather Review, 128*(8), 2905-2919.

Hamill, T. M., Whitaker, J. S., & Snyder, C. (2001). Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Monthly Weather Review, 129*(11), 2776-2790.

He, Y., Wang, B., Liu, M., Liu, L., Yu, Y., Liu, J., Li, R., Zhang, C., Xu, S., Huang, W., Liu, Q., Wang, Y., & Li, F. (2017). Reduction of initial shock in decadal predictions using a new initialization strategy. *Geophysical Research Letters, 44*(16), 8538-8547, DOI: 10.1002/2017GL074028.

He, Y., Wang, B., Liu, L., Huang, W., Xu, S., Liu, J., Wang, Y., Li, L., Huang, X., Peng, Y., Lin, Y., & Yu, Y. (2020a). A DRP-4DVar-based coupled data assimilation system with a simplified off-line localization technique for decadal predictions. *Journal of Advances in Modeling Earth Systems, 12*, e2019MS001768. https://doi.org/10.1029/2019MS001768.

He, Y., Wang, B., Huang, W., Xu, S., Wang, Y., Liu, L., Li, L., Liu, J., Yu, Y., Lin, Y., Huang, X., & Peng, Y. (2020b). A new DRP-4DVar-based coupled data assimilation system for decadal predictions using a fast online localization technique. *Climate Dynamics, 54*(19), 3541-3559, DOI: 10.1007/s00382-020-05190-w.

Houtekamer, P. L., & Mitchell, H. L. (1998). Data assimilation using an ensemble Kalman filter technique. *Monthly Weather Review, 126*, 796-811.

Houtekamer, P. L., & Mitchell, H. L. (2001). A sequential ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review, 129*(1), 123-137.

Houtekamer, P. L., Mitchell, H. L., Pellerin, G., Buehner, M., Charron, M., Spacek, L., & Hansen, B. (2005). Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. *Monthly Weather Review, 133*(3), 604-620.

Kleist, D. T., & Ide, K. (2015a). An OSSE-based evaluation of hybrid variational-ensemble data assimilation for the NCEP GFS. Part I: System description and 3D-hybrid results. *Monthly Weather Review, 143*, 433-451.

Kleist, D. T., & Ide, K. (2015b). An OSSE-based evaluation of hybrid variational–ensemble data assimilation for the NCEP GFS. Part II: 4DEnVar and hybrid variants. *Monthly Weather Review, 143*(2), 452-470.

Koizumi, K., Ishikawa, Y., & Tsuyuki, T. (2005). Assimilation of precipitation data to the JMA mesoscale model with a four-dimensional variational method and its impact on precipitation forecasts. *Sola, 1*, 45-48.

Kuhl, D. D., Rosmond, T. E., Bishop, C. H., Mclay, J., & Baker, N. L. (2013). Comparison of hybrid ensemble/4dvar and 4dvar within the NAVDAS-AR data assimilation framework. *Monthly Weather Review, 141*(8), 2740-2758.

Le Dimet, F. X., & Talagrand, O. (1986). Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus A: Dynamic Meteorology and Oceanography, 38*(2), 97-110.

Lewis, J. M., & Derber, J. C. (1985). The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus A: Dynamic Meteorology and Oceanography, 37*(4), 309-322.

Li, F., Wang, B., He Y., et al. (2021a). Important role of North Atlantic air–sea coupling in the interannual predictability of summer precipitation over the eastern Tibetan Plateau. *Climate Dynamics, 56*, 1433-1448.

Li, F., Wang, B., He Y., et al. (2021b). Improved decadal predictions of East Asian summer monsoon with a weakly coupled data assimilation scheme. *International Journal of Climatology,* 1-22.

Liu, C., Xiao, Q., & Wang, B. (2008). An ensemble-based four-dimensional variational data assimilation scheme. Part I: Technical formulation and preliminary test. *Monthly Weather Review, 136*(9), 3363-3373.

Liu, C., Xiao, Q., & Wang, B. (2009). An ensemble-based four-dimensional variational data assimilation scheme. Part II: Observing system simulation experiments with advanced research WRF (ARW). *Monthly Weather Review, 137*(5), 1687-1704.

Liu, J., & Wang, B. (2011). Rainfall assimilation using a new four-dimensional variational method: A single-point observation experiment. *Advances in Atmospheric Sciences, 28*(4), 735-742.

Lorenc, A. C. (2003). The potential of the ensemble Kalman filter for NWP—a comparison with 4D-Var. *Quarterly Journal of the Royal Meteorological Society, 129*(595), 3183-3203.

Lorenc, A. C., Bowler, N. E., Clayton, A. M., Pring, S. R., & Fairbairn, D. (2015). Comparison of hybrid-4DEnVar and hybrid-4DVar data assimilation methods for global NWP. *Monthly Weather Review, 143*(1), 212-229.

Qiu, C., Zhang, L., & Shao, A. (2007). An explicit four-dimensional variational data assimilation method. *Science in China Series D: Earth Sciences, 50*(8), 1232-1240.

Rabier, F., Jarvinen, H., Klinker, E. Mahfouf J. F., & Simmons, A. (2000). The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Quarterly Journal of the Royal Meteorological Society, 126*(564), 1143-1170.

Rawlins, F., Ballard, S. P., Bovis, K. J., Clayton, A. M., Li, D., Inverarity, G. W., Lorenc, A. C., & Payne, T. J. (2007). The Met Office global four-dimensional variational data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society, 133*(623), 347-362.

Raynaud, L., Berre, L., & Desroziers, G. (2011). An extended specification of flow-dependent background error variances in the Météo-France global 4D-Var system. *Quarterly Journal of the Royal Meteorological Society, 137*(656), 607-619.

Shen, S., Liu, J., & Wang, B. (2015). Evaluation of the historical sampling error for global models. *Atmospheric and Oceanic Science Letters, 8*(5), 250–256.

Shi, P., Lu, H., Leung, L., He, Y., Wang, B., Yang, K., et al. (2021). Significant land contributions to interannual predictability of East Asian summer monsoon rainfall. *Earth's Future, 9*, DOI: 10.1029/E2020EF001762.

Simmons, A. J., & Hollingsworth, A. (2002). Some aspects of the improvement in skill of numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society, 128*, 647-677.

Su, Y., Shen, X., zhang, H., & Liu, Y. (2020). A study on the three-dimensional reference atmosphere in GRAPES_GFS: Constructive reference state and real forecast experiment (in Chinese). *Acta Meteorologica Sinica, 78*(6):962-971. DOI: 10.11676/qxxb2020.075.

Tian, X., Xie, Z., & Dai, A. (2008). An ensemble-based explicit four-dimensional variational assimilation method. *Journal of Geophysical Research: Atmospheres, 113*, D21124. DOI: 10.1029/2008JD010358

Tian, X., Xie, Z., and Sun, Q. (2011). A POD-based ensemble four-dimensional variational assimilation method. *Tellus, 63A*(4), 805-816.

Wang, B., Liu, J. J., Wang, S., Cheng, W., Liu, J., Liu, C., Xiao, Q., & Kuo, Y. (2010a). An economical approach to four-dimensional variational data assimilation. *Advances in Atmospheric Sciences, 27*(4), 715-727.

Wang, B., Liu, J., Lu, B., & Tan, X. (2010b). An orthogonal expansion of filtering function in localization and its application to solution of CNOP. *Geophysical Research Abstracts, 12*, EGU2010-6786, EGU General Assembly 2010.

Wang, B., Liu, J., Liu, L, Xu, S., & Huang, W. (2018). An approach to localization for ensemble-based data assimilation. *Plos One, 13*(1), e0191088. https://doi.org/10.1371/journal.pone.0191088.

Wang, X., Barker, D. M., Snyder, C., & Hamill, T. M. (2008a). A hybrid ETKF-3DVAR data assimilation scheme for the WRF model. Part I: Observing system simulation experiment. *Monthly Weather Review, 136*(12), 5116-5131.

Wang, X., Barker, D. M. , Snyder, C., & Hamill, T. M. (2008b). A hybrid ETKF–3DVAR data assimilation scheme for the WRF model. Part II: Real observation experiments. *Monthly Weather Review, 136*(12), 5132-5147.

Wang, X., Parrish, D., Kleist, D., & Whitaker, J. (2013). GSI 3DVAR-based ensemble–variational hybrid data assimilation for NCEP global forecast system: Single-resolution experiments. *Monthly Weather Review, 141*(11), 4098-4117.

Wang, X., & Lei, T. (2014). GSI-based four-dimensional ensemble–variational (4DEnsVar) data assimilation: Formulation and single-resolution experiments with real data for NCEP global forecast system. *Monthly Weather Review, 142*(9), 3303-3325.

Whitaker, J. S., Hamill, T. M., Wei, X., Song Y., & Toth Z. (2008). Ensemble data assimilation with the NCEP global forecast system. *Monthly Weather Review, 136*, 463-482.

Whitaker, J. S., Compo, G. P., & Thépaut, J. N. (2009). A comparison of variational and ensemble-based data assimilation systems for reanalysis of sparse observations. *Monthly Weather Review, 137*(6), 1991-1999.

Xue, J., Zhuang, S., Zhu, G., Zhang, H., Liu, Z., Liu, Y., & Zhuang, Z. (2008). Scientific design and preliminary results of three-dimensional variational data assimilation system of GRAPES. *Chinese Science Bulletin, 53*(22), 3446-3457.

Zhang, L., Liu, Y. Z., Liu, Y., Gong, J., Lu, H., Jin, Z., Tian, W., Liu, G., Zhou, B., & Zhao, B. (2019). The operational global four-dimensional variational data assimilation system at the China Meteorological Administration. *Quarterly Journal of the Royal Meteorological Society, 145*, 1882–1896.

Zhang, M., & Zhang, F. (2012). E4dvar: Coupling an ensemble Kalman filter with four-dimensional variational data assimilation in a limited-area weather prediction model. *Monthly Weather Review, 140*(2), 587-600.

Zhao, J., & Wang, B. (2010). Sensitivity of the DRP-4DVar performance to perturbation samples obtained by two different methods. *Journal of Meteorological Research, 24*(5), 527-538.

Zhao, Y., Wang, B., & Liu, J. J. (2012). A DRP-4DVar data assimilation scheme for typhoon initialization using sea level pressure data. *Monthly Weather Review, 140*(4), 1191-1203.

Zupanski, M. (1996). A preconditioning algorithm for four-dimensional variational data assimilation. *Monthly Weather Review, 124*(11), 2562-2573.

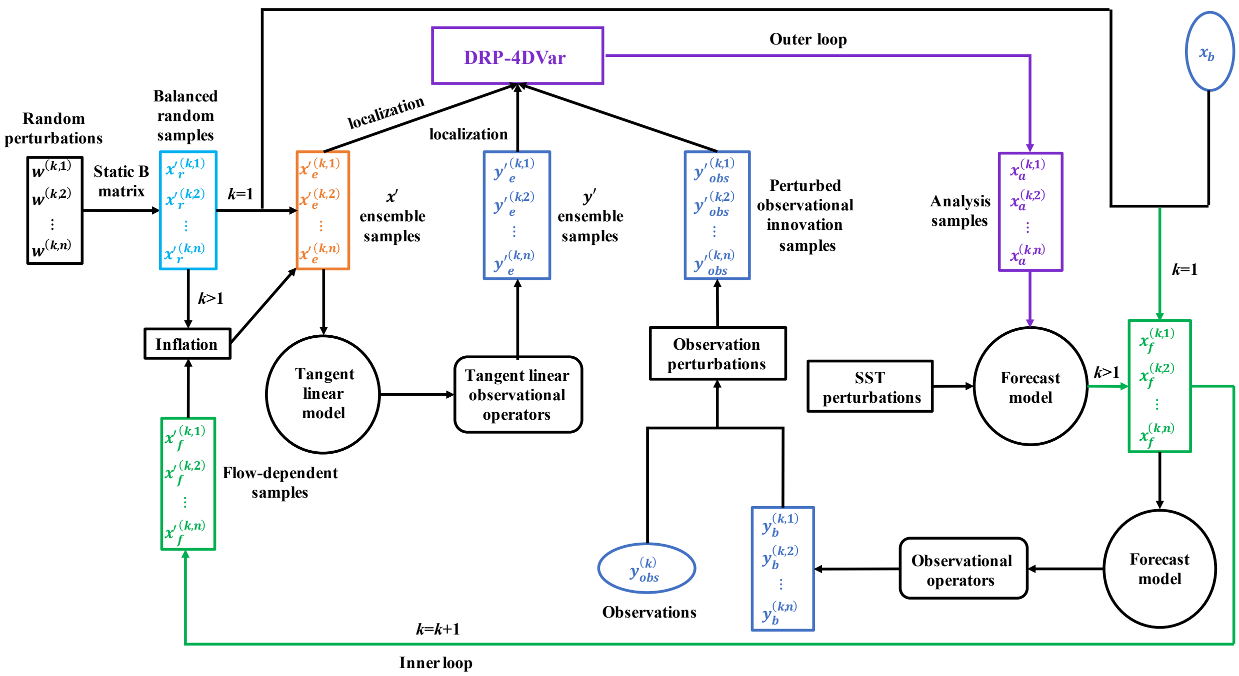


Figure 1. Schematic of the DRP-4DVar system used in this study.

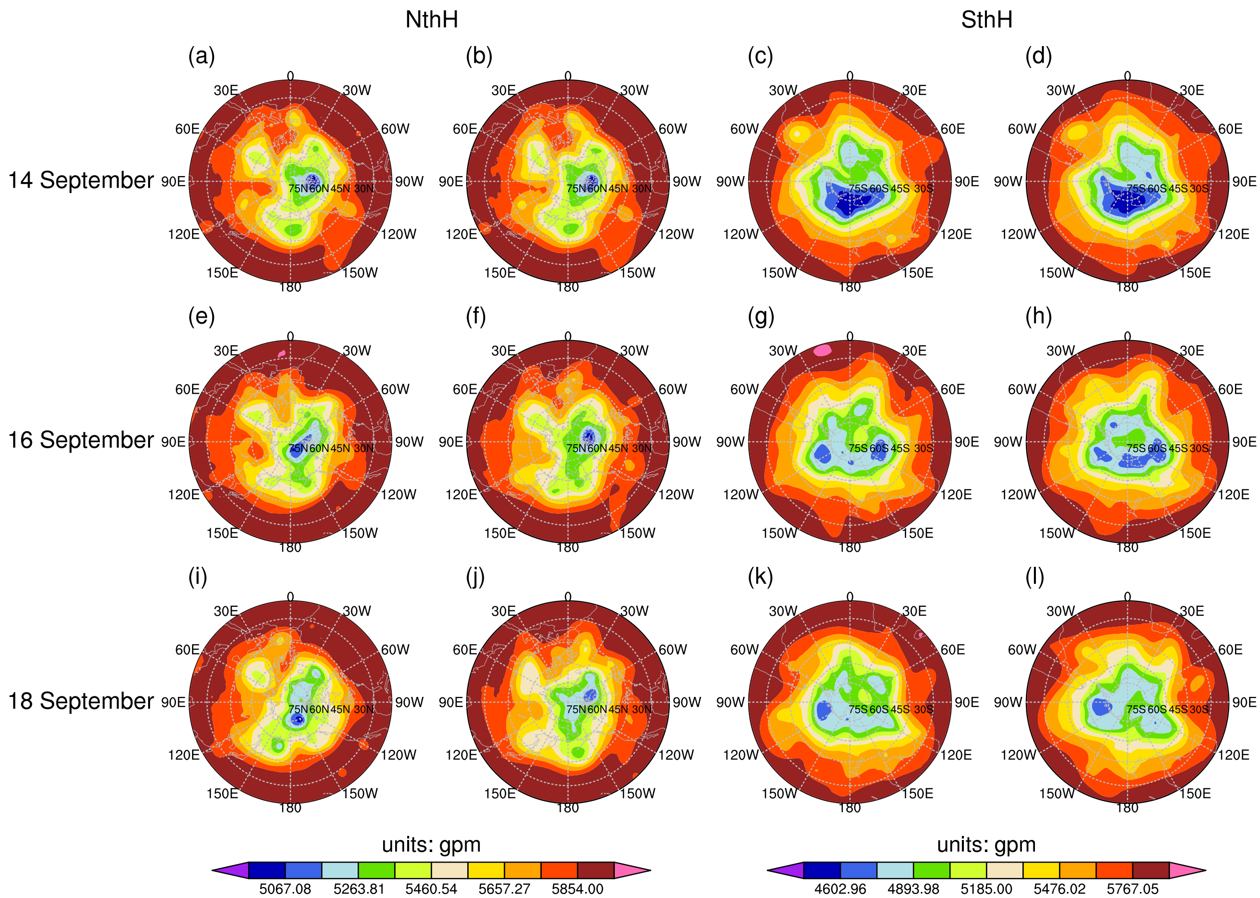


Figure 2. The 500hPa geopotential height at 1200 UTC on 14 September 2016 (top), 16 September 2016 (middle) and 18 September 2016 (bottom) from the ERA-Interim reanalysis (left) and the "truth" state (middle left) in the Northern Hemisphere (20°N~90°N), and the results in the Southern Hemisphere (20°S~90°S; the ERA-Interim reanalysis, middle right; the "truth" state, right) are plotted.

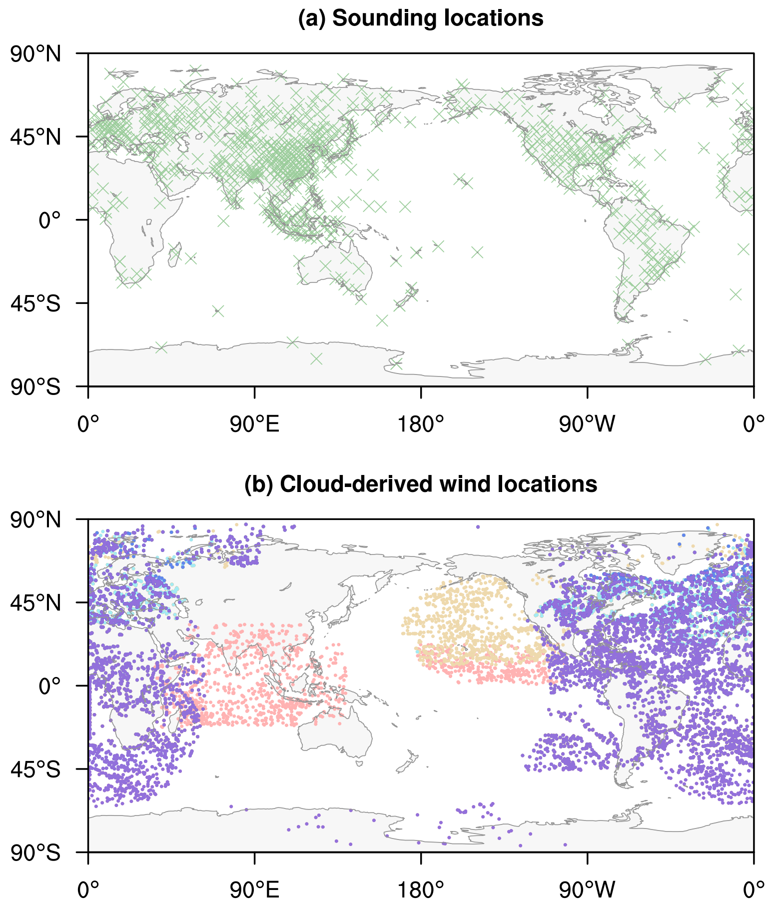


Figure 3. Locations of (a) sounding and (b) cloud-derived wind observations covered the period from 0900 UTC 13 September 2016 to 1500 UTC 13 September 2016. Different colored dots in (b) indicate different sampling times.

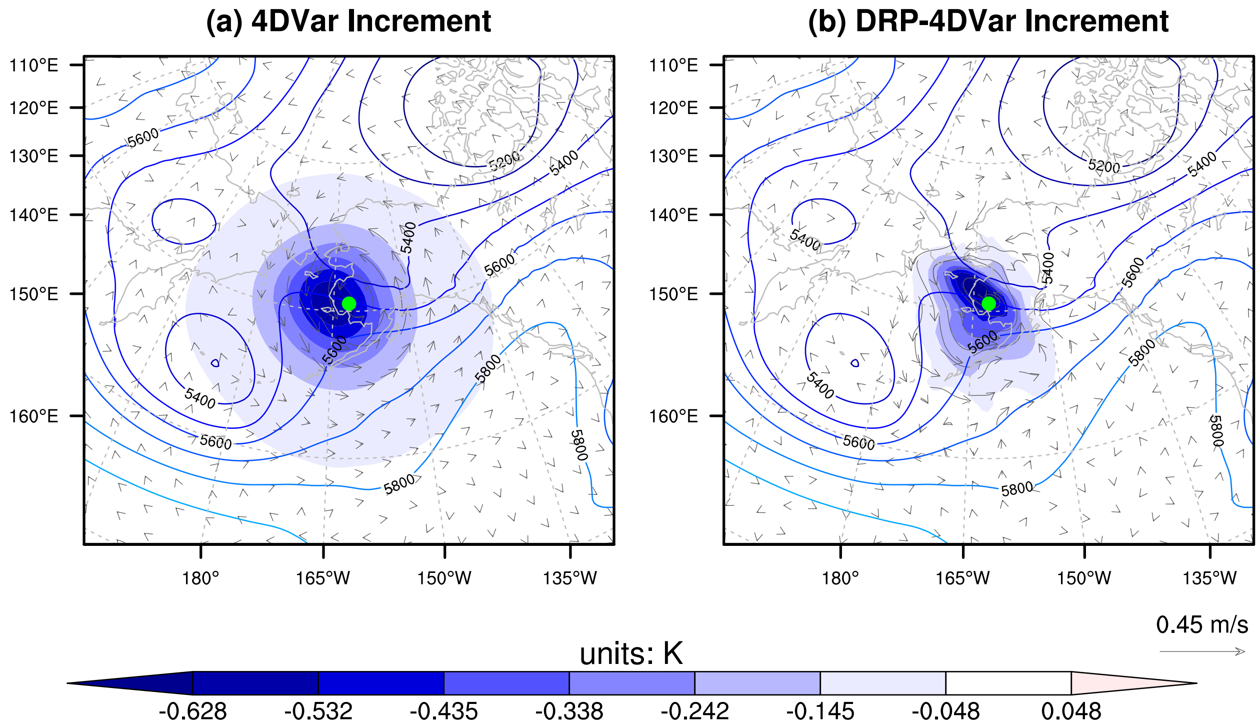


Figure 4. The temperature (shading; units: K) and vector wind (vector; units: m/s) analysis increments at the beginning of the assimilation window from (a) 4DVar and (b) DRP-4DVar on the model level closest to the single temperature observation assimilated, which locates at 500hPa (marked with a green dot). The solid contour is the 500hPa background field geopotential height (units: gpm) valid at the middle of the assimilation window, when the single-point observation is located.

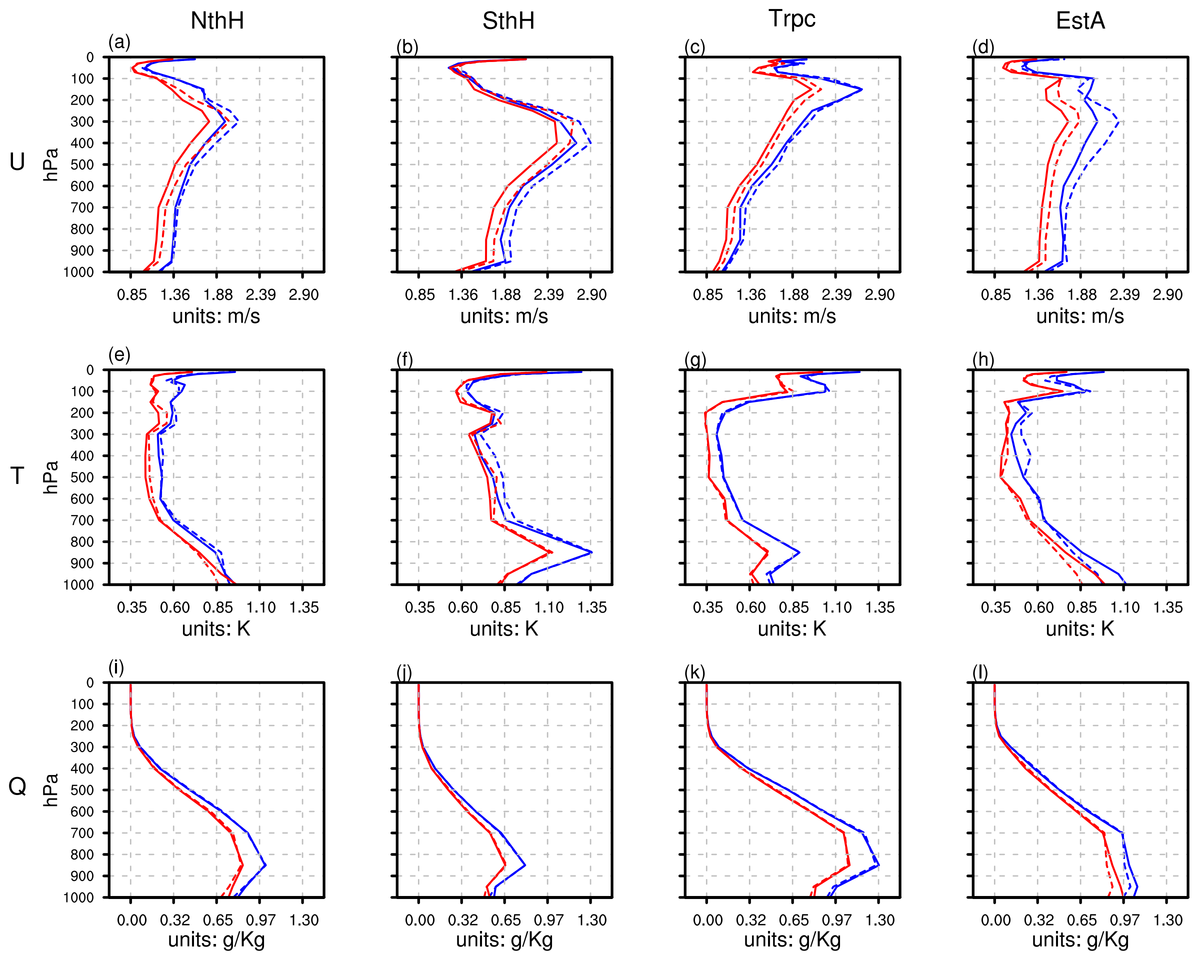


Figure 5. Vertical profiles of the ARMSE (verified relative to the "truth" state) of the background (dashed line) and analysis (solid line) fields of the zonal wind (top; units: m/s), temperature (middle; units: K) and specific humidity (bottom; units: g/Kg) in the Northern Hemisphere (20°N~90°N; left), Southern Hemisphere (20°S~90°S; middle left), Tropics (middle right), and East Asian(right). The blue and red lines show the 4DVar and DRP-4DVar results, respectively.

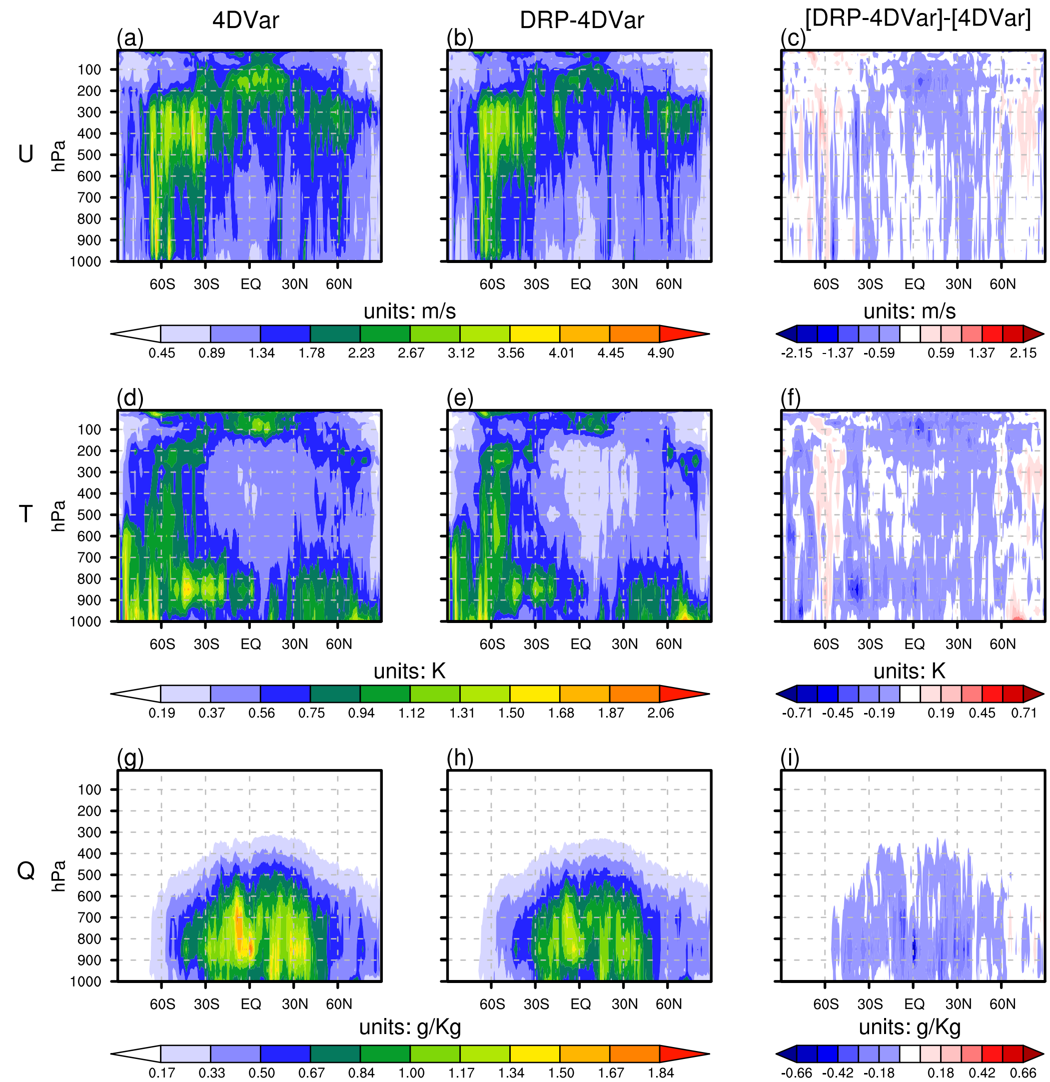


Figure 6. The pressure versus latitude plots of the ARMSEs (verified relative to the "truth" state) of the zonal wind (top; units: m/s), temperature (middle; units: K) and specific humidity (bottom; units: g/Kg) analyses of 4DVar (left), DRP-4DVar (middle) and the ARMSE differences between DRP-4DVar and 4DVar (right), respectively.

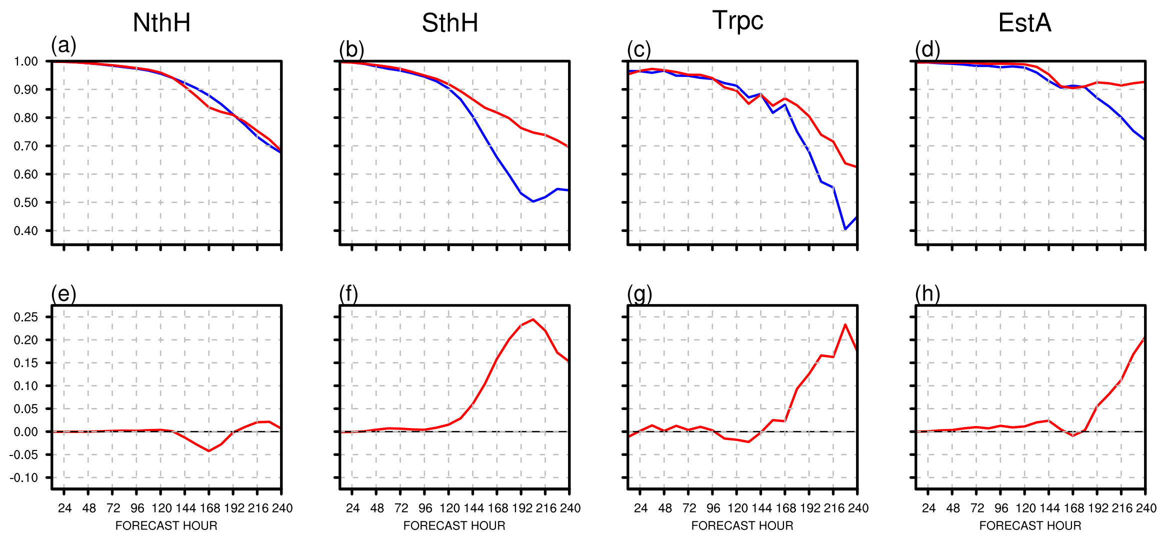


Figure 7. The anomaly correlation coefficients (ACCs) of the DRP-4DVar-based (red line) and 4DVar-based (blue line) 10-day forecasts of the 500hPa geopotential height against the "truth" state in the (a) Northern Hemisphere (20°N~90°N), (b) Southern Hemisphere (20°S~90°S), (c) Tropics, and (d) East Asian. The corresponding ACC differences between DRP-4DVar and 4DVar (red line) are also plotted in (e-h).

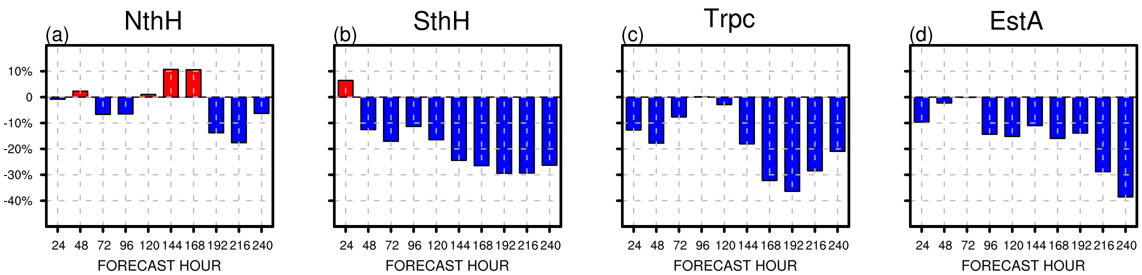


Figure 8. The relative change rates of ARMSE (RCRA) of the 500hPa geopotential height forecasts from 4DVar to DRP-4DVar in the (a) Northern Hemisphere (20°N~90°N), (b) Southern Hemisphere (20°S~90°S), (c) Tropics, and (d) East Asian.

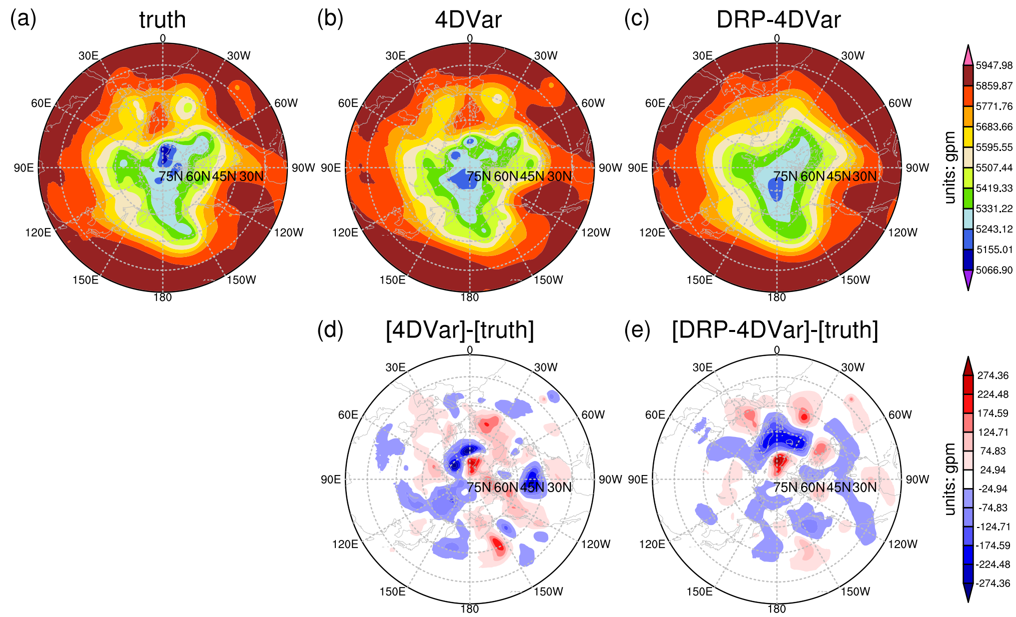
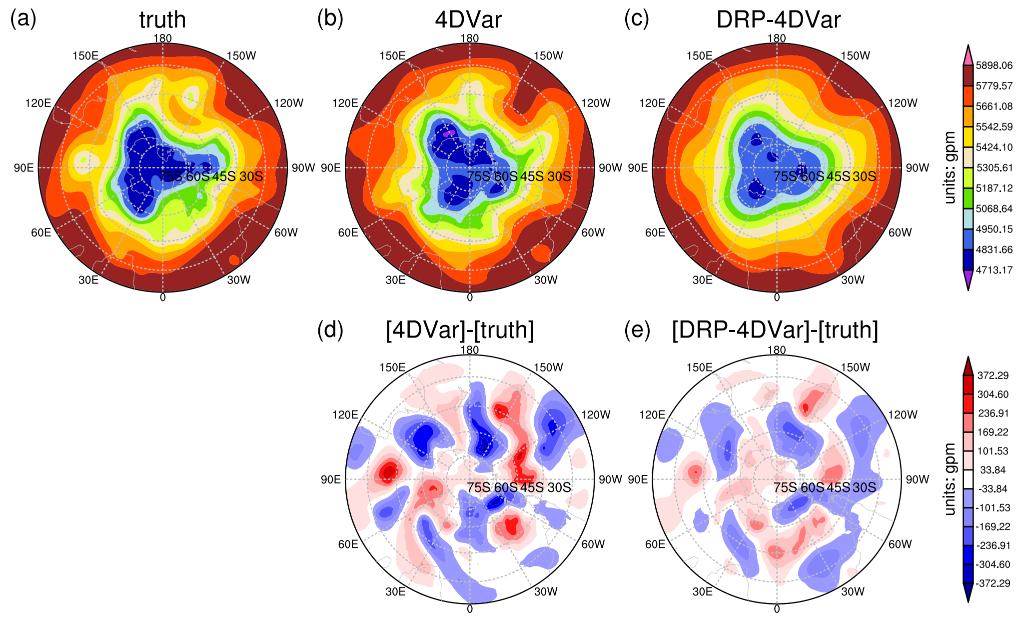


Figure 9. The horizontal distributions of the 216-h forecast of the 500hPa geopotential height in the Northern Hemisphere (20°N~90°N) for (a) the "truth" state, (b) 4DVar and (c) DRP-4DVar. The differences (d) between the 4DVar-based forecast and the "truth" state, and (e) between the DRP-4DVar-based forecast and the "truth" state are also plotted, respectively.



984

985 **Figure 10**. As in Figure 9, but showing the results in the Southern Hemisphere (20°S~90°S).

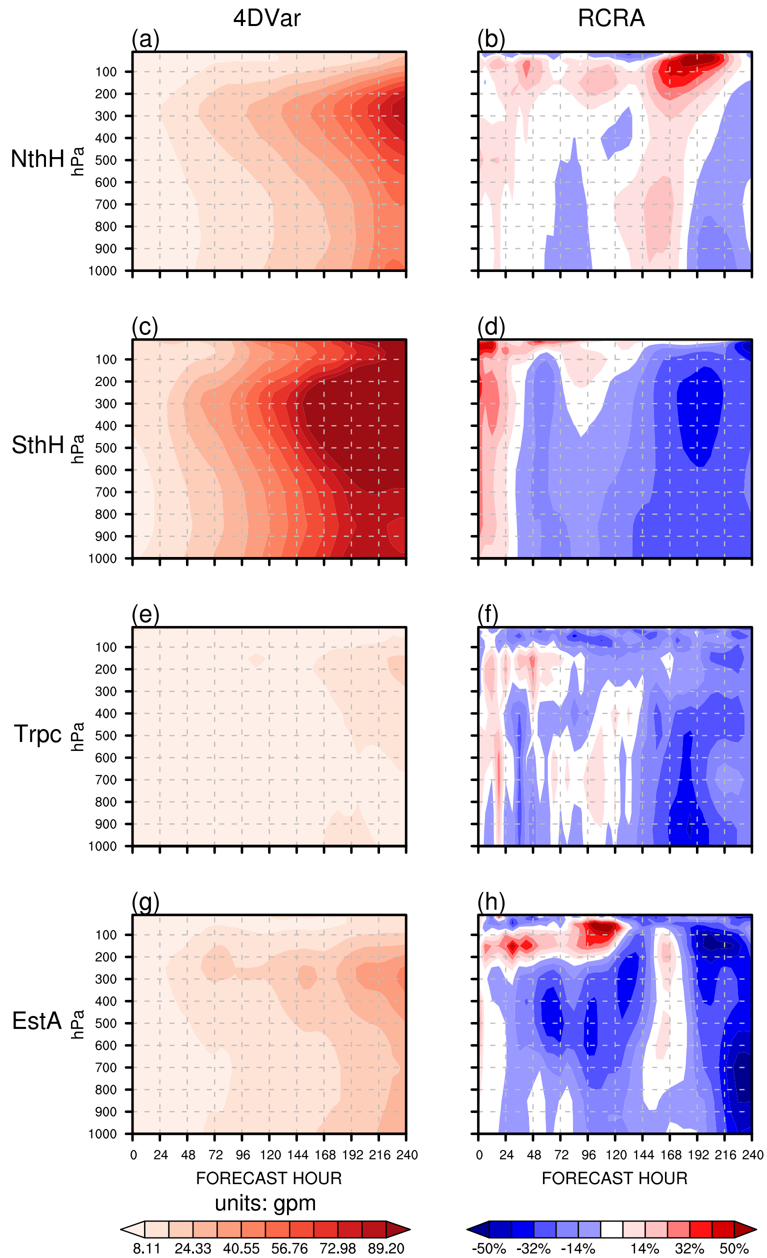
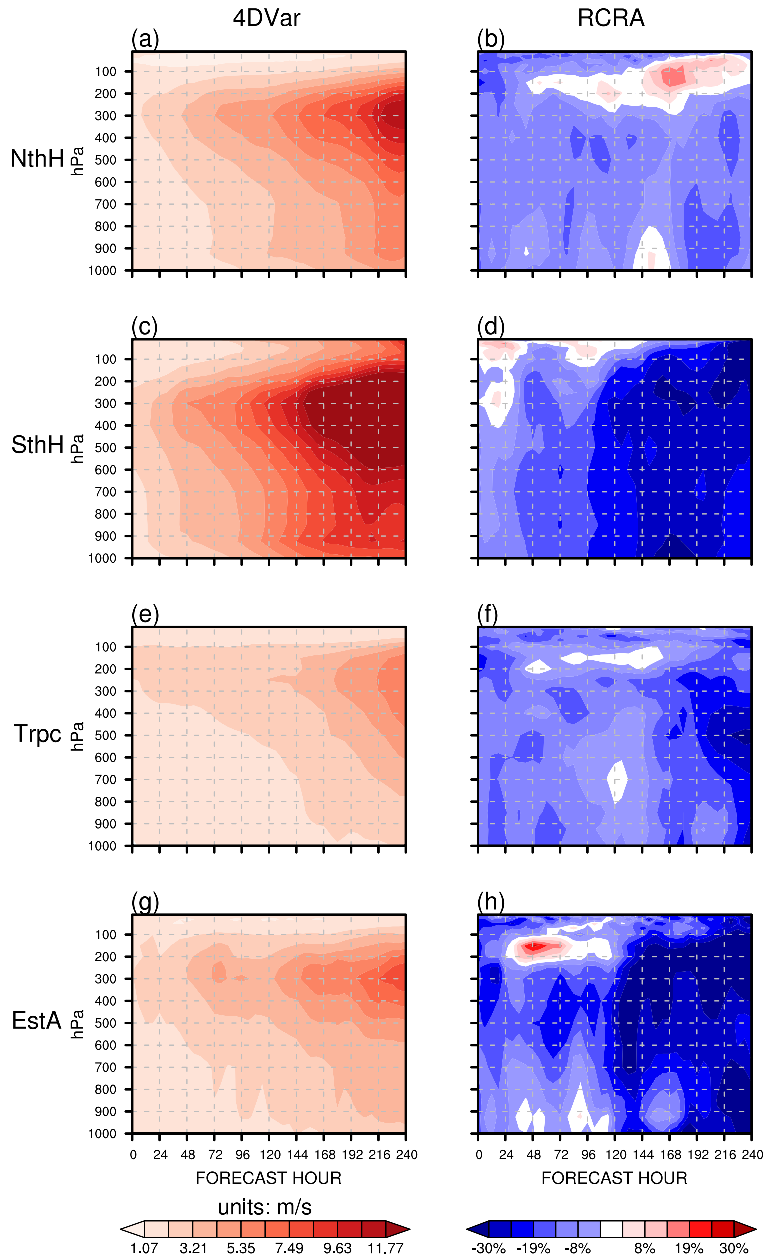


Figure 11. The ARMSEs of the geopotential height forecasts (units: gpm) initiated from the 1200 UTC 13 September 2016 analysis of the 4DVar experiment as a function of lead time (left) in the (a) Northern Hemisphere (20°N~90°N), (c) Southern Hemisphere (20°S~90°S), (e) Tropics, and (g) East Asian. The relative change rate of ARMSE (RCRA) from 4DVar to DRP-4DVar (b, d, f, h) are plotted.



992

993 **Figure 12**. As in Figure 11, but showing the results of the zonal wind forecasts.

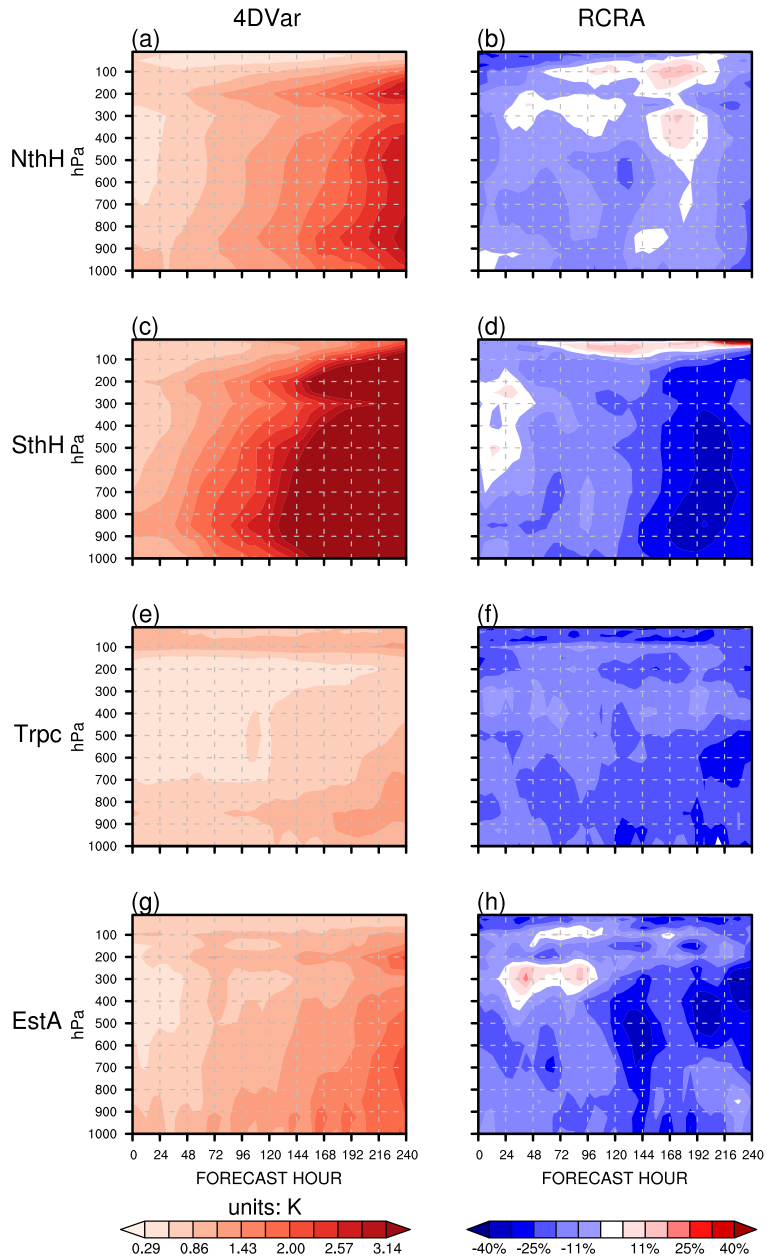
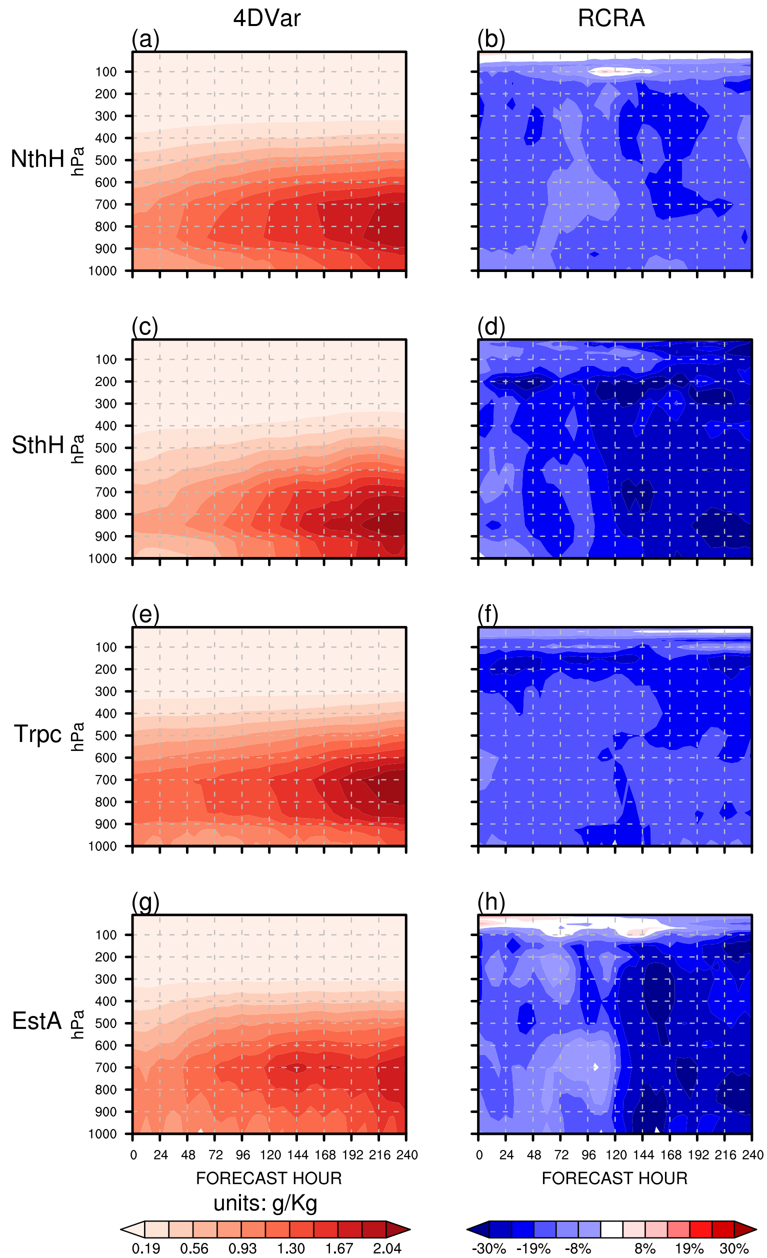


Figure 13. As in Figure 11, but showing the results of the temperature forecasts.

996 

997 **Figure 14**. As in Figure 11, but showing the results of the specific humidity forecasts.

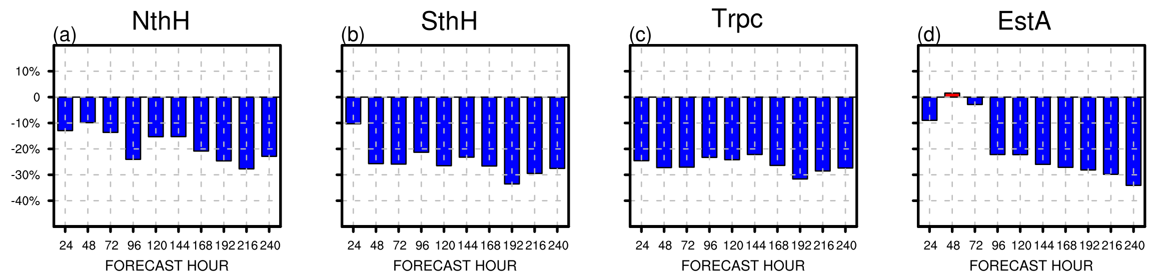


Figure 15. As in Figure 8, but showing the results of the 24-h accumulated precipitation forecasts.