ADVANCES IN ATMOSPHERIC SCIENCES, VOL. 35, MAY 2018, 518–530

• **Original Paper** •

A Prototype Regional GSI-based EnKF-Variational Hybrid Data Assimilation System for the Rapid Refresh Forecasting System: Dual-Resolution Implementation and Testing Results

Yujie PAN[1,3], Ming XUE[*1,2,3], Kefeng ZHU[2,3], and Mingjun WANG[2,3]

[1]*Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters/Key Laboratory of Meteorological Disaster, Ministry of Education/Joint International Research Laboratory of Climate and Environment Change, Nanjing University of Information Science and Technology, Nanjing* 210044*, China*

[2]*Key Laboratory of Mesoscale Severe Weather/Ministry of Education and School of Atmospheric Sciences, Nanjing University, Nanjing* 210093*, China*

[3]*Center for Analysis and Prediction of Storms, University of Oklahoma, Norman, Oklahoma* 73072*, USA*

(Received 26 April 2017; revised 8 October 2017; accepted 27 October 2017)

ABSTRACT

A dual-resolution (DR) version of a regional ensemble Kalman filter (EnKF)-3D ensemble variational (3DEnVar) coupled hybrid data assimilation system is implemented as a prototype for the operational Rapid Refresh forecasting system. The DR 3DEnVar system combines a high-resolution (HR) deterministic background forecast with lower-resolution (LR) EnKF ensemble perturbations used for flow-dependent background error covariance to produce a HR analysis. The computational cost is substantially reduced by running the ensemble forecasts and EnKF analyses at LR. The DR 3DEnVar system is tested with 3-h cycles over a 9-day period using a 40/∼13-km grid spacing combination. The HR forecasts from the DR hybrid analyses are compared with forecasts launched from HR Gridpoint Statistical Interpolation (GSI) 3D variational (3DVar) analyses, and single LR hybrid analyses interpolated to the HR grid. With the DR 3DEnVar system, a 90% weight for the ensemble covariance yields the lowest forecast errors and the DR hybrid system clearly outperforms the HR GSI 3DVar. Humidity and wind forecasts are also better than those launched from interpolated LR hybrid analyses, but the temperature forecasts are slightly worse. The humidity forecasts are improved most. For precipitation forecasts, the DR 3DEnVar always outperforms HR GSI 3DVar. It also outperforms the LR 3DEnVar, except for the initial forecast period and lower thresholds.

Key words: dual-resolution 3D ensemble variational data assimilation system, Rapid Refresh forecasting system

Citation: Pan, Y. J., M. Xue, K. F. Zhu, and M. J. Wang, 2018: A prototype regional GSI-based EnKF-variational hybrid data assimilation system for the Rapid Refresh forecasting system: Dual-resolution implementation and testing results. *Adv. Atmos. Sci.*, **35**(5), 518–530, https://doi.org/10.1007/s00376-017-7108-0.

1. Introduction

Studies have shown that background error covariance (BEC) plays an important role in atmospheric data assimilation (DA), and especially for the mesoscale and convective scale where weather systems are more transient and intermittent (Zhang et al., 2006; Meng and Zhang, 2007; Liu and Xue, 2008; Ancell et al., 2014). In typical 3D variational (3DVar) systems, static BEC is usually used in the cost function, which does not contain flow-dependent spatial covariance (Lorenc, 1986); any cross-covariances usually only reflect static, large-scale balances (Barker et al., 2004; Descombes et al., 2015). For smaller scale flows, flow-dependent spatial covariance and cross-covariances tend to have increasing importance.

The ensemble Kalman filter (EnKF), originally proposed by Evensen (1994), uses flow-dependent BEC derived from an ensemble of forecasts. EnKF has become increasingly popular in recent years and has been applied to operational (e.g., Buehner et al., 2010a, b) or prototype numerical weather prediction (NWP) systems (e.g., Hamill et al., 2011; Wang et al., 2013). However, due to the relatively small ensemble size dictated by available computing resources, sampling error in the ensemble-derived BEC is usually significant (Hamill et al., 2001; Miyoshi et al., 2014; Anderson, 2016). Covariance localization is typically used to help alleviate the problem, which can affect flow balances (Hamill et al., 2001; Greybush et al., 2010). As an alternative way of alleviating the problem, Hamill and Snyder (2000) proposed a hybrid algorithm, in which a weighted average of the static and flow-dependent covariances is used within a 3DVar

* Corresponding author: Ming XUE
Email: mxue@ou.edu

framework; they found that the hybrid system outperforms EnKF when the ensemble size is small or when the model error is large. An alternative computationally more efficient implementation of the hybrid idea, called the extended control variable (ECV) method, was later proposed by Lorenc (2003). The ECV method extends the original control vector in the 3DVar cost function by adding an additional term—the extended control variable term preconditioned upon the square root of the ensemble covariance. Wang et al. (2007) showed that the method of Hamill and Snyder (2000) using a simple weighted average and the ECV method are mathematically equivalent. With the ECV approach, a hybrid 3D ensemble variational (3DEnVar) algorithm is relatively easy to implement within an existing 3DVar framework. Recent studies have demonstrated that forecasts initialized from analyses of a hybrid algorithm are usually better than those from traditional 3DVar, and are comparable to or better than those from pure EnKF (Wang et al., 2009; Li et al., 2012; Zhang and Zhang, 2012; Wang et al., 2013; Zhang et al., 2013; Pan et al., 2014; Schwartz and Liu, 2014). This approach has been used in several global NWP (Kuhl et al., 2013; Wang et al., 2013; Kleist and Ide, 2015) and regional mesoscale systems (e.g., Wang et al., 2008a, b; Schwartz and Liu, 2014; Schwartz et al., 2015; Wu et al., 2017).

Most of the hybrid DA studies mentioned above employed single-resolution (SR) configurations, in which the deterministic EnVar analysis and the ensemble analyses and forecasts are performed at the same resolution. For a high-resolution (HR) DA system, SR configuration can be computationally expensive, especially for operational purposes. For these reasons, dual-resolution (DR) hybrid schemes have been proposed, which combine HR background forecasts and hybrid analysis with BECs derived from low-resolution (LR) ensemble forecasts that are typically provided by a cycled EnKF system (Buehner et al., 2010b; Hamill et al., 2011; Clayton et al., 2013; Kleist and Ide, 2015).

For regional models, one DR hybrid example was recently reported in Schwartz et al. (2015), based on Weather Research and Forecasting (WRF) DA systems. Schwartz et al. (2015) compared the WRF SR and DR hybrid analyses and forecasts over a ~3.5-week period. They found that the 45/15-km DR analyses were completed around three times faster and required about one quarter of the disk space of the 15-km SR analyses. Moreover, forecasts launched from the 15-km SR hybrid analyses had no significant differences from those launched from the 45/15-km DR hybrid analyses, although forecasts from the SR system captured more fine-scale features. The DR hybrid system consumes much less computational resource than the 15-km SR hybrid system.

The Rapid Refresh (RAP) forecasting system is an hourly-updated operational DA/prediction system of the U.S. National Weather Service using a 13-km horizontal grid spacing, and it replaced the operational Rapid Update Cycle (Benjamin et al., 2004) system as a regional operational analysis and forecast system in 2012. RAP uses the WRF-ARW model (Skamarock and Klemp, 2008) for forecasting and the Gridpoint Statistical Interpolation (GSI) analysis system (Wu et al., 2002; Kleist et al., 2009) for DA. GSI 3DVar was used until February 2014, when it was replaced by a hybrid 3DEnVar using flow-dependent covariances derived from the Global Forecast System (GFS) EnKF system. An updated version was implemented in August 2016 using 75% flow-dependent covariances derived from a GFS 80-member ensemble run at an ~30-km grid spacing (Benjamin et al., 2016). In a sense, the operational RAP hybrid DA system is using a DR algorithm, although the ensemble forecasts are from a completely different model, and the GFS EnKF ensemble forecasts are only available four times a day; therefore, many hourly RAP hybrid analyses share forecasts from the same cycle. Such a configuration is clearly not optimal. Thus, a self-consistent EnKF DA system for RAP itself, running at LR, is desirable for maximum consistency and computational cost saving.

In fact, a regional GSI-based EnKF system using the EnSRF (ensemble square-root filter) (Whitaker and Hamill, 2002) had already been established for RAP in Zhu et al. (2013), and tested at a grid spacing (40 km) three times that of the operational RAP. The same RAP operational data stream, with 3-h intervals over a 9-day period, was assimilated, and 18-h deterministic forecasts were evaluated. Results showed that the EnKF system was consistently better than the parallel GSI 3DVar system. Building on the EnKF system tested in Zhu et al. (2013), Pan et al. (2014) implemented a coupled EnKF-3DEnVar hybrid system and evaluated it at the same 40-km resolution as Zhu et al. (2013) for the same test period. With equal weights given to the static and flow-dependent covariances, the hybrid system outperformed GSI 3DVar, and was either comparable to or slightly better than the EnKF system (Pan et al., 2014).

The EnKF and 3DEnVar hybrid systems implemented for RAP in Zhu et al. (2013) and Pan et al. (2014) were tested at a reduced resolution with a 40-km grid spacing because running the ensemble analyses and forecasts at the full ~13-km resolution is expensive. With a DR hybrid system, a single hybrid analysis is performed at the native ~13-km grid spacing to provide initial conditions for the RAP deterministic forecast (the operational RAP is deterministic at this time), using ensemble forecasts produced at the 40-km grid spacing at a much lower cost. Testing and evaluating an ~13-km hybrid system coupled with EnKF cycles at the 40-km grid spacing is the main goal of this paper.

The rest of the paper is organized as follows: section 2 describes the implementation of the coupled DR EnKF-3DEnVar hybrid system for RAP; the design of the testing experiments is given in section 3; the results and comparisons of the experiments are discussed in section 4; section 5 provides conclusions and additional discussion.

2. GSI-based DR EnKF-3DEnVar hybrid system

The GSI-based DR EnKF-3DEnVar hybrid system for RAP is based on the EnKF system reported in Zhu et al.

(2013) and the 3DEnVar hybrid system reported in Pan et al. (2014); both were tested at a reduced SR with a 40-km grid space. Two-way and one-way coupling can be applied in a DR hybrid system (Clayton et al., 2013; Schwartz et al., 2015). For one-way coupling, EnKF only provides the LR ensemble covariance for HR hybrid 3DEnVar analyses. In the two-way DR hybrid, an additional step is performed that re-centers the LR ensemble about the HR analysis. However, Clayton et al. (2013), Wang et al. (2013) and Schwartz et al. (2015) found relatively small impacts of performing such a re-centering. Furthermore, keeping the DR system one-way coupled with the LR EnKF system makes the intercomparisons cleaner. Thus, one-way coupling is employed in this study.

Figure 1 shows the flowchart for the one-way-coupled DR EnKF-3DEnVar hybrid cycles used in this paper. As with the SR hybrid system documented in Pan et al. (2014), the GSI system is used for observation processing, including data quality control, thinning and calculation of the innovations. The EnKF system directly ingests observation innovations processed by GSI and produces an ensemble of analyses. The EnKF analyses and forecasts are run on the 40-km grid. In the DR system, a single deterministic forecast and 3DEnVar hybrid analysis are produced on the HR grid of ~13-km grid spacing in each cycle, using the 40-km ensemble forecasts from the EnKF cycles for flow-dependent BECs. In the rest of this paper, LR refers to the 40-km grid spacing, and HR refers to the ~13-km grid spacing.

Specifically, the DR hybrid algorithm is described below, with the presentation of the general algorithm mostly following Pan et al. (2014) and Schwartz et al. (2015). The analysis increment vector, $\boldsymbol{\delta x}$, of the DR hybrid can be expressed as

$$\boldsymbol{\delta x} = \boldsymbol{\delta x}_1 + \boldsymbol{S D \alpha} \,, \tag{1}$$

where $\boldsymbol{\delta x}_1$ is the HR analysis increment vector associated with the static background covariance with a length of n_h, and $\boldsymbol{S D \alpha}$ is the increment associated with flow-dependent covariances from the K-member ensemble. Equation (1) includes an interpolation operator $\boldsymbol{S}$, that interpolates the $\boldsymbol{D\alpha}$ from LR to HR space, while it is an identity matrix for an SR 3DEnVar algorithm. $\boldsymbol{D}$ is a $n_l \times (Kn_l)$ matrix defined as $\boldsymbol{D} = [\mathrm{diag}(\boldsymbol{x}'_{l1}) \quad \mathrm{diag}(\boldsymbol{x}'_{l2}) \quad \ldots \quad \mathrm{diag}(\boldsymbol{x}'_{lK})]$, where n_l is the dimension of the state vector on the LR grid, and l indicates the lower resolution. diag() is an operator that converts vector $\boldsymbol{x}'_{li}$ into a diagonal matrix with size $n_l \times n_l$ (Wang, 2010; Schwartz et al., 2015). The subscript i indicate the ith member of the ensemble. $\boldsymbol{\alpha}$ is an extended control variable with column vector length of Kn_l formed by concatenating extended control variables for each ensemble member $\boldsymbol{\alpha}_K$ as

$$\boldsymbol{\alpha} = \begin{bmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \\ \vdots \\ \boldsymbol{\alpha}_K \end{bmatrix} . \tag{2}$$

The corresponding cost function to obtain the analysis increment $\boldsymbol{\delta x}_1$ and $\boldsymbol{\alpha}$ is expressed as

$$\begin{aligned} J(\boldsymbol{\delta x}_1, \boldsymbol{\alpha}) &= \beta_1 J_\mathrm{b} + \beta_2 J_\mathrm{e} + J_\mathrm{o} \\ &= \frac{1}{2}\beta_1 \boldsymbol{\delta x}_1^\mathrm{T} \boldsymbol{B}^{-1} \boldsymbol{\delta x}_1 + \frac{1}{2}\beta_2 \boldsymbol{\alpha}^\mathrm{T} \boldsymbol{A}^{-1} \boldsymbol{\alpha} \\ &\quad + \frac{1}{2}[\boldsymbol{H\delta x} - \boldsymbol{y}']^\mathrm{T} \boldsymbol{R}^{-1} [\boldsymbol{H\delta x} - \boldsymbol{y}'] \,, \end{aligned} \tag{3}$$

where J_o is observation term, $\boldsymbol{H}$ is the tangent-linear observation operator, $\boldsymbol{y}'$ is the observation innovation vector, $\boldsymbol{B}$ is the static BEC, and $\boldsymbol{R}$ is the observation error covariance. $\boldsymbol{A}$ is a $Kn_l \times Kn_l$ block diagonal matrix that prescribes the covariance localization scale for the flow-dependent covariance derived from the ensemble forecasts. β_1 and β_2 in front of J_b and J_e are the weights given to static and ensemble BECs,

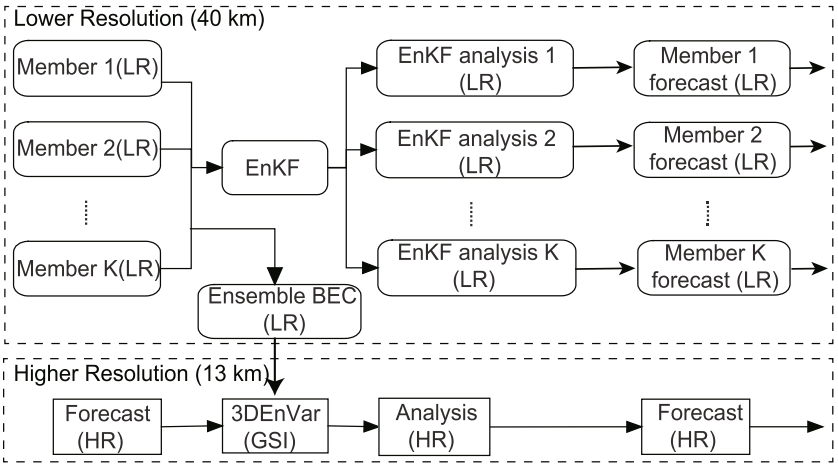

Fig. 1. Flowchart of a full GSI-based EnKF-3DEnVar DR hybrid DA cycle with one-way coupling between the EnKF (upper portion at 40-km horizontal grid spacing) and 3DEnVar DR hybrid analysis (lower portion at 13-km horizontal grid spacing). The arrows pointing from EnKF downward to 3DEnVar indicates the prevision of 40-km ensemble perturbations from EnKF to hybrid analyses. GSI is used as the DA system for 3DEnVar.

respectively, and they are constrained by

$$\frac{1}{\beta_1}+\frac{1}{\beta_2}=1\ . \tag{4}$$

The gradients of the cost function with respect to $\boldsymbol{\delta x}_1$ and $\boldsymbol{\alpha}$ have the form

$$\nabla_{\boldsymbol{\delta x}_1} J=\beta_1\boldsymbol{B}^{-1}\boldsymbol{\delta x}_1+\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}(\boldsymbol{H\delta x}-\boldsymbol{y}')\ , \tag{5}$$

$$\nabla_{\boldsymbol{\alpha}} J=\beta_2\boldsymbol{A}^{-1}\boldsymbol{\alpha}+\boldsymbol{D}^{\mathrm{T}}\boldsymbol{S}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}(\boldsymbol{H\delta x}-\boldsymbol{y}')\ , \tag{6}$$

where $\boldsymbol{D}^{\mathrm{T}}$ and $\boldsymbol{S}^{\mathrm{T}}$ are the adjoints of $\boldsymbol{D}$ and $\boldsymbol{S}$ in Eq. (1), respectively. $\boldsymbol{S}^{\mathrm{T}}$ is applied to $\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}(\boldsymbol{H\delta x}-\boldsymbol{y}')$, while $\boldsymbol{S}$ is applied to $\boldsymbol{D\alpha}$ from the LR space to HR space.

3. Experimental design

3.1. *Model and domain configuration*

In the RAP hybrid DA system, WRF is used as the forecast model. The Model Evaluation Tools (Brown et al., 2009) developed by the Developmental Testbed Center is used for forecast verification.

As stated earlier, the DR hybrid system uses a 40/~13-km horizontal grid spacing combination. The LR domain at 40-km grid spacing for the ensemble covers North America with 207×207 horizontal grid points (bold box in Fig. 2a), while the HR domain at ~13-km grid spacing has 616×616 horizontal grid points covering roughly the same domain (the HR domain is not plotted in Fig. 2). Both domains have 50 vertical levels extending up to 10 hPa at the model top using terrain-following hydrostatic-pressure-based vertical coordinates that stretch with height (Skamarock et al., 2008). The static background error statistics calculated based on the NCEP North American Model forecasts using the National Meteorological Center method, as provided in the GSI system (Hu et al., 2016), are used in this study. The error statistics are latitude and sigma-level dependent only; they are interpolated to the analysis grid within GSI. The flow-dependent BECs are derived from the ensemble forecasts provided by the EnKF system at the 40-km grid spacing.

3.2. *Observations for assimilation and verification*

As in Pan et al. (2014), the operational data stream of RAP excluding satellite radiance data is assimilated in the DR hybrid system. The distributions of the data at 0000 UTC May 8 are shown in Fig. 2. Eighteen-hour forecasts from the analyses are verified against surface and sounding data in the HR domain. The surface data verified include surface pressure, 2-m relatively humidity (RH), 2-m temperature (T), 10-m zonal and meridional wind components (U and V, respectively); the sounding observations include RH, T, U and V.

3.3. *Verification techniques*

The root-mean square error (RMSE) is used to evaluate the forecasts, and the bootstrap resampling method (Candille et al., 2007; Buehner and Mahidjiba, 2010; Schwartz and Liu, 2014) following Pan et al. (2014) is used to determine the statistical significance of error differences. The RMSEs are calculated against the observations at certain levels and forecast hours first, and then aggregated over all cycles for specific forecast hours for skill evaluation.

To assess the statistical significance, bootstrap resampling is performed. New samples are created by randomly drawing from the dataset 3000 times, allowing the same data to be drawn more than once. With the resample, we calculate the aggregated RMSEs along with a two-tailed confidence interval from 5% to 95%. As in Pan et al. (2014), RMSE differences are calculated between a specific experiment and its benchmark first. The bootstrap method is then applied to the RMSE differences with confidence intervals from 5% to 95% to determine the significance of improvement. When all confidence intervals of the RMSE differences are below/above zero, the experiment is significantly better/worse than the benchmark experiment at a 90% confidence level. Additional discussion on the use of the bootstrap method for calculating the statistical significance of forecast differences can be found in Pan et al. (2014), Schwartz and Liu (2014), and Xue et al. (2013).

The Gilbert skill score (GSS) (Gandin and Murphy, 1992) is used to evaluate the precipitation forecast skills of 12-h deterministic forecasts against the 4-km NCEP Stage IV precipitation data in the CONUS (Conterminous United States) domain in which the data are available (Lin and Mitchell, 2005).

3.4. *Experimental design*

As in Zhu et al. (2013) and Pan et al. (2014), the same test period from 8–16 May 2010 is used, which contained active episodes of convection. All DA experiments start at 0000 UTC 8 May and end at 2100 UTC 16 May with continuous three-hourly cycles. The initial fields and boundary conditions are interpolated from operational GFS analyses and forecasts. Random perturbations are created by the random CV3 option in the WRF DA system (Barker, 2005; Barker et al., 2012) and added to the GFS analysis initial condition at 0000 UTC 8 May 2010 to start the ensemble forecasts for the EnKF and the GFS forecasts to create perturbed ensemble boundary conditions (Torn et al., 2006).

Pan et al. (2014) suggested that the BEC weighting factor, as one of the important tuning parameters in a hybrid algorithm, has a great impact on the performance of the hybrid DA system. To examine the performance of the DR hybrid system and its sensitivity to the weighting factor, three DR hybrid experiments—namely, HyDR05, HyDR09 and HyDR10—are run using 50%, 90% and 100% weights given to the ensemble BEC, respectively. Because HyDR09 produces the best analyses and forecasts among all DR hybrid experiments, it is also called the DR hybrid control experiment, named HyDR_Ctl (Table 1).

The well-tuned 3DEnVar experiment (Hybrid1W_Ctl) at the 40-km grid from Pan et al. (2014) is adopted as a benchmark, and the same well-tuned EnKF experiment with 40 members from Pan et al. (2014) is also used in this study to provide the LR ensemble perturbations for the DR hybrid

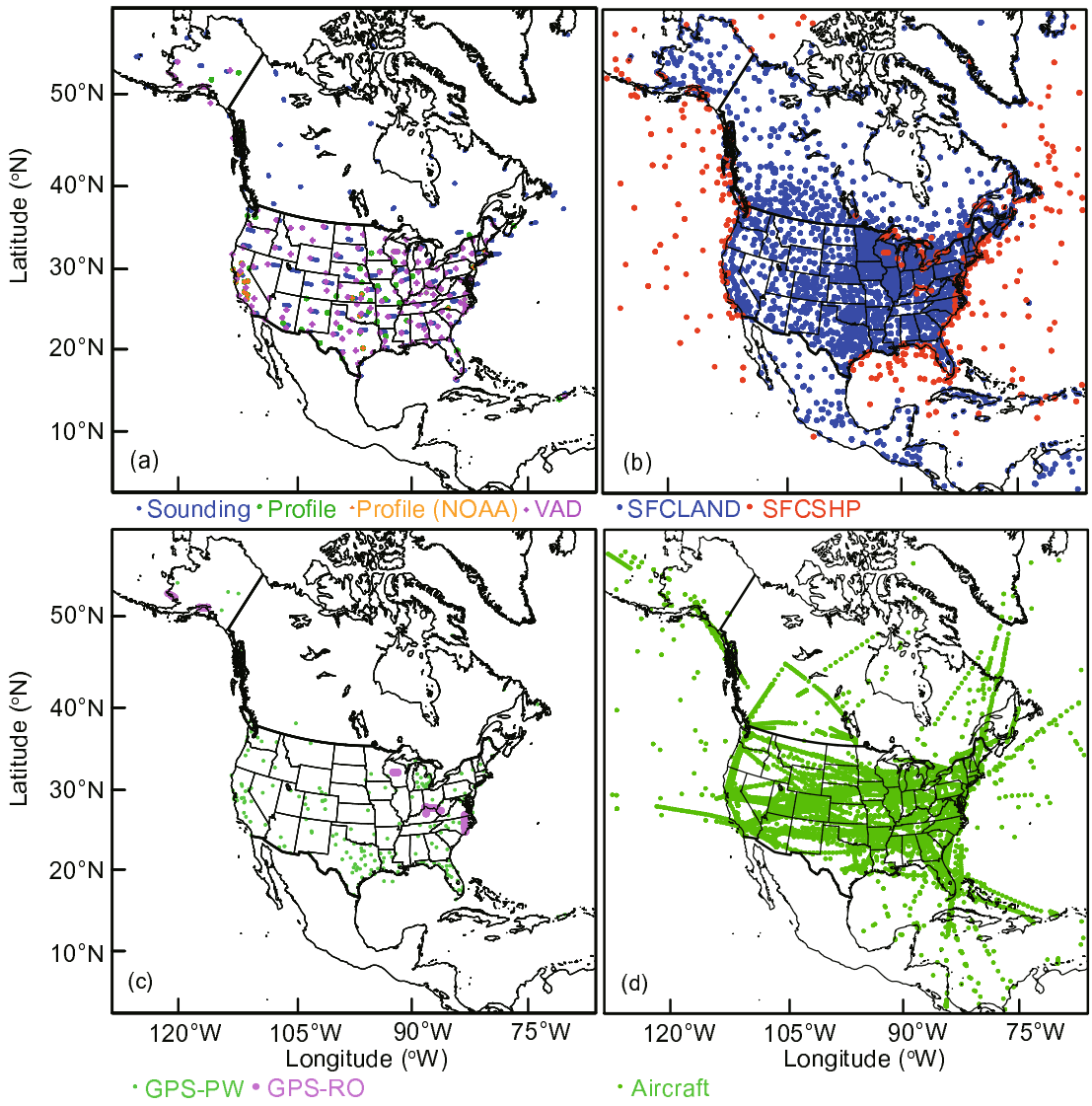

Fig. 2. Example of the horizontal distributions of (a) sounding, profile and VAD (Velocity Azimuth Display), (b) surface stations over land and for ships, (c) GPS-PW (Precipitable Water) and GPS-RO (Radio Occultation), and (d) aircraft observations at 0000 UTC 8 May.

Table 1. List of DA and forecast experiments. In the experiment names, "Hy" and "Var" indicate the hybrid and 3DVar DA methods, respectively. "LR", "HR" and "DR" following the DA method indicate 40 km, ~ 13 km and 40/ ~ 13 km, respectively. The digits such as "09" indicate the weight of the ensemble covariance. "HRF" indicates that forecasts are performed at HR.

Experiment	Ensemble covariance weighting factor ($1/\beta_2$)	Horizontal localization scale (km)	Vertical localization scale [in ln(p)]	For initial conditions or background forecasts	Analysis/ ensemble horizontal grid spacing (km)	Forecast horizontal grid spacing (km)
HyDR_Ctl/HyDR09	0.9	300	0.3	DR hybrid analyses	~ 13/40	~ 13
HyDR05	0.5	300	0.3	DR hybrid analyses	~ 13/40	~ 13
HyDR10	1.0	300	0.3	DR hybrid analyses	~ 13/40	~ 13
HyLR_HRF	As in Pan et al. (2014)		—	Analyses of Hybrid1W_Ctl of Pan et al. (2014)	40/40	~ 13
VarHR_HRF	—	—	—	GSI 3DVar analyses on HR grid	~ 13/ ~ 13	~ 13

experiments. Hybrid1W_Ctl from Pan et al. (2014) employed half and half static/flow-dependent covariance, a horizontal covariance localization scale of 300 km, and a vertical covariance localization scale of 0.3 in terms of natural logarithm of pressure. Experiment HyLR_HRF (Table 1) involves forecasts initialized from interpolated fields from the analyses of Hybrid1W_Ctl every cycle. A HR GSI 3DVar DA experiment, named VarHR_HRF, is also run at the ~13-km resolution (Table 1). In other words, HR forecasts from the LR hybrid control experiments and the HR 3DVar analyses are used as references for evaluating forecasts from DR DA experiments.

The analyses of DR DA may benefit from the HR background forecasts, which contain more detailed flow structures. The configurations of HyDR05 are the same as HyLR_HRF, except that the deterministic background forecasts of HyDR05 are performed on the ~13-km grid instead of the 40-km grid. In HyLR_HRF, forecasts are run at ~13-km grid spacing from interpolated 40-km analyses of Hybrid1W_Ctl. The comparison between HyDR05 and HyLR_HRF isolates the impact of the increased background forecast resolution.

For variational minimization in either 3DVar or 3DEnVar, two outer-loop iterations and 100 inner-loop iterations are used. Evaluations are mainly based on forecasts on the ~13-km grid, launched from either the HR analyses or fields interpolated from LR 40-km analyses. All experiments are listed in Table 1.

4. Results and discussion

In Pan et al. (2014), the skills of EnKF, 3DEnVar hybrid and 3DVar at a single reduced 40-km resolution were compared. The results showed 3DEnVar hybrid using 50% ensemble covariance significantly outperformed GSI 3DVar for all variables through the entire 18-h forecast period, while 3DEnVar hybrid had a comparable performance to EnKF overall. In this section, the performance of the DR hybrid system and its sensitivity to the weighting factor of BECs and the resolution of background forecasts are examined. In section 4.4, precipitation forecast skills are evaluated.

4.1. *Sensitivity to the covariance weighting factor in DR hybrid experiments*

In Pan et al. (2014), the lowest RMSEs were obtained when using 50% ensemble BEC in their 40-km control hybrid experiment, Hybrid1W_Ctl. At an ~13-km grid spacing, smaller-scale features can be captured, which tend to be more transient and hence more flow-dependent. The analysis may benefit from a higher weight for the ensemble covariance. Experiments HyDR05, HyDR09 (also named HyDR_Ctl) and HyDR10 are compared to examine the impact of flow-dependent covariance in the DR hybrid system.

The aggregated 3-h forecast RMSEs verified against sounding data are shown in Fig. 3. The RMSEs at each pressure level were obtained by averaging values within a layer 50 hPa above and below that pressure from all cycles, except for the topmost and lowest levels. The 3-h forecasts are also used as the background in each DA cycle, and their errors can be used as a proxy for measuring the DA quality. The results show that HyDR09 has the smallest RMSEs for RH, U and V at almost all levels. For T, the RMSEs from HyDR09 are higher than those from HyDR05 above 800 hPa. The performance of HyDR10 is comparable to or worse than HyDR05 for RH below 600 hPa, and for T, U and V at all levels. These results indicate that, with the DR hybrid 3DEnVar system, when the grid spacing of the hybrid analysis as well as the background forecast is decreased from the 40-km used in Pan et al. (2014) to ~13-km, optimum results are obtained when the weight for the ensemble BECs is 90% (among the weights examined), instead of the 50% for the SR LR case. This may be because of the increased level of flow dependency of the background errors at HR, as suggested earlier. Raising the weighting factor for the flow-dependent covariances means that more mesoscale information can be involved in the DA. However, the forecasting skill of T to the weighting factor is opposite to RH, U and V at the middle to upper levels.

4.2. *Comparison of DR hybrid DA with HR 3DVAR*

In this section, we compare the performance of experiments HyDR_Ctl (i.e., HyDR09) using hybrid 3DEnVar with VarHR_HRF, which uses the pure 3DVar DA method run at HR (see Table 1).

The 9-day aggregated RMSEs of the 3-h forecasts verified against sounding data at all levels are shown in Fig. 4. As shown in Fig. 4, VarHR_HRF underperforms HyDR_Ctl, with its errors being significantly larger for RH, U and V at most levels, while the errors for T are comparable.

The overall domain- and level-aggregated RMSEs verified against sounding and surface data are shown in Fig. 5 and Fig. 6, respectively, for analyses (hour 0) and forecasts at 3-h intervals up to 18 hours. HyDR_Ctl significantly outperforms VarHR_HRF at the analysis and forecast for all variables throughout the entire forecast period. The RMSEs of all variables are noticeably lower in the analyses than in the forecasts, and forecast errors increase quickly in the first three hours before becoming more stable thereafter; such rapid error growth is likely associated with fast small-scale error growth.

Overall, the DR coupled EnKF-3DEnVar hybrid scheme significantly outperforms the 3DVar scheme for all variables at all forecast hours when verified against soundings and surface observations. The results suggest the efficacy of using a DR configuration for a hybrid DA system.

4.3. *Impact of HR background forecast*

The impacts of the HR background forecast are investigated by comparing HyLR_HRF with HyDR05, in which the only differences lie with the resolution of the background forecasts. HyDR_Ctl is also included in this section to assess the impacts of HR background forecasts and flow-dependent

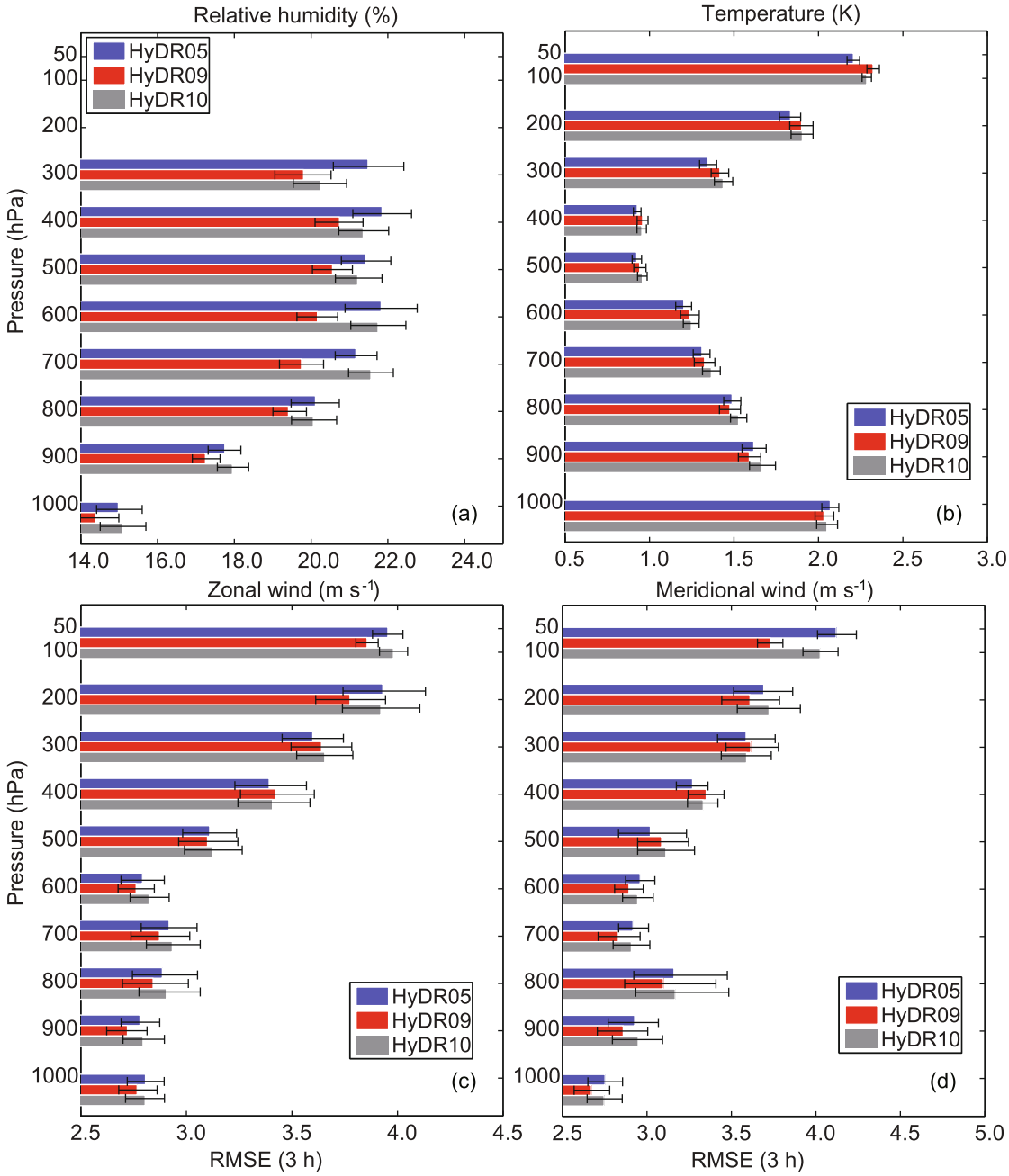

Fig. 3. Aggregated 3-h forecast RMSEs along with confidence error bars at different height levels verified against sounding data for (a) RH, (b) T, (c) U, and (d) V for experiments HyDR05, HyDR09, and HyDR10. The error bars represent the two-tailed 90% confidence interval (5% on the left and 95% on the right) using the bootstrap distribution method.

covariance.

The 3-h forecast RMSEs verified against sounding data (Fig. 4) show that HyDR05 underperforms HyLR_HRF for RH and performs comparably for T, U and V at most levels. When using a higher weighting factor of 90% for flow-dependent covariances in HyDR_Ctl, the RMSEs are smaller than those from HyDR05 for RH at all levels, except for T at 1000–800 hPa, and U and V at 500–300 hPa. These results suggest that HyDR_Ctl benefits from the HR with 90% flow-dependent covariances.

The comparisons among HyLR_HRF, HyDR_Ctl and HyDR05 of the domain-aggregated RMSEs for the analyses and forecasts up to 18 hours against sounding data are shown in Fig. 5. The RMSEs of HyDR05 are comparable or slightly worse than those from HyLR_HRF, while the RMSEs of HyDR_Ctl are significantly smaller than those of HyLR_HRF for all variables except T. At the lower resolution of 40 km, the best analyses and forecasts were obtained in Pan et al. (2014) when equal weights were given to the static and flow-dependent covariances in the hybrid DA. As the grid resolu-

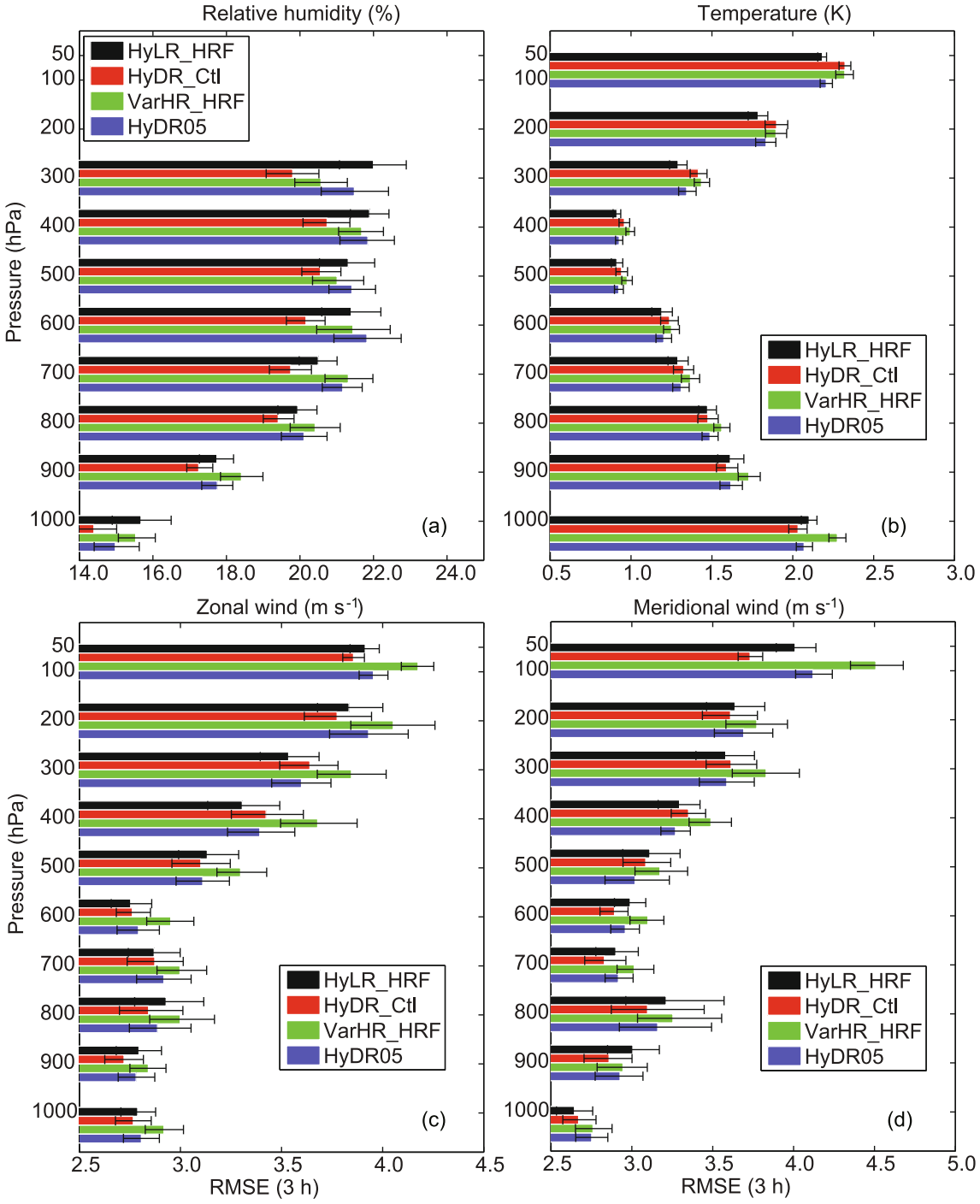

Fig. 4. As in Fig. 3 but for experiments HyLR_HRF, HyDR_Ctl, VarHR_HRF and HyDR05. The error bars indicate the two-tailed 90% confidence interval using the bootstrap method with 5% on the left and 95% on the right.

tion increases, the smooth static covariance becomes less appropriate, which explains why a higher ensemble covariance weight of 90% used in HyDR_Ctl is beneficial.

The impacts of the HR background forecasts are further examined by verifying analyses and forecasts up to 18 hours against surface data (Fig. 6). HyDR_Ctl has significantly smaller RMSEs than HyLR_HRF for 2-m RH and 10-m U and V from the analysis time and throughout the entire forecast period, and for surface pressure except at a few forecast hours. Large differences are found in the RH errors between the HR 3DVar/DR hybrid and the LR 3DEnVar hybrid (Fig. 6b), suggesting that for the surface moisture field DA can benefit significantly from the increased background resolution, given the better resolution of terrain and mesoscale boundary layer structures. For 2-m T, smaller errors at the analysis time in HyDR_Ctl and VarHR_HRF than those in HyLR_HRF indicate a better fit of the analyses to surface T observations. However, the forecast errors of T in HyDR_Ctl

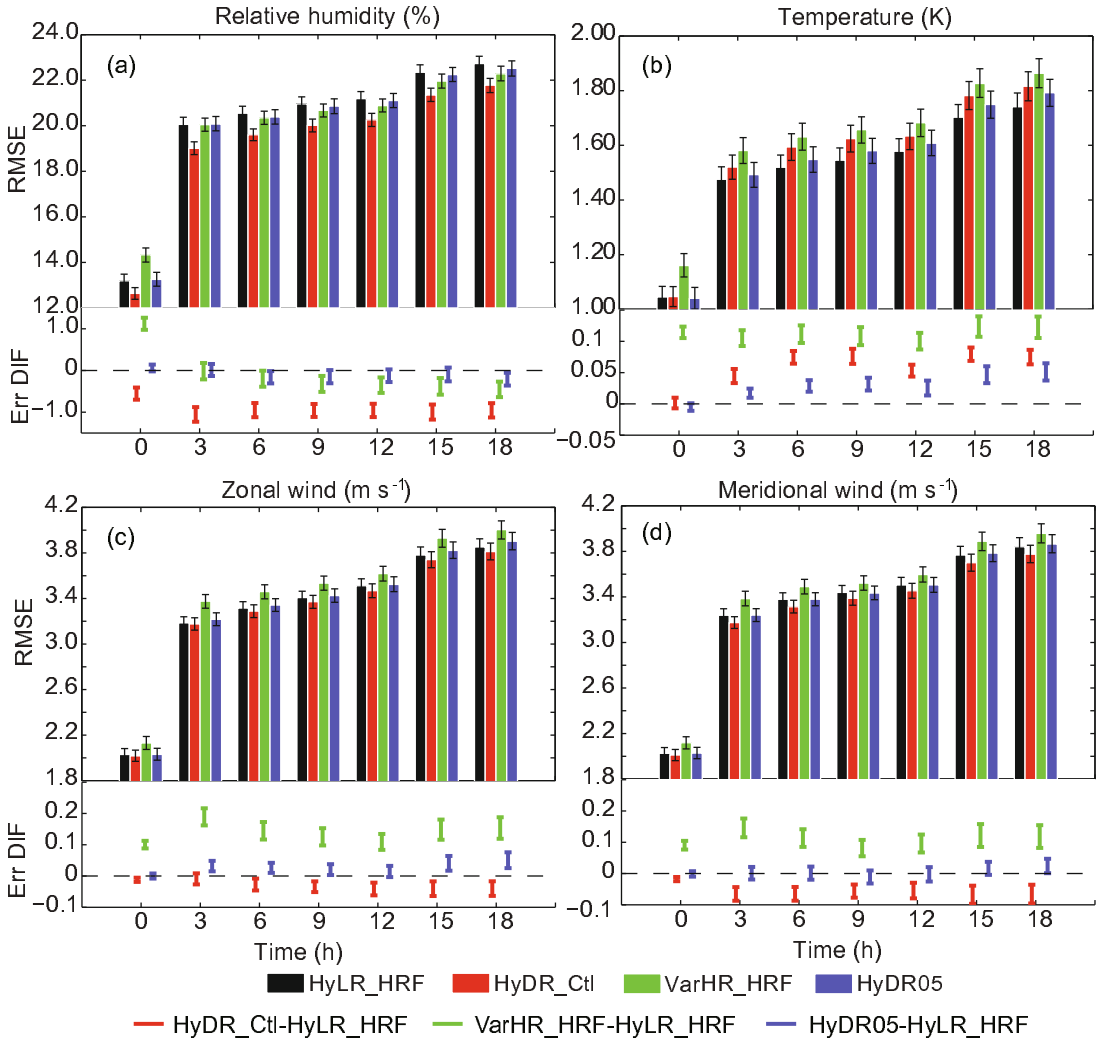

Fig. 5. The bar chart in each frame shows the RMSEs of forecasts verified against sounding data, aggregated over the entire domain and over the nine-day period. The lower panel shows the 90% confidence interval of the RMSE differences between HyDR_Ctl and VarHR_HRF or HyDR05 and HyLR_HRF for (a) RH, (b) T, (c) U, and (d) V, for different forecast hours. If the interval does not include zero, the difference is statistically significant at the 90% confidence level. The error bars in the histograms represent the two-tailed 90% confidence interval with 5% at the bottom and 95% on the top using the bootstrap distribution method.

and VarHR_HRF become larger after three hours of forecasting than those in HyLR_HRF.

The results seem to suggest that the humidity and wind fields benefit more from the higher background resolution with increasing flow-dependent covariance, while this is not necessarily the case for the temperature forecasts, at least when verified against conventional data in terms of the RMSEs. Experiments with various combinations of resolutions used in the analysis and forecasting steps shed some light on such complex behaviors (not shown), but are not enough to fully answer the questions. Results also imply a need for multi-scale DA algorithms that explicitly treat observations and background errors representing different scales (Li et al., 2015) and use scale-dependent (Buehner and Shlyaeva, 2015) and/or multi-scale covariance localization (Miyoshi and Kondo, 2013).

4.4. *Precipitation forecast skill*

In this section, the precipitation forecasts from HyLR_HRF, HyDR_Ctl, VarHR_HRF and HyDR05 are verified against the 4-km NCEP Stage IV precipitation data. The GSS (Gandin and Murphy, 1992), also known as the equitable threat score, is calculated, as in Pan et al. (2014), for the 0.1, 1.25 and 2.5 mm h^{-1} thresholds.

The GSSs are shown in Fig. 7. That HyDR_Ctl outperforms VarHR_HRF for all thresholds and all forecast hours suggests that the analysis method is important for precipitation forecasting skill. The results are consistent with those of Schwartz (2016), who examined DR hybrid DA with a 20/4-km grid combination.

With the HR forecasts, HyDR05 has better skill than HyLR_HRF after five hours at the threshold of 0.1 mm h^{-1}, and at one to eight hours at the threshold of 2.5 mm h^{-1}.

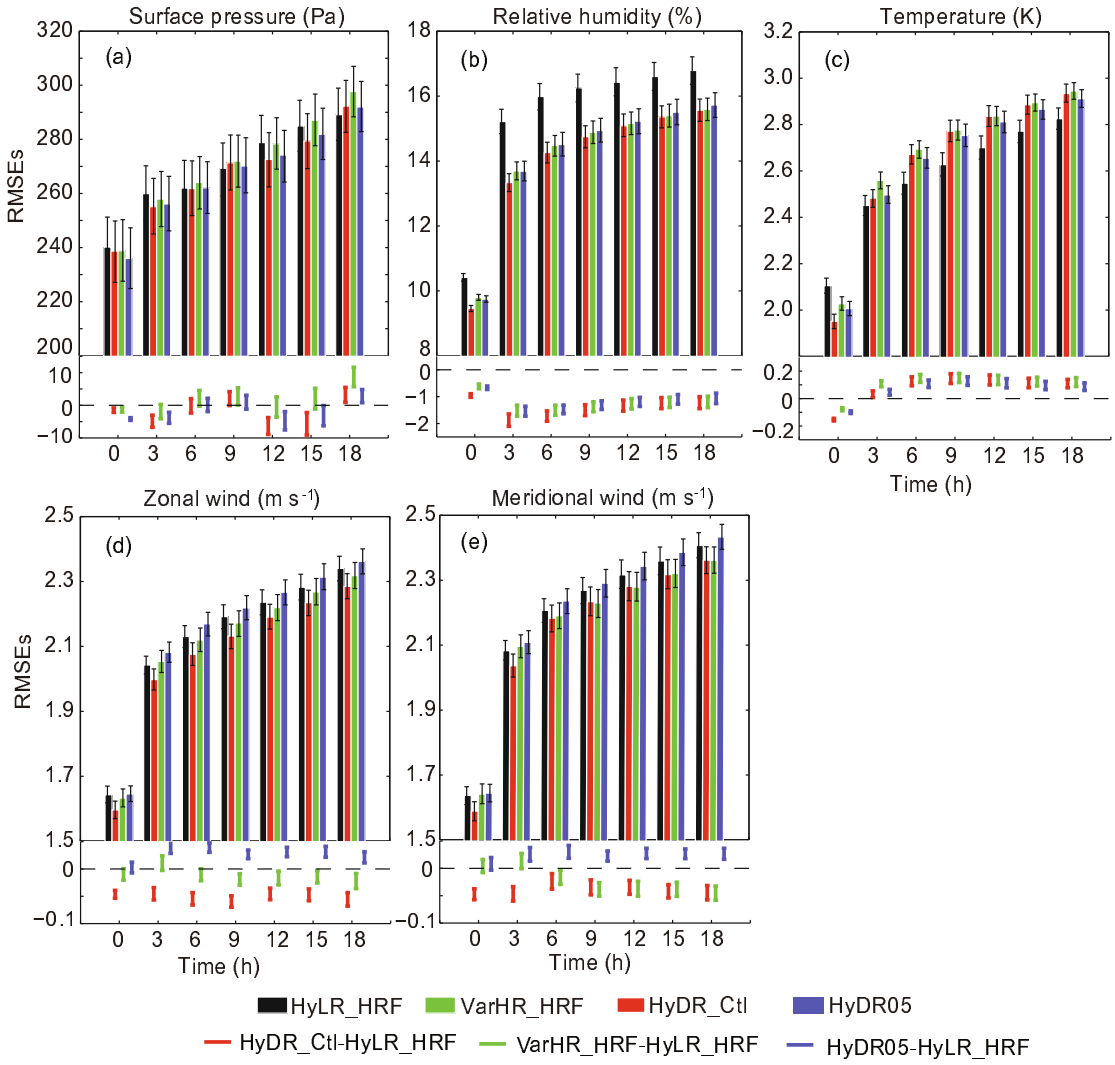

Fig. 6. The bar chart in the upper panel of each frame shows the RMSEs of forecasts verified against surface station observations, aggregated over the entire domain and over the nine-day period, for (a) surface pressure, (b) 2-m RH, (c) 2-m T, (d) 10-m U, and (e) 10-m V for different forecast hours. Confidence error bars represent the two-tailed 90% confidence interval (5% at the bottom and 95% on the top) using the bootstrap distribution method. The lower panel of each frame shows the 90% confidence interval of the RMSE differences between HyDR_Ctl, VarHR_HRF or HyDR05 and HyLR_HRF.

With more flow-dependent covariance being used, HyDR_Ctl shows the best skill among all experiments after 10 hours at the 0.1 mm h^{-1} threshold, and generally all hours at the 1.25 and 2.5 mm h^{-1} thresholds.

The results indicate that, for precipitation, especially heavier precipitation, there is a clear benefit to running the hybrid DA at HR (relative to the LR hybrid), and to using ensemble-derived flow-dependent covariance (relative to 3DVar). The improved precipitation forecasts are consistent with reduced errors in the analyses and forecasts of humidity.

5. Summary and conclusions

Based on the NCEP operational GSI variational framework, a coupled EnKF-3DEnVar DR hybrid DA system, using a 40/~13-km combination, is tested for the RAP model configuration. The LR ensemble is provided by a single LR EnKF system developed and tested in Zhu et al. (2013), and the single LR EnKF-3DEnVar hybrid system tested in Pan et al. (2014) is used as a benchmark to evaluate the performance of the coupled DR hybrid system. As in Zhu et al. (2013) and Pan et al. (2014), a nine-day period from 8–17 May 2010, containing active convective systems, is used to examine the performance of the DR hybrid DA system. The conventional data stream used by the operational RAP system is assimilated. The analyses and forecasts are verified against surface and sounding observations, while HR precipitation forecasts are verified against Stage IV precipitation data.

A 90% weight for the ensemble covariance produces the best forecasts with the DR hybrid system (HyDR_Ctl), while

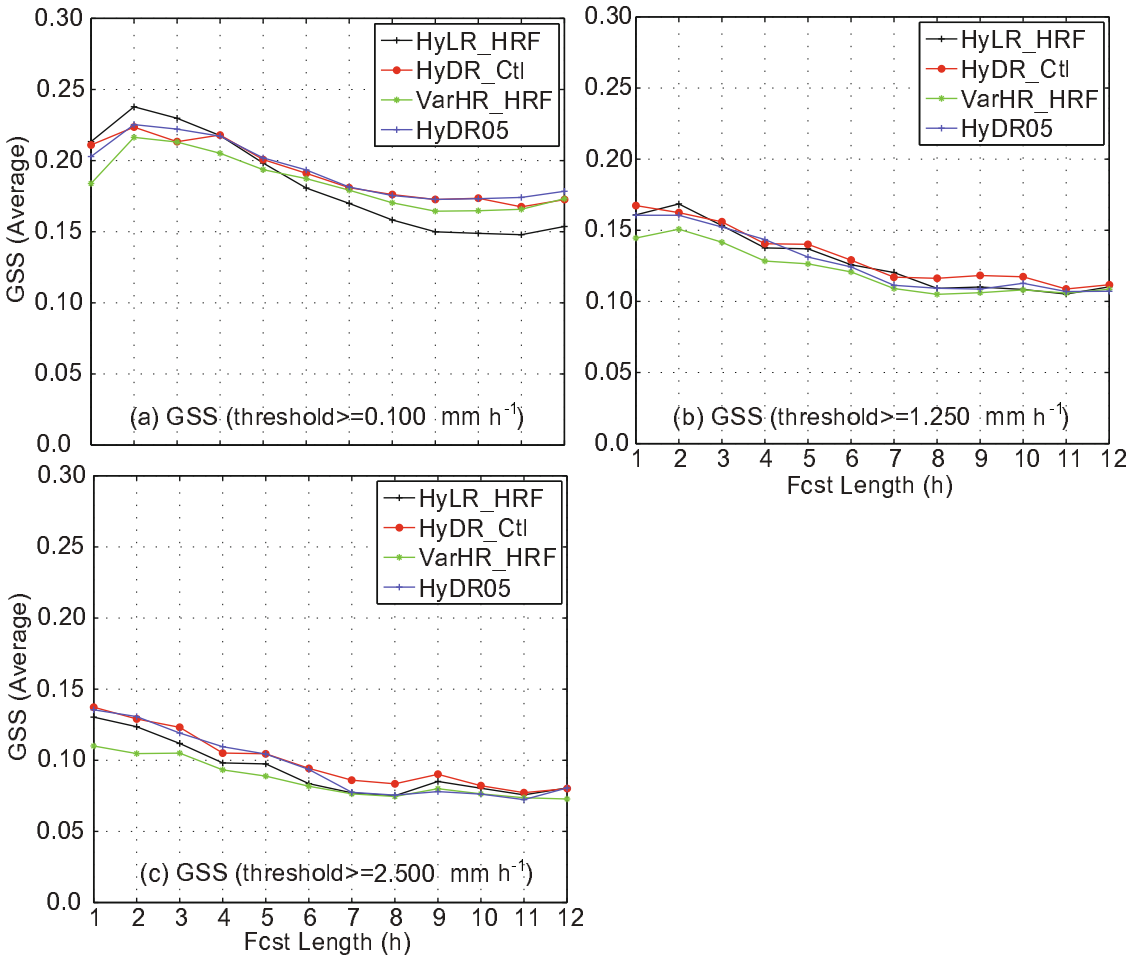

Fig. 7. Aggregated precipitation GSSs of 13-km forecasts as a function of forecast length for thresholds of (a) 0.1 mm h^{-1}, (b) 1.25 mm h^{-1} and (c) 2.5 mm h^{-1}.

a 50% percent weight given to the flow-dependent ensemble covariance is found to have the best performance with the 40-km LR hybrid system in Pan et al. (2014). The comparisons among HyLR_HRF, HyDR_Ctl and HyDR05 suggest that the hybrid analyses and forecasts can benefit from the HR background (3-h deterministic HR forecast) when the weight given to the ensemble covariance is larger. In comparison, the impacts of the HR background forecasts are limited when 50% static covariance is used. By increasing the weight for the ensemble covariance to 90% within the DR hybrid system, the humidity and wind fields are improved throughout the 18 hours of the forecast, and the improvements to the humidity fields are the largest. However, the response of temperature forecasting skill to the weighting factor is opposite to other variables at the middle to upper levels. The exact reasons require further investigation.

The overall benefits of performing the hybrid analyses at HR while still keeping the EnKF cycles at LR to reduce the computational cost are clear, especially for the precipitation forecasting skill. Such benefits can be larger when the grid resolution becomes convection-allowing or convection-resolving. Corresponding to the largest improvement of RH, the precipitation forecast skill from the DR hybrid system is higher than that from the HR 3DVar method, suggesting that the analysis method is as important as the analysis resolution for convection-allowing predictions. These results are consistent with Schwartz (2016), in which 4-km convection-allowing forecasts using 20/4-km DR hybrid DA based on the GSI system were examined. Their precipitation forecasts over the first 12 hours from 4-km 3DVar and hybrid 3DEnVar were better than the forecasts from corresponding downscaled 20-km analyses. All precipitation forecasts from their 4-km hybrid analyses were more skillful than those from their 4-km 3DVar analyses.

Acknowledgements. This work was primarily supported by the National Natural Science Foundation of China (Grant Nos. 41730965, 41775099 and 2017YFC1502104), and PAPD (the Priority Academic Program Development of Jiangsu Higher Education Institutions).

REFERENCES

Ancell, B. C., C. F. Mass, K. Cook, and B. Colman, 2014: Comparison of surface wind and temperature analyses from an ensemble Kalman filter and the NWS real-time mesoscale analy-

sis system. *Wea. Forecasting*, **29**, 1058–1075, https://doi.org/10.1175/WAF-D-13-00139.1.

Anderson, J. L., 2016: Reducing correlation sampling error in Ensemble Kalman Filter data assimilation. *Mon. Wea. Rev*., **144**, 913–925, https://doi.org/10.1175/MWR-D-15-0052.1.

Barker, D., and Coauthors, 2012: The weather research and forecasting model's community variational/ensemble data assimilation system: WRFDA. *Bull. Amer. Meteor. Soc*., **93**, 831–843, https://doi.org/10.1175/BAMS-D-11-00167.1.

Barker, D. M., 2005: Southern high-latitude ensemble data assimilation in the Antarctic mesoscale prediction system. *Mon. Wea. Rev*., **133**, 3431–3449, https://doi.org/10.1175/MWR3042.1.

Barker, D. M., W. Huang, Y.-R. Guo, A. J. Bourgeois, and Q. N. Xiao, 2004: A three-dimensional variational data assimilation system for MM5: Implementation and initial results. *Mon. Wea. Rev*., **132**, 897–914, https://doi.org/10.1175/1520-0493(2004)132<0897:ATVDAS>2.0.CO;2.

Benjamin, S. G., and Coauthors, 2004: An hourly assimilation forecast cycle: The RUC. *Mon. Wea. Rev*., **132**, 495–518, https://doi.org/10.1175/1520-0493(2004)132<0495:AHACTR>2.0.CO;2.

Benjamin, S. G., and Coauthors, 2016: A north American hourly assimilation and model forecast cycle: The rapid refresh. *Mon. Wea. Rev*., **144**, 1669–1694, https://doi.org/0.1175/MWR-D-15-0242.1.

Brown, B., J. H. Gotway, R. Bullock, E. Gilleland, T. Fowler, D. Ahijevych, and T. Jensen, 2009: The Model Evaluation Tools (MET): Community tools for forecast evaluation. *25th International Conference on Interactive Information and Processing Systems* (*IIPS*) *for Meteorology*, *Oceanography*, *and Hydrology*, Paper 9A. 6, Phoenix, AZ, American Meteor Society.

Buehner, M., and A. Mahidjiba, 2010: Sensitivity of global ensemble forecasts to the initial ensemble mean and perturbations: Comparison of EnKF, singular vector, and 4D-var approaches. *Mon. Wea. Rev*., **138**, 3886–3904, https://doi.org/10.1175/2010MWR3296.1.

Buehner, M., and A. Shlyaeva, 2015: Scale-dependent background-error covariance localisation. *Tellus A*, **67**, 28027, https://doi.org/10.3402/tellusa.v67.28027.

Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, 2010a: Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part I: Description and single-observation experiments. *Mon. Wea. Rev*., **138**, 1550–1566, https://doi.org/10.1175/2009MWR3157.1.

Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, 2010b: Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part II: One-month experiments with real observations. *Mon. Wea. Rev*., **138**, 1567–1586, https://doi.org/10.1175/2009MWR3158.1.

Candille, G., C. Côté, P. L. Houtekamer, and G. Pellerin, 2007: Verification of an ensemble prediction system against observations. *Mon. Wea. Rev*., **135**, 2688–2699, https://doi.org/10.1175/MWR3414.1.

Clayton, A. M., A. C. Lorenc, and D. M. Barker, 2013: Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office. *Quart. J. Roy. Meteor. Soc*., **139**, 1445–1461, https://doi.org/10.1002/qj.2054.

Descombes, G., T. Auligné, F. Vandenberghe, D. M. Barker, and J. Barré, 2015: Generalized background error covariance matrix model (GEN_BE v2.0). *Geoscientific Model Development*, **8**, 669–696, https://doi.org/10.5194/gmd-8-669-2015.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res*., **99**, 10 143–10 162, https://doi.org/10.1029/94JC00572.

Gandin, L. S., and A. H. Murphy, 1992: Equitable skill scores for categorical forecasts. *Mon. Wea. Rev*., **120**, 361–370, https://doi.org/10.1175/1520-0493(1992)120<0361:ESSFCF>2.0.CO;2.

Greybush, S. J., E. Kalnay, T. Miyoshi, K. Ide, and B. R. Hunt, 2010: Balance and ensemble Kalman filter localization techniques. *Mon. Wea. Rev*., **139**, 511–522, https://doi.org/10.1175/2010MWR3328.1.

Hamill, T. M., and C. Snyder, 2000: A hybrid ensemble Kalman filter - 3D variational analysis scheme. *Mon. Wea. Rev*., **128**, 2905–2919, https://doi.org/10.1175/1520-0493(2000)128<2905:AHEKFV>2.0.CO;2.

Hamill, T. M., J. S. Whitaker, and C. Snyder, 2001: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev*., **129**, 2776–2790, https://doi.org/10.1175/1520-0493(2001)129<2776:DDFOBE>2.0.CO;2.

Hamill, T. M., J. S. Whitaker, M. Fiorino, and S. G. Benjamin, 2011: Global ensemble predictions of 2009's tropical cyclones initialized with an ensemble Kalman filter. *Mon. Wea. Rev*., **139**, 668–688, https://doi.org/10.1175/2010MWR3456.1.

Hu, M., H. Shao, D. Stark, K. Newman, C. Zhou, and X. Zhang, 2016: Grid-Point Statistical Interpolation (GSI) User's Guide Version 3.5. Developmental Testbed Center, 141 pp. [Available online at http://www.dtcenter.org/com-GSI/users/docs/index.php]

Kleist, D. T., and K. Ide, 2015: An OSSE-based evaluation of hybrid variational–ensemble data assimilation for the NCEP GFS. Part II: 4DEnVar and hybrid variants. *Mon. Wea. Rev*., **143**, 452–470, https://doi.org/10.1175/MWR-D-13-00350.1.

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W.-S. Wu, and S. Lord, 2009: Introduction of the GSI into the NCEP global data assimilation system. *Wea. Forecasting*, **24**, 1691–1705, https://doi.org/10.1175/2009WAF2222201.1.

Kuhl, D. D., T. E. Rosmond, C. H. Bishop, J. McLay, and N. L. Baker, 2013: Comparison of hybrid ensemble/4DVar and 4DVar within the NAVDAS-AR data assimilation framework. *Mon. Wea. Rev*., **141**, 2740–2758, https://doi.org/10.1175/MWR-D-12-00182.1.

Li, Y. Z., X. G. Wang, and M. Xue, 2012: Assimilation of radar radial velocity data with the WRF hybrid ensemble-3DVAR system for the prediction of hurricane Ike (2008). *Mon. Wea. Rev*., **140**, 3507–3524, https://doi.org/10.1175/MWR-D-12-00043.1.

Li, Z. J., J. C. McWilliams, K. Ide, and J. D. Farrara, 2015: A multiscale variational data assimilation scheme: Formulation and illustration. *Mon. Wea. Rev*., **143**, 3804–3822, https://doi.org/10.1175/MWR-D-14-00384.1.

Lin, Y., and K. E. Mitchell, 2005: The NCEP Stage II/IV hourly precipitation analyses: Development and applications. *19th Conference on Hydrology*, Paper 1. 2, American Meteor Society, San Diego, CA.

Liu, H. X., and M. Xue, 2008: Prediction of convective initiation and storm evolution on 12 June 2002 during IHOP_2002. Part I: Control simulation and sensitivity experiments. *Mon.*

Wea. Rev., **136**, 2261–2283, https://doi.org/10.1175/2007MWR2161.1.

Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **112**, 1177–1194, https://doi.org/10.1002/qj.49711247414.

Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP-a comparison with 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **129**, 3183–3204, https://doi.org/10.1256/qj.02.132.

Meng, Z. Y., and F. Q. Zhang, 2007: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part II: Imperfect model experiments. *Mon. Wea. Rev.*, **135**, 1403–1423, https://doi.org/10.1175/MWR3101.1.

Miyoshi, T., and K. Kondo, 2013: A multi-scale localization approach to an ensemble Kalman filter. *SOLA*, **9**, 170–173, https://doi.org/10.2151/sola.2013-038.

Miyoshi, T., K. Kondo, and T. Imamura, 2014: The 10,240-member ensemble Kalman filtering with an intermediate AGCM. *Geophys. Res. Lett.*, **41**, 5264–5271, https://doi.org/10.1002/2014GL060863.

Pan, Y. J., K. F. Zhu, M. Xue, X. G. Wang, M. Hu, S. G. Benjamin, S. S. Weygandt, and J. S. Whitaker, 2014: A GSI-based coupled EnSRF–En3DVar hybrid data assimilation system for the operational rapid refresh model: Tests at a reduced resolution. *Mon. Wea. Rev.*, **142**, 3756–3780, https://doi.org/10.1175/MWR-D-13-00242.1.

Schwartz, C. S., 2016: Improving large-domain convection-allowing forecasts with high-resolution analyses and ensemble data assimilation. *Mon. Wea. Rev.*, **144**, 1777–1803, https://doi.org/10.1175/MWR-D-15-0286.1.

Schwartz, C. S., and Z. Q. Liu, 2014: Convection-permitting forecasts initialized with continuously cycling limited-area 3DVAR, ensemble Kalman filter, and "hybrid" variational–ensemble data assimilation systems. *Mon. Wea. Rev.*, **142**, 716–738, https://doi.org/10.1175/MWR-D-13-00100.1.

Schwartz, C. S., Z. Q. Liu, and X.-Y. Huang, 2015: Sensitivity of limited-area hybrid variational-ensemble analyses and forecasts to ensemble perturbation resolution. *Mon. Wea. Rev.*, **143**, 3454–3477, https://doi.org/10.1175/MWR-D-14-00259.1.

Skamarock, W. C., and J. B. Klemp, 2008: A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. *J. Comput. Phys.*, **227**, 3465–3485, https://doi.org/10.1016/j.jcp.2007.01.037.

Skamarock, W. C., and Coauthors, 2008: A Description of the Advanced Research WRF Version 3. NCAR Technical Note NCAR/TN-475+STR, 7-8, https://doi.org/10.5065/D68S4MVH.

Torn, R. D., G. J. Hakim, and C. Snyder, 2006: Boundary conditions for limited-area ensemble Kalman filters. *Mon. Wea. Rev.*, **134**, 2490–2502, https://doi.org/10.1175/MWR3187.1.

Wang, X. G., 2010: Incorporating ensemble covariance in the Gridpoint Statistical Interpolation variational minimization: A mathematical framework. *Mon. Wea. Rev.*, **138**, 2990–2995, https://doi.org/10.1175/2010MWR3245.1.

Wang, X. G., C. Snyder, and T. M. Hamill, 2007: On the theoretical equivalence of differently proposed ensemble/3DVAR hybrid analysis schemes. *Mon. Wea. Rev.*, **135**, 222–227, https://doi.org/10.1175/MWR3282.1.

Wang, X. G., D. M. Barker, C. Snyder, and T. M. Hamill, 2008a: A hybrid ETKF-3DVAR data assimilation scheme for the WRF model. Part II: Real observation experiment. *Mon. Wea. Rev.*, **136**, 5132–5147, https://doi.org/10.1175/2008MWR2445.1.

Wang, X. G., D. M. Barker, C. Snyder, and T. M. Hamill, 2008b: A hybrid ETKF-3DVAR data assimilation scheme for the WRF model. Part I: Observing system simulation experiment. *Mon. Wea. Rev.*, **136**, 5116–5131, https://doi.org/10.1175/2008MWR2444.1.

Wang, X. G, T. M. Hamill, J. S. Whitaker, and C. H. Bishop, 2009: A comparison of the hybrid and EnSRF analysis schemes in the presence of model errors due to unresolved scales. *Mon. Wea. Rev.*, **137**, 3219–3232, https://doi.org/10.1175/2009MWR2923.1.

Wang, X. G., D. Parrish, D. Kleist, and J. Whitaker, 2013: GSI 3DVar-based ensemble–variational hybrid data assimilation for NCEP global forecast system: Single-resolution experiments. *Mon. Wea. Rev.*, **141**, 4098–4117, https://doi.org/10.1175/MWR-D-12-00141.1.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924, https://doi.org/10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2.

Wu, W.-S., R. J. Purser, and D. F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.*, **130**, 2905–2916, https://doi.org/10.1175/1520-0493(2002)130<2905:TDVAWS>2.0.CO;2.

Wu, W. S., D. F. Parrish, E. Rogers, and Y. Lin, 2017: Regional ensemble–variational data assimilation using global ensemble forecasts. *Wea. Forecasting*, **32**, 83–96, https://doi.org/10.1175/WAF-D-16-0045.1.

Xue, M., J. Schleif, F. Y. Kong, K. W. Thomas, Y. H. Wang, and K. F. Zhu, 2013: Track and intensity forecasting of hurricanes: Impact of convection-permitting resolution and global ensemble Kalman filter analysis on 2010 Atlantic season forecasts. *Wea. Forecasting*, **28**, 1366–1384, https://doi.org/10.1175/WAF-D-12-00063.1.

Zhang, F. Q., Z. Y. Meng, and A. Aksoy, 2006: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part I: Perfect model experiments. *Mon. Wea. Rev.*, **134**, 722–736, https://doi.org/10.1175/MWR3101.1.

Zhang, F. Q., M. Zhang, and J. Poterjoy, 2013: E3DVar: coupling an ensemble Kalman filter with three-dimensional variational data assimilation in a limited-area weather prediction model and comparison to E4DVar. *Mon. Wea. Rev.*, **141**, 900–917, https://doi.org/10.1175/MWR-D-12-00075.1.

Zhang, M., and F. Q. Zhang, 2012: E4DVar: Coupling an ensemble Kalman filter with four-dimensional variational data assimilation in a limited-area weather prediction model. *Mon. Wea. Rev.*, **140**, 587–600, https://doi.org/10.1175/MWR-D-11-00023.1.

Zhu, K. F., Y. J. Pan, M. Xue, X. G. Wang, J. S. Whitaker, S. G. Benjamin, S. S. Weygandt, and M. Hu, 2013: A regional GSI-based ensemble Kalman filter data assimilation system for the rapid refresh configuration: Testing at reduced resolution. *Mon. Wea. Rev.*, **141**, 4118–4139, https://doi.org/10.1175/MWR-D-13-00039.1.