A 4DEnVar-based Ensemble Four-Dimensional Variational (En4DVar) Hybrid Data Assimilation System for Global NWPs: System Description and Primary Tests

Zhu Shujun[1], Wang Bin[2], Zhang Lin[3], Liu J. J.[2], Liu Yongzhu[4], Gong Jiandong[5], Xu Shiming[1], Wang Yong[1], Huang Wenyu[6], Liu Li[7], He Yujun[8], Wu Xiangjun[3], Zhao Bin[9], and Chen Fajing[4]

[1]Department of Earth System Science, Tsinghua University
[2]LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences
[3]Research and Development Division, Numerical Weather Prediction Center of China Meteorological Administration
[4]Numerical Weather Prediction Center of China Meteorological Administration
[5]China Meteorological Administration
[6]Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing, China
[7]Tsinghua University
[8]Chinese Academy of Sciences
[9]CMA Earth System Modeling and Prediction Centre, China Meteorological Administration

November 16, 2022

Abstract

This study developed an ensemble four-dimensional variational (En4DVar) hybrid data assimilation (DA) system. Different from most of the available En4DVar systems that adopted ensemble Kalman Filter class or ensemble DA approaches to produce ensemble covariances for their hybrid background error covariances (BECs), it used a four-dimensional ensemble-variational (4DEnVar) system to obtain the ensemble covariance. The localization scheme for 4DEnVar applied orthogonal functions to decompose the correlation matrix so that it was implemented easily and rapidly. In terms of analysis quality and forecast skill, the En4DVar system was evaluated in the single-point observation experiments and observing system simulation experiments (OSSEs) with sounding and cloud-derived wind observations, using its standalone four-dimensional variational (4DVar) and 4DEnVar components as references. The single-point observation experiments visually verified the explicit flow-dependent characteristic of the BEC due to the introduction of the ensemble covariance from the 4DEnVar system. The OSSE-based sensitivity experiments revealed different contributions of the weight for the ensemble covariance in the En4DVar system to the forecasts in the Northern and Southern Extratropics and Tropics. A much higher weight for the ensemble covariance in a properly inflated hybrid covariance helped En4DVar produce the most reasonable analysis. The forecast initialized by En4DVar is overall better than by 4DVar and 4DEnVar, although the quality of En4DVar analysis is between those of 4DVar and 4DEnVar ensemble mean analyses. It indicates that the flow-dependent ensemble covariance provided by 4DEnVar dominantly contributes to the improvements in the En4DVar-initialized forecast, with certain but necessary constraint from the balanced climatological covariance.

A 4DEnVar-based Ensemble Four-Dimensional Variational (En4DVar) Hybrid Data Assimilation System for Global NWPs: System Description and Primary Tests

S. J. Zhu[1], B. Wang[2,1,5,6], L. Zhang[3,4], J. J. Liu[2,5], Y. Z. Liu[3,4], J. D. Gong[3,4], S. M. Xu[1], Y. Wang[1], W. Y. Huang[1], L. Liu[1], Y. J. He[2], X. J. Wu[3,4], B. Zhao[3,4], and F. J. Chen[3,4]

[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]CMA Earth System Modeling and Prediction Centre, China Meteorological Administration, Beijing, 100081, China,

[4]State Key Laboratory of Severe Weather, China Meteorological Administration, Beijing, 100081, China,

[5]Innovation Group 311020008, Southern Marine Science and Engineering Guangdong Laboratory, Zhuhai, China,

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

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

Key Points:

- An En4DVar data assimilation system with a hybrid background error covariance was developed for global numerical weather predictions
- The hybrid covariance is realized by linearly combining the climatological covariance of 4DVar and the ensemble covariance of 4DEnVar
- The En4DVar-initialized forecast is improved relative to 4DVar- and 4DEnVar-initialized forecasts

Abstract

This study developed an ensemble four-dimensional variational (En4DVar) hybrid data assimilation (DA) system. Different from most of the available En4DVar systems that adopted ensemble Kalman Filter class or ensemble DA approaches to produce ensemble covariances for their hybrid background error covariances (BECs), it used a four-dimensional ensemble-variational (4DEnVar) system to obtain the ensemble covariance. The localization scheme for 4DEnVar applied orthogonal functions to decompose the correlation matrix so that it was implemented easily and rapidly. In terms of analysis quality and forecast skill, the En4DVar system was evaluated in the single-point observation experiments and observing system simulation experiments (OSSEs) with sounding and cloud-derived wind observations, using its standalone four-dimensional variational (4DVar) and 4DEnVar components as references. The single-point observation experiments visually verified the explicit flow-dependent characteristic of the BEC due to the introduction of the ensemble covariance from the 4DEnVar system. The OSSE-based sensitivity experiments revealed different contributions of the weight for the ensemble covariance in the En4DVar system to the forecasts in the Northern and Southern Extratropics and Tropics. A much higher weight for the ensemble covariance in a properly inflated hybrid covariance helped En4DVar produce the most reasonable analysis. The forecast initialized by En4DVar is overall better than by 4DVar and 4DEnVar, although the quality of En4DVar analysis is between those of 4DVar and 4DEnVar ensemble mean analyses. It indicates that the flow-dependent ensemble covariance provided by 4DEnVar dominantly contributes to the improvements in the En4DVar-initialized forecast, with certain but necessary constraint from the balanced climatological covariance.

Plain Language Summary

The balance constraint and flow dependence characteristic of the background error covariance (BEC) are two key factors to affect the analysis quality of a data assimilation (DA) system. The balanced BEC used in the four-dimensional variational (4DVar) DA approach has no explicit flow dependence, while the flow-dependent BEC estimated by an ensemble-based DA method rarely keeps balanced. The hybrid of the 4DVar approach and an ensemble class DA method can achieve these two important characteristics of BEC. In this study, a hybrid DA system

called ensemble 4DVar (En4DVar) system was developed. It has two unique features. First, it used a four-dimensional ensemble-variational (4DEnVar) system to dynamically provide the ensemble covariance, which differs from most of the available En4DVar systems that estimate their dynamic covariances with the ensemble Kalman Filter class approaches or ensemble of 4DVars method. Second, the ensemble covariance was localized in the sample space using a limited number of leading eigenvectors of the correlation function. In the single-point observation experiments and observing system simulation experiments, the new En4DVar system exhibited obvious flow-dependent characteristic and higher forecast skill than both the 4DVar and 4DEnVar systems although its analysis error is between those of the latter two.

1 Introduction

Data assimilation (DA) can provide the optimal analysis of the atmospheric state for the numerical weather prediction (NWP) model combining the background forecasts and observations. The quality of the analysis relies on proper estimates of the background and observation error covariances. In recent decades, several advanced research centers have adopted the four-dimensional variational (4DVar) DA approach for the global NWP model (Rabier et al., 2000; Gauthier & Thépaut, 2001; Kadowaki, 2005; Rawlins et al., 2007; Gauthier et al., 2007; Zhang et al., 2019). 4DVar can include the constraints of both the dynamics and physics to the analysis (Rabier et al., 2000; Wang et al., 2010a), but its background error covariance (BEC) is a highly parameterized model based on the homogeneous and isotropic assumptions (Walsak and Cullen, 2014). Although the flow dependence of the BEC is implicitly realized within the assimilation window by the tangent linear model (TLM) and the adjoint model (ADM), this information cannot be developed into the next assimilation window (Lorenc, 2003). Therefore, the static climatological covariance model used by 4DVar is difficult to match with the time-variant atmospheric state (Buehner et al., 2010a). In parallel with the development of the variational approach, the ensemble Kalman filter (EnKF) approach, which estimates the BEC using ensemble forecasts, has also been noted for its timesaving and the ability to provide the explicit flow-dependent BEC spanning the assimilation windows (Evensen, 1994). This approach has been applied to many models (Whitaker et al., 2008, 2009; Buehner et al., 2010a, 2010b), but the sampling error and the underestimation of the BEC during the assimilation cycle limit its performance.

To overcome the limitation of using static covariance model in 4DVar, attempts have been made to incorporate the flow-dependent information extracted from the ensemble forecasts into the standard variational framework. This is typically realized by linearly combining the climatological covariance and the ensemble covariance (Hamill & Snyder, 2000), or using the control variables preconditioned by the square root of the ensemble covariance to extend the variational control variables (Lorenc, 2003), or using ensemble information to estimate the parameters of the covariance model (Bonavita et al., 2012). Upon these advances, the hybrid DA system including ensemble information is gradually replacing the standard 4DVar system (Raynaud et al., 2011; Bonavita et al., 2012; Clayton et al., 2013; Lorenc et al., 2015; Wang et al., 2013; Kleist & Ide, 2015a, 2015b). The European Centre for Medium-range Weather Forecasts

(ECMWF) and Météo-France used the ensemble of 4DVars to provide the dynamic BECs for the hybrid DA system and achieve higher forecast skills (Bonavita et al., 2012; Raynaud et al., 2011). The Met Office combined the climatological covariance and the ensemble covariance obtained by the ensemble transformed Kalman filter (ETKF) method to construct hybrid DA system (Clayton et al., 2013). The use of the TLM and ADM is retained in the aforementioned hybrid DA systems, which is referred to as the En4DVar method. Another method called 4DEnVar, which avoids the use of the ADM, has also been implemented in some prediction centers. The National Centers for Environmental Prediction (NCEP), and the Met Office utilized the ensemble covariance obtained from the ensemble Kalman filter (EnKF) class methods, namely the ensemble square root filter (EnSRF) method, and the ensemble transform Kalman Filter (ETKF) method to construct the hybrid 4DEnVar system, respectively (Kleist & Ide, 2015b; Bowler et al., 2017a).

Recently, there have been interests in using the 4DEnVar system to estimate the ensemble covariance for hybrid DA systems. The Met Office verified that the hybrid DA system can be benefited from using the En-4DEnVar system instead of using the ETKF system (Bowler et al., 2017a). Some 4DEnVar methods that can reduce the dimension of the optimization problem from the model space to a subspace composed by a set of basis vectors and can avoid the ADM have been proposed (Qiu et al., 2007; Tian et al., 2008; Wang et al., 2010a). The dimension-reduced projection 4DVar (DRP-4DVar) method is one of the representative 4DEnVar methods. In contrast to the extensive and successful applications of the DRP-4DVar method 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., 2021a, 2021b), its applications in global medium-range weather forecasts are very lacking except an early primary tests on it based on a three-dimensional variational (3DVar) DA system (Shen et al., 2015) and a very recent effort under a 4DVar framework (Zhu et al., 2022).

Using the original ensemble covariance directly may result in sampling error and rank deficiency problems, and localization techniques can alleviate these problem (Hamill et al., 2001). The localization of the ensemble covariance in hybrid DA systems is usually implemented upon unbalanced variables in model space using the transformations copied from the variational system (Clayton et al., 2013; Lorenc, 2015). In addition to this implementation of localization (e.g., spectral filtering transformation), ensemble-sample-based subspace localization approaches using

orthogonal functions to decompose the correlation function were developed (Liu et al., 2009; Buehner et al., 2010a, 2010b; Bishop et al., 2011; Kuhl et al., 2013; Wang et al., 2010b, 2018).

This study focuses on developing an En4DVar hybrid DA system for the global NWP model. It has two unusual aspects compared with other hybrid DA systems. On one hand, a 4DEnVar system based on the DRP-4DVar method, which uses the statistical relationship between the model space and the observation space to avoid the use of ADM, was utilized to estimate the ensemble covariance for the En4DVar system. On the other hand, a limited number of leading eigenvectors of the correlation function was employed to localize the ensemble covariance instead of localization transformations from the climatological covariance model.

The purpose of this study is to provide a mathematical derivation and system introduction to the En4DVar hybrid DA system. The performance of this system was also evaluated with single-point observation experiments and observing system simulation experiments (OSSEs), using its standalone 4DVar and 4DEnVar components with identical settings as references. The theoretical basis of the En4DVar system is presented in Section 2. The designs of the system and experiments are introduced in Section 3. The results are analyzed in Section 4. Finally, the conclusions and discussions are given in the last section.

2 Methodology

2.1 Climatological covariance

The global 4DVar system (Zhang et al., 2019) used in this study minimizes the following incremental formulation (Courtier, 2004) defined within the DA window $[t_0, t_n]$ by introducing $\delta x = x - x^g$, and obtained the optimal analysis increment at the initial time t_0:

$$J[\delta x(t_0)] = \frac{1}{2}\{\delta x(t_0) - [x^b(t_0) - x^g(t_0)]\}^T B_0^{-1}\{\delta x(t_0) - [x^b(t_0) - x^g(t_0)]\} + \frac{1}{2}\sum_{i=0}^{n}[\mathbf{H}_i\delta x(t_i) - d_i]^T R_i^{-1}[\mathbf{H}_i\delta x(t_i) - d_i]. \quad (1)$$

Here, x and δx represent the state vector and its increment, respectively; x^b and x^g denote the state vectors of background and first guess, separately; B and R are severally the BEC matrix (B-matrix) and observation error covariance matrix, respectively; $d_i = y_i^o - H_i[x^g(t_i)]$ and y_i^o represent the observation innovation and observation itself at time t_i, separately; H_i and $\mathbf{H}_i$ are the nonlinear observation operator and its tangent linear operator at time t_i,

correspondingly; M_i and $\mathbf{M}_i$ mean the nonlinear forecast model and its tangent linear model that integrate from the analysis time to time t_i, severally; and $\delta x(t_i) = \mathbf{M}_i \delta x(t_0)$ it indicates the forecast state vector at time t_i initialized from the increment state.

Introducing the preconditioning transformation $\delta x = Uv$, the condition number of the Hessian matrix of Eq. (1) can be improved. Thus, the background term of Eq. (1) can be simplified as $\frac{1}{2}[v(t_0)]^T[v(t_0)]$. Here, U is the square root of the climatological B-matrix B, and v is the preconditioned state variable vector.

The BEC utilized in the standard 4DVar system is a highly parameterized model:

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

U includes multiple transformation processes:

$$U = U_p s^u U_h U_v. \tag{3}$$

Here, U_h and U_v represent the horizontal and vertical transformations of the background error correlation transformation matrix, respectively. s^u denotes the diagonal matrix consisting of the background root mean square error of the approximately uncorrelated variables (e.g., stream function, unbalanced velocity potential, unbalanced non-dimensional pressure and specific humidity), which is usually obtained by the National Meteorological Center (NMC) method and zonally averaging. U_p is the physical transformation operator that converts the approximately uncorrelated variables into the model variables (Zhang et al., 2019).

Although the standard 4DVar system realizes the implicit flow-dependent characteristic of the BEC within the assimilation window using the TLM and ADM, this characteristic cannot carry over to the next assimilation window. Thus, the abovementioned climatological covariance model can only simply describes the mean state of BEC, but cannot represent the evolution of BEC with the weather state in detail.

2.2 Ensemble covariance

Unlike the static climatological covariance, the ensemble covariance can realize the explicit flow-dependence of the BEC using the ensemble forecasts initialized from different initial conditions (ICs).

175 The ICs to initialize the ensemble forecasts are provided by the 4DEnVar system (Zhu et
176 al., 2022). This system was developed using the DRP-4DVar method that solves the optimization
177 problem on the subspace composed of a limited number of ensemble samples (Wang et al., 2010a).
178 The 4DEnVar system uses a set of IC perturbation samples in model space $\delta X =$
179 $[\delta x_1, \delta x_2, \cdots, \delta x_K]$ (K is the ensemble size), and the corresponding observational perturbation
180 samples $\delta Y = [\delta y_1, \delta y_2, \cdots, \delta y_K]$ calculated by δX to define the projection matrices:

$$\begin{cases} p_x = \frac{1}{\sqrt{K-1}}\left[\delta x_1 - \overline{\delta x}, \delta x_2 - \overline{\delta x}, \cdots, \delta x_K - \overline{\delta x}\right] \\ p_y = \frac{1}{\sqrt{K-1}}\left[\left[\delta y_1 - \overline{\delta y}, \delta y_2 - \overline{\delta y}, \cdots, \delta y_K - \overline{\delta y}\right]\right] \end{cases}. \tag{4}$$

181 Here, $\overline{\delta x}$ and $\overline{\delta y}$ are the ensemble means of the IC perturbation samples in model space and
182 observational perturbation samples in observation space, respectively. δy_k is calculated using the
183 TLM and tangent linear observation operators. δx and $\mathbf{H}_i \delta x(t_i)$ can be projected onto the
184 subspace composed of the aforementioned projection matrices as the following:

$$\begin{cases} \delta x = p_x \alpha \\ \mathbf{H}_i \delta x(t_i) = p_y(t_i)\alpha \end{cases}, \tag{5}$$

185 where $\alpha = \alpha(t_0) = \left(\alpha_1(t_0), \alpha_2(t_0), \cdots, \alpha_N(t_0)\right)^T$ is the vector composed of the linear
186 combination coefficients.

187 Substituting Eq. (5) into Eq. (1), a new cost function defined in the subspace is obtained:

$$J[\alpha] = \frac{1}{2}[\alpha]^T[\alpha] + \frac{1}{2}\sum_{i=0}^{n}\left[p_y(t_i)\alpha - d_i\right]^T R_i^{-1}\left[p_y(t_i)\alpha - d_i\right]. \tag{6}$$

188 The optimal analysis can be obtained by minimizing Eq. (6) and using the first transformation in
189 Eq. (5).

190 The ensemble covariance can be estimated using the ensemble forecasts initialized from
191 the previous ensemble analyses produced by the 4DEnVar system:

$$B_e = p_x p_x^T. \tag{7}$$

192 The limited ensemble size, which is much smaller than the dimension of the model space may
193 easily result in the spurious correlation and rank deficiency problems in the BEC. Thus, a
194 localization technique should be conducted upon the ensemble covariance (Hamill et al., 2001).

The localized ensemble BEC can be implemented as the Schür product between the ensemble BEC and the localization correlation matrix C:

$$B_e = (p_x p_x^T) \circ C, \tag{8}$$

where $\circ$ represents the Schür (element-by-element) product of two matrices with the same dimension. The localization correlation matrix C is designed to reduce the spurious correlation in the ensemble covariance and as a result to increase the rank of BEC.

In implementation, the zonal (meridional) correlation model is defined adopting the GC correlation function (Gaspari & Cohn, 1999) with respect to the dimensionless distance r in zonal (meridional) direction:

$$C(r) = \begin{cases} -\frac{1}{4}r^5 + \frac{1}{2}r^4 + \frac{5}{8}r^3 - \frac{5}{3}r^2 + 1, 0 \le r \le 1 \\ \frac{1}{12}r^5 - \frac{1}{2}r^4 + \frac{5}{8}r^3 + \frac{5}{3}r^2 - 5r + 4 - \frac{2}{3}r^{-1}, 1 < r \le 2 \\ 0, 2 < r \end{cases}, \tag{9}$$

and the vertical correlation model is defined using the following correlation function with respect to the non-dimensional logarithmic pressure distance r:

$$C(r) = \frac{1}{1.0 + K_p r^2}. \tag{10}$$

Due to using the localized ensemble covariance in Eq. (8) directly may lead to lots of computational costs, the localization correlation matrix C is usually decomposed into transformations that can be copied from the variational system (Clayton et al., 2013; Lorenc, 2015). The localization of the ensemble covariance may lead to the reduction of sampling errors through producing extended ensemble samples based on the Schür products between the ensemble forecasts with small ensemble size and a number of leading eigenvectors of C, without a lot of extra complex modeling. Therefore, the localization can be implemented in an extended-ensemble-sample-based subspace. Through approximately decomposing the correlation matrix into a number of leading eigenvectors and ignoring their time-variation, the extended IC perturbation samples in model space and the corresponding observational perturbation samples are obtained:

$$\begin{cases} Ep_x = \left[\left(p_{x,1} \circ \boldsymbol{\rho}_{x,1}, \cdots, p_{x,1} \circ \boldsymbol{\rho}_{x,L}\right), \cdots, \left(p_{x,N} \circ \boldsymbol{\rho}_{x,1}, \cdots, p_{x,N} \circ \boldsymbol{\rho}_{x,L}\right)\right] \\ Ep_y = \left[\left(p_{y,1} \circ \boldsymbol{\rho}_{y,1}, \cdots, p_{y,1} \circ \boldsymbol{\rho}_{y,L}\right), \cdots, \left(p_{y,N} \circ \boldsymbol{\rho}_{y,1}, \cdots, p_{y,N} \circ \boldsymbol{\rho}_{y,L}\right)\right] \end{cases}, \tag{11}$$

where $\boldsymbol{\rho}_{x,j}$ and $\boldsymbol{\rho}_{y,j}(j = 1,2,\cdots,L)$ are the selected leading eigenvectors in model and observation spaces according to the cumulative contribution of variance, respectively. The leading eigenvectors can be further decomposed into zonal, meridional and vertical components.

The leading eigenvectors are obtained using the empirical orthogonal function (EOF) in the zonal and vertical directions, and the sine function in the meridional direction:

$$\begin{cases} \boldsymbol{\rho}_{x,j} = \boldsymbol{E}_{x,j}\left(\lambda_{x,j}\right)^{1/2} \\ \boldsymbol{\rho}_{y,j} = \boldsymbol{E}_{y,j}\left(\lambda_{y,j}\right)^{1/2} \end{cases}. \tag{12}$$

Here, $\boldsymbol{E}_{x,j}$ and $\boldsymbol{E}_{y,j}$ are the eigenvectors of the correlation functions in the model space and observation space, and $\lambda_{x,j}$ and $\lambda_{y,j}$ are the corresponding eigenvalues, respectively.

Redefining an $K \times L$-dimensional vector $\beta = \beta(t_0)$ that consists of the combination coefficients of the extended ensemble samples, a newer cost function is defined in the extended subspace composed of the extended ensemble samples:

$$J[\beta] = \frac{1}{2}[\beta]^T[\beta] + \frac{1}{2}\sum_{i=0}^{n}\left[Ep_y(t_i)\beta - d_i\right]^T R_i^{-1}\left[Ep_y(t_i)\beta - d_i\right]. \tag{13}$$

The optimal analysis can be obtained by minimizing Eq. (13) and using the transformation $\delta x = Ep_x\beta$. Moreover, the localized ensemble BEC can be expressed as:

$$B_e = Ep_x Ep_x^T. \tag{14}$$

2.3 Hybrid covariance

The hybrid covariance can be obtained by a linear combination of the climatological covariance and the ensemble covariance:

$$B = \gamma_c B_c + \gamma_e B_e. \tag{15}$$

Here, γ_c and γ_e denote the scalar weights of the climatological and ensemble covariances, respectively. Both the climatological and ensemble covariances of the En4DVar system have a time-variant feature inside the assimilation window, but the way to achieve the time-variant feature is different. The climatological (ensemble) covariance realizes the implicit flow dependence through the TLM and ADM (the statistical relationship between the model space and the observation space). However, they greatly differ in the inclusion of balance constraint and characteristic of explicit flow dependence spanning the assimilation windows. The climatological

covariance includes the balance constraint but fails to store and use its implicit evolution information for the next assimilation window, while the ensemble covariance explicitly evolves with state from one window to the next but includes little balance constraint. Therefore, a hybrid of these covariances may be an efficient way to keep the balance and characteristic of explicit flow dependence in the BEC.

To achieve a hybrid BEC, we use the extended control variable approach (Lorenc, 2003), which is mathematically equivalent to directly combining the covariances linearly (Wang et al., 2007). The climatological and ensemble contributions to the increment δx are multiplied by $\sqrt{\gamma_c}$ and $\sqrt{\gamma_e}$, respectively:

$$\delta x = \sqrt{\gamma_c}\delta x_c + \sqrt{\gamma_e}\delta x_e, \tag{16}$$

where $\delta x_c = Uv$, $\delta x_e = Ep_x\beta$.

Then, the new cost function can be obtained:

$$\begin{aligned} J[v(t_0), \beta(t_0)\,] = {} & \tfrac{1}{2}[v(t_0)]^T[v(t_0)] + \tfrac{1}{2}[\beta(t_0)]^T[\beta(t_0)] + \\ & \tfrac{1}{2}\textstyle\sum_{i=0}^{n}[\mathbf{H}_i\mathbf{M}_i\delta x(t_0) - d_i]^T R_i^{-1}[\mathbf{H}_i\mathbf{M}_i\delta x(t_0) - d_i]. \end{aligned} \tag{17}$$

To minimize Eq. (13), $v(t_0)$ and $\alpha(t_0)$ needs to satisfy $\begin{cases} \left[\frac{\partial J}{\partial v(t_0)}\right]^T = 0 \\ \left[\frac{\partial J}{\partial \beta(t_0)}\right]^T = 0 \end{cases}$. The calculation of the gradient for the variational component in Eq. (17) requires the use of the ADM, while that for the ensemble component can avoid the use of the ADM by adopting the transformation $\mathbf{H}_i\delta x(t_i) = Ep_y(t_i)\beta$:

$$\left[\frac{\partial J}{\partial \beta(t_0)}\right]^T = \beta(t_0) + \sqrt{\gamma_e}\textstyle\sum_{i=0}^{n}\left[Ep_y(t_i)\right]^T R_i^{-1}[\mathbf{H}_i\mathbf{M}_i\delta x(t_0) - d_i]. \tag{18}$$

3 Experimental design

3.1 Configuration of En4DVar system

The En4DVar system developed in this study realized the hybrid BEC by a linear combination of the static climatological covariance from the standard 4DVar system and the flow-dependent ensemble covariance from the 4DEnVar system. The ensemble covariance in the

En4DVar system is estimated by the low-resolution ensemble samples, which is produced by previously running 60-member 4DEnVar system. The 4DVar and 4DEnVar systems all adopt a dual-resolution framework with $0.5° \times 0.5°$ for the outer loop and $1.0° \times 1.0°$ for the inner loop. The flowchart of the En4DVar system is given in Figure 1.

Unlike the ensemble prediction systems that use singular vectors as initial perturbations without including any observations, the 4DEnVar system itself produces the initial perturbations that considers both the model constraints and observational information. The IC samples of the 4DEnVar system in the first assimilation window are generated by the "randomcv" method (Baker, 2005), which generates an ensemble of reasonable and balanced samples by the control variable transformation $\delta x = Uv$. To alleviate the problems of sampling error and underestimation in BEC during the assimilation cycle, localization, inflation, and observation and SST perturbation techniques are employed. Flow-dependent ensemble samples are generated by assimilating perturbed observations at every 6-hour assimilation window during the assimilation cycle (Figure 1a).

In the En4DVar system, 4DVar uses the same 6-hour assimilation window, and the analysis time is located at the beginning of the assimilation window as 4DEnVar does. Therefore, at the beginning of each assimilation window of 4DVar, the flow-dependent ensemble samples in model space and the corresponding observational ensemble samples at different time slots can be extracted from the 4DEnVar system to estimate the ensemble covariance for the En4DVar system (Figure 1b). A localization scheme with a filter radius of 7° in the horizontal directions and a filter parameter K_p of 3 in the vertical direction is designed based on the theoretical basis introduced in Section 2.2, which is utilized to remove the spurious remote correlations in the ensemble covariance. The optimal problem of the ensemble component of the En4DVar is solved in a subspace consisting of extended samples obtained by the Schür product between the ensemble samples and the leading eigenvectors of the localization correlation function (Figure 1a).

3.2 Experiment design

To verify the impact of introducing the ensemble covariance into the 4DVar system, six sets of experiments were conducted based on the En4DVar, standard 4DVar and 4DEnVar systems using identical settings, respectively, including three sets of assimilation cycle experiments and three sets of deterministic 10-day forecast experiments. The experiments based on the latter two

DA systems were used as the references for comparisons and evaluations. The analyses from the 4DEnVar system used the ensemble means of its 60 analysis samples.

All experiments were conducted and evaluated based on an OSSE. The forecast model used in these experiments are the GRAPES-GFS model version V3.0 with three different horizontal resolutions (i.e., 0.25° × 0.25°, 0.5° × 0.5° and 1.0° × 1.0°) and the same 87 vertical levels (Su et al., 2020). The 4DVar system of the GRAPES-GFS model was in use in operational global forecasts, which was demonstrated to outperform the operational 3DVar system in terms of both the analysis error and medium-range forecast skill (Zhang et al., 2019). OSSE allows an objective evaluation of DA system (Wang et al., 2008; Wang et al., 2010a; Kleist & Ide, 2015a , 2015b). Here, the GRAPES-GFS version with the resolution of 0.25° × 0.25° was used to generate an uninterrupted free run from 0600 UTC 10 September to 1200 UTC 22 October, 2016, initialized from the ERA-5 reanalysis, and the results of this run starting from the second day were used as the "truth" state of the OSSE. The run for the first day was adopted to eliminate the effects of spin-up, and the simulated geopotential height and precipitation of the run for the first 8 days were verified using the ERA-Interim reanalysis. The "observations" were the "truth" state transformed from the model space to the observation space by the observation operators of the sounding and cloud-derived wind superimposed with normal distribution random perturbations of which the expectations equal zero and the standard deviations adopt the observation errors. In addition, the 15-hour forecast by the GRAPES-GFS version with the resolution of 0.5°×0.5° initialized from the 6-hour forecast of the ERA-Interim dataset was used as the background field of three sets of assimilation cycle experiments for the first assimilation window. For the 4DEnVar systems in the first assimilation window, 60 ensemble background fields were produced using the aforesaid background plus 60 random perturbations whose ensemble mean is equal to zero. The results of the experiments for the first 2 days were used for spin-up, and those from day 3 onwards for about one month covering the period from 0900 UTC 13 September 2016 to 0900 UTC 11 October 2016 were analyzed. To facilitate comparisons with the reanalysis data, the analyses at the middle of each assimilation window were produced with 3-h forecasts initialized from the corresponding analyses at the beginning of the window for the evaluations and comparisons. Restarted from these analyses, three sets of 10-day deterministic forecast experiments were carried out and the forecast skills were evaluated and compared.

Moreover, in order to visually verify the performances and flow-dependent characteristics of the DA systems, the results of the single-point observation experiments in the assimilation window (0900 UTC 13 September 2016 to 1500 UTC 13 September 2016) on the third day of the OSSE were used. Three single-point observation experiments based on the En4DVar system with three linear combinations of the variational and ensemble components adopted the same background, single-point observation, observation error (0.95°K) and horizontal filter radius (15°) of localization. The three combinations included $\gamma_c = 0.8$ and $\gamma_e = 0.2$, $\gamma_c = 0.5$ and $\gamma_e = 0.5$ and $\gamma_c = 0.2$ and $\gamma_e = 0.8$ in Eq. (16), respectively. The background used the ensemble mean of the background samples of the 4DEnVar system at the beginning of this assimilation window. The single-point observation was selected from the sounding temperature observations in the OSSE valid at 1200 UTC 13 September 2016, located upstream of the top of the short-wave ridge (60.77°S, 161.83°W) in the middle troposphere (500hPa), with an observation innovation of -1.53K.

4 Results

4.1 Single-point observation experiments

To demonstrate the impact of introducing the ensemble covariance provided by the 4DEnVar system, three experiments assimilating the same single-point observation by the En4DVar system with different linear combinations for its hybrid covariance were conducted. The analysis increments produced in these experiments are shown in Figure 2. All the experiments show maximum negative temperature increments near the observation at the beginning of the assimilation window resulting from assimilating the single-point observation, implying lower temperatures in three analysis fields than in the background field. There are also clear cyclonic wind responses around the maximum negative temperature increments, which indicates that the BEC of all experiments satisfy certain balance constraints. In addition, it is satisfactory that no significant spurious correlations are exhibited in the experiments due to the localization of the ensemble BEC (Figure 2a-2c).

In addition to examining the basic assimilation capability, the single-point observation experiments can also visualize the flow-dependent characteristic of the BEC. On one hand, by visualizing the evolution of the increments over time, we can find that the center of the maximum

346 negative value of all experiments moves to the southeast of the observation. At the end of the
347 assimilation window, the movement of the center of the maximum negative value and the
348 extension of the increment along the gradient of the geopotential height are most obvious. This is
349 reasonable considering that the background field in the vicinity of the observation is the northwest
350 flow, and indicates that the BECs in all the experiments have flow-dependent characteristics. On
351 the other hand, the analysis increments in the experiments with higher weight to the ensemble
352 covariance extend more strongly along the gradient of the geopotential height and show more
353 obvious asymmetric distributions. This highly weighted ensemble covariance helps the En4DVar
354 system acquire a more obvious flow-dependent characteristic than the lowly weighted one.

355 4.2 Observing system simulation experiments

356 4.2.1 Evaluation metric

357 To evaluate the performance of the En4DVar system, its analysis quality and forecast skill
358 were measured by comparing the analysis and forecast against the "truth" state. The root mean
359 square error (RMSE) is usually utilized to measure the analyses and forecast errors. Here, we
360 adopted a metric called anomaly root mean square error (ARMSE; He et al., 2021a) instead of
361 RMSE to exclude the systematic bias and only measure the random error that cannot be corrected
362 like the systematic bias. This metric is defined as:

$$ARMSE = \sqrt{\frac{\sum_{n=1}^{N} w_{(n)} \times \left(M_{(n)} - truth_{(n)} - bias\right)^2}{\sum_{n=1}^{N} w_{(n)}}}, \tag{19}$$

363 where $w_{(n)}$ is the weighted coefficients at the n-th grid point, N is the grid size, $M_{(n)}$ and $truth_{(n)}$
364 denote the forecast (or analysis) and the "truth" state at the n-th grid point, $bias$ represents the
365 mean error ($bias = \frac{\sum_{n=1}^{N} w_{(n)} \times \left(M_{(n)} - truth_{(n)}\right)}{\sum_{n=1}^{N} w_{(n)}}$). We use the anomaly correlation coefficient (ACC) metric
366 to assess the similarity between the anomalies of the forecast (or analysis) and the "truth" state.
367 The statistically significant results of the differences of the metrics between different experiments
368 were given by the paired t-test. In addition, to facilitate comparisons of the abovementioned
369 metrics, we divided the globe into three regions including the Northern Extratropics (20°N~90°N,
370 180°W~180°E; NH-X), Southern Extratropics (20°S~90°S, 180°W~180°E; SH-X) and Tropics
371 (20°S~20°N, 180°W~180°E; TR).

372 4.2.2 Weights of climatological and ensemble covariances in the hybrid covariance

373 In this subsection, the weight coefficients of the climatological and ensemble covariances
374 in the hybrid covariance were determined based on two sets of sensitivity experiments. These
375 sensitivity experiments aim to investigate the effect of different linear combinations of two
376 covariances in the hybrid covariance with and without inflation on the analysis quality of the
377 En4DVar system, respectively. Without inflation, the sum of the weight coefficients of two
378 covariances is equal to 1, i.e., $\gamma_c + \gamma_e = 1$. In this case, the first set of sensitivity experiments were
379 conducted using a small, a medium and a large weight coefficients of the ensemble covariance,
380 i.e., ($\gamma_c = 0.8, \gamma_e = 0.2$), ($\gamma_c = 0.5, \gamma_e = 0.5$) and ($\gamma_c = 0.2, \gamma_e = 0.8$), respectively. The
381 scorecards of the ACC and ARMSE for the forecasts initialized by the En4DVar system with the
382 aforesaid three linear combinations against the 4DVar-initialized forecast are shown in Figure 3.
383 The En4DVar system with these three combinations shows significant improvements on the first
384 4 lead forecast days and comparable performances on the last 5 lead days in the Northern
385 Extratropics compared with the 4DVar system. In Tropics, its improvements extend to almost all
386 the lead forecast days. In the Southern Extratropics, however, performance degradations occurs in
387 the lead days 2-5. In summary, introducing the ensemble covariance from the 4DEnVar system
388 into the 4DVar system benefits the short-term forecasts in the Northern Extratropics and the
389 forecasts on almost all lead days in Tropics, although it degrades the short-term forecasts in the
390 Southern Extratropics. Without inflation, the En4DVar system with the same weight for both the
391 climatological and ensemble covariances performs best, indicating equal importance of the balance
392 constraint and flow dependence in assimilation analysis.

393 Moreover, four inflated linear combinations in which the sum of weight coefficients of two
394 covariances is larger than 1 were used in the En4DVar system to seek further improvements of the
395 forecasts, especially in the Southern Extratropics. They include ($\gamma_c = 0.8, \gamma_e = 0.25$), ($\gamma_c =$
396 $0.6, \gamma_e = 0.5$), ($\gamma_c = 0.5, \gamma_e = 0.6$) and ($\gamma_c = 0.25, \gamma_e = 0.8$). Figure 4 displays the scorecards of
397 the ACC and ARMSE for the En4DVar-initialized forecasts with the above combinations against
398 the 4DVar-initialized forecasts, from which it is found that the En4DVar system with all
399 combinations except ($\gamma_c = 0.5, \gamma_e = 0.6$) in the Southern Extratropics outperforms the 4DVar
400 system in the globe. It is suggested that linear combinations with proper inflations perform better
401 than those without any inflations (Figures 3 and 4). In particular, some of the inflated combinations

shows improvements in the Southern Extratropics. This is probably due to the underestimation of the hybrid covariance without inflation although the ensemble covariance has already been inflated in the 4DEnVar system, and therefore proper inflations are required to allow more observational information to be incorporated into ICs for better forecast skills.

Finally, the inflated combination in which the ensemble covariance is dominant, i.e., ($\gamma_c = 0.25, \gamma_e = 0.8$), was chosen as the optimal combination to construct the hybrid covariance of the En4DVar system. The reasons include three aspects. Firstly, the combinations without inflation have worse performances than those with inflation in the Southern Extratropics (Figures 3 and 4) and thus they are impossibly the optimal choice. Secondly, among the four inflated combinations, the combination ($\gamma_c = 0.5, \gamma_e = 0.6$) can be excluded, because it performs worst in the Southern Extratropics although it has comparable performance to the others in the Northern Extratropics and Tropics (Figure 4). Thirdly, the combinations ($\gamma_c = 0.6, \gamma_e = 0.5$) and ($\gamma_c = 0.25, \gamma_e = 0.8$) are better than ($\gamma_c = 0.8, \gamma_e = 0.25$) except some degradations in their initialized short-term forecasts in the Southern Extratropics (Figures 5a-5b). At last, the combination ($\gamma_c = 0.25, \gamma_e = 0.8$) basically outperforms ($\gamma_c = 0.6, \gamma_e = 0.5$), although it degrades the short-term forecasts in the Southern Extratropics compared with the latter (Figure 5c). In addition, it slightly improves the medium-range forecasts compared with the 4DVar system (Figure 4d).

The dominant role of the ensemble covariance in the hybrid BEC of the En4Var system in this study quite differs from those in the available En4DVar systems in which the climatological covariances are usually dominant (Clayton et al., 2013; Lorenc et al., 2015), indicating the high quality of the initial perturbations generated by the 4DEnVar system that allows increasing the weighting of the ensemble covariance in an En4DVar system (Clayton et al., 2013; Bowler et al., 2017a).

4.2.3 Analysis quality

After the weight coefficients of the climatological and ensemble covariances were determined, the En4DVar system was completed and came into evaluations in terms of analysis quality. The analysis errors of the En4DVar, 4DVar and 4DEnVar systems verified against the "truth" state were compared and evaluated to investigate whether the En4DVar system further improved the analysis quality on the basis of both 4DVar and 4DEnVar systems.

First, the vertical profiles of the ARMSE differences of the analysis and background between the En4DVar and 4DVar systems and between the En4DVar and 4DEnVar systems are provided in Figure 6. In general, the ARMSE differences between the En4DVar and 4DVar analyses and backgrounds keep negative for zonal wind, temperature and specific humidity in the globe on almost all vertical layers. It is indicated that En4DVar fully outperforms 4DVar in terms of both analysis and background errors except the degraded performances in zonal wind and temperature at the vertical layers above 100hPa in the Southern Extratropics, with more significant improvements in the analysis than in the background (Figures 6d-6l). For geopotential height, En4DVar performs slightly better than 4DVar in the Northern Extratropics and on the middle and upper layers in the Tropics, and slightly (significantly) worse on the middle (upper) layers in the Southern Extratropics and on the lower layers in the Tropics (Figures 6a-6c). In contrast, the ARMSE differences between the En4DVar and 4DEnVar ensemble mean analysis and background errors have various signs for different variables at different vertical layers. The analysis and background of geopotential height in En4DVar outperform the 4DEnVar ensemble mean, especially on the upper layers in the Southern Extratropics (Figure 6a-6c). The error differences of zonal wind are positive on the layers below 400hPa (200hPa) in the Northern and Southern Extratropics (Tropics), indicating less observational information incorporated by the En4DVar than by the 4DEnVar ensemble mean on these layers (Figures 6d-6f). For temperature, the En4DVar performs worse than the 4DEnVar ensemble mean on the lower layers except the layers near surface in the Northern and Southern Extratropics and on the layers above 200hPa in the Northern Extratropics and Tropics, but better on the layers between 500hPa and 200hPa (above 500hPa) in the Northern Extratropics and Tropics (Southern Extratropics) (Figures 6g-6i). The analysis and background of specific humidity in En4DVar are inferior to those in 4DEnVar ensemble mean (Figures 6j-6l). In conclusion, the analysis and background of En4DVar draw closer to the observations than those of 4DVar during the assimilation cycle by introducing flow-dependent information produced by 4DEnVar although they are not as close to the observations as those of 4DEnVar ensemble mean.

The pressure-latitude distributions of the ARMSEs in the analysis fields of geopotential height, zonal wind, temperature and specific humidity were further analyzed (Figure 7). The error maxima in the En4DVar analyses of geopotential height occur in the stratosphere near 60°S (Figure 7a). The error maxima in the zonal wind analyses related to jets are mainly concentrated

462 in the stratosphere and middle and upper troposphere near 60°S, which extend toward lower
463 troposphere (Figure 7d). The error maxima in the temperature analyses occur near the surface from
464 the poles to the equator and extend to the middle and upper troposphere and stratosphere near 60°S
465 (Figure 7g). In addition, the error structure of the specific humidity analyses showed a semicircular
466 region with large values at low and middle latitudes in the lower and middle troposphere (Figure
467 7j). The ARMSE differences between the En4DVar and 4DVar analyses of geopotential height,
468 zonal wind, temperature and specific humidity are shown in the middle column of Figure 7. The
469 differences are negative on almost all vertical layers at all latitudes for all four variables except the
470 stratosphere at high latitudes in the Southern Hemisphere for geopotential height, zonal wind and
471 temperature, which indicate the almost all decrease of the errors in the En4DVar analyses of the
472 four variables compared with 4DVar. However, from the right column of Figure 7 that displays
473 the ARMSE differences between the En4DVar and 4DEnVar analyses of the aforesaid four
474 variables, it can be observed that the En4DVar analyses of zonal wind and specific humidity are
475 not as close to the observations as the 4DEnVar ensemble mean analyses in general, except the
476 improvements in the zonal wind analyses at the upper troposphere and stratosphere in the Southern
477 Hemisphere and at high latitudes in the Northern Hemisphere. The En4DVar system improves the
478 temperature analyses near the poles and at the middle troposphere near the equator, but degrades
479 them above the upper troposphere and at the low troposphere near the equator, compared with
480 4DEnVar ensemble mean. Overall, among the aforesaid three kinds of analyses, the 4DEnVar
481 ensemble mean analyses of all the variables except geopotential height are the closest to the
482 observations, which may possibly benefit from the flow dependence of the ensemble covariance,
483 while the 4DVar analyses are the farthest from the observations, which may probably result from
484 the balance constraint included in the climatological covariance. The En4DVar analyses of all the
485 variables except geopotential height are between the above two kinds of analyses, which may be
486 attributed to the use of the hybrid covariance with both the flow dependence and balance constraint.
487 In addition, the En4DVar analyses of geopotential height are closer to the observations than the
488 4DEnVar ensemble mean analyses in most regions, indicating that the hybrid covariance allows
489 larger improvements in the analyses of geopotential height than in those of other variables (Figure
490 7c). Given the roles of the flow dependence and balance constraint in the analysis quality, their
491 impacts on the forecast skill are more concerned.

492 4.2.4 Forecast skill

493 A good fit of the analysis to the observations does not necessarily lead to an improved
494 forecast. Therefore, we measured the forecast error to verify the impact of improved or degraded
495 analysis on the forecast skill.

496 Figure 8 shows the comparison results of the ACC metric for the 500hPa potential height
497 forecast. The forecast initialized by En4DVar performs better than (comparable to) that initialized
498 by 4DVar after (within) two (four) days in the Northern (Southern) Extratropics (Figures 8a-8b),
499 and significantly better on all the lead days in the Tropics (Figure 8c). As for the 4DEnVar-
500 initialized forecast, its improvements against the 4DVar-initialized forecast are only sighted on
501 part of the lead days, including the lead days 5-8 (7-10) in the Northern (Southern) Extratropics
502 and the lead days 6-10 in the Tropics. Significant (slight) degradations are found on the lead days
503 2-4 (5-6) in the Southern Extratropics and on the lead days 1-3 (4-5) in the Tropics. The 9- and 10-
504 day forecasts in the Northern Extratropics also become worse. Moreover, Figure 8 also indicates
505 that the En4DVar-intialized forecast outperforms the 4DEnVar-intialized forecast in general
506 except the lead days 7, 8 and 10 in the Southern Extratropics. It is concluded that En4DVar is
507 overall better than both 4DVar and 4DEnVar in terms of the ACC-based forecast skill. This
508 conclusion is also true when the forecast skill is measured by ARMSE (Figure 9), except the
509 obvious degradations of En4DVar relative to 4DVar (4DEnVar) before the lead day 5 (on the lead
510 days 7-8) in the Southern (Northern) Extratropics. The most likely reason for the improvements in
511 the forecast of 500hPa geopotential height initialized by En4DVar against those initialized by both
512 4DVar and 4DEnVar is its alleviation of excessively or insufficiently incorporating observations
513 into the IC analysis through the hybrid covariance that combines the flow-dependent ensemble
514 covariance from 4DEnVar and the climatological covariance with balance constraint from 4DVar.

515 The En4DVar system can also improve the geopotential height forecasts on other vertical
516 levels. Large ARMSEs in the En4DVar-initialized forecast are mainly distributed in the
517 stratosphere and upper troposphere in the Northern and Southern Extratropics and increase with
518 the lead time (Figures 10a and 10d), which are consistent with the analysis error distribution.
519 These forecast errors are significantly reduced for most lead forecast days compared with those in
520 the 4DVar-initialized forecast, although they seriously increase in the upper stratosphere on all the
521 lead days and extend to the lower troposphere on the first 3 lead days in the Southern Extratropics

(Figures 10e). This is consistent with the increased analysis errors located on the upper layers (Figures 6b and 7b). The errors also decrease compared with those in the 4DEnVar-initialized forecast, except the slight (large) increase at the layers between 800hPa and 500hPa on the lead days 6-9 in the Northern Extratropics and in the middle and lower troposphere (stratosphere) on the lead days 7-10 in the Southern Extratropics (Figures 10c and 10f). Significant improvements can be found in the stratosphere in both the Northern and Southern Extratropics on the lead days 1-9 and 1-6, respectively, and in the upper and middle troposphere in the Southern Extratropics on the lead days 4-6. In the Tropics, the errors of the En4DVar-initialized forecast are much smaller than in the Northern and Southern Extratropics, and En4DVar slightly outperforms or performs comparably to both 4DVar and 4DEnVar (Figures 10g-10i).

The error structure of the En4DVar-initialized zonal wind forecast is similar to that of the geopotential height forecast, with large errors in the stratosphere and upper troposphere in the Northern and Southern Extratropics and increasing with the lead time (Figures 11a and 11d). The performance of En4DVar for the zonal wind forecast is similar to that for the geopotential height forecast, with significantly reduced errors on most lead forecast days compared with both 4DVar and 4DEnVar (the middle and right column of Figure 11). Compared with the forecast errors in the En4DVar-initialized geopotential height on the first lead day in the Southern Extratropics and Tropics that are not reduced against the 4DVar-initialized one, the forecast errors in zonal wind are decreased (Figures 11e and 11h). This is consistent with the reduced analysis errors of zonal wind by En4DVar relative to those by 4DVar in contrast with the undiminished analysis errors of geopotential height (Figure 6e).

The error structure of the En4DVar-initialized temperature forecast is slightly different from those of the geopotential height and zonal wind forecasts. There are large ARMSEs in both the stratosphere and the lower and middle troposphere in the Northern and Southern Extratropics, which increase with the lead time (Figure 12a and 12d). These errors are reduced in the Northern Extratropics on all lead days compared with those in the 4DVar- and 4DEnVar-initialized forecasts, except the last two lead days when the errors are slightly larger than those in the 4DVar-initialized forecast (Figures 12b-12c). In the Southern Extratropics, the En4DVar-initialized forecast is better (worse) than the 4DVar-initialized forecast mainly on the last 4 lead days below (above) 200hPa (Figure 12e), and is better than the deterministic forecast initialized from the 4DEnVar ensemble mean analysis at all vertical layers on all the lead days except the upper and low troposphere on

553 the last 3 lead days (Figure 12f). The significant improvements by En4DVar are located at the
554 vertical layers between 900hPa and 100hPa on the lead days 4-6 compared with 4DEnVar. In the
555 Tropics, the performance of En4DVar for the temperature forecast is similar to those for the
556 geopotential height and zonal wind forecasts, with much smaller AMSEs than in the Northern and
557 Southern Extratropics and slightly smaller AMSEs than in both the 4DVar- and 4DEnVar-
558 initialized forecasts overall (Figures 12g-12i).

559 The error structure of the En4DVar-initialized specific humidity forecast is slightly
560 different from those of other variables. There are large errors at the layers between 900hPa and
561 700hPa, which increase with lead time (Figures 13a, 13d and 13g). On one hand, the En4DVar-
562 initialized forecast is better than the 4DVar-initialized forecast, except the last 2 lead days at the
563 layers near surface in the Northern Extratropics. On the other hand, the En4DVar-initialized
564 forecast outperforms the 4DEnVar-initialized forecast except in the first few hours due to the larger
565 analysis errors of En4DVar than that of 4DEnVar. The significant improvements by En4DVar are
566 located in the lower troposphere compared with both 4DVar and 4DEnVar.

567 Overall, the change from 4DVar to En4DVar significantly (slightly) improves the
568 (medium-range) forecasts in the Tropics and Northern Extratropics (Southern Extratropics)
569 (Figure 4d). In particular, it leads to a sustained improvement in the middle troposphere in the
570 Tropics and the differences of ACC and ARMSE are statistically significant on almost all the lead
571 days. This is similar to the findings in Kleist et al. (2015a) on the change from 3DVar to 3D Hybrid.
572 Similarly, the change from 4DEnVar to En4DVar also significantly improves the forecasts at all
573 vertical levels in the Tropics on almost all the lead days (Figure 14). Significant improvements can
574 also be found in the short-term forecasts in the Northern Extratropics. In addition, the
575 improvements in the Southern Hemisphere are basically more significant than in the Northern
576 Hemisphere.

577 **5 Summary and discussion**

578 In this study, a new En4DVar hybrid DA system for the GRAPES-GFS model was
579 proposed and implemented using the extended control variable method based on the original
580 4DVar system. Most operational centers in the world have chosen to use an ensemble of 4DVars
581 or EnKF or something like this to estimate the ensemble covariance for En4DVar hybrid DA
582 systems. Bowler et al. (2017a, 2017b) also used the En-4DEnVar system instead of the ETKF

583 system to provide ensemble forecasts to the hybrid DA systems and obtained encouraging results,
584 but there are fewer related studies. Here, a 4DEnVar system was adopted to estimate the ensemble
585 covariance for the new En4DVar system. This system gave a dominant weight to the ensemble
586 covariance and a small weight to the balanced climatological covariance of 4DVar. The weight
587 coefficients of two covariances were determined based on the sensitivity experiments. In addition,
588 we applied an ensemble-sample-based subspace localization scheme, which is different from the
589 localization transformations in the climatological covariance, to reduce the spurious remote
590 correlations in the ensemble covariance. This scheme is not only easy to implement, but also allows
591 rapid generation of expanded samples thus improving the quality of the ensemble covariance and
592 benefiting the optimal analysis.

593 We conducted single-point observation experiments and OSSE-based one-month DA cycle
594 experiments and batch forecast experiments to systematically evaluate the analysis quality and
595 forecast skill of the new system using the 4DVar and 4DEnVar systems as references. Single-point
596 observation experiments show that introducing the ensemble covariance can help the En4DVar
597 system acquire the explicit flow-dependent characteristic. We also investigated the roles of the
598 ensemble and climatological covariances in the improvements of the En4DVar-initialized forecast.
599 The results show that a properly inflated hybrid covariance with a dominant weight to the ensemble
600 covariance leads to the optimal improvements in the forecast, indicating a more important role of
601 the flow dependence included in the ensemble covariance than that of the balance constraint
602 incorporated in the climatological covariance. According to the above investigation, the weight
603 coefficients of the climatological and ensemble covariances in the hybrid covariance of the
604 En4DVar system were finally determined, which were 0.25 and 0.8, respectively.

605 Using the optimal hybrid covariance, the performance of En4DVar on the analysis was
606 evaluated. The analysis error of the En4DVar system is smaller than that of the 4DVar system, but
607 basically larger than that of the 4DEnVar system. It is encouraging that the 500hPa geopotential
608 height forecast initialized from the En4DVar analysis overall outperforms those initialized from
609 both the 4DVar analysis and 4DEnVar ensemble mean analysis in terms of the ACC-based and
610 ARMSE-based forecast skills. The En4DVar system performs similarly for the geopotential height
611 forecast on other vertical layers. It also improves the forecasts of other variables in general relative
612 to both 4DVar and 4DEnVar. Overall, on one hand, the upgrade from 4DVar to En4DVar
613 significantly improves the forecasts in the Tropics and Northern Extratropics, and also slightly

ameliorates the medium-range forecasts in the Southern Extratropics. On the other hand, although the ICs from the En4DVar analysis are as not close to the observations as those from the 4DEnVar ensemble mean analysis, the En4DVar-initialized forecast generally performs better than the 4DEnVar-initialized forecast, significantly in the middle and upper troposphere in the Tropics and in the short-term forecasts in the Northern and Southern Extratropics. It is suggested that a much more accurate IC may not lead to a farther improvement of forecast unless the generation of this IC includes some balance constraint.

Future improvements of the En4DVar system will depend on the increase of the quality of the ensemble covariance and the use of more observations. To increase the quality of the ensemble covariance, the 4DEnVar system needs further development and updating. For example, the observational ensemble samples can be generated using nonlinear forecast models instead of TLMs, which can both improve the qualities of the ensemble samples and save time in model integration. The inclusion of the balance constraint into the 4DEnVar system may also ameliorate the ensemble covariance, which may be implemented through an advanced localization technique upon the separated unbalanced variables. In addition, the assimilation of satellite observations may also effectively improve the ensemble covariance and the En4DVar analysis.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (2018YFC1506703). 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., 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.

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.

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

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

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

676 Gauthier, P., & Thépaut, J. N. (2001). Impact of the digital filter as a weak constraint in the
677 preoperational 4dvar assimilation system of Météo-France. *Monthly Weather Review*, 129(8),
678 2089-2102.

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

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

684 Hamill, T. M., Whitaker, J. S., & Snyder, C. (2001). Distance-dependent filtering of background
685 error covariance estimates in an ensemble Kalman filter. *Monthly Weather Review*, 129(11),
686 2776-2790.

687 He, Y., Wang, B., Liu, M., Liu, L., Yu, Y., Liu, J., Li, R., Zhang, C., Xu, S., Huang, W., Liu, Q.,
688 Wang, Y., & Li, F. (2017). Reduction of initial shock in decadal predictions using a new
689 initialization strategy. *Geophysical Research Letters*, 44(16), 8538-8547, DOI:
690 10.1002/2017GL074028.

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

695 He, Y., Wang, B., Huang, W., Xu, S., Wang, Y., Liu, L., Li, L., Liu, J., Yu, Y., Lin, Y., Huang,
696 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.

Kadowaki, T. (2005). A 4-dimensional variational assimilation system for the JMA Global Spectrum Model. *CAS/JSC Research activities in atmospheric and oceanic modelling*, 34, 117-118.

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.

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.

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. (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.

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

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

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

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

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

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

741 Shi, P., Lu, H., Leung, L., He, Y., Wang, B., Yang, K., et al. (2021a). Significant land contributions
742 to interannual predictability of East Asian summer monsoon rainfall. *Earth's Future*, 9, DOI:
743 10.1029/E2020EF001762.

744 Shi P F, Wang B, He Y J, Lu H, Yang K, et al. (2021b). Contributions of weakly coupled data
745 assimilation-based land initialization to interannual predictability of summer climate over
746 Europe. *Journal of Climate*, 10.1175/JCLI-D-20-0506.1.

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

750 Tian, X., Xie, Z., & Dai, A. (2008). An ensemble-based explicit four-dimensional variational
751 assimilation method. *Journal of Geophysical Research: Atmospheres*, 113, D21124. DOI:
752 10.1029/2008JD010358.

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

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

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

762 Wang, X., Snyder, C., & Hamill, T. M. (2007). On the theoretical equivalence of differently
763 proposed ensemble–3dvar hybrid analysis schemes. *Monthly Weather Review*, 135(1), 222-227.

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

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

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

772 Whitaker, J. S., Compo, G. P., & Thépaut, J. N. (2009). A comparison of variational and ensemble-
773 based data assimilation systems for reanalysis of sparse observations. *Monthly Weather Review*,
774 137(6), 1991-1999.

775 Wlasak, M. A., & Cullen, M. J. P. (2014) Modelling static 3-D spatial background error
776 covariances – the effect of vertical and horizontal transform order. *Advances in Science &*
777 *Research*, 11, 63-67.

778 Zhang, L., Liu, Y. Z., Liu, Y., Gong, J., Lu, H., Jin, Z., Tian, W., Liu, G., Zhou, B., & Zhao, B.
779 (2019). The operational global four-dimensional variational data assimilation system at the
780 China Meteorological Administration. *Quarterly Journal of the Royal Meteorological Society*,
781 145, 1882–1896.

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

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

786 Zhu, S., Wang, B., Zhang, L., Liu, J., Liu, Y., Gong, J., Xu, S., Wang, Y., Huang, W., Li, L., He,
787 Y. & Wu, X. (2022). A Four-Dimensional Ensemble-Variational (4DEnVar) Data Assimilation
788 System for Global NWPs: System Description and Primary Tests. *Journal of Advances in*
789 *Modeling Earth Systems*, revised.

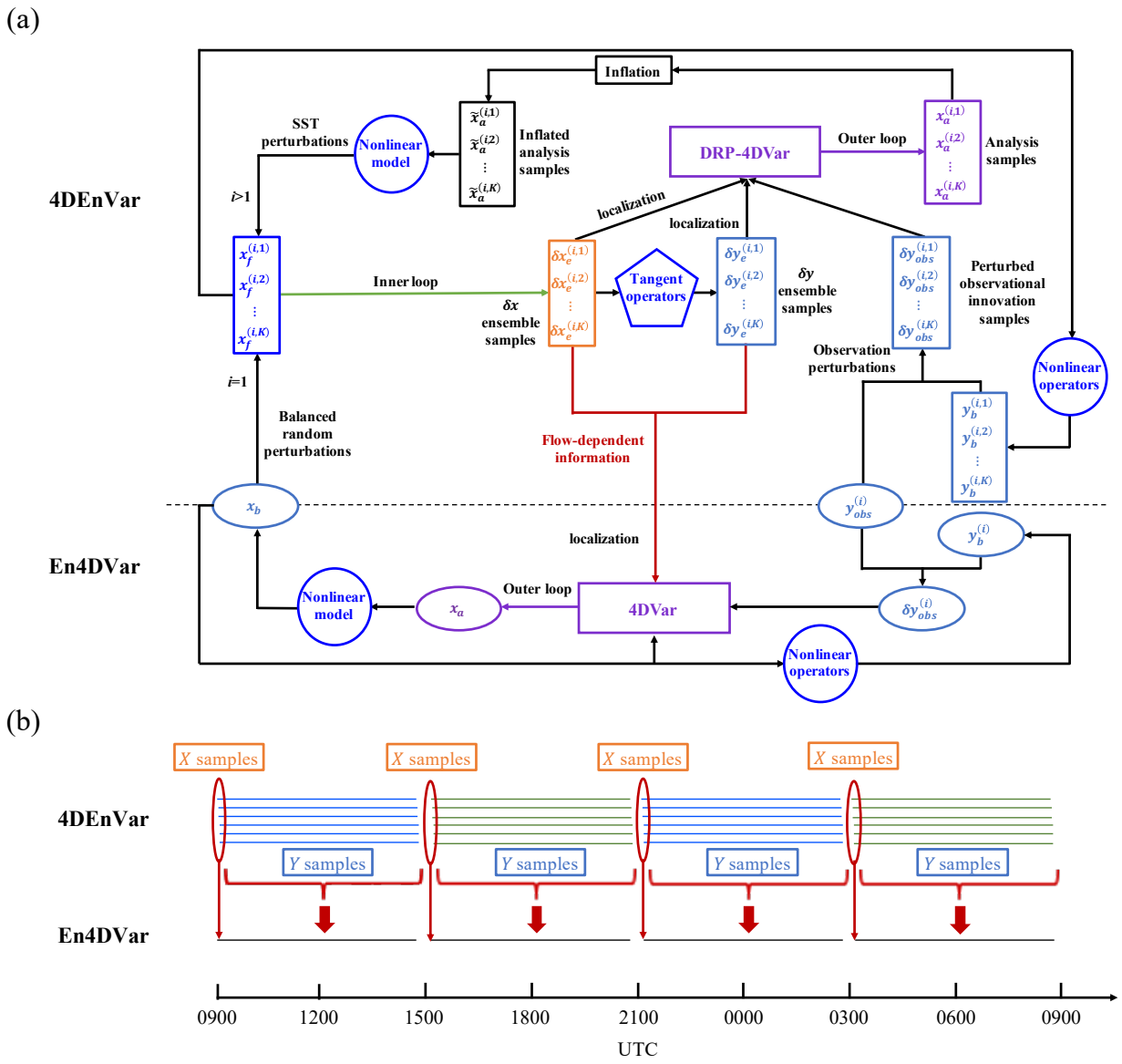

Figure 1. Schematic of the En4DVar system developed in this study. The key components and techniques of the En4DVar system are shown in (a), and the exchange of information between the 4DEnVar system and the En4DVar system on the timeline during the assimilation cycle is shown in (b).

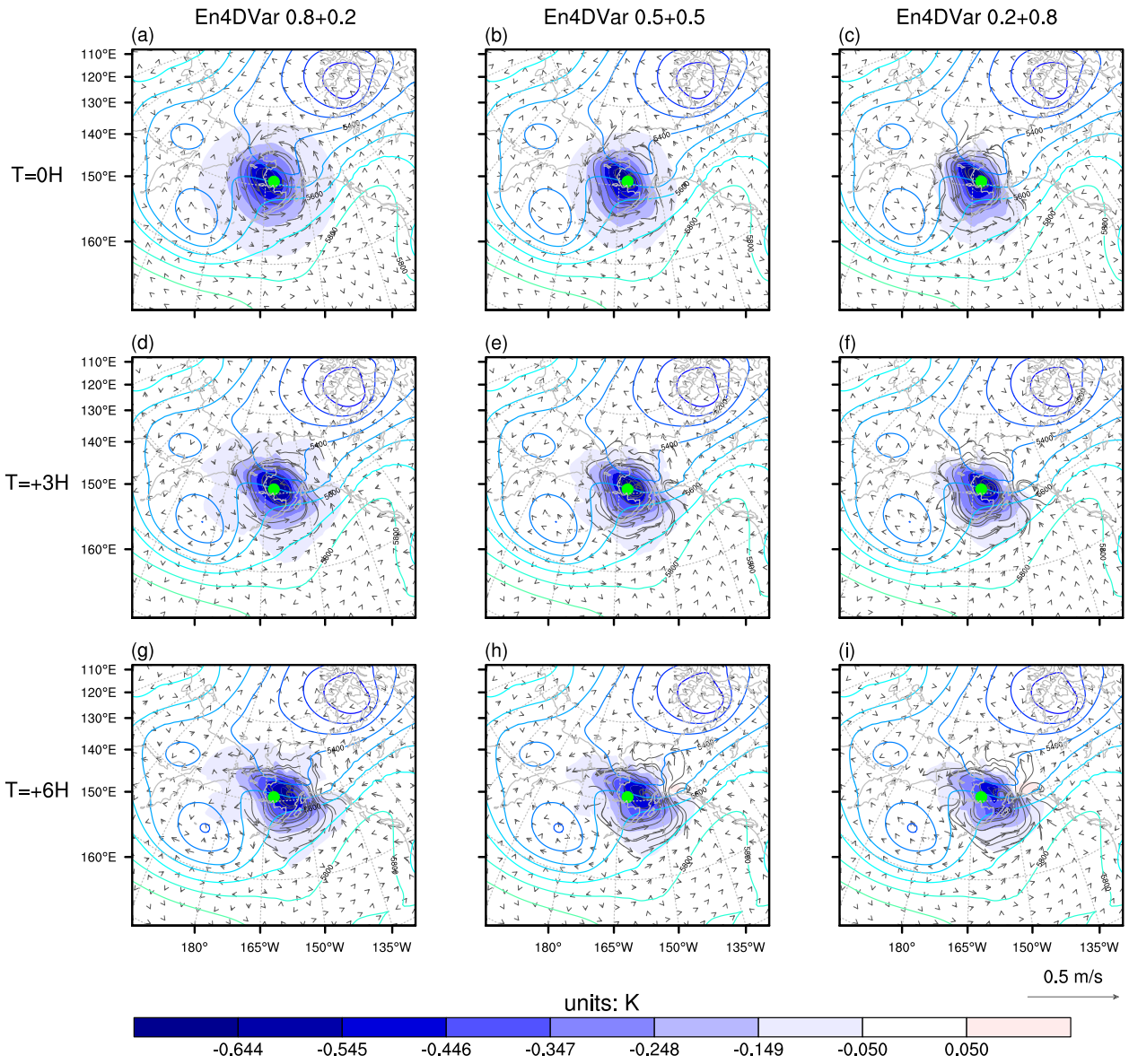

Figure 2. The temperature (shading; units: K) and vector wind (vector; units: m/s) analysis increments valid at the (top) beginning, (middle row) middle, and (bottom) end of the assimilation window resulting from assimilating the single-point temperature observation valid at the middle of the assimilation window for the En4DVar system adopting the linear combination coefficients of (left) $(\gamma_c = 0.8, \gamma_e = 0.2)$, (middle column) $(\gamma_c = 0.5, \gamma_e = 0.5)$ and (right) $(\gamma_c = 0.2, \gamma_e = 0.8)$ on the model level closest to the assimilated observation (marked with a green dot). The solid contour is the 500hPa background field geopotential height (units: gpm) valid at the (top) beginning, (middle row) middle, and (bottom) end of the assimilation window.

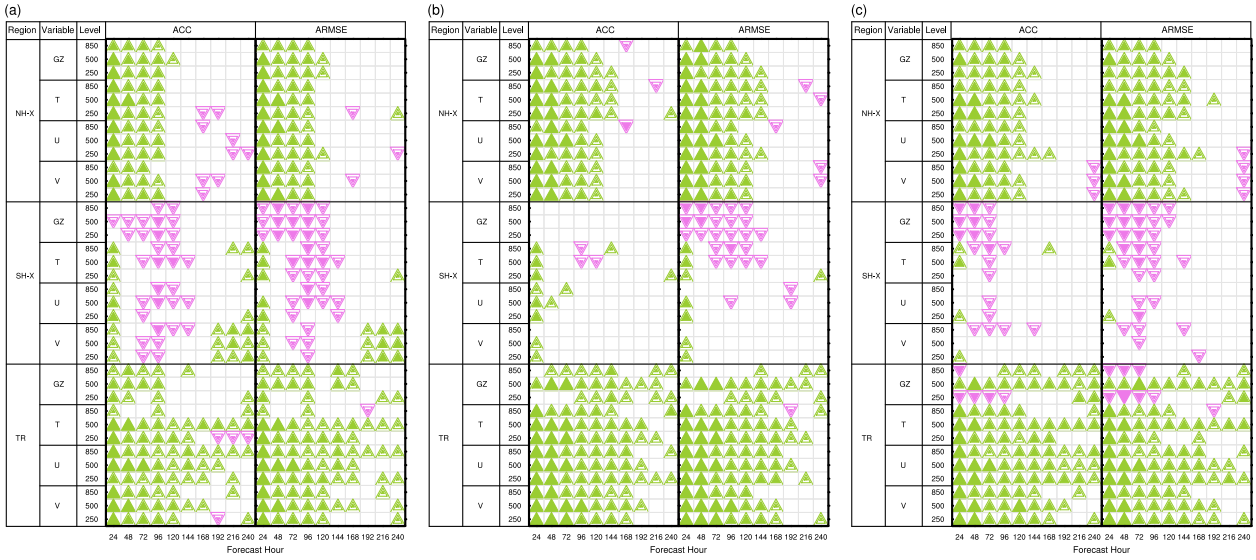

Figure 3 The scorecard of the ACC and ARMSE calculated for the forecasts initialized by the En4DVar system adopting the linear combination coefficients of (a) $(\gamma_c = 0.8, \gamma_e = 0.2)$, (b) $(\gamma_c = 0.5, \gamma_e = 0.5)$ and (c) $(\gamma_c = 0.2, \gamma_e = 0.8)$ against those initialized by the 4DVar system with identical settings, respectively. If the former forecast has a significantly higher / smaller (lower / larger) ACC / ARMSE than the latter, a green upward-pointing (red downward-pointing) triangle is marked. The large, medium and small filled triangles respectively indicate that the differences of ACC or ARMSE are greater than 3 times, between 1 times and 3 times, and between 0.5 times and 1 times the t value of the 95% confidence level, i.e., fairly significant, significant, and insignificant. No triangles are shown when the differences are less than 0.5 times the t value of the 95% confidence level, indicating equivalent.

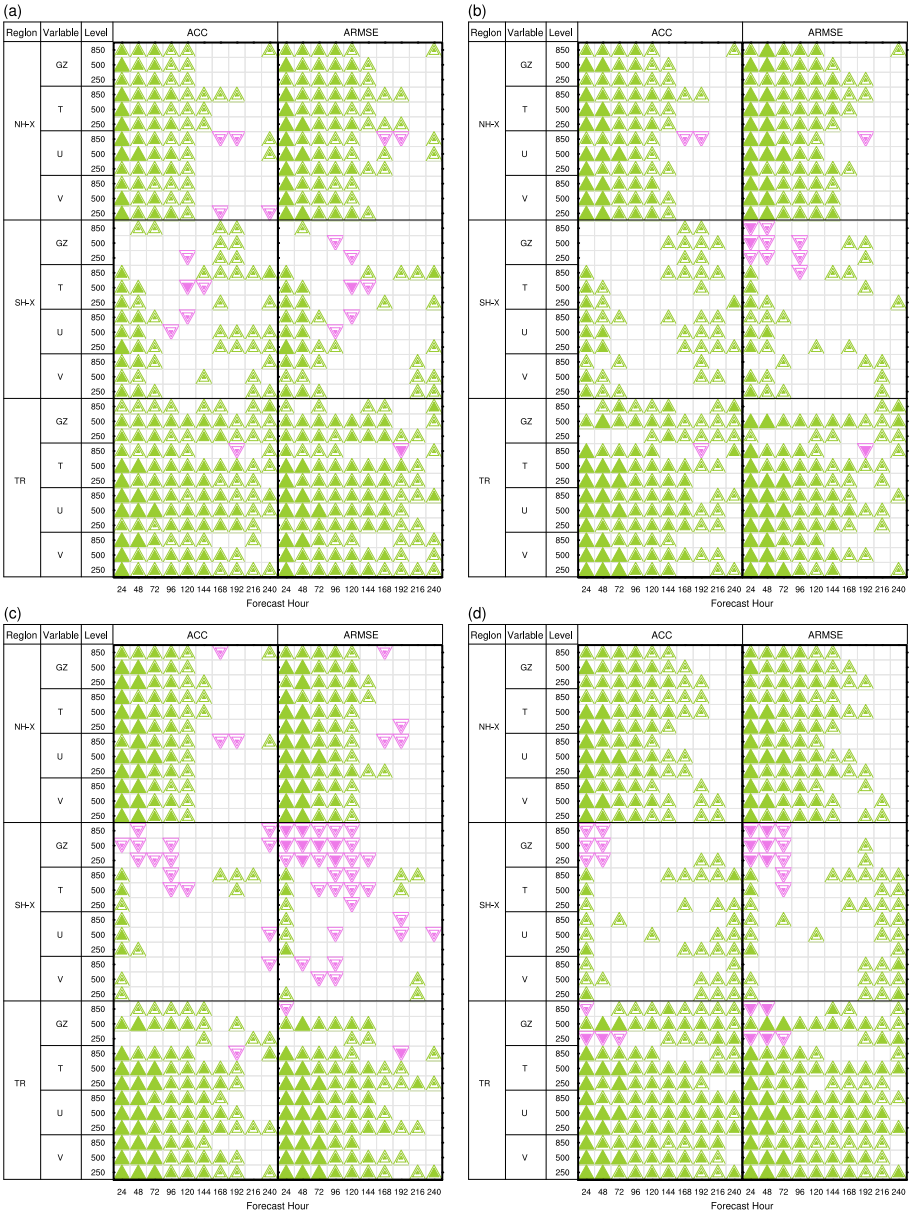

Figure 4. Same as Figure 3, except for the linear combinations of (a) ($\gamma_c = 0.8, \gamma_e = 0.25$), (b) ($\gamma_c = 0.6, \gamma_e = 0.5$), (c) ($\gamma_c = 0.5, \gamma_e = 0.6$) and (d) ($\gamma_c = 0.25, \gamma_e = 0.8$).

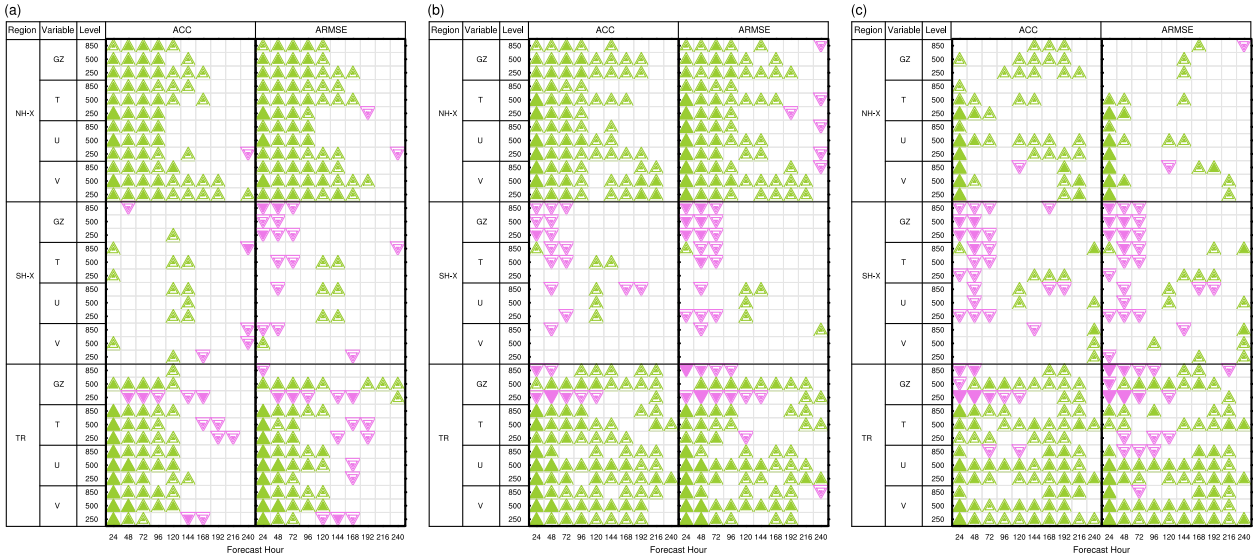

Figure 5. Same as Figure 3, except for the forecasts initialized by the En4DVar system adopting the linear combinations of (a) $(\gamma_c = 0.6, \gamma_e = 0.5)$ and (b) $(\gamma_c = 0.25, \gamma_e = 0.8)$ against those using the linear combination of $(\gamma_c = 0.8, \gamma_e = 0.25)$, and (c) $(\gamma_c = 0.25, \gamma_e = 0.8)$ against those using the linear combination of $(\gamma_c = 0.6, \gamma_e = 0.5)$.

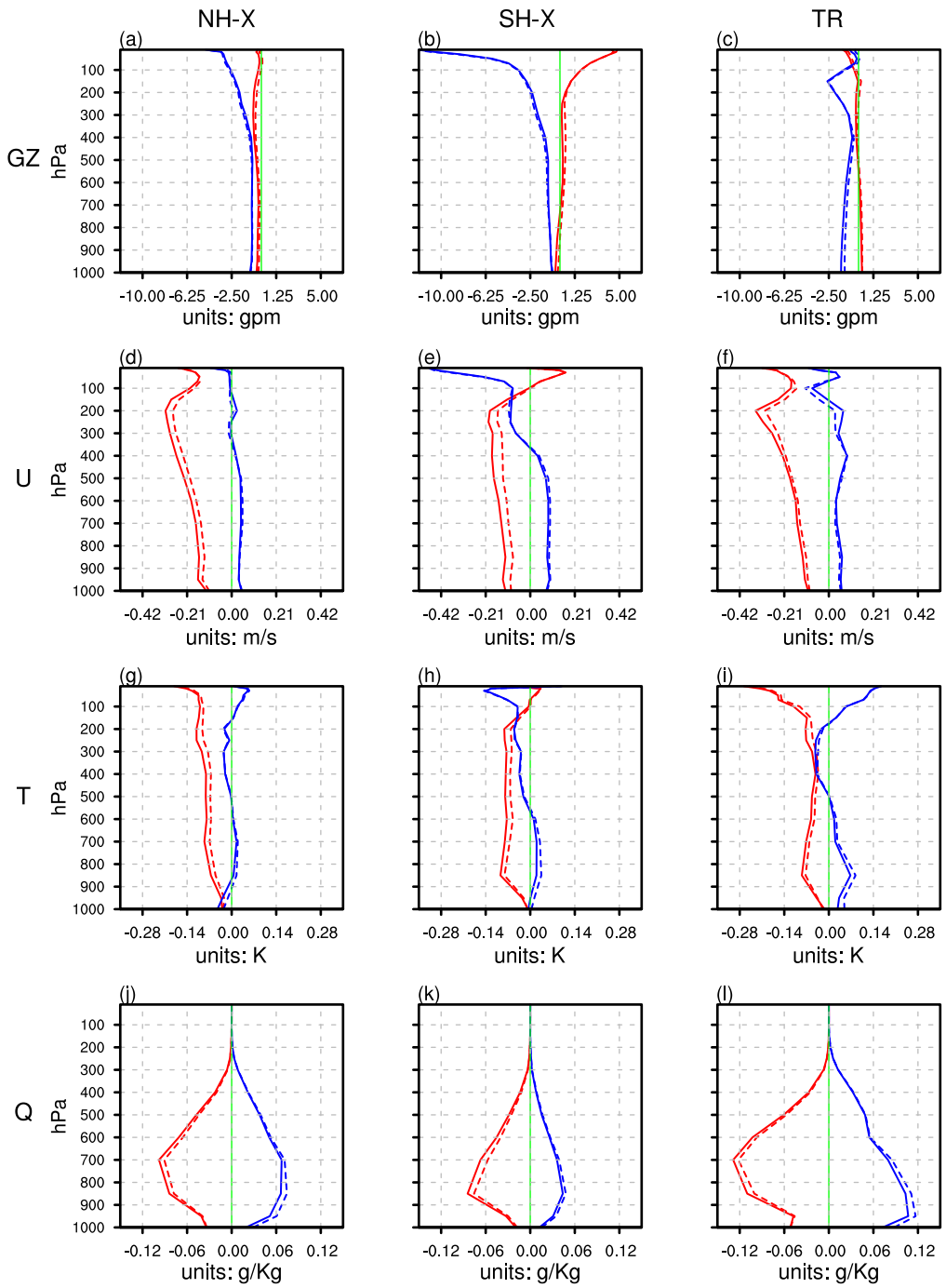

Figure 6. Vertical profiles of the ARMSE (verified against the "truth" state) differences of the background (dashed line) and analysis (solid line) fields of the geopotential height (row 1; units: gpm), zonal wind (row 2; units: m/s), temperature (row 3; units: K) and specific humidity (row 4; units: g/Kg) in the Northern Extratropics (20°N~90°N, 180°W~180°E; left), Southern Extratropics (20°S~90°S, 180°W~180°E; middle column) and Tropics (20°S~20°N, 180°W~180°E; right). The red (blue) lines show the ARMSE differences between the En4DVar and 4DVar (4DEnVar ensemble mean) background and analyses.

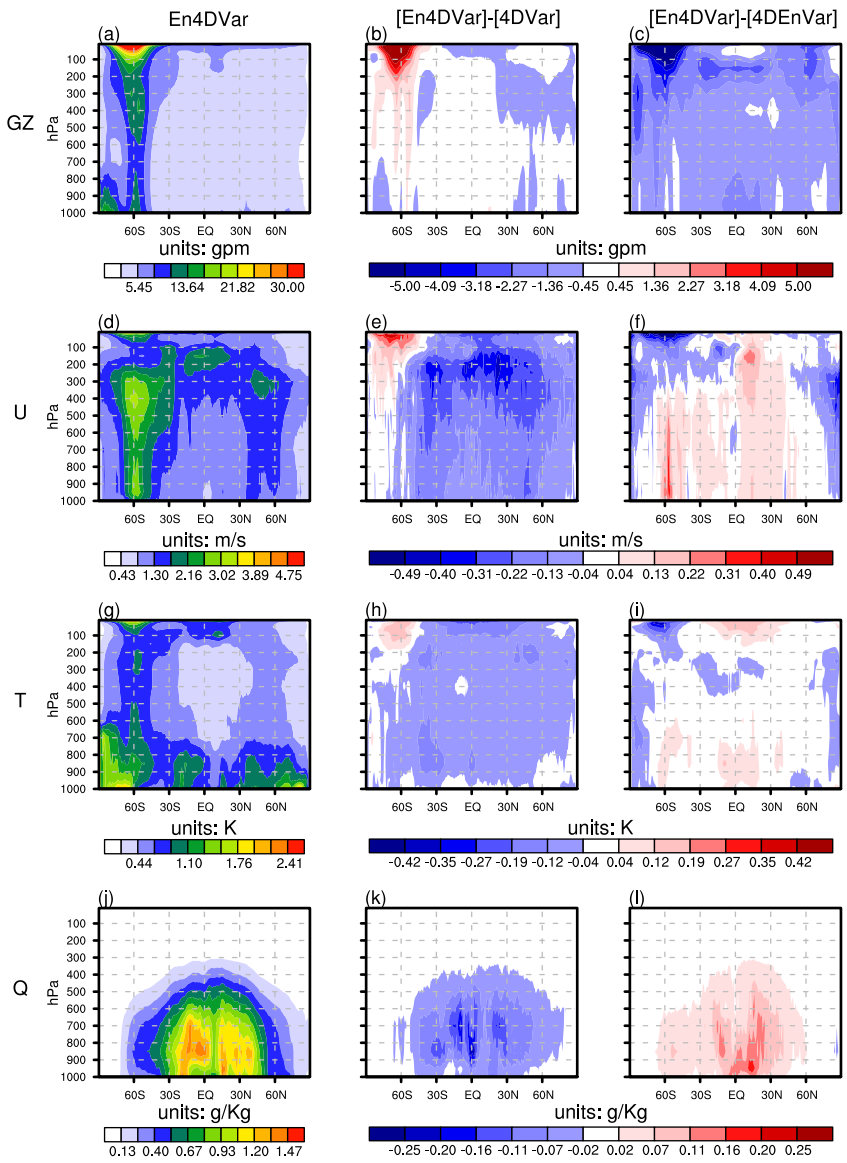

Figure 7. The pressure versus latitude plots of the ARMSEs (verified relative to the "truth" state) of the geopotential height (row 1; units: gpm), zonal wind (row 2; units: m/s), temperature (row 3; units: K) and specific humidity (row 4; units: g/Kg) analyses from the En4DVar (left), and the ARMSE differences between En4DVar and 4DVar (middle column) and between En4DVar and 4DEnVar (left), respectively.

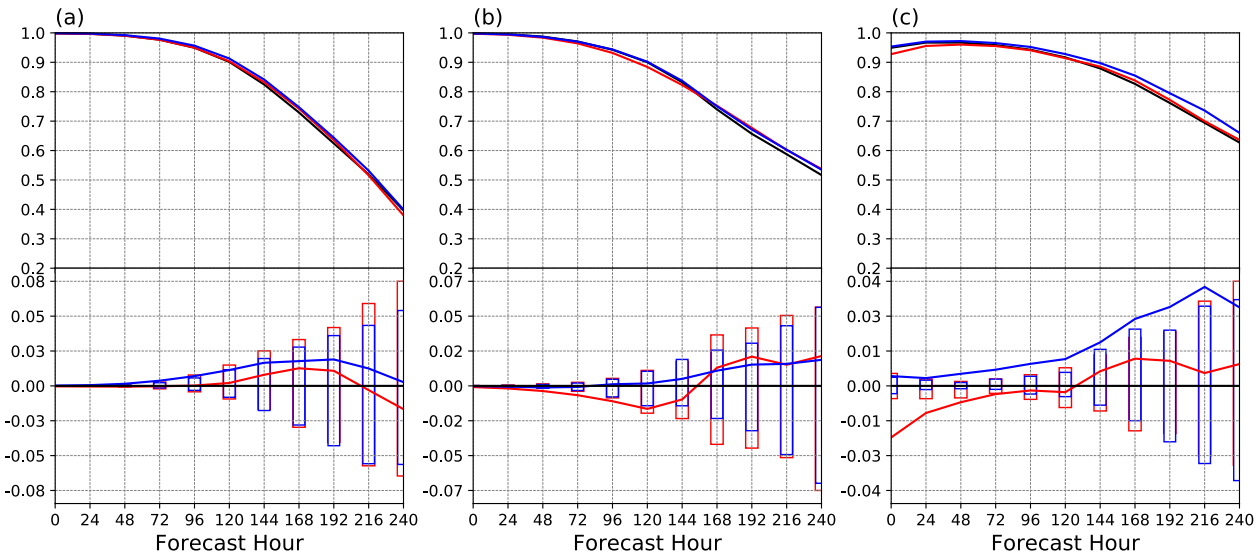

Figure 8. The anomaly correlation coefficients (ACCs) of the forecasts of the 500hPa geopotential height respectively initialized from the 4DVar (black line), 4DEnVar ensemble mean (red line) and En4DVar analyses (blue line) against the "truth" state in the (a) Northern Extratropics (20°N~90°N; 180°W~180°E), (b) Southern Extratropics (20°S~90°S; 180°W~180°E) and (c) Tropics (20°S~20°N, 180°W~180°E). The corresponding ACC differences between the forecasts initialized from the 4DEnVar ensemble mean and 4DVar analyses (red line) and between those initialized from the En4DVar and 4DVar analyses (blue line), with the 95% confidence threshold, are also plotted in the bottom.

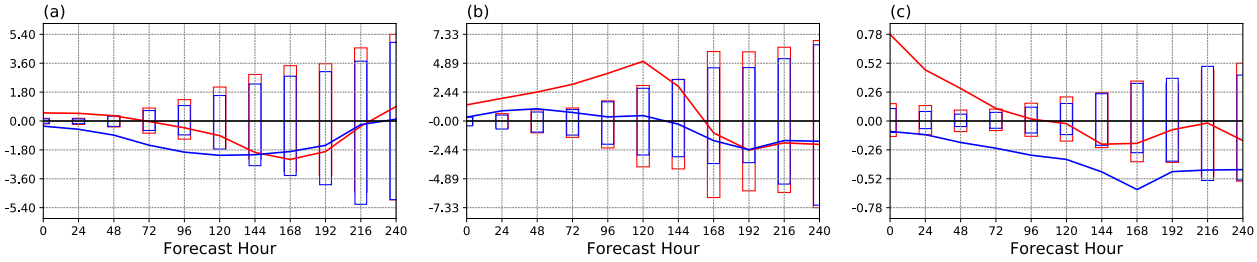

Figure 9. The ARMSE differences between the forecasts of the 500hPa geopotential height initialized from the 4DEnVar ensemble mean and 4DVar analyses (red line) and between those initialized from the En4DVar and 4DVar analyses (blue line) against the "truth" state in the (a) Northern Extratropics (20°N~90°N; 180°W~180°E), (b) Southern Extratropics (20°S~90°S; 180°W~180°E) and (c) Tropics (20°S~20°N, 180°W~180°E). The bar charts shows the 95% confidence threshold.

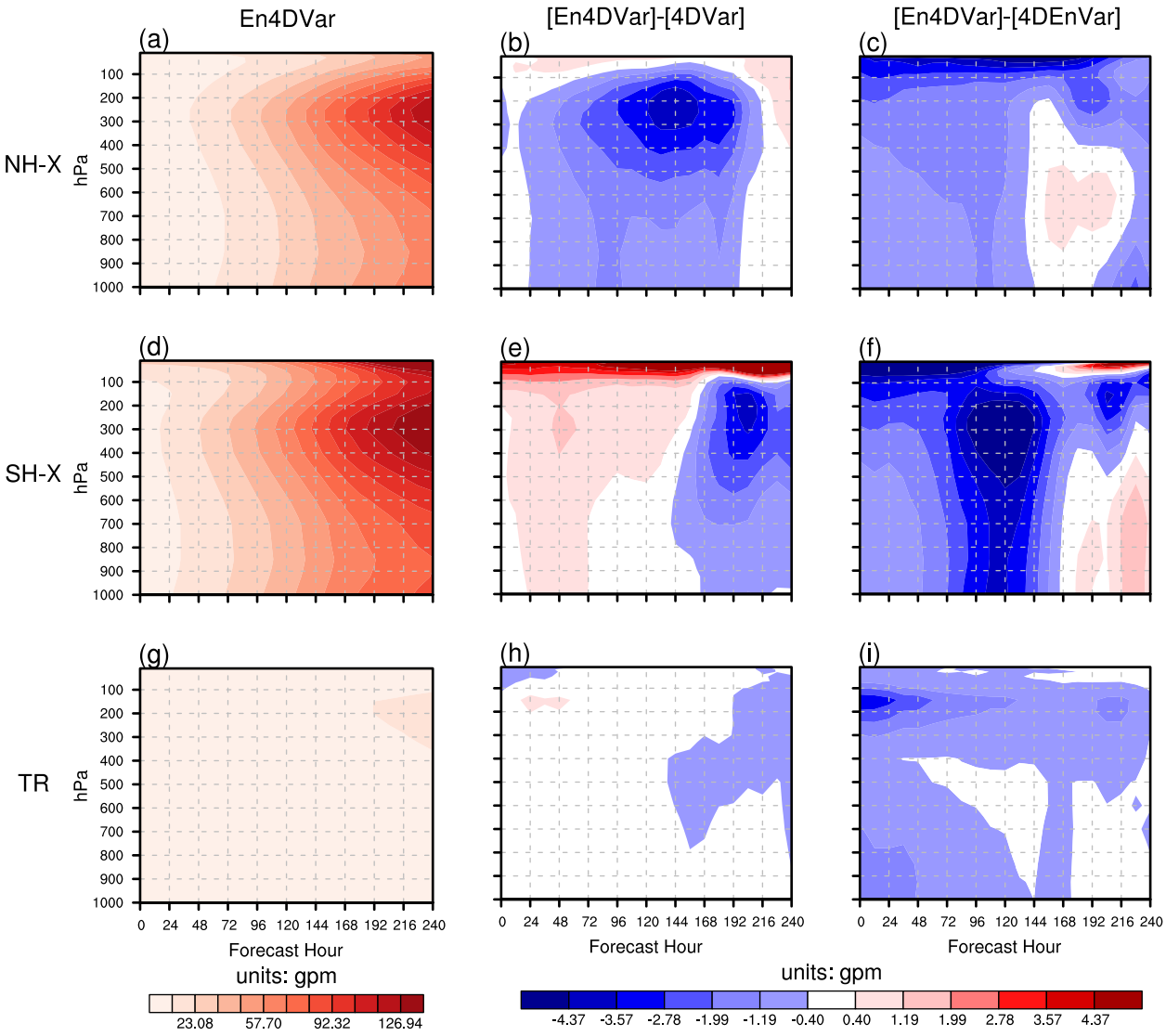

Figure 10. The ARMSEs of the geopotential height forecasts (units: gpm) initiated from the 1200 UTC analyses of the En4DVar experiment as a function of lead time (left column) in the (a) Northern Extratropics (20°N~90°N; 180°W~180°E), (d) Southern Extratropics (20°S~90°S; 180°W~180°E) and (g) Tropics (20°S~20°N, 180°W~180°E). The differences of ARMSE between the En4DVar- and 4DVar-(4DEnVar-)initialized (deterministic) forecasts are plotted in the middle (right) column.

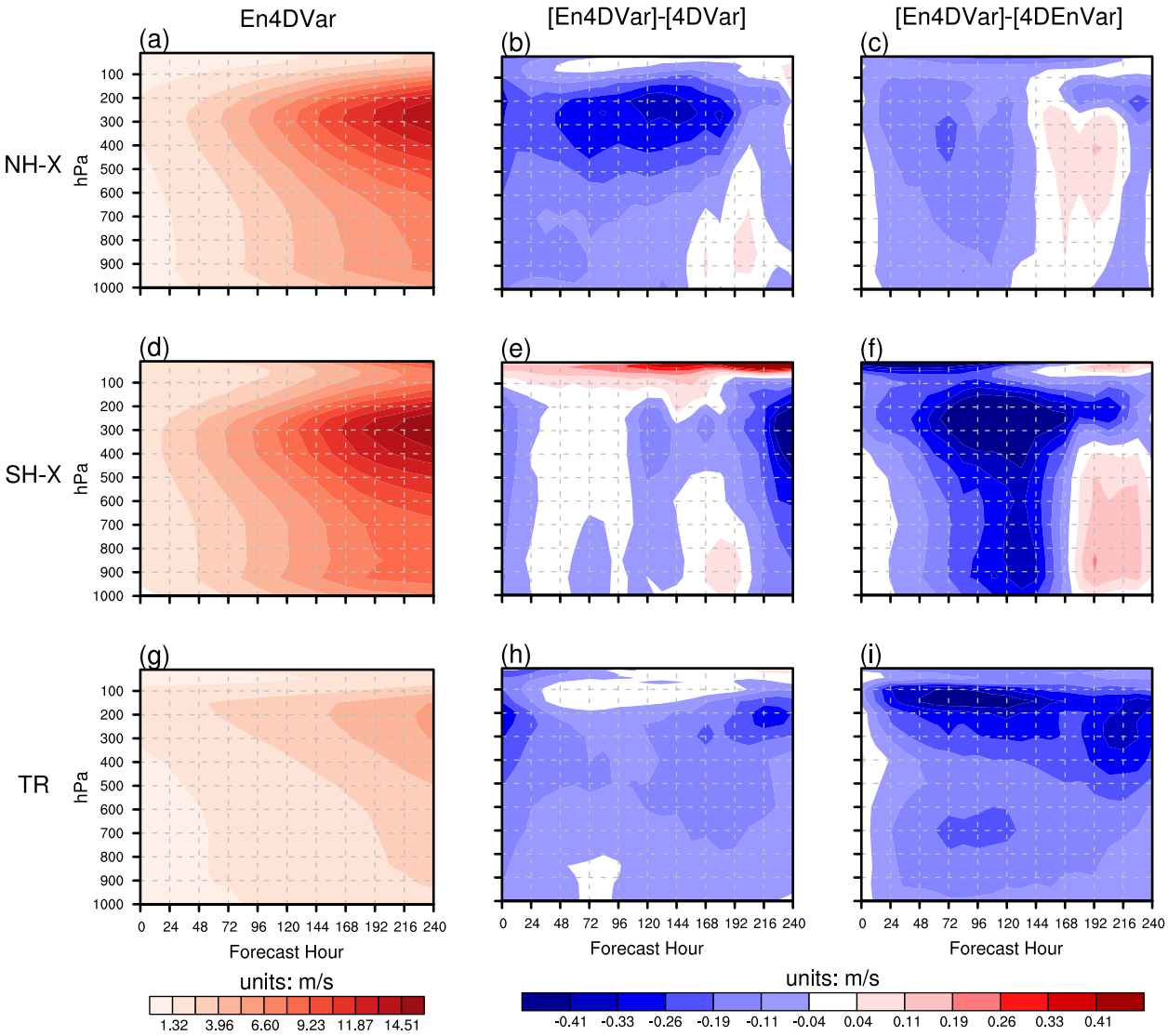

Figure 11. Sam as Figure 10, except for the zonal wind forecasts (units: m/s).

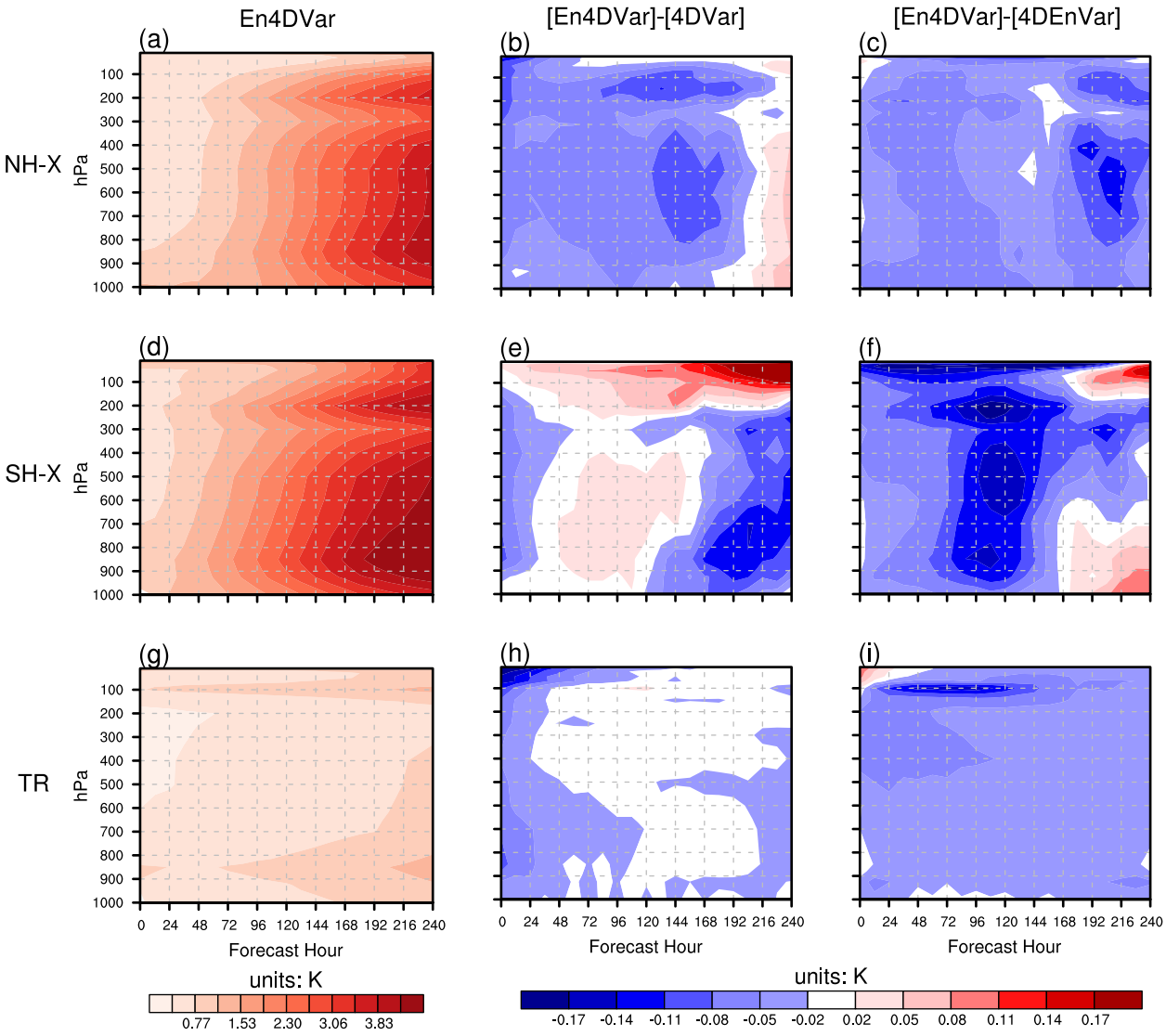

Figure 12. Sam as Figure 10, except for the temperature forecasts (units: K).

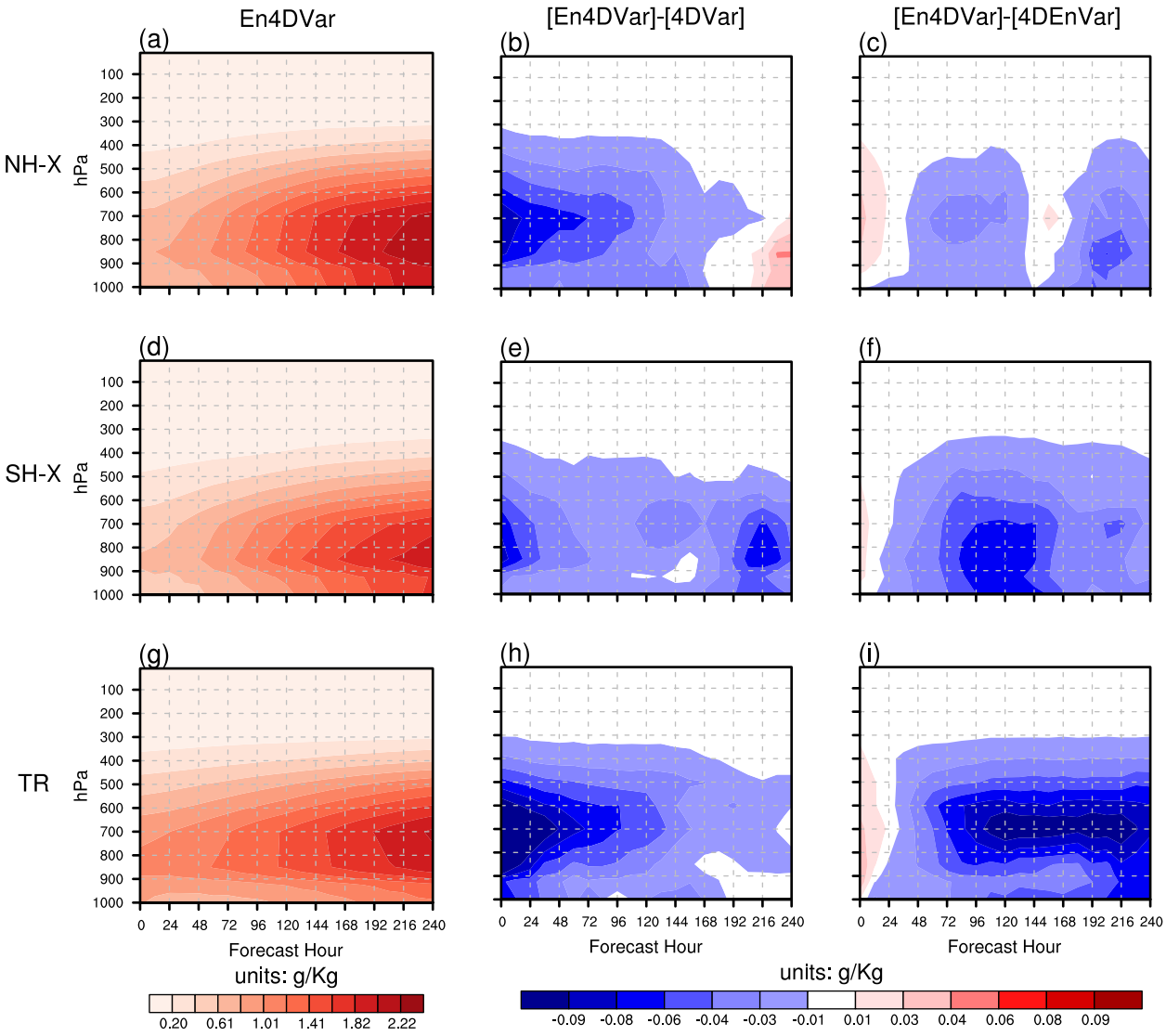

Figure 13. Sam as Figure 10, except for the specific humidity forecasts (units: g/kg).

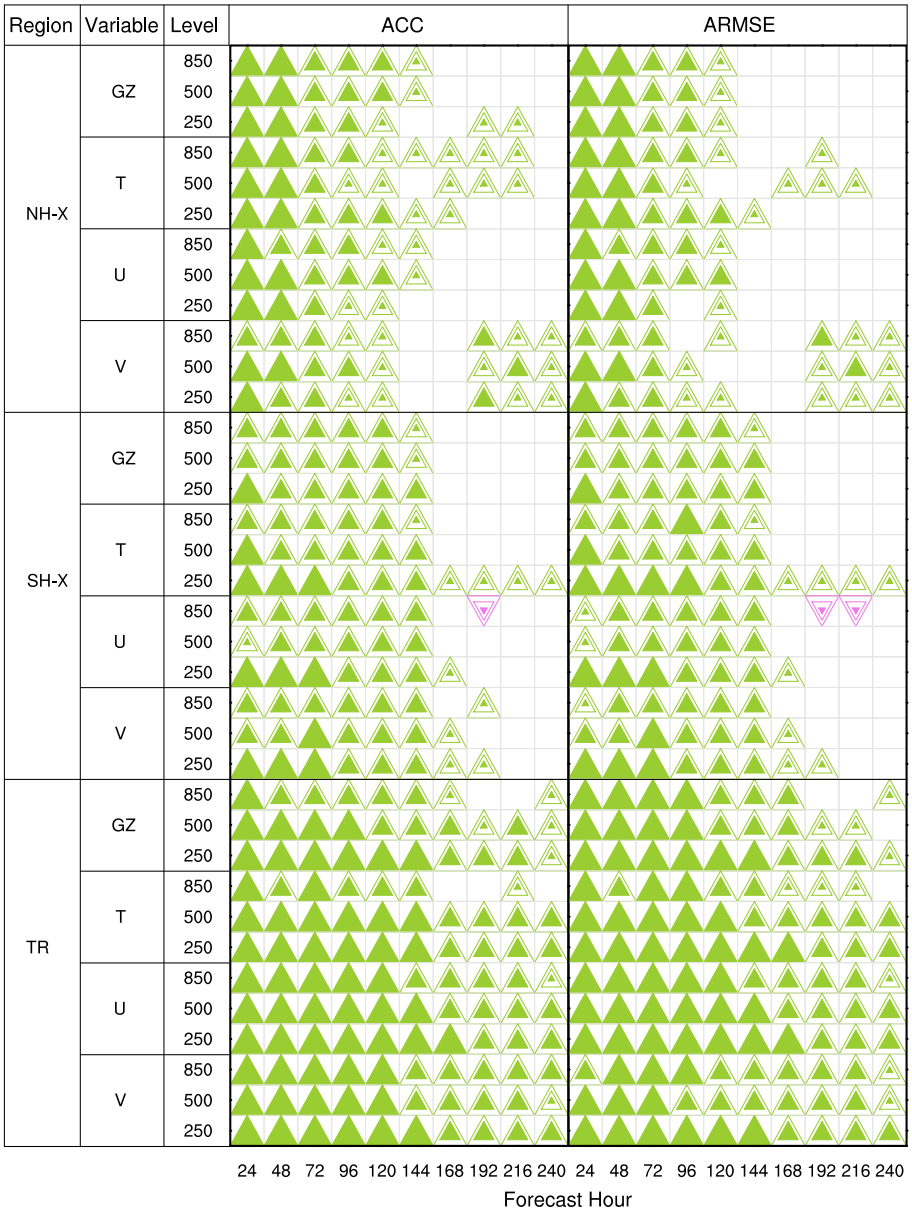

Figure 14. Same as Figure 3, except for the forecasts initialized by the En4DVar system adopting the linear combination coefficients of $(\gamma_c = 0.25, \gamma_e = 0.8)$ against the 4DEnVar-initialized deterministic forecasts.