

A GSI-Based Coupled EnSRF–En3DVar Hybrid Data Assimilation System for the Operational Rapid Refresh Model: Tests at a Reduced Resolution

YUJIE PAN,* KEFENG ZHU,⁺ MING XUE,[#] XUGUANG WANG,[@] MING HU,[&] STANLEY G. BENJAMIN,**
STEPHEN S. WEYGANDT,⁺⁺ AND JEFFREY S. WHITAKER⁺⁺

* Center for Analysis and Prediction of Storms, Norman, Oklahoma and Nanjing University of Information Science and Technology, Nanjing, China

⁺ Center for Analysis and Prediction of Storms, Norman, Oklahoma and Nanjing University, Nanjing, China

[#] Center for Analysis and Prediction of Storms and School of Meteorology, University of Oklahoma, Norman, Oklahoma and Nanjing University, Nanjing, China

[@] Center for Analysis and Prediction of Storms and School of Meteorology, University of Oklahoma, Norman, Oklahoma

[&] NOAA/Earth System Research Laboratory and Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, Colorado

** NOAA/Earth System Research Laboratory, Boulder, Colorado,

⁺⁺ Nanjing University, Nanjing, China

(Manuscript received 30 July 2013, in final form 9 May 2014)

ABSTRACT

A coupled ensemble square root filter–three-dimensional ensemble-variational hybrid (EnSRF–En3DVar) data assimilation (DA) system is developed for the operational Rapid Refresh (RAP) forecasting system. The En3DVar hybrid system employs the extended control variable method, and is built on the NCEP operational gridpoint statistical interpolation (GSI) three-dimensional variational data assimilation (3DVar) framework. It is coupled with an EnSRF system for RAP, which provides ensemble perturbations. Recursive filters (RF) are used to localize ensemble covariance in both horizontal and vertical within the En3DVar. The coupled En3DVar hybrid system is evaluated with 3-h cycles over a 9-day period with active convection. All conventional observations used by operational RAP are included. The En3DVar hybrid system is run at $1/3$ of the operational RAP horizontal resolution or about 40-km grid spacing, and its performance is compared to parallel GSI 3DVar and EnSRF runs using the same datasets and resolution. Short-term forecasts initialized from the 3-hourly analyses are verified against sounding and surface observations. When using equally weighted static and ensemble background error covariances and 40 ensemble members, the En3DVar hybrid system outperforms the corresponding GSI 3DVar and EnSRF. When the recursive filter coefficients are tuned to achieve a similar height-dependent localization as in the EnSRF, the En3DVar results using pure ensemble covariance are close to EnSRF. Two-way coupling between EnSRF and En3DVar did not produce noticeable improvement over one-way coupling. Downscaled precipitation forecast skill on the 13-km RAP grid from the En3DVar hybrid is better than those from GSI 3DVar analyses.

1. Introduction

Three-dimensional variational (3DVar; Lorenc 1986) and four-dimensional variational (4DVar; Le Dimet and Talagrand 1986; Talagrand and Courtier 1987) data assimilation (DA) methods have been used successfully at operational numerical weather prediction (NWP)

centers for more than two decades (e.g., Parrish and Derber 1992; Courtier et al. 1998; Rabier et al. 2000). Typically, static, flow-independent background error covariance (BEC) is used in the background term of the variational cost function. Neglecting the flow-dependent nature of the background error is a key deficiency, especially within a 3DVar framework where the NWP model is not directly used to incorporate model dynamics into the DA system (e.g., Parrish and Derber 1992; Purser et al. 2003a). This deficiency becomes more severe for mesoscale and convective-scale DA where even fewer state variables (cf. the full set) are directly

Corresponding author address: Dr. Ming Xue, Center for Analysis and Prediction of Storms, University of Oklahoma, 120 David L. Boren Blvd., Norman, OK 73072.
E-mail: mxue@ou.edu

observed and large-scale balance relationships, which are often built into 3DVar systems, become invalid (e.g., Gao et al. 2004; Ge et al. 2012). While some efforts had been made to build spatially inhomogeneous, anisotropic BEC into 3DVar frameworks (e.g., Purser et al. 2003b; Wu et al. 2002), major issues exist on how to determine the flow-dependent covariances and how to efficiently introduce them into a variational DA framework.

The ensemble Kalman filter (EnKF) algorithm, as initially developed by Evensen (1994) and Burgers et al. (1998), offers an alternative to the variational formulation. The EnKF employs the Monte Carlo sampling approach, where an ensemble of model forecasts is used to provide and evolve flow-dependent covariances, while the filter updates the ensemble states using an optimal weight through a least squares approach. Many subsequent studies have refined the filter algorithm by addressing a number of issues that are often related to the sampling error associated with the use of relatively small ensembles that is necessitated by practical computational constraints (e.g., Burgers et al. 1998; Houtekamer and Mitchell 1998; Hamill et al. 2001; Anderson 2001; Whitaker and Hamill 2002; Evensen 2003). Because of their ability to estimate flow-dependent BECs and to evolve them through assimilation cycles, and their relative ease of implementation, the ensemble DA methods (Bishop et al. 2001; Anderson 2001; Hunt et al. 2007; Whitaker and Hamill 2002; Evensen 1994) have gained much popularity within both the research and operational communities in recent years. The ensemble filters have been used in operational global forecast systems to provide ensemble-based BEC (e.g., Raynaud et al. 2011; Bonavita et al. 2012; Hamill et al. 2011b; Wang et al. 2013) as well as initial conditions for ensemble forecasts (e.g., Houtekamer et al. 2005; Whitaker et al. 2008; Hamill et al. 2011a). The application of EnKF to mesoscale models has also enjoyed encouraging successes (e.g., Fujita et al. 2007; Meng and Zhang 2007; Bonavita et al. 2008) while for the convective scale, EnKF has shown great ability in dealing with complex, nonlinear physical processes (e.g., Tong and Xue 2005) that may even involve two-moment microphysics parameterization (e.g., Xue et al. 2010; Jung et al. 2012; Putnam et al. 2014). Accurate representation of microphysical processes is especially important at the convective scale.

While EnKF provides a way of estimating flow-dependent BEC, the estimated covariance matrix is severely rank deficient as a result of the much smaller ensemble sizes used compared to the degrees of freedom of typical NWP model state (Houtekamer and Mitchell 1998; Hamill and Snyder 2000). The use of much larger ensembles is often computationally impractical while determining optimal localization that alleviates the rank

deficiency problem in concert with appropriate covariance inflation makes tuning expensive (Anderson 2007, 2012). Another approach that can help alleviate this problem is to combine the full-rank static BEC with the rank-deficient ensemble BEC, creating a so-called hybrid¹ algorithm.

Hamill and Snyder (2000) were the first to propose a 3DVar-based hybrid scheme in which the static BEC in a 3DVar system was replaced by a linear combination of the static and ensemble-derived BEC. The system was tested with a low-resolution quasigeostrophic model and simulated data in a perfect model setting. By running the hybrid analysis system multiple times with perturbed observations, the system is able to provide an ensemble of analyses. It was found that the analysis performs the best when BEC is estimated almost fully from the ensemble, especially when the ensemble size was large (100 in their case). When the ensemble is smaller, the system benefits from a lesser weighting given to the ensemble-based covariances. Wang et al. (2009) also found that a hybrid system based on an ensemble transform Kalman filter (ETKF) is more robust than EnKF for a two-layer primitive equation model when the ensemble size is small and when the model error is large. The hybrid formulation in these studies requires explicit evaluation and storage of the ensemble covariances, which is very expensive for full NWP models.

Lorenc (2003) proposed an elegant, alternative hybrid formulation, in which the control variables of the regular variational cost function are augmented by extended control variables (ECV), which are preconditioned upon the square root of ensemble covariance. The ECV formulation involves adding an additional term to the variational cost function for the ECVs, which has a similar form as the original background term, and is therefore relatively easy to implement based on an existing variational DA framework. Wang et al. (2007) proved that the ECV formulation is mathematically equivalent to that of Hamill and Snyder (2000). The potential for the hybrid system to perform better than a pure EnKF when the ensemble size is relatively small makes it attractive for operational implementation where computational constraint is often a significant issue. A variational framework used by the hybrid scheme also makes it easier to include additional equation constraints in the cost function (e.g., Ge et al. 2012; Kleist et al. 2009b). Furthermore, for observations whose forward operators are nonlocal, such as those of satellite radiance data, the state-space-based covariance localization used in the hybrid formulation is

¹ In this study, we use the word “hybrid” to refer to a combination of the static and ensemble-derived flow-dependent covariances (i.e., the hybrid covariance).

potentially advantageous (Campbell et al. 2010). As suggested by Lorenc (2003) and Buehner et al. (2010a, b), both (traditional) 3DVar and 4DVar can be formulated to use the ensemble covariance with the extended control variable method, and we call such ensemble-variational formulations En3DVar and En4DVar,² respectively, or EnVar in general.

Buehner (2005) implemented the ECV hybrid approach within the Canadian operational global 3DVar framework, and found that the hybrid scheme produced comparable or better forecasts than those initialized using 3DVar. Buehner et al. (2010a, b) further compared the performances of the coupled EnKF–En3DVar and EnKF–En4DVar with the pure 3DVar and 4DVar for global forecasts. Based on the variational DA framework of the Advanced Research Weather Research and Forecasting Model (WRF-ARW; Skamarock et al. 2005), Wang et al. (2008a, b) implemented the ECV-based hybrid, coupling it with an ETKF (Bishop et al. 2001) that is used to update the ensemble perturbations (which we call ETKF–En3DVar hybrid). This WRF hybrid DA system was further applied for tropical cyclone DA (Wang 2011; Li et al. 2012). Most recently, Zhang and Zhang (2012) coupled a mesoscale EnKF system with WRF 4DVar through the WRF hybrid DA framework (hence EnKF–En4DVar hybrid but they called it E4DVar), and Zhang et al. (2013) further compared the performances of EnKF–En3DVar (they called it E3DVar) and EnKF–En4DVar hybrid for mesoscale applications. Mizzi (2012) reported results testing the GSI-based En3DVar hybrid, using ETKF, local ensemble transform Kalman filter (LETKF), and the regular EnKF for ensemble perturbation updating, respectively, and WRF-ARW as the prediction model, for a hurricane period. In general, the introduction of flow-dependent ensemble covariance into 3DVar or 4DVar improves the forecast results. In fact, for the National Centers for Environmental Prediction (NCEP) operational Global Forecast System (GFS), an EnKF–En3DVar hybrid DA system (Whitaker et al. 2011; Hamill et al. 2011b) based on an EnKF and

the operational gridpoint statistical interpolation (GSI) 3DVar (Kleist et al. 2009a) was developed and operationally implemented in 2012, replacing GSI 3DVar. Wang et al. (2013) reported the testing results from the GSI-based En3DVar hybrid system for GFS at a reduced resolution.

It has been a general decision at NCEP that the hybrid DA approach will be applied to its regional models as well, including the North American Mesoscale Model (NAM) and the recently implemented (on 1 May 2012) Rapid Refresh (RAP) system, the replacement to the Rapid Update Cycle (RUC; Benjamin et al. 2004). Toward this end, an EnKF system was recently established for the RAP and tested at a reduced resolution by Zhu et al. (2013, hereafter Z13) using the operational observation data stream of RAP. The same as for the GFS EnKF system, the ensemble square root filter (EnSRF) algorithm of Whitaker and Hamill (2002) was used in Z13. As one of the deterministic ensemble filter algorithms, EnSRF avoids sampling issues associated with the use of “perturbed observations” (Whitaker and Hamill 2002; Tippett et al. 2003).

In Z13, short-range (up to 18 h) forecasts from 3-hourly EnSRF analyses over a 9-day period were found to be consistently better than forecasts from corresponding GSI 3DVar analyses, in terms of both model state forecasts and precipitation forecast skill scores. The primary goal of this current work is to extend the work Z13 by establishing and testing a coupled EnSRF–En3DVar hybrid DA system for RAP that can potentially be implemented operationally. As the first step, we test and evaluate the hybrid DA system running at $1/3$ of the native resolution of operational RAP; running the EnSRF DA system at this reduced resolution is dictated by the limited operational computing resources in the near future. Moreover, running the En3DVar hybrid analyses at the same resolution facilitates easy and direct comparisons with the EnSRF results, and also provides us with a benchmark for a future dual-resolution implementation. With a dual-resolution implementation, the En3DVar analyses will be run at a higher resolution, using the reduced-resolution ensemble perturbations (as is with the operational GFS hybrid DA system). In this paper, we focus on documenting and comparing the results obtained from all three systems (i.e., the GSI 3DVar, EnSRF, and En3DVar hybrid) at the reduced, 40-km grid spacing.

The rest of the paper is organized as follows. The coupled EnSRF–En3DVar hybrid system for RAP is first described in section 2. Experimental setup and testing results are discussed in sections 3 and 4, respectively. Downscaled precipitation forecasts on the 13-km RAP grid, starting from interpolated 40-km En3DVar hybrid,

²Here, En4DVar is an extension of the traditional 4DVar scheme to include the use of ensemble-derived background error covariance through the extended control variable method. The scheme still involves the use of an adjoint model. Liu et al. (2008) proposed an alternative algorithm that does not involve the use of a model adjoint, and En4DVar was used to refer to their algorithm. In Liu and Xiao (2013) their algorithm is renamed 4D_{En}Var, to better differentiate the algorithm from traditional 4DVar. Our current usage is also consistent with the conventions used by papers at the recent WMO Data Assimilation Symposium of Maryland, held in October 2013.

EnSRF, and GSI 3DVar analyses, are compared in [section 5](#). Finally, [section 6](#) provides conclusions and an additional discussion.

2. GSI-based EnSRF–En3DVar hybrid system for Rapid Refresh

a. The Rapid Refresh system

The operational hourly-updated RUC system was designed to improve short-range weather forecasting through frequent updating of initial conditions with the latest observations ([Benjamin et al. 2004](#)). The RAP is a replacement of the RUC system and is based on the nonhydrostatic WRF-ARW dynamic core ([Skamarock et al. 2005](#)). RAP became operational at NCEP on 1 May 2012 using the GSI 3DVar as the data assimilation system. In February 2014, the system was upgraded to incorporate BEC derived from ensemble perturbations from the 80-member GFS EnSRF system that feeds the operational hybrid En3DVar DA system of the GFS model. The GSI is a unified DA framework for both global and regional models ([Kleist et al. 2009a](#)). The horizontal grid spacing of RAP is ~ 13 km and has 50 vertical levels extending up to 10 hPa at the model top. Compared to the RUC, the RAP system is capable of assimilating more observations, including satellite radiance data, and has a larger domain which covers North America in its entirety. The physics options used by the operational RAP include the Grell-G3 cumulus parameterization, Thompson microphysics, the Rapid Radiative Transfer Model (RRTM) longwave radiation, Goddard shortwave radiation, Mellor–Yamada–Janjić (MYJ) turbulent mixing, and the RUC–Smirnova land surface model. Details on these schemes can be found in [Benjamin et al. \(2009\)](#).

As with the RUC, the RAP employs a digital filter initialization (DFI) to reduce high-frequency noise during the initial period of model integration. In the operational RAP system, twice DFI (TDFI; [Lynch and Huang 1992](#)), which applies the DFI twice, once on the adiabatic backward time integration and once on the full-physics forward time integration, is used. Considering that for high-resolution applications where diabatic processes are more important, adiabatic integration can introduce significant errors, [Z13](#) chose to employ the digital filter launching (DFL) procedure ([Lynch and Huang 1994](#)) instead in their EnSRF system for RAP. DFL applies the DFI only once, on the forward integration time series. In this study, the same procedure is followed by the EnSRF and En3DVar hybrid experiments. In our tests with 3-hourly cycles reported in this paper, the DFL employs a 40-min filter window centered at 20 min of

forecast time, and used a Dolph filter ([Lynch 1997](#)) with a cutoff half-width of 20 min.

b. The coupled EnSRF–En3DVar hybrid system for RAP

As mentioned earlier, our En3DVar hybrid system is based on the operational GSI 3DVar system for RAP and it uses the operational data stream of RAP. To facilitate direct comparisons with the RAP EnSRF and GSI 3DVar systems as reported in [Z13](#), we run our hybrid tests also at the reduced resolution of ~ 40 -km grid spacing with 3-hourly assimilation cycles instead of the ~ 13 -km grid spacing and hourly cycles of the operational RAP. The use of the reduced-resolution EnSRF system is due to the expected constraint in available operational computational resources. The choice of 3-hourly cycles is to enable us to run a larger number of experiments and for more rapid prototyping of the system. The running of the continuously cycled experiments over a 9-day period is computationally expensive in terms of both CPU and storage requirements. Extensive experimentation and tuning were required to arrive at quasi-optimal configurations of the RAP EnSRF system, including configurations of covariance inflation and localization. For future operational implementation, it is desirable to run the En3DVar at the native RAP resolution, while using lower-resolution EnSRF perturbations in a dual-resolution model to save computational cost. The implementation and testing of the dual-resolution coupled hybrid system for RAP will be done in the future.

A one-way coupled EnSRF–En3DVar hybrid system is made up of four key steps: 1) GSI-based observation processing that includes both quality control and calculation of a full set of observation innovations; 2) EnSRF analyses using the innovations calculated by the GSI and the background ensemble forecasts to yield an ensemble of analyses; 3) an En3DVar analysis using the background ensemble forecasts from the EnSRF cycle for flow-dependent covariance estimation; and 4) carrying out ensemble forecasts from the EnSRF ensemble analyses and a single control forecast from the En3DVar hybrid analysis to the next analysis time.

[Figure 1](#) shows a flowchart for both one-way and two-way coupled EnSRF–En3DVar analysis–forecast cycle as employed in this paper. For one-way-coupled En3DVar, the EnSRF system provides the background ensemble forecast perturbations to the ECV-based En3DVar hybrid variational analysis, but does not recenter the EnSRF analyses on the En3DVar analysis. Two-way coupling includes an additional step that recenters the EnSRF analysis ensemble on the En3DVar control analysis (the thick black arrows and bold black box in [Fig. 1](#)). The

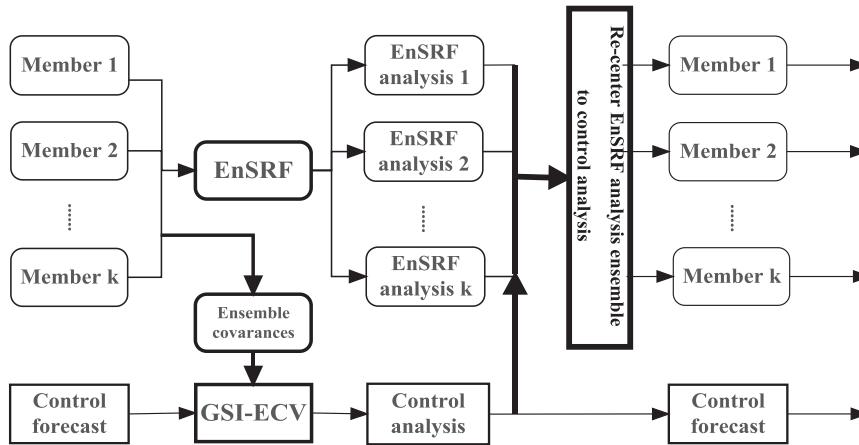


FIG. 1. Flowchart of a full GSI-based EnSRF–En3DVar hybrid data assimilation cycle, with one-way or two-way coupling between the EnSRF (top portion) and En3DVar hybrid control analysis (bottom portion denoted En3DVar). The thick upward-pointing arrow indicates the feedback of the En3DVar hybrid analysis to the EnSRF in the two-way coupling procedure, when the En3DVar hybrid control analysis is used to replace the ensemble mean of the EnSRF analyses.

two-way coupling implicitly assumes that the En3DVar control analysis is better than the EnSRF ensemble mean analysis, and the recentering should help prevent the divergence between the EnSRF and En3DVar analyses so that the ensemble perturbations can sample the control forecast uncertainty well. Divergence between the two systems tends to become more serious when continuous cycles are run for a long period of time.

As pointed out earlier, the GSI-based En3DVar hybrid analysis is achieved using the ECV method (Wang 2010). Within this framework, the analysis increment $\delta\mathbf{x}$ is a sum of two terms, defined as

$$\delta\mathbf{x} = \delta\mathbf{x}_1 + \sum_{k=1}^K (\mathbf{a}_k \circ \mathbf{x}'_k), \quad (1)$$

where $\delta\mathbf{x}_1$ is the analysis increment associated with static BEC \mathbf{B} and the second term on the right-hand side is the increment associated with the ensemble covariance; \mathbf{x}'_k is the k th ensemble background perturbation normalized by $\sqrt{K-1}$, where K is the ensemble size. Vectors \mathbf{a}_k ($k = 1, \dots, K$) in the second term are the extended control variables. The analysis increment $\delta\mathbf{x}$ is obtained by minimizing the following cost function:

$$\begin{aligned} J(\delta\mathbf{x}_1, \mathbf{a}) &= \beta_1 J_b + \beta_2 J_e + J_o \\ &= \frac{1}{2} \beta_1 \delta\mathbf{x}_1^T \mathbf{B}^{-1} \delta\mathbf{x}_1 + \frac{1}{2} \beta_2 \mathbf{a}^T \mathbf{A}^{-1} \mathbf{a} \\ &\quad + \frac{1}{2} [\mathbf{y}_o - H(\mathbf{x}_b + \delta\mathbf{x})]^T \mathbf{R}^{-1} [\mathbf{y}_o - H(\mathbf{x}_b + \delta\mathbf{x})], \end{aligned} \quad (2)$$

which gives the solutions of partial increment $\delta\mathbf{x}_1$ and ECV \mathbf{a} . Vector \mathbf{a} is formed by concatenating K vectors \mathbf{a}_k . Compared to a traditional 3DVar cost function, a weighted sum of J_b and J_o is replaced by the sum of weighted J_b and J_e terms and J_o , where J_b is the traditional background term associated with static covariance \mathbf{B} , J_o is the observation term as in traditional 3DVar, and J_e is the additional term associated with flow-dependent covariance for the ECV. Weighting factors β_1 and β_2 are placed in front of J_b and J_e terms, respectively, and they are constrained by

$$\frac{1}{\beta_1} + \frac{1}{\beta_2} = 1 \quad (3)$$

to conserve the total variances in current GSI hybrid implementation.

The ECVs are constrained by a block-diagonal matrix \mathbf{A} , which defines the ensemble covariance localization (Lorenz 2003; Wang et al. 2007). In the GSI-based En3DVar hybrid implementation, the horizontal and vertical covariance localizations, or the effects of matrix \mathbf{A} in Eq. (2), are achieved by applying recursive filter transforms (Hayden and Purser 1995), analogous to the treatment of \mathbf{B} in Eq. (2). The parameters in the recursive filter will determine the correlation length scale in \mathbf{A} as a precondition and, therefore, prescribe the covariance localization length scale for the ensemble covariance. The vertical covariance localization scale (CLS) is measured in either scaled height (the natural log of pressure) or the number of model levels while the horizontal CLS is measured either in kilometers or number of

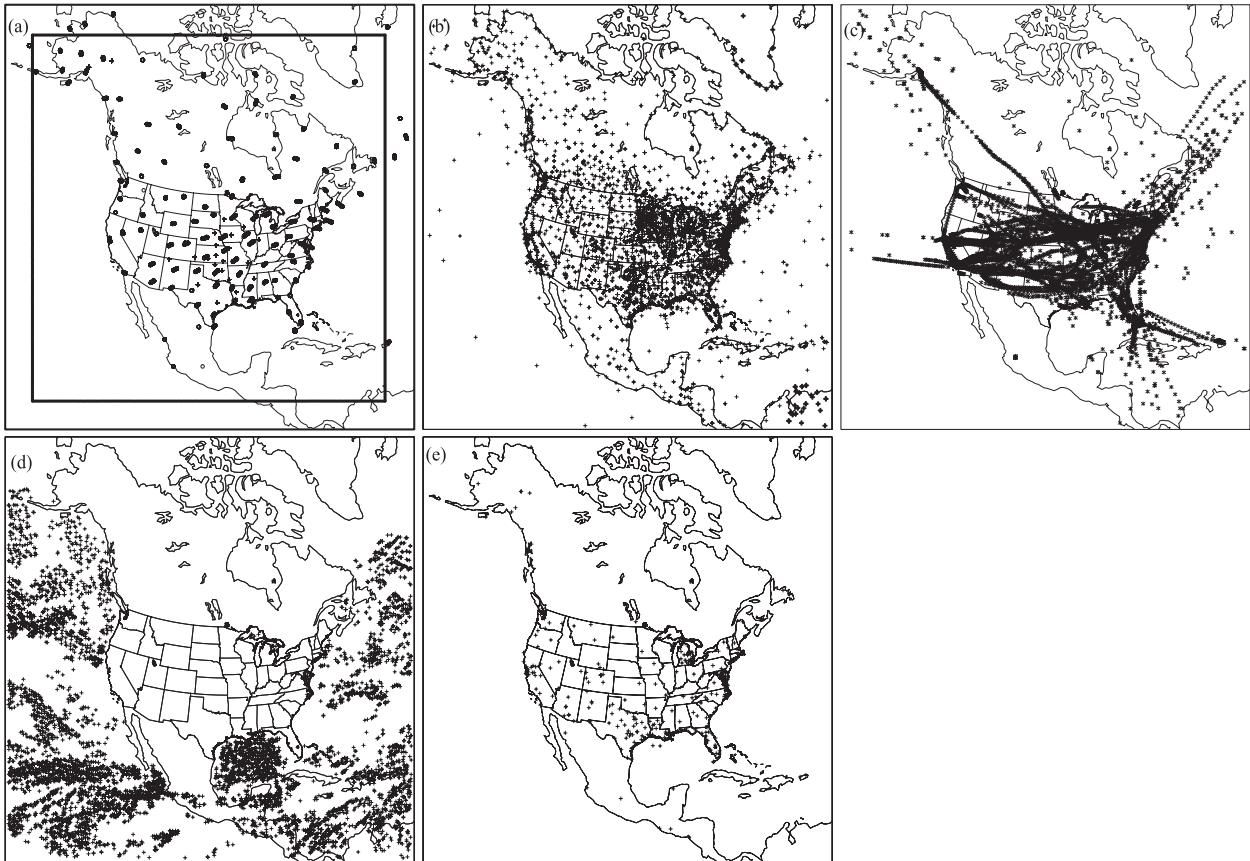


FIG. 2. Example of the horizontal distributions of observation at 0000 UTC 14 May: (a) sounding (circles) and profile (pluses); (b) surface stations over land and for ships, buoys, etc.; (c) aircraft observations; (d) satellite retrieval winds; and (e) GPS precipitable water (PW) data. The small box in (a) is the domain used by 13-km forecasts. (a)–(d) Adopted from Z13.

grid points in GSI. In this study, the natural log of pressure is used for the vertical, and kilometer is used for the horizontal localization.

Apart from the variational minimization of the En3DVar hybrid cost function given by Eq. (2), a major component of the overall coupled EnSRF–En3DVar hybrid DA system is an ensemble DA system that provides the perturbations. In our case, the EnSRF system used is the one described in Z13. The configuration settings follow the control experiment of Z13. To facilitate fair comparisons between the En3DVar hybrid and EnSRF experiments, the CLSs in the En3DVar hybrid system are specified to match the CLSs used by the EnSRF as closely as possible in the control experiments, and the vertical and horizontal scales are measured in the natural log of pressure and kilometers, respectively. The e -folding distance from the Gaspari and Cohn (1999) localization function is $\sqrt{2}\sqrt{0.3}S_{GC}$ (where S_{GC} is cutoff radii in the EnSRF), while an e -folding distance from the recursive filter is $2\sqrt{2}S_{RF}$ (Barker et al. 2004; Wang et al. 2008b) (where S_{RF} is recursive filter

covariance localization length scale). Thus, to keep the same e -folding distance for both EnSRF and En3DVar, the cutoff radii in the EnSRF S_{GC} can be converted to the recursive filter localization length scale S_{RF} in hybrid according to

$$S_{RF} = \sqrt{0.15}S_{GC}/\sqrt{2}. \quad (4)$$

3. Experiment designs

a. Model, observations, ensemble configuration, and verification techniques

The test period, model domains, and boundary conditions used in this study are the same as in Z13. The DA experiments at ~ 40 -km grid spacing are run in continuous 3-hourly cycles throughout the 9-day retrospective testing period from 8 to 16 May 2010; the cycles start at 0000 UTC 8 May 2010 and end at 2100 UTC 16 May 2010. The 40-km model domain (as shown in Fig. 2) covers North America with 207×207 grid points.

A slightly smaller domain at ~ 13 -km grid spacing, as indicated by the bold rectangle in Fig. 2a, is used for forecasts at the native RAP resolution and for precipitation verification. The domains have 50 vertical levels. The 18-h deterministic forecasts (after applying DFL) are launched every 3 h from the En3DVar hybrid control analyses as well as EnSRF ensemble mean analyses on the 40-km domain. The 3-hourly ensemble forecasts are produced within the assimilation cycles of EnSRF, which are fed into the En3DVar control analysis (Fig. 1). Two outer loops and 50 iterations, the same as in operational RAP GSI 3DVar, were utilized for all the En3DVar and GSI 3DVar experiments. The 13-km deterministic forecasts start from interpolated 40-km analyses at 0000 and 1200 UTC for precipitation forecast evaluation. The lateral boundary conditions for both grids come from operational GFS forecasts. Perturbations created using the random-CV method in the WRF 3DVar (Barker et al. 2004) are added to GFS forecast boundary conditions for the ensemble forecasts and to the GFS analysis initial condition at 0000 UTC 8 May 2010 to start the initial ensemble of EnSRF.

The observations used in this study are the same as those used in the operational RAP except for the exclusion of satellite radiance data. The real-time RAP system collects data from 1.5 h before and 0.5 h after the time of analysis. However, for 0000 and 1200 UTC it waits half an hour longer for more data (such as sounding data) to arrive. In our tests, the datasets assimilated at 3-hourly intervals are the datasets collected and used by the operational hourly RAP system; as a result, observations that arrived in real time outside the 2-h (2.5 h for 0000 and 1200 UTC) windows are not used. They include surface observations (land reporting stations, mesonets, ships, buoys, etc.), upper-air observations [radiosondes, aircrafts, wind profilers, velocity–azimuth display (VAD) data, and satellite retrieval winds], and GPS precipitable water (PW), the same as in Z13 except for the exclusion of PW data there. The exclusion of the PW data in Z13 was due to an initial problem with the EnSRF code, which has since been fixed. The distributions of most major observation types are shown in Fig. 2. The satellite radiance data are not included in the experiments reported here because our preliminary tests suggested that the bias correction remains an important issue within the system that would require careful treatment for positive impacts. Our most recent tests with the radiance data using the EnSRF show small positive impacts, and the results will be reported separately in the future. Initial studies of EnKF for NCEP GFS global model also excluded satellite radiance data (Whitaker et al. 2008).

The short-range deterministic forecasts from the En3DVar, EnSRF ensemble mean, and GSI 3DVar analyses are verified against surface and sounding observations. The Model Evaluation Tools (MET) developed by the Developmental Testbed Center (DTC; Brown et al. 2009) are employed here. MET contains comprehensive verification metrics for both deterministic and probabilistic forecasts. Root-mean-square error (RMSE) is used as the primary verification metric for the 40-km deterministic forecasts here. The RMSEs for temperature T , relative humidity RH, and wind components U and V are calculated against upper air soundings, and those for surface pressure P , 2-m RH, 2-m T , and 10-m U and V are calculated against surface observations.

The statistical significance of RMSEs is determined by using bootstrap resampling (Candille et al. 2007; Buehner and Mahidjiba 2010; Schwartz and Liu 2014). The RMSEs from all cycles are randomly selected 3000 times, and for these samples, the mean is calculated, along with a two-tailed 90% confidence interval from 5% to 95%. To determine whether the improvements from En3DVar on GSI 3DVar are statistically significant, the mean RMSE differences between En3DVar and GSI 3DVar together with a 90% confidence interval are computed and plotted in each figure. The RMSE differences from all cycles are also randomly selected 3000 times, and for these samples, a two-tailed 90% confidence interval from 5% to 95% is calculated. The same technique is also applied to the differences between En3DVar experiments and EnSRF_Ctl to determine whether the improvement of En3DVar over EnSRF is statistically significant. That the bounds of a 90% confidence interval between the forecast pair are all lower than zero means RMSEs from the first experiment are always lower than the second one at the 90% confidence level; therefore, the improvement from the first experiment over the second one is statistically significant at the 90% confidence level. Conversely, that zero is included within the bounds of the 90% confidence level denotes statistically insignificant situations (Schwartz and Liu 2014; Xue et al. 2013).

For the 12-hourly forecasts on the 13-km grid, the Gilbert skill score (GSS; Gandin and Murphy 1992), also known as the equitable threat score (ETS), and frequency bias (BIAS) are used to verify precipitation forecasts against NCEP stage-IV precipitation data (Lin and Mitchell 2005). The error and skill scores are aggregated over all forecasts within the 9-day test period. The same evaluation procedure was used in Z13 although they only presented the GSSs.

b. Assimilation experiments

Experiments performed in this study are listed in Table 1. First, well-tuned En3DVar hybrid one-way

TABLE 1. List of data assimilation experiments. In the horizontal and vertical localization columns, the diagonal arrow (\nearrow) means increasing with height.

Expt group	Expt (including alternative names)	Ensemble covariance weighting factor ($1/\beta_2$)	Horizontal cutoff radius for hybrid/EnSRF (km)	Vertical cutoff radius for hybrid/EnSRF in $\ln(p)$	EnSRF–En3DVar coupling
Control expt	GSI3DVar	—	700 \nearrow 1050	—	—
	EnSRF_Ctl	—	700 \nearrow 1050	RH and T: 1.1/4 \nearrow 1.1/2 U and V: 1.1/2 \nearrow 1.1 PS and PW: 1.6	—
	Hybrid1W_Ctl /Hybrid05	0.5	~1095	1.1	One way
Sensitivity expt on covariance weighting factors	Hybrid2W_Ctl	0.5	~1095	1.1	Two way
	Hybrid01	0.1	~1095	1.1	One way
	Hybrid09	0.9	~1095	1.1	One way
Sensitivity expt on localization scales	Hybrid10	1.0	~1095	1.1	One way
	Hybrid_HS	0.5	~701	1.1	One way
	Hybrid_HL	0.5	~1300	1.1	One way
Sensitivity expt on height- and observation-type-dependent localization scales	Hybrid_VS	0.5	~1095	0.36	One way
	Hybrid_VL	0.5	~1095	1.8	One way
	EnSRF_Con	—	1095	1.1	—
	Hybrid_Con	1.0	~1095	1.1	One way
	(Perturbations from EnSRF_CON)				
	Hybrid_HD	1.0	700 \nearrow 1050	1.1/2 \nearrow 1.1	One way
Sensitivity expt on localization scales	Hybrid3G	1.0	700 \nearrow 1050	RH and T: 1.1/4 \nearrow 1.1/2 U and V: 1.1/2 \nearrow 1.1 PS and PW: 1.6 (observations are assimilated in 3 groups)	One way
	EnSRF3G	—	700 \nearrow 1050	RH and T: 1.1/4 \nearrow 1.1/2 U and V: 1.1/2 \nearrow 1.1 PS and PW: 1.6 (observations are assimilated in 3 groups)	—

(Hybrid1W_Ctl) and two-way coupled (Hybrid2W_Ctl), EnSRF (EnSRF_Ctl) control, and GSI 3DVar (GSI3DVar) experiments are compared. The EnSRF control experiment, EnSRF_Ctl, uses 40 ensemble members and corresponds to experiment EnKF_CtrHDL from Z13 except for additional GPS PW data in this study, and uses a single suite of physics parameterizations in the ensemble to keep the setup simple (so that the EnSRF, GSI 3DVar, and the En3DVar experiments all use the same set of physics in the forecast model). The En3DVar hybrid control experiment assigns equal weights ($1/\beta_2 = 0.5$) to the static and ensemble BECs. The EnSRF codes and configurations are the same as the EnSRF control experiment in Z13, except for the exclusion of GPS PW data there. A combination of static and adaptive covariance inflation is applied in EnSRF as in Z13.

There are mainly two sets of tunable parameters in the En3DVar hybrid scheme. One set is the covariance

weighting factors, which define the weights placed on the BECs. Four sensitivity experiments test the relative weights given to the static and ensemble BECs, with $1/\beta_2 = 0.1, 0.5, 0.9,$ and 1.0 (Hybrid01, Hybrid05, Hybrid09, and Hybrid10) corresponding to 10%, 50%, 90%, and 100% weight given to the ensemble BEC, respectively.

The other set of tunable parameters includes the horizontal and vertical CLSs applied to the covariances. For weighting factor $1/\beta_2 = 0.5$ with one-way coupling, we test three horizontal CLSs $S_h = 192, 300,$ and 356 km in Hybrid_HS, Hybrid1W_Ctl, and Hybrid_HL, respectively [corresponding to cutoff radii of 700, 1095, and 1300 km according to Eq. (4)]; three vertical CLSs $S_v = -0.1, -0.3,$ and -0.5 [corresponding to cutoff radii of 0.36, 1.1, and 1.8 according to Eq. (4)] are tested in Hybrid_VS, Hybrid1W_Ctl, and Hybrid_VL, respectively. The minus sign is due to the use of $\ln(p)$ as the length measure. To facilitate the comparison with control

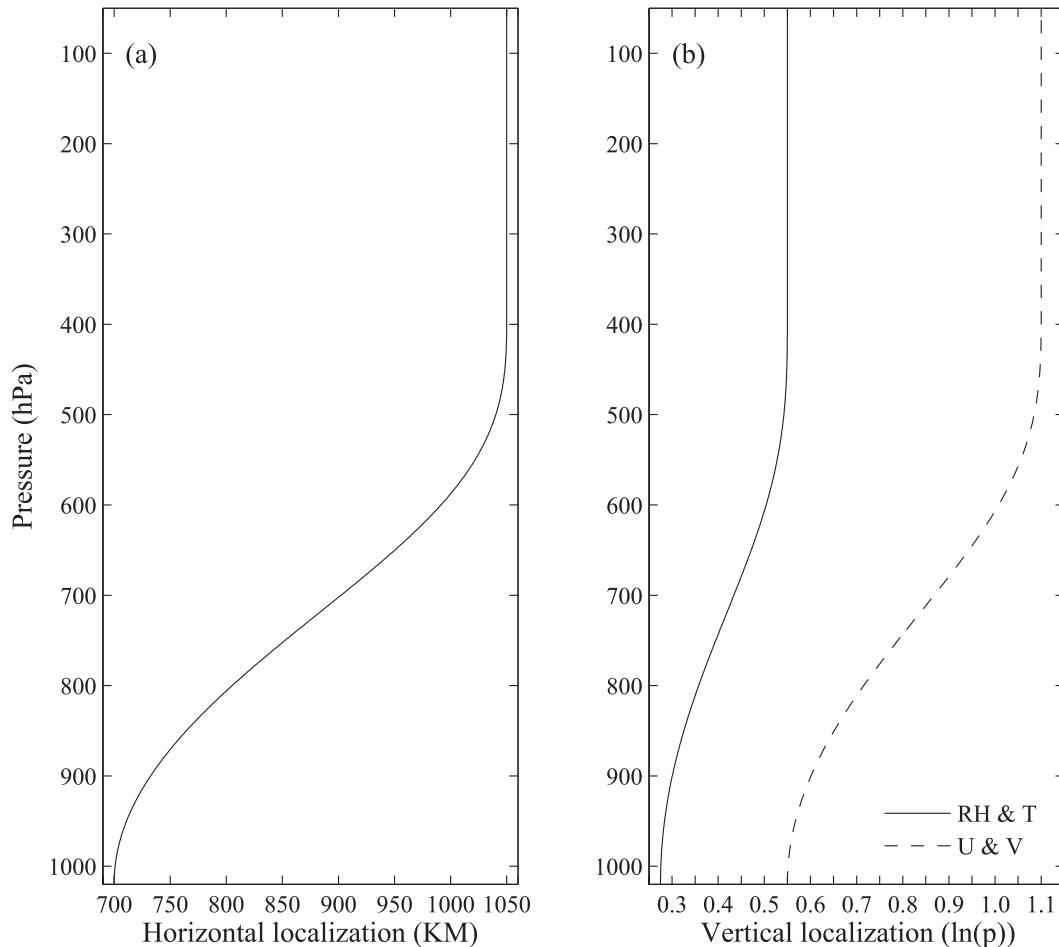


FIG. 3. Profiles of (a) horizontal and (b) vertical cutoff radii for the EnSRF control experiment. The horizontal axis is the cutoff radius of a given observation at a particular vertical position given in pressure. The vertical axis is the vertical position of observations given by the pressure.

experiment Hybrid1W_Ctl, the mean domain-average RMSE difference, defined as

$$D = \frac{1}{N} \sum_{k=1}^N (\text{RMSE}_{\text{Hybrid}^*}^k - \text{RMSE}_{\text{Benchmark}}^k), \quad (5)$$

where N is the total number of cycles and k refers to the k th cycle, is calculated between experiment Hybrid*; the benchmark experiment is Hybrid1W_Ctl here and Hybrid* refers to one of En3DVar sensitivity experiments.

All CLSs used in the En3DVar hybrid experiments described above are constant with height. However, the cutoff radii used in the well-tuned EnSRF control experiment of Z13 (EnKF_CtrHDL in their paper) are height- and observation-type dependent based on the vertical position of the observations. These localization settings are shown in Fig. 3. The horizontal cutoff radius r_{cut} at the model top is 1.5 times the value at the surface

for all state variables; as shown in Fig. 3a, r_{cut} increases from 700 km at the surface to 1050 km at the model top. The vertical cutoff radius $\ln(p_{\text{cut}})$ is not only height dependent, but also observation-type dependent. For RH and T observations (solid line in Fig. 3b), the vertical cutoff radii at the model top and surface are set to a quarter of 1.1 and half of 1.1, respectively. For wind observations (dashed line in Fig. 3b), $\ln(p_{\text{cut}})$ is twice as large as that for RH and T observations. For surface pressure observations and GPS PW data (which are most strongly linked to low-level moisture), their vertical localization radii are set to a constant value of 1.6. These settings were used in the control experiment of Z13, and their choices were guided by the correlation scales found in the NMC-method-derived error statistics used by GSI 3DVar and were further tuned based on sensitivity experiments.

In the En3DVar system, height-dependent localization is straightforward to implement, but not

observation-type-dependent localization, because unlike the serial EnSRF scheme, En3DVar analyzes all observations simultaneously and the localization is performed in the state instead of the observation space (Campbell et al. 2010). Theoretically, if the localization treatment was the same for the EnSRF ensemble mean analysis as for the En3DVar analysis and when the ensemble-derived covariance is used at 100%, the results from the two algorithms should be very close. We observed differences between such EnSRF and En3DVar analyses in our experiments, and want to see if localization is the main cause for these differences. We are interested in finding out if the height- and observation-dependent covariance localization treatments would potentially improve the performance of En3DVar as in EnSRF. These are examined in the next four experiments (EnSRF_Con, Hybrid_Con, Hybrid_HD, and Hybrid3G), all performed with 100% ensemble covariance and all used one-way coupling.

EnSRF_Con and Hybrid_Con use constant horizontal and vertical localization radii that are close to the corresponding cutoff radii used by EnSRF_Ctl at the model top. Unlike the other En3DVar experiments, the ensemble perturbations for Hybrid_Con were provided by EnSRF_Con instead of EnSRF_Ctl for consistency.

In Hybrid_HD, the height-dependent horizontal CLSs are chosen to match the height-dependent cutoff radii of EnSRF_Ctl closely, while the vertical CLSs for all variables are chosen to be the same as that for wind observations in EnSRF_Ctl (Table 1).

The only way to apply different localization to different observations in En3DVar is to break the analysis into multiple steps, with each step analyzing a subset or a subgroup of observations. To do this, the corresponding EnSRF analysis that provides the ensemble perturbations also needs to be broken up into multiple steps and the EnSRF and En3DVar need to be run in alternating order. Doing so significantly increases the overall computational costs for operational implementation, since the disk I/O associated with the reading and writing of ensembles and with the cost function minimizations are done multiple times per analysis cycle, but is doable in a research mode. Toward this end, experiments EnSRF3G and Hybrid3G are run, where each analysis is broken into three steps, with each step analyzing one of the three groups of observations consisting of 1) RH and T , 2) U and V , and 3) and PS and GPS PW data, respectively. Within each step, the EnSRF ensemble analysis is followed by an En3DVar hybrid analysis step using the latest EnSRF-updated ensemble perturbations.

Because the EnSRF includes both static and adaptive covariance inflation (Z13), it is difficult to maintain the

TABLE 2. List of mean domain average absolute RMSE difference pairs.

Name	Hybrid*	Benchmark
1GC1	Hybrid_Con	EnSRF_Ctl
1GC2	Hybrid_Con	EnSRF_Con
1GHD	Hybrid_HD	EnSRF_Ctl
3GHD	Hybrid3G	EnSRF3G

same amount and effects of inflation when each EnSRF analysis is broken into three steps. Applying the static inflation every EnSRF substep can overinflate the covariance, while applying it only in the last step would change the overall behavior of the filter. Because our primary goal here is to determine if the difference between the EnSRF and En3DVar analyses (with 100% ensemble covariance) is primarily caused by the observation-based localization, to avoid the above issue, we run EnSRF3G without any covariance inflation and examine the RMSE differences between the EnSRF and En3DVar analyses. We just need to find out if the En3DVar hybrid analyses are closer to the EnSRF analyses when observation-type-dependent localization is similarly used in the En3DVar through the split-step procedure.

The mean domain average absolute RMSE difference, defined as

$$DB = \frac{1}{N} \sum_{k=1}^N |\text{RMSE}_{\text{Hybrid}^*}^k - \text{RMSE}_{\text{Benchmark}}^k|, \quad (6)$$

is used to measure how close the En3DVar and EnSRF analyses are. The differences between Hybrid_Con and EnSRF_Con, Hybrid_Con and EnSRF_Ctl, Hybrid_HD and EnSRF_Ctl, and Hybrid3G and EnSRF3G (Table 2) will be calculated to examine the impacts of constant localization, height-dependent localization, and observation-dependent localization, respectively. The statistical significance of DB is also determined by using bootstrap resampling. The DBs at cycles are randomly selected 3000 times. For this sample, a mean is calculated, along with a two-tailed 90% confidence interval from 5% to 95%. If the error bars from the experiments pair do not overlap, the differences between En3DVar and EnSRF are significantly reduced at the 90% confidence level.

4. Results of experiments

a. GSI 3DVar, EnSRF, and En3DVar hybrid control experiments

The RAP system had been run experimentally in real time for several years at the National Oceanic and Atmospheric Administration (NOAA) Earth System Research

Laboratory (ESRL) before being officially implemented at NCEP in May 2012. In this study, we borrow from a recent configuration of the experimental 13-km RAP for our 40-km grid spacing tests.

In this section, we present and compare the results from the En3DVar hybrid one-way (Hybrid1W_Ctl) and two-way coupled (Hybrid2W_Ctl), EnSRF (EnSRF_Ctl) control experiments, and those of the GSI 3DVar (GSI3DVar) experiment. Similar to Z13, single observation tests were first performed to examine the general behaviors of the En3DVar system and compared to the EnSRF results. The tests used the analysis of EnSRF_Ctl after 5 days of 3-hourly cycles as background and the analysis increments appear reasonable. To save space, the results are not shown here.

The RMSEs of the 3-h forecasts at different height levels verified against sounding data are shown in Fig. 4. These forecasts were launched from the GSI 3DVar, EnSRF ensemble mean, and En3DVar hybrid analyses. The RMSE for each pressure level was calculated by averaging values obtained from all cycles within a layer 50 hPa above and below that pressure, except for the lowest and topmost levels. The RMSEs of EnSRF_Ctl are overall lower than those of GSI 3DVar except for the temperature at the upper levels where the error can be ~ 0.1 K greater. The performances of one-way and two-way coupled En3DVar hybrid schemes are very close. With half static and half flow-dependent covariances in these experiments, Hybrid1W_Ctl and Hybrid2W_Ctl outperform GSI 3DVar, and are also generally better than EnSRF_Ctl except for RH above 500 hPa, V at 100 hPa, and T below 900 hPa.

The average RMSEs for all levels over the entire domain are shown in Fig. 5 for forecast hours 3–18. Generally, both EnSRF and En3DVar hybrid significantly outperform GSI 3DVar for all the variables throughout the forecast period at the 90% confidence level (the intervals of error differences do not include zero). For RH, the average RMSEs of En3DVar hybrid are slightly higher than those of EnSRF_Ctl by 9 h, which appears to be related to the larger errors at 3 h at the upper levels (Fig. 4a). The RMSEs become slightly smaller after 9 h. However, the improvement of En3DVar hybrid over EnSRF for RH is not statistically significant. For T and U , the domain-averaged RMSEs of En3DVar hybrid are significantly and consistently smaller than those of GSI 3DVar and EnSRF throughout the forecast period (Fig. 5). For V , the errors of the En3DVar and EnSRF are very similar and are all clearly lower than those of GSI 3DVar. The reason that En3DVar performs better than EnSRF for U may relate to the dominance of the east–west flows that may increase the validity of the static covariance. Overall, the

En3DVar hybrid outperforms GSI 3DVar and EnSRF for T and V for the 18 h of the forecast.

Figure 6 shows the average RMSEs for 3–18-h forecasts against surface observations. For 2-m T and 10-m U , the EnSRF and En3DVar outperform the GSI 3DVar at all forecast hours significantly, with the EnSRF significantly outperforming the En3DVar hybrid at most forecast hours. For 2-m RH and 10-m V , EnSRF occasionally underperforms GSI 3DVar slightly, but at most forecast hours it is better. The En3DVar hybrid schemes improve over EnSRF further, enough to ensure better or equal performance than GSI 3DVar for all hours, and more clearly so for RH. For surface pressure, EnSRF underperforms GSI 3DVar initially, but becomes better after 9 h; throughout the forecast period, the En3DVar hybrid outperforms both GSI 3DVar and EnSRF significantly. In general, there is little difference between the one-way and two-way En3DVar hybrid schemes. If the cycles were run for a much longer time period, a larger divergence between the EnSRF and En3DVar hybrid may develop in a one-way coupling mode. In that case, two-way coupling would then show a bigger advantage. When the En3DVar hybrid runs at a higher resolution than the EnSRF in a dual-resolution mode, there may also be more of a benefit from two-way coupling.

Overall, the En3DVar hybrid schemes significantly outperform GSI 3DVar for all the variables at all forecast hours for sounding and surface observations. Compared to EnSRF, their performances are comparable, or even better for some variables. The results indicate the benefit of combining the static and flow-dependent covariances. In the next section, the sensitivity to the covariance weighting factors is examined.

Finally, one may have concern that the 9-day cycled assimilation period is not long enough for the ensemble DA system to spin up (over the course of evaluating and testing our EnSRF and En3DVar hybrid systems, we had run over 100 cycled experiments so extending the experiment period would be expensive). To answer this question, we examine how the short-range forecast errors evolve through the 9-day period. Figure 7 shows the domain-averaged 3-h forecast RMSEs verified against sounding data at 0000 and 1200 UTC through the test period. We can see that the relative performances of GSI3DVar, EnSRF, and En3DVar hybrid do not change much throughout the 9-day period, even in the earlier days. These results indicate that the ensemble system had spun up rather quickly.

b. Sensitivity to covariance weighting factors

We perform a set of four 1-way-coupled En3DVar hybrid experiments with $1/\beta_2 = 0.1, 0.5, 0.9, \text{ and } 1.0$,

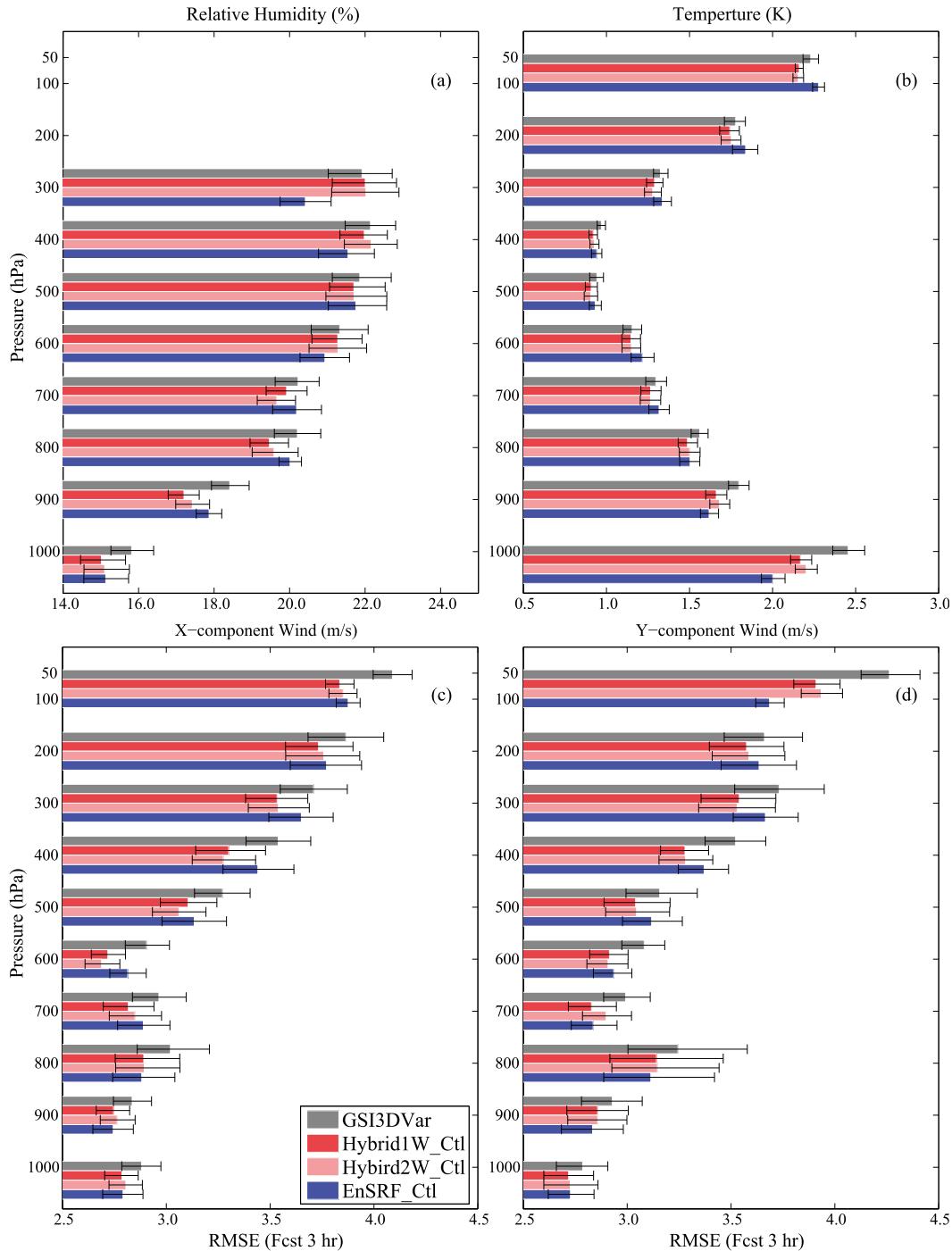


FIG. 4. Mean 3-h forecast RMSEs at different height levels verified against sounding data for (a) RH, (b) T , (c) U , and (d) V for the labeled experiments. Error bars represent the two-tailed 90% confidence interval [(left) 5% and (right) 95%] using the bootstrap distribution method.

which are the weights given to the ensemble covariance. The one with $1/\beta_2 = 0.5$, called Hybrid05 here, is the same as experiment Hybrid1W_Ctl discussed earlier (Table 1). The 3-h forecast RMSEs at different height

levels verified against sounding data are shown in Fig. 8 for these four experiments, GSI 3DVar, and EnSRF_Ctl. It can be seen that the En3DVar hybrid and EnSRF schemes generally outperform GSI 3DVar,

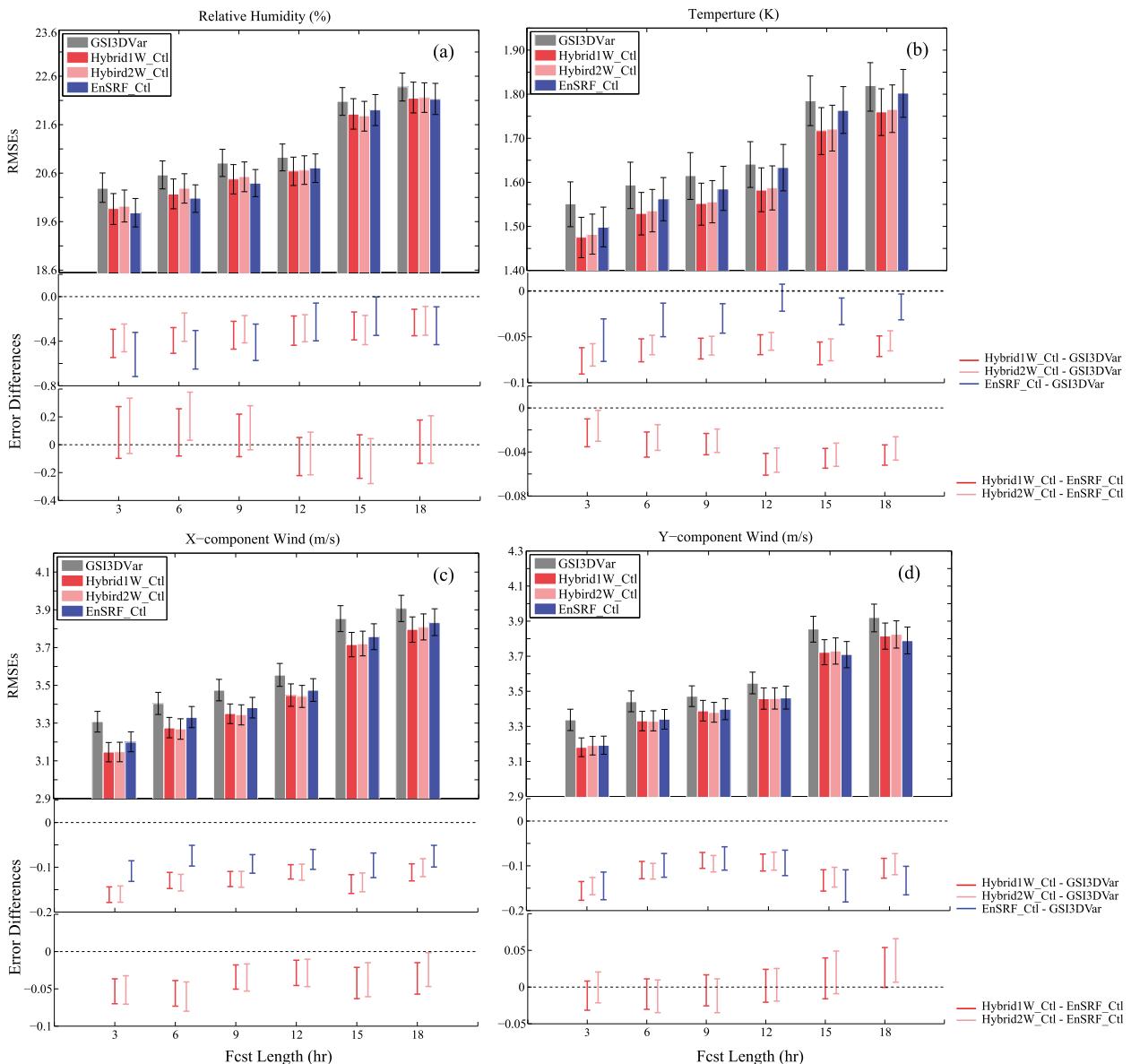


FIG. 5. The 9-day and domain-averaged forecast RMSEs verified against sounding data (top part of each panel) and the 90% confidence interval of the RMSE differences between En3DVar hybrid experiments and GSI 3DVar/EnSRF_Ctl (bottom part of each panel) for (a) RH, (b) T , (c) U , and (d) V , for different forecast hours.

except for RH from 700 to 400 levels for Hybrid09 and Hybrid10 (i.e., the hybrid scheme with 90% or 100% ensemble covariances). Introducing 10% static covariance into the En3DVar hybrid framework reduces the error slightly (comparing Hybrid09 to Hybrid10 for RH), and further increasing it to 50% brings the RH errors below those of GSI 3DVar at all levels (Fig. 8a).

The average RMSEs for all levels over the entire domain are shown in Fig. 9 for forecast hours 3–18. All En3DVar hybrid experiments significantly outperform GSI 3DVar for all variables throughout the entire

forecast period at the 90% confidence level, except for the RH of Hybrid10 after 9 h. The errors of Hybrid05 are about the lowest among all En3DVar hybrid experiments, while errors of Hybrid10 are the greatest and significantly worse than those of EnSRF_Ctl. RMSE differences between Hybrid01 and EnSRF_Ctl are generally smaller than those between Hybrid09 and EnSRF_Ctl for T , U , and V .

Overall, introducing 10% ensemble covariance into the variational framework in Hybrid01 has a much larger impact (cf. Hybrid01 to GSI) than adding 10%

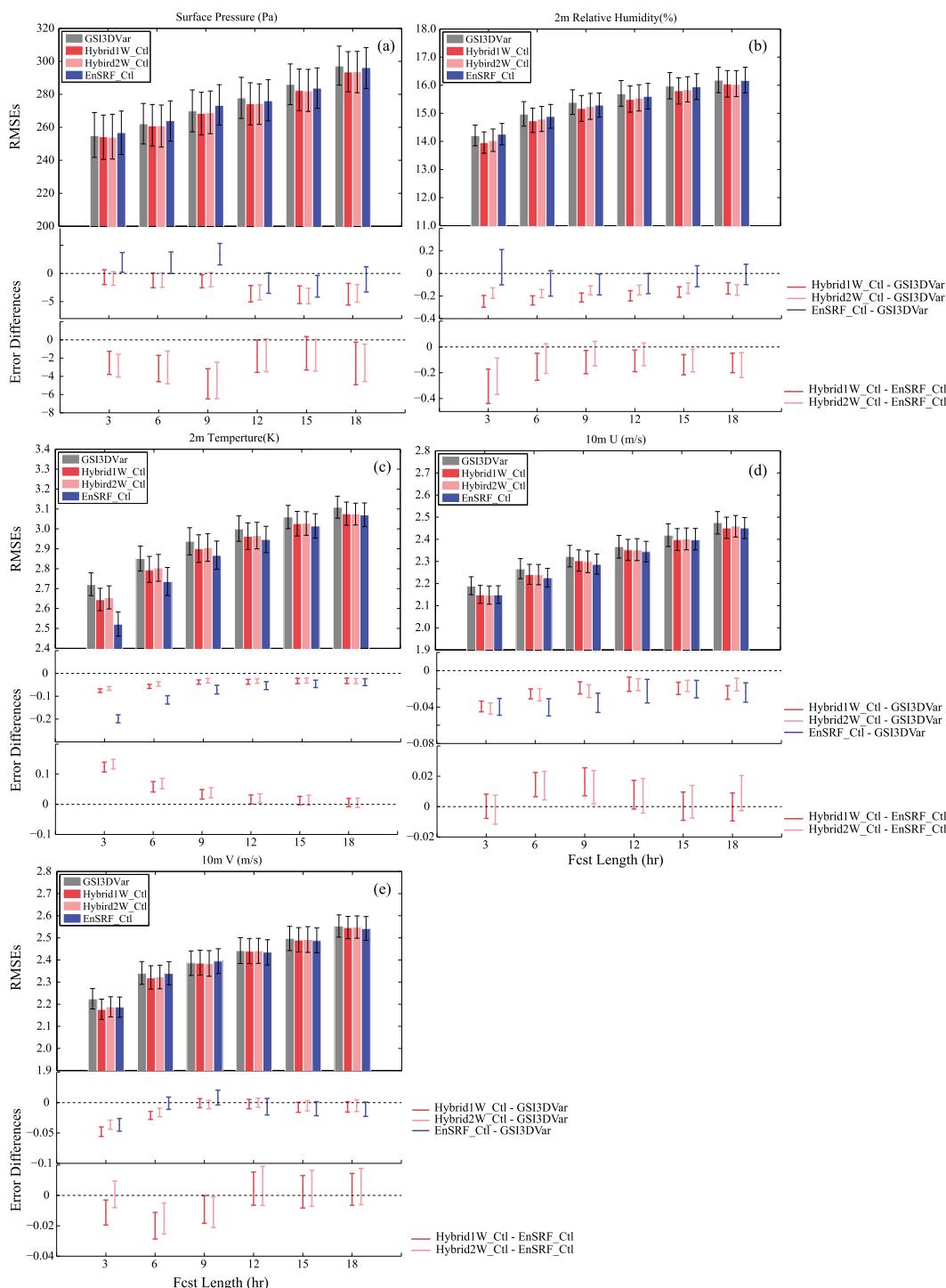


FIG. 6. The 9-day and domain-averaged forecast RMSEs verified against surface observations (top part of each panel) and the 90% confidence interval of the RMSE differences between En3DVar hybrid experiments and GSI 3DVar/EnSRF_Ctl (bottom part of each panel) for (a) surface pressure, (b) 2-m RH, (c) 2-m temperature, (d) 10-m U , and (e) 10-m V for different forecast hours. The horizontal axis is forecast hour. The error bars in domain-averaged forecast RMSEs represent the two-tailed 90% confidence interval.

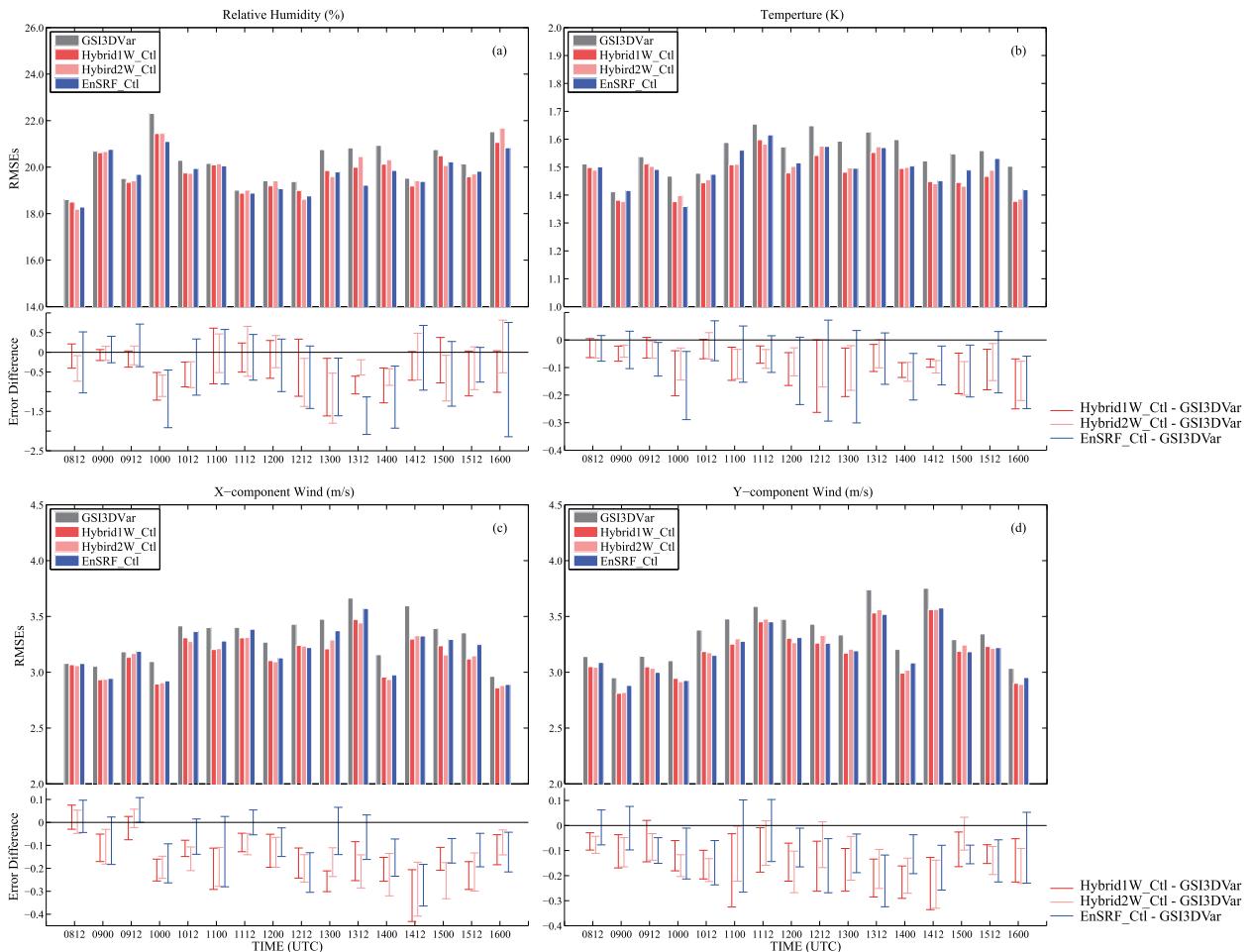


FIG. 7. Domain-averaged 3-h forecast RMSEs (top part of each panel) verified against sounding data at 0000 and 1200 UTC through test period and the 90% confidence interval of RMSE differences (bottom part of each panel) between the En3DVar hybrid and EnSRF experiments and GSI 3DVar for (a) RH, (b) T , (c) U , and (d) V .

static covariance into the En3DVar hybrid framework (cf. Hybrid09 to Hybrid10), and the errors of Hybrid01 are generally between those of Hybrid05 and GSI and are closer to those of Hybrid05, especially for wind fields. Hybrid05 gives the smallest errors on average.

It can also be noticed from Fig. 8 that EnSRF_Ctl outperforms Hybrid10, except for wind between 500- and 200-hPa levels. As pointed out earlier, if covariance localization treatments were the same in EnSRF and Hybrid10, their results should be very close given that the ensemble covariance is used at 100% in both cases. The use of height- and observation-type-dependent covariance localization in the EnSRF but not in the En3DVar hybrid is suspected to be the main cause of the differences. It had been found to help improve the EnSRF results in Z13, but is not used within the En3DVar hybrid. In the following section, we examine whether doing something similar within the En3DVar

hybrid framework can improve the En3DVar hybrid results too.

c. Sensitivity to ensemble covariance localization

In Z13, several tests with the horizontal and vertical covariance localizations were performed. In this paper, the EnSRF experiment uses the same configuration as experiment EnKF_CtrlHDL of Z13, with height- and observation-type-dependent localization radii. For the En3DVar analysis, covariance localization also requires tuning. Because En3DVar realizes covariance localization in the state or gridpoint space, it is impossible to use observation-type-dependent localization unless different observations are analyzed separately.

In this section, we first look at the experiments that use smaller or larger horizontal and vertical CLS than those used in Hybrid1W_Ctl. For weighting factor

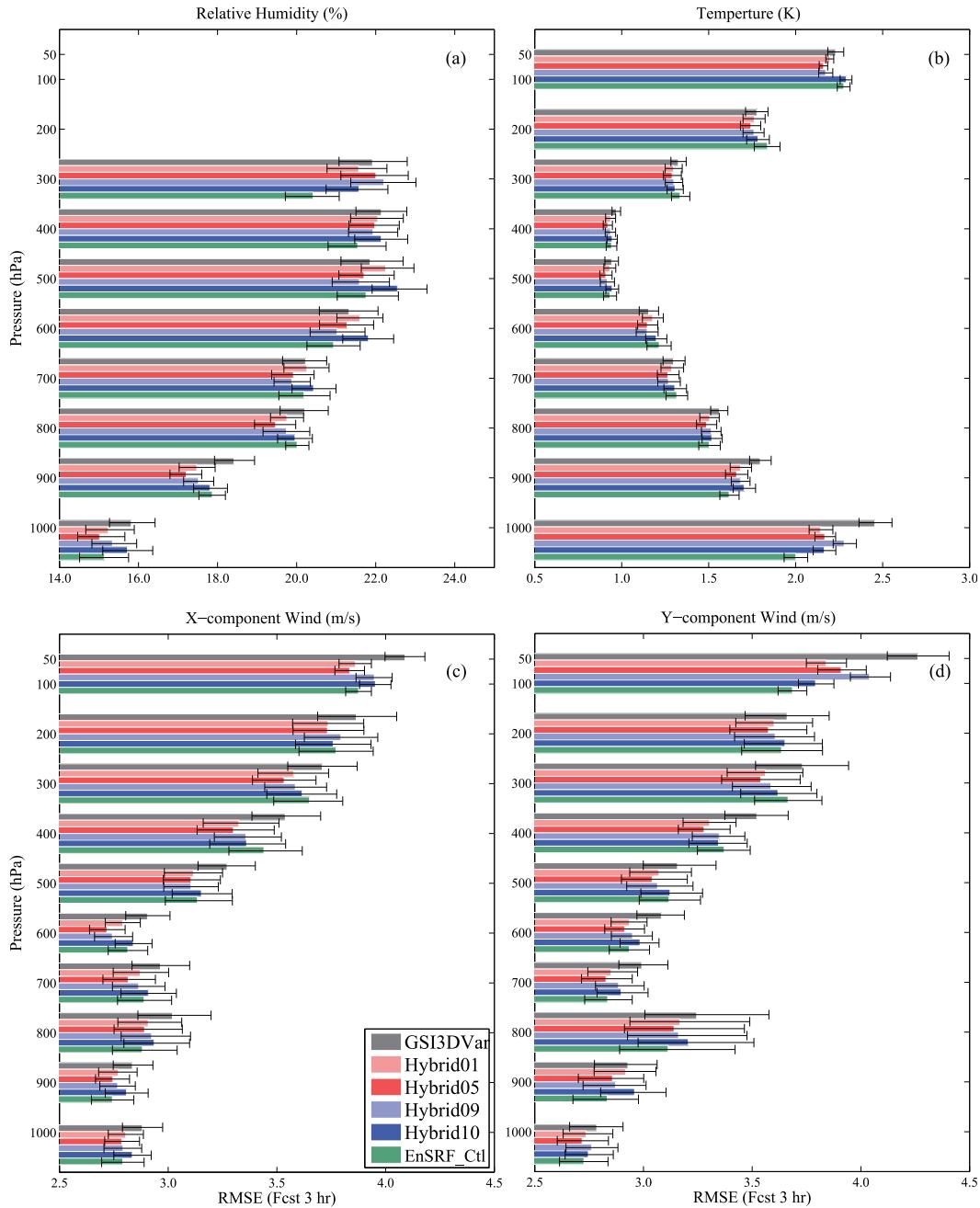


FIG. 8. Mean 3-h forecast RMSEs at different height levels verified against sounding data for (a) RH, (b) T , (c) U , and (d) V for experiments GSI 3DVar, Hybrid01, Hybrid05, Hybrid07, Hybrid10, and EnSRF_Ctl.

$1/\beta_2 = 0.5$, we test three horizontal CLSs, $S_h = 700$, 1095, and 1300 km (in Hybrid_HS, Hybrid1W_Ctl, and Hybrid_HL, respectively) and three vertical CLSs $S_v = 0.36$, 1.1, and 1.8 (for Hybrid_VS, Hybrid1W_Ctl, and Hybrid_VL, respectively). The domain-averaged forecast RMSE differences between 3-h forecasts and those of Hybrid1W_Ctl are shown in Fig. 10. When the CLSs increase or decrease from those of control experiment,

the En3DVar hybrid performs worse for almost all variables, except for T when the horizontal CLS is increased (Fig. 10). However, even though reduced CLSs are not preferred according to Fig. 10, the RH errors are reduced at levels above 800 hPa when using reduced CLSs (not shown), suggesting that we may be able to benefit from the use of observation-type and/or height-dependent CLSs, as in the case of EnSRF (Z13). Doing

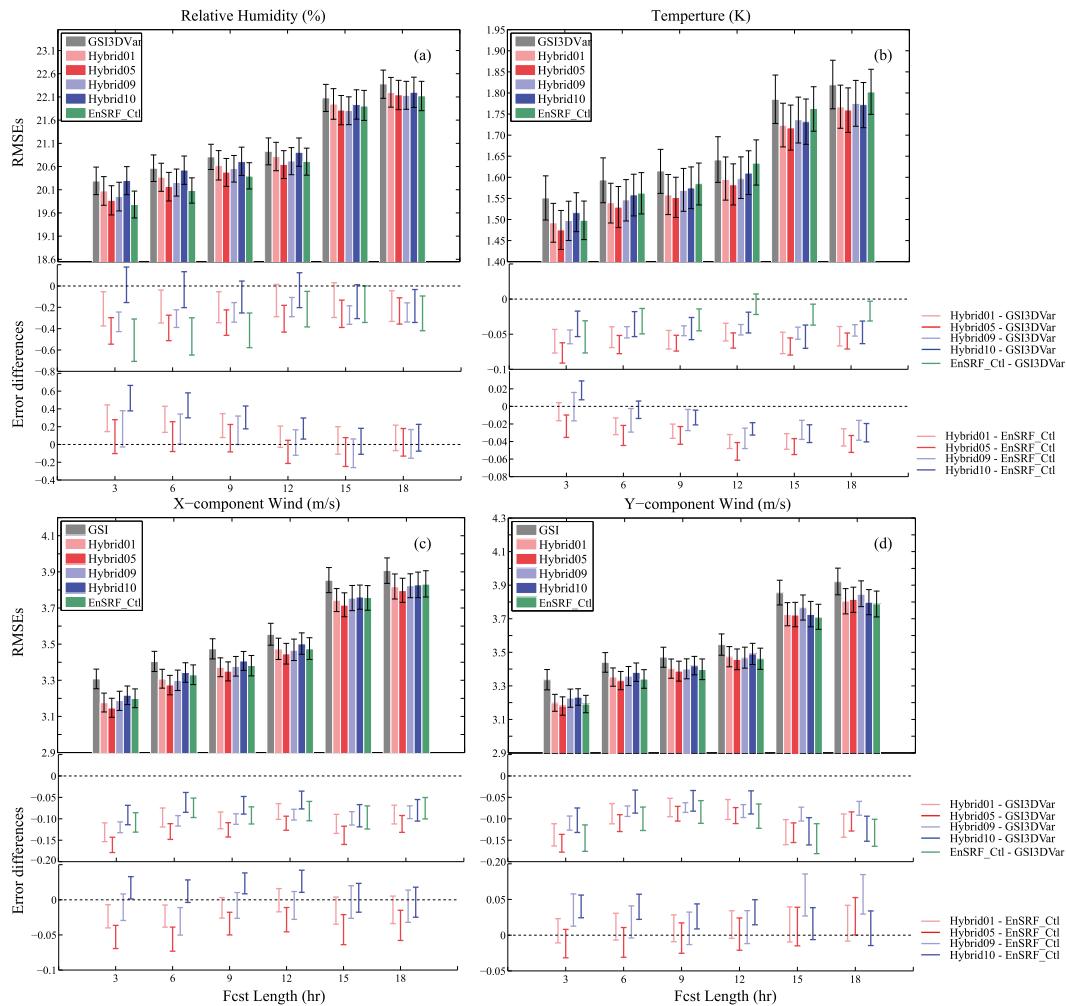


FIG. 9. As in Fig. 5, but for experiments GSI 3DVar, Hybrid01, Hybrid05, Hybrid07, Hybrid10, and EnSRF_Ctl.

so may also help further improve the En3DVar hybrid performance.

For a fair comparison with EnSRF, EnSRF_Con and Hybrid_Con, which use the same constant horizontal and vertical cutoff radii/length scales, are compared when ensemble covariance is used at 100%. As shown in Fig. 11, the significantly greater RMSEs of EnSRF_Con than EnSRF_Ctl suggest the height- and observation-type-dependent localization is a key reason for improving the performance of EnSRF. To see the effect of height-dependent localization (observation-type-dependent localization is not possible in a single step analysis) on En3DVar, we introduce it into Hybrid_HD, which uses 100% ensemble covariance. As shown in Fig. 11, Hybrid_HD outperforms Hybrid_Con and is much closer to EnSRF_Ctl for RH, U , and V . For wind, Hybrid_HD is even slightly better than EnSRF_Ctl at the middle levels (Fig. 11), while Hybrid_Con is poorer

than EnSRF_Ctl at all levels. For RH, EnSRF_Ctl still has smaller RMSEs than hybrid_HD above 700 hPa. The greater RMSEs from EnSRF_Con suggest that the smaller localization radii used in EnSRF_Ctl at the higher levels are beneficial.

The cutoff radii used in the EnSRF_Ctl are also observation-type dependent. As discussed earlier, since the En3DVar algorithm analyzes all observations simultaneously by variational minimization in the state space, making it impossible to apply observation-type-dependent localization within a single analysis step. Experiments EnSRF3G and Hybrid3G break each analysis cycle into three substeps of coupled EnSRF–En3DVar analyses, with each step analyzing a subgroup of observations that share the same height-dependent localization scales. Here, we use the absolute RMSE differences between pairs of En3DVar and EnSRF experiments (Table 2) together with the 90% confidence interval determined

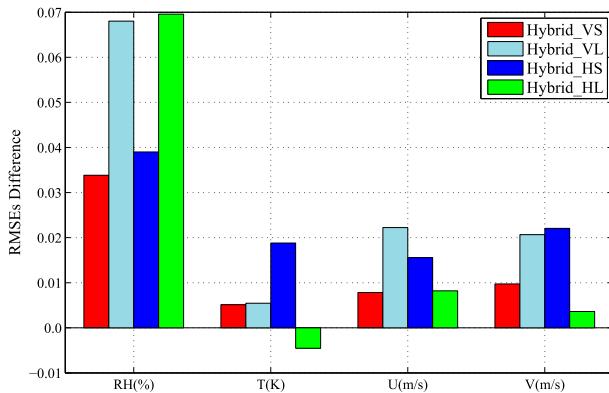


FIG. 10. Mean absolute forecast RMSE differences between different experiments and Hybrid1W_Ctl, verified against sounding data, for 3-h forecast averaged over the 9-day forecast period over the entire model domain.

by the bootstrap resampling procedure to determine the statistical significance of the differences. When the error bars from different experiment pairs do not overlap, the RMSE differences between En3DVar and EnSRF are considered statistically significant. As shown in Fig. 12, the RMSE differences between Hybrid_HD and EnSRF_Ctl (labeled 1GHD) are reduced, by about $\frac{1}{4}$ to $\frac{1}{3}$ for RH, U , and V compared to the differences between Hybrid_Con and EnSRF_Ctl (labeled 1GC1). The reduction in the difference is smaller but still statistically significant for T . When constant localization is used in both EnSRF_Con and Hybrid_Con, the differences (labeled 1GC2) are also reduced compared to 1GC1 case, although the differences are slightly larger than the 1GHD case.

When observation-type-dependent localization is used in Hybrid3G and EnSRF3G, the differences between them (3DHD) are even smaller for RH, U , and V . For RH (which has the largest RMSE difference between En3DVar and EnSRF according to Fig. 11a), the RMSE difference of RH is about 0.5% versus the 1.25% for the constant localization case. The reductions for T , U , and V are smaller but still evident. Clearly, the differences between En3DVar and EnSRF are much smaller when height- and observation-type-dependent localization is used in both algorithms, especially for humidity.

Figure 13 shows the RMSE differences together with the 90% confidence interval at different height levels. In reference to Fig. 11, those levels where domain average absolute RMSE differences between Hybrid_HD and EnSRF_Ctl (1GHD) are greater than those between Hybrid_Con and EnSRF_Ctl (1GC1) correspond to the levels where Hybrid_HV outperforms Hybrid_Con, given that EnSRF_Ctl is generally the best among the

three experiments. The RMSE differences of 1GC2 are also smaller than 1GC1, but the constant localization degrades the performance of EnSRF and En3DVar for almost all the variables and at all levels in reference to Fig. 11. For 3GHD, the average absolute RMSE differences are the smallest for RH at all levels, for T above 800 hPa and for U and V above 600 hPa. For U and V , the 3GHD differences are slightly larger below 700 hPa than 1GHD and clearly smaller than 1GC1. These results show that when similar height- and observation-type-dependent covariance localization is used in the En3DVar framework using 100% ensemble covariance, differences between EnSRF and En3DVar are significantly reduced, and such localization treatment generally brings the En3DVar results closer to the better EnSRF results. The reduction in the RMSE differences for RH is greater than those for T , U , and V . Because the humidity field tends to contain smaller scale structures than other fields, it appears to benefit from tighter localization more when using height- and observation-type-dependent localization. However, because there are still differences between the EnSRF and En3DVar algorithms, some differences still exist between their results, as indicated by the green bars in Fig. 12. When the ensemble covariance is used at 50%, height-dependent localization did not improve the En3DVar hybrid results as much as in the 100% case (not shown).

In summary, the use of height-dependent localization in the En3DVar hybrid framework when using full ensemble covariance improves the resulting model forecasts at almost all levels and forecast hours. Height- and observation-type-dependent localizations used in EnSRF are responsible for about half of the differences between the EnSRF and the En3DVar with full ensemble covariance. Unfortunately, observation-type-dependent localization is difficult or expensive to implement with the En3DVar algorithm.

5. Precipitation forecast skills on 13-km grid

In this section, precipitation forecasts on the 13-km grid initialized from the 40-km GSI 3DVar, EnSRF_Ctl ensemble mean, Hybrid1W_Ctl, and Hybrid2W_Ctl analyses (Table 3) are compared. Considering extensive CPU and storage requirements, we launched the forecasts only twice a day at 0000 and 1200 UTC. The precipitation forecasts are verified against the NCEP stage-IV precipitation data. GSSs calculated for the 0.1, 1.25, and 2.5 mm h⁻¹ thresholds are calculated as in Z13.

The GSSs and BIASs for the forecasts are shown in Fig. 14. Both EnSRF and En3DVar hybrid outperform GSI 3DVar on average for all forecast hours and thresholds

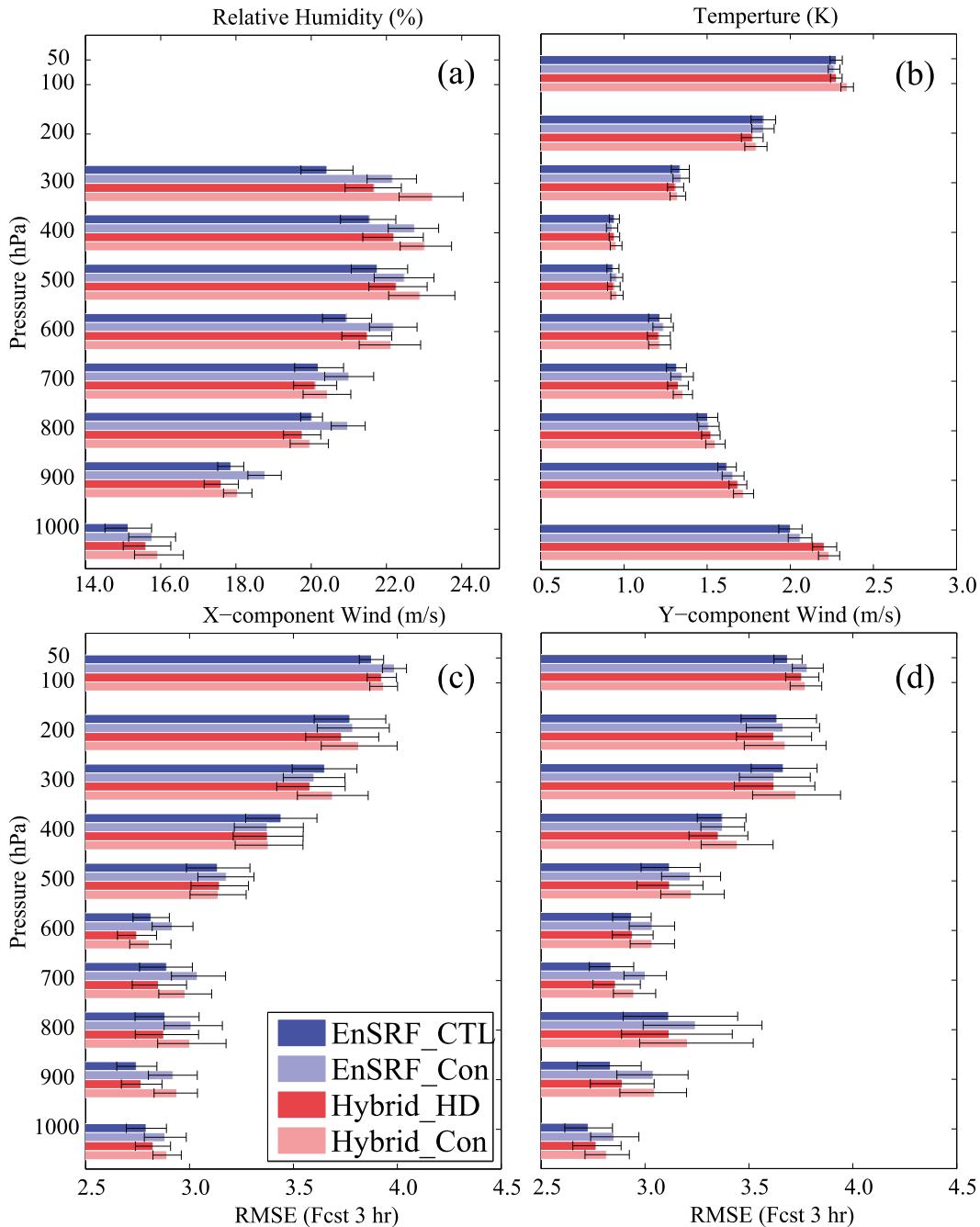


FIG. 11. As in Fig. 4, but for experiments EnSRF_Ctl, EnSRF_Con, Hybrid_HD, and Hybrid_Con.

shown. EnSRF13 has higher GSSs than Hybrid for 0.1 mm h^{-1} after 3 h. For greater thresholds of 1.25 and 2.5 mm h^{-1} , forecasts of Hybrid1W13 are comparable to EnSRF13 by 7 h, and are better than Hybrid2W13 during the first four hours, which is consistent with the domain-averaged RMSEs of RH shown in Fig. 5a. Figure 14 show that EnSRF generally has the highest positive BIASs. The hybrid schemes have the lowest BIASs

in the first 5 h, and values between those of GSI 3DVar and EnSRF after 5 h. The BIAS differences are relatively small and BIAS is always positive for both thresholds examined.

From sensitivity experiments, we learned that the constant localization used in En3DVar hybrid is one of the reasons for the deterioration of humidity forecasts compared to EnSRF. To help further understand the

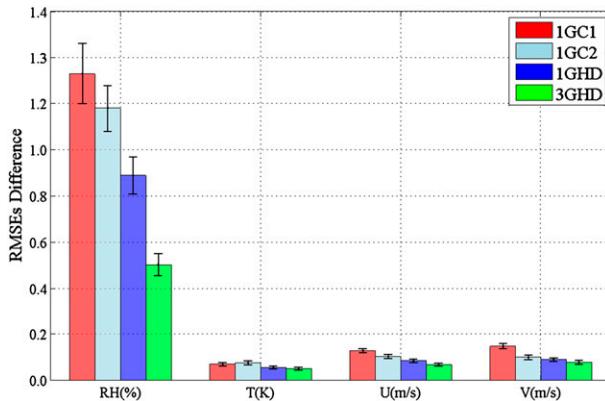


FIG. 12. The 9-day and domain-averaged absolute 3-h forecast RMSE differences verified against sounding data, where 1GC1 means the difference between Hybrid_Con and EnSRF_Ctl, 1GC2 means the difference between Hybrid_Con and EnSRF_Con, 1GHD means the difference between Hybrid_HD and EnSRF_Ctl, and 3GHD means the difference between Hybrid3G and EnSRF3G. The error bars represent the two-tailed 90% confidence interval.

impact of localization on precipitation forecast skill, the forecasts on the 13-km grid initialized from EnSRF_Ctl and EnSRF_Con ensemble mean, and Hybrid_Con analyses are also compared. Without height- and observation-type-dependent localization in EnSRF13-Con, its GSSs are lower than those of EnSRF13 in the first 9 h and are close to those of Hybrid1W13Con from 4 to 7 h, for the 0.1 mm h^{-1} threshold (Fig. 15), indicating the role of height- and observation-type-dependent localization. The differences among the experiments are smaller for larger thresholds (not shown).

6. Summary and discussion

In this paper, a coupled EnSRF–En3DVar hybrid data assimilation system based on the NCEP operational GSI variational framework is established and tested for the Rapid Refresh (RAP) forecasting system. It uses a recently developed, well-tuned, 40-member EnSRF system, as documented in Z13, to update and provide the ensemble perturbations. A 9-day spring period starting from 8 May 2010 that contains active convection is used to examine the performance of the system through comparisons with parallel experiments using EnSRF and GSI 3DVar. The En3DVar hybrid, EnSRF, and GSI 3DVar experiments use the same observational datasets as the operational RAP system except for the exclusion of satellite radiance data. The experiments are performed at a reduced resolution of ~ 40 -km grid spacing with 3-hourly assimilation cycles rather than at the native 13-km grid spacing with hourly cycles of the operational RAP. The systems are evaluated based on

forecast RMSEs verified against surface observations and upper-air sounding data for 3–18-h forecasts. The effects of static and ensemble covariance weighting factors and covariance localization configurations are examined through sensitivity experiments.

With equal weighting for the ensemble and static covariances, the En3DVar hybrid scheme outperforms GSI 3DVar for all standard variables at all levels with statistical significance, and is slightly better than EnSRF, especially for later forecast hours, except for moisture during the earlier forecast hours. Apparently, the En3DVar hybrid scheme benefits from the combined use of static and ensemble covariances. Introducing 10% flow-dependent covariance into the standard 3DVar framework has a much bigger positive impact than including 10% static covariance in the En3DVar framework. The forecasts from En3DVar analyses with 100% ensemble covariance and constant covariance localization scales are worse than those from pure EnSRF analyses using height- and observation-type-dependent covariance localization, especially for relative humidity. The height-dependent localization scheme in which the horizontal localization cutoff radii increase with height, and the observation-type-dependent localization scheme in which the cutoff radii for relative humidity and temperature observations are set to be smaller than those for winds led to smaller forecast RMSEs for the pure EnSRF, especially at the high and low levels. Using similar height-dependent localization, En3DVar with 100% ensemble covariance became much closer to pure EnSRF. When using similar observation-type-dependent covariance localization in En3DVar, by running the coupled EnSRF–En3DVar analyses in three steps with each analyzing a subset of observation variables (in a similar way as in EnSRF), the results of En3DVar, with 100% ensemble covariance, become even closer to those of EnSRF. The benefit of height- and observation-type-dependent localization is negligible when the ensemble covariance is used at 50%. The multistep EnSRF–En3DVar analysis procedure is, unfortunately, not very practical because of the much increased computational costs. It is straightforward for pure EnSRF because the algorithm is serial, where observations are assimilated sequentially.

Previous studies (e.g., Hamill and Snyder 2000; Wang et al. 2008b) had found that the En3DVar hybrid system is more robust than EnSRF when the ensemble size is small or model error is large. In our study, the EnSRF, En3DVar hybrid one-way and two-way using 20 instead of 40 ensemble members for control experiments were also run and compared. In such a case, EnSRF and En3DVar are both degraded, but the En3DVar hybrid is now consistently better than EnSRF for all variables and

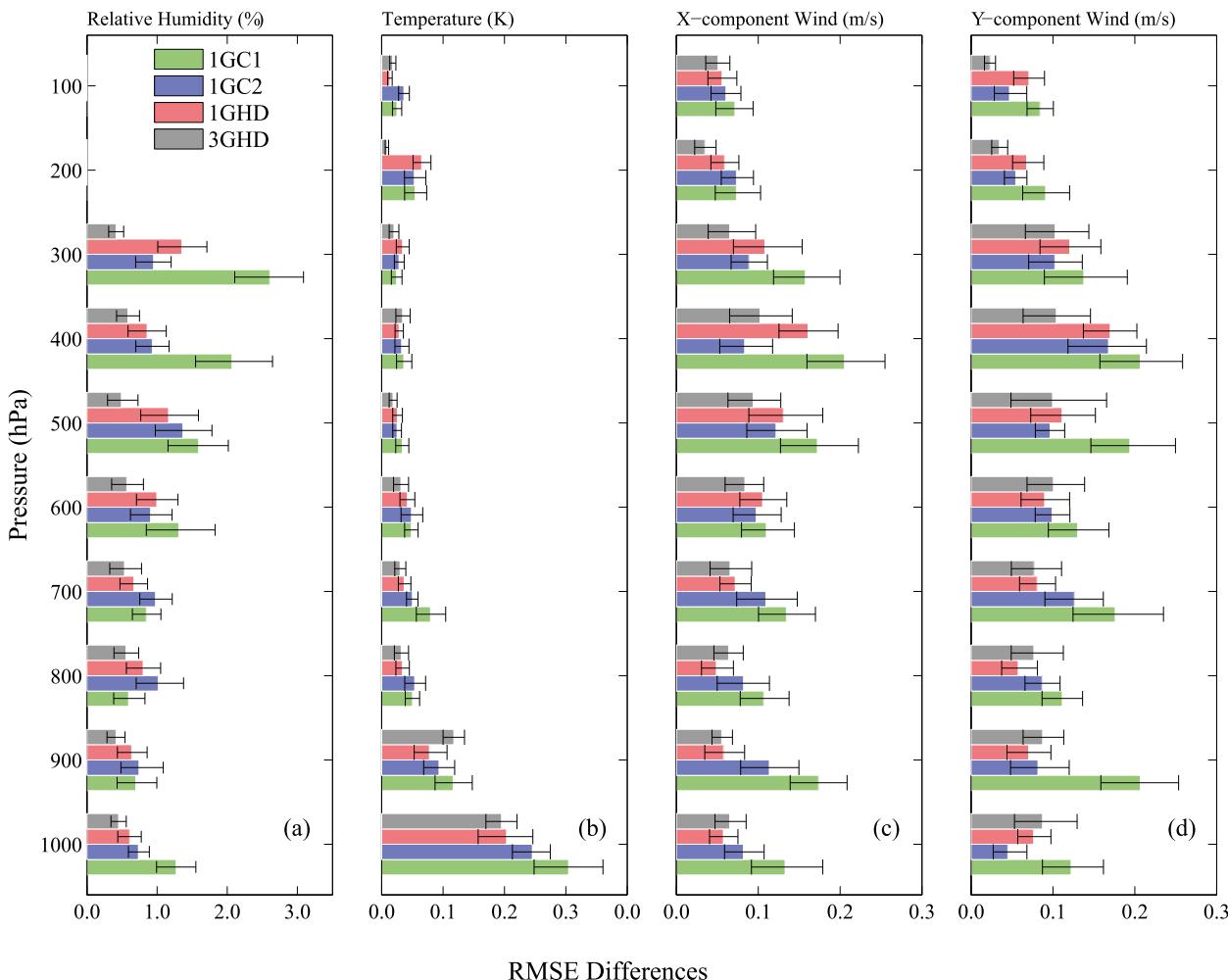


FIG. 13. The 9-day and domain-averaged absolute RMSE differences between Hybrid_Con and EnSRF_Ctl (labeled 1GC1), Hybrid_Con and EnSRF_Con (labeled 1GC2), Hybrid_HD and EnSRF_Ctl (labeled 1GHD), and Hybrid3G and EnSRF3G (labeled 3GHD) for (a) RH, (b) T , (c) U , and (d) V at different height levels. The error bars represent the two-tailed 90% confidence interval.

all forecast hours (not shown). That indicates a larger benefit of the static covariance when the ensemble is small, agreeing with earlier findings. While further specific tuning of the 20-member EnSRF may improve the results somewhat, we do not expect the general conclusion to change.

Deterministic forecasts were launched on a 13-km grid from interpolated 40-km En3DVar hybrid control, EnSRF ensemble mean, and GSI 3DVar analyses at 0000 and 1200 UTC of each day. Hourly accumulated precipitation is better predicted in the En3DVar hybrid and EnSRF experiments than GSI 3DVar, but for light precipitation, En3DVar hybrid does not perform as well as EnSRF, which is consistent with the slightly worse humidity forecasts of En3DVar. When constant covariance localization is used in EnSRF, its precipitation forecast skills become closer to those of En3DVar using

100% ensemble covariance, indicating the covariance localization difference between the EnSRF and En3DVar as a key cause for the performance differences.

Despite the encouraging results, the En3DVar hybrid system we have established for RAP still has room for further improvement. Adding satellite and radar data and examining their impacts are among the desired

TABLE 3. List of precipitation forecast name and corresponding data assimilation experiments on 40-km grid.

Precipitation forecast name	Data assimilation expt
EnSRF13	EnSRF_Ctl
Hybrid1W13	Hybrid1W_Ctl
Hybrid2W13	Hybrid2W_Ctl
EnSRF13Con	EnSRF_Con
Hybrid13Con	Hybrid_Con

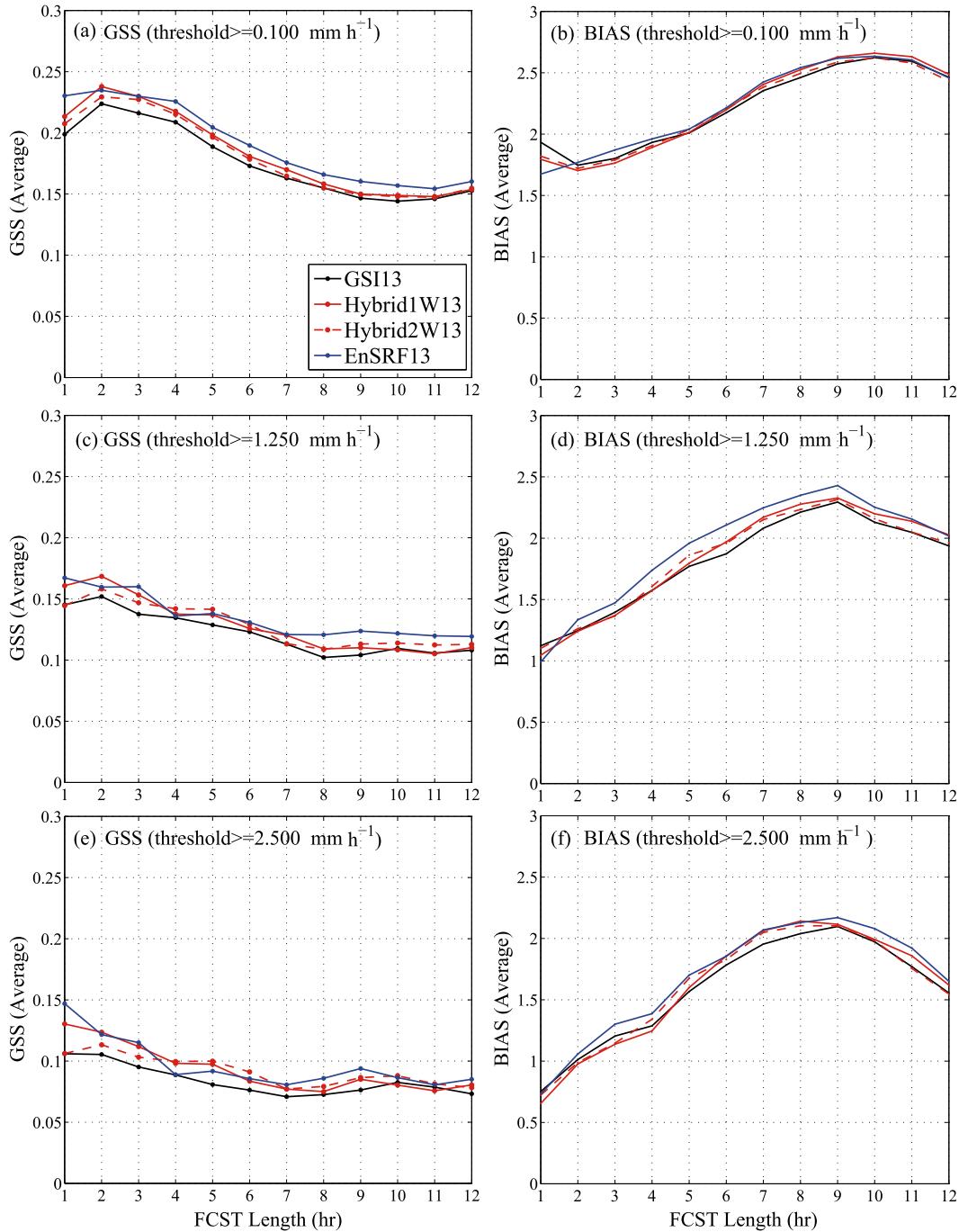


FIG. 14. Average precipitation GSSs and BIASs of 13-km forecasts as a function of forecast length for thresholds (a),(b) 0.1; (c),(d) 1.25; and (e),(f) 2.5 mm h⁻¹ for control experiments.

tasks, as is a dual-resolution implementation where the En3DVar is run at the native RAP resolution. These aspects are being pursued and the results will be reported in future papers. A further refined and tested version, initially in a dual-resolution mode coupling a 13-km En3DVAR with a 40-km EnSRF, will likely become operational in the future, replacing the GFS

EnSRF system used in the recently implemented operational RAP hybrid DA system.

Acknowledgments. This work was primarily supported by the FAA Aviation Weather Research Program through Grant DOC-NOAA NA080AR4320904, and by MOST of China Grant 2013CB430100. Additional

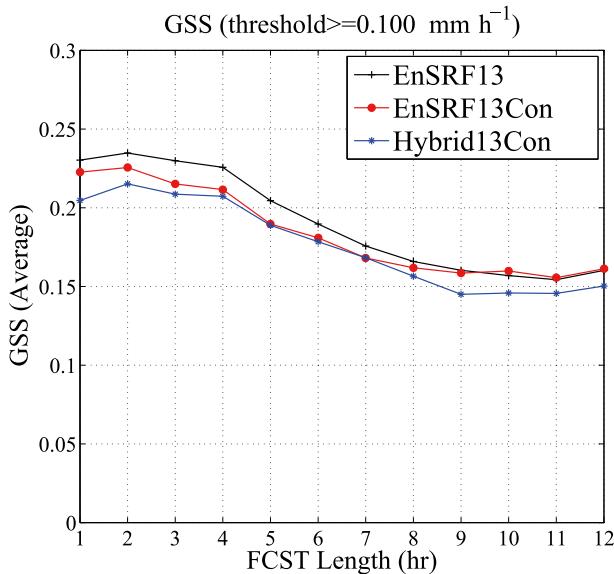


FIG. 15. As in Fig. 14, but for EnKF13, EnKF13Con, and Hybrid13Con.

support was provided by NSF Grants AGS-0802888, OCI-0905040, AGS-0941491, AGS-1046171, and AGS-1046081, and by the NOAA Warn-on-Forecast program. The fourth author is also supported by NOAA THORPEX NA08OAR4320904, NASA NIP NNX10AQ78G, and NOAA HFIP Grant NA12NWS4680012.

REFERENCES

- Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884–2903, doi:10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2.
- , 2007: An adaptive covariance inflation error correction algorithm for ensemble filters. *Tellus*, **59A**, 210–224, doi:10.1111/j.1600-0870.2006.00216.x.
- , 2012: Localization and sampling error correction in ensemble Kalman filter data assimilation. *Mon. Wea. Rev.*, **140**, 2359–2371, doi:10.1175/MWR-D-11-00013.1.
- 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, doi: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, doi:10.1175/1520-0493(2004)132<0495:AHACTR>2.0.CO;2.
- , and Coauthors, 2009: Technical review of rapid refresh/RUC project. NOAA/ESRL/GSD internal review, 3 Nov 2009. [Available online at http://ruc.noaa.gov/pdf/RR-RUC-TR_11_3_2009.pdf.]
- Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.*, **129**, 420–436, doi:10.1175/1520-0493(2001)129<0420:ASWTET>2.0.CO;2.
- Bonavita, M., L. Torrisi, and F. Marcucci, 2008: The ensemble Kalman filter in an operational regional NWP system: Preliminary results with real observations. *Quart. J. Roy. Meteor. Soc.*, **134**, 1733–1744, doi:10.1002/qj.313.
- , L. Isaksen, and E. Hölm, 2012: On the use of EDA background error variances in the ECMWF 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **138**, 1540–1559, doi:10.1002/qj.1899.
- 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 Conf. on Int. Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology*, Phoenix, AZ, Amer. Meteor. Soc., 9A.6. [Available online at <https://ams.confex.com/ams/pdfpapers/151349.pdf>.]
- Buehner, M., 2005: Ensemble-derived stationary and flow-dependent background-error covariances: Evaluation in a quasi-operational NWP setting. *Quart. J. Roy. Meteor. Soc.*, **131**, 1013–1043, doi:10.1256/qj.04.15.
- , 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, doi:10.1175/2010MWR3296.1.
- , 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, doi:10.1175/2009MWR3157.1.
- , —, —, —, and —, 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, doi:10.1175/2009MWR3158.1.
- Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998: Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.*, **126**, 1719–1724, doi:10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2.
- Campbell, W. F., C. H. Bishop, and D. Hodyss, 2010: Vertical covariance localization for satellite radiances in ensemble Kalman filters. *Mon. Wea. Rev.*, **138**, 282–290, doi:10.1175/2009MWR3017.1.
- 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, doi:10.1175/MWR3414.1.
- Courtier, P., and Coauthors, 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: formulation. *Quart. J. Roy. Meteor. Soc.*, **124**, 1783–1807, doi:10.1002/qj.49712455002.
- 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, doi:10.1029/94JC00572.
- , 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.*, **53**, 343–367, doi:10.1007/s10236-003-0036-9.
- Fujita, T., D. J. Stensrud, and D. C. Dowell, 2007: Surface data assimilation using an ensemble Kalman filter approach with initial condition and model physics uncertainties. *Mon. Wea. Rev.*, **135**, 1846–1868, doi:10.1175/MWR3391.1.
- Gandin, L. S., and A. H. Murphy, 1992: Equitable skill scores for categorical forecasts. *Mon. Wea. Rev.*, **120**, 361–370, doi:10.1175/1520-0493(1992)120<0361:ESSFCF>2.0.CO;2.
- Gao, J.-D., M. Xue, K. Brewster, and K. K. Droegemeier, 2004: A three-dimensional variational data analysis method with recursive filter for Doppler radars. *J. Atmos. Oceanic*

- Technol.*, **21**, 457–469, doi:10.1175/1520-0426(2004)021<0457:ATVDAM>2.0.CO;2.
- Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757, doi:10.1002/qj.49712555417.
- Ge, G., J. Gao, M. Xue, and K. K. Droegemeier, 2012: Diagnostic pressure equation as a weak constraint in a storm-scale three dimensional variational radar data assimilation system. *J. Atmos. Oceanic Technol.*, **29**, 1075–1092, doi:10.1175/JTECH-D-11-00201.1.
- Hamill, T. M., and C. Snyder, 2000: A hybrid ensemble Kalman filter–3D variational analysis scheme. *Mon. Wea. Rev.*, **128**, 2905–2919, doi:10.1175/1520-0493(2000)128<2905:AHEKFF>2.0.CO;2.
- , 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, doi:10.1175/1520-0493(2001)129<2776:DDFOBE>2.0.CO;2.
- , —, M. Fiorino, and S. G. Benjamin, 2011a: Global ensemble predictions of 2009's tropical cyclones initialized with an ensemble Kalman filter. *Mon. Wea. Rev.*, **139**, 668–688, doi:10.1175/2010MWR3456.1.
- , —, D. T. Kleist, M. Fiorino, and S. G. Benjamin, 2011b: Predictions of 2010's tropical cyclones using the GFS and ensemble-based data assimilation methods. *Mon. Wea. Rev.*, **139**, 3243–3247, doi:10.1175/MWR-D-11-00079.1.
- Hayden, C. M., and R. J. Purser, 1995: Recursive filter objective analysis of meteorological fields: Applications to NESDIS operational processing. *J. Appl. Meteor.*, **34**, 3–15, doi:10.1175/1520-0450-34.1.3.
- Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796–811, doi:10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2.
- , —, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen, 2005: Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. *Mon. Wea. Rev.*, **133**, 604–620, doi:10.1175/MWR-2864.1.
- Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230**, 112–126, doi:10.1016/j.physd.2006.11.008.
- Jung, Y., M. Xue, and M. Tong, 2012: Ensemble Kalman filter analyses of the 29–30 May 2004 Oklahoma tornadic thunderstorm using one- and two-moment bulk microphysics schemes, with verification against polarimetric data. *Mon. Wea. Rev.*, **140**, 1457–1475, doi:10.1175/MWR-D-11-00032.1.
- Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W.-S. Wu, and S. Lord, 2009a: Introduction of the GSI into the NCEP global data assimilation system. *Wea. Forecasting*, **24**, 1691–1705, doi:10.1175/2009WAF2222201.1.
- , —, —, —, R. M. Errico, and R. Yang, 2009b: Improving incremental balance in the GSI 3DVAR analysis system. *Mon. Wea. Rev.*, **137**, 1046–1060, doi:10.1175/2008MWR2623.1.
- Le Dimet, F. X., and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus*, **38A**, 97–110, doi:10.1111/j.1600-0870.1986.tb00459.x.
- Li, Y., X. Wang, and M. Xue, 2012: Assimilation of radar radial velocity data with the WRF ensemble-3DVAR hybrid system for the prediction of Hurricane Ike (2008). *Mon. Wea. Rev.*, **140**, 3507–3524, doi:10.1175/MWR-D-12-00043.1.
- Lin, Y., and K. E. Mitchell, 2005: The NCEP stage II/IV hourly precipitation analyses: Development and applications. *19th Conf. on Hydrology*, San Diego, CA, Amer. Meteor. Soc., 1.2. [Available online at https://ams.confex.com/ams/Annual2005/techprogram/paper_83847.htm.]
- Liu, C., and Q. Xiao, 2013: An ensemble-based four-dimensional variational data assimilation scheme. Part III: Antarctic applications with Advanced Research WRF using real data. *Mon. Wea. Rev.*, **141**, 2721–2739, doi:10.1175/MWR-D-12-00130.1.
- , —, and B. Wang, 2008: An ensemble-based four-dimensional variational data assimilation scheme. Part I: Technical formulation and preliminary test. *Mon. Wea. Rev.*, **136**, 3363–3373, doi:10.1175/2008MWR2312.1.
- Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **112**, 1177–1194, doi:10.1002/qj.49711247414.
- , 2003: The potential of the ensemble Kalman filter for NWP—A comparison with 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **129**, 3183–3204, doi:10.1256/qj.02.132.
- Lynch, P., 1997: The Dolph–Chebyshev window: A simple optimal filter. *Mon. Wea. Rev.*, **125**, 655–660, doi:10.1175/1520-0493(1997)125<0655:TDCWAS>2.0.CO;2.
- , and X.-Y. Huang, 1992: Initialization of the HIRLAM model using a digital filter. *Mon. Wea. Rev.*, **120**, 1019–1034, doi:10.1175/1520-0493(1992)120<1019:IOTHMU>2.0.CO;2.
- , and —, 1994: Diabatic initialization using recursive filters. *Tellus*, **46A**, 583–597, doi:10.1034/j.1600-0870.1994.t01-4-00003.x.
- Meng, Z., and F. 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, doi:10.1175/MWR3352.1.
- Mizzi, A. P., 2012: Comparison of the GSI/ETKF regional hybrid using WRF/ARW with GSI/LETKF and GSI/EnKF regional hybrids. *16th Conf. on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface*, New Orleans, LA, Amer. Meteor. Soc., 9.3. [Available online at <https://ams.confex.com/ams/92Annual/webprogram/Paper199538.html>.]
- Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763, doi:10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2.
- Purser, R. J., W.-S. Wu, D. F. Parrish, and N. M. Roberts, 2003a: Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: Spatially homogeneous and isotropic Gaussian covariances. *Mon. Wea. Rev.*, **131**, 1524–1535, doi:10.1175/1520-0493(2003)131<1524:NAOTAO>2.0.CO;2.
- , —, —, and —, 2003b: Numerical aspects of the application of recursive filters to variational statistical analysis. Part II: Spatially inhomogeneous and anisotropic general covariances. *Mon. Wea. Rev.*, **131**, 1536–1548, doi:10.1175/2543.1.
- Putnam, B. J., M. Xue, Y. Jung, N. A. Snook, and G. Zhang, 2014: The analysis and prediction of microphysical states and polarimetric variables in a mesoscale convective system using double-moment microphysics, multi-network radar data, and the ensemble Kalman filter. *Mon. Wea. Rev.*, **142**, 141–162, doi:10.1175/MWR-D-13-00042.1.
- Rabier, F., H. Jarvinen, E. Klinker, J.-F. Mahfouf, and A. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Quart. J. Roy. Meteor. Soc.*, **126**, 1143–1170, doi:10.1002/qj.49712656415.

- Raynaud, L., L. Berre, and G. Desroziers, 2011: An extended specification of flow-dependent background error variances in the Meteo-France global 4D-Var system. *Quart. J. Roy. Meteor. Soc.*, **137**, 607–619, doi:10.1002/qj.795.
- Schwartz, C. S., and Z. 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, doi:10.1175/MWR-D-13-00100.1.
- Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. D. Powers, 2005: A description of the advanced research WRF version 2. NCAR Tech. Note NCAR/TN-468+STR, 88 pp. [Available online at http://www.mmm.ucar.edu/wrf/users/docs/arw_v2.pdf.]
- Talagrand, O., and P. Courtier, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. Part I: Theory. *Quart. J. Roy. Meteor. Soc.*, **113**, 1311–1328, doi:10.1002/qj.49711347812.
- Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble square root filters. *Mon. Wea. Rev.*, **131**, 1485–1490, doi:10.1175/1520-0493(2003)131<1485:ESRF>2.0.CO;2.
- Tong, M., and M. Xue, 2005: Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS experiments. *Mon. Wea. Rev.*, **133**, 1789–1807, doi:10.1175/MWR2898.1.
- Wang, X., 2010: Incorporating ensemble covariance in the Grid-point Statistical Interpolation variational minimization: A mathematical framework. *Mon. Wea. Rev.*, **138**, 2990–2995, doi:10.1175/2010MWR3245.1.
- , 2011: Application of the WRF hybrid ETKF–3DVAR data assimilation system for hurricane track forecasts. *Wea. Forecasting*, **26**, 868–884, doi:10.1175/WAF-D-10-05058.1.
- , C. Snyder, and T. M. Hamill, 2007: On the theoretical equivalence of differently proposed ensemble/VAR hybrid analysis schemes. *Mon. Wea. Rev.*, **135**, 222–227, doi:10.1175/MWR3282.1.
- , D. M. Barker, C. Snyder, and T. M. Hamill, 2008a: A hybrid ETKF–3DVAR data assimilation scheme for the WRF model. Part I: Observing system simulation experiment. *Mon. Wea. Rev.*, **136**, 5116–5131, doi:10.1175/2008MWR2444.1.
- , —, —, and —, 2008b: A hybrid ETKF–3DVAR data assimilation scheme for the WRF model. Part II: Real observation experiment. *Mon. Wea. Rev.*, **136**, 5132–5147, doi:10.1175/2008MWR2445.1.
- , 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, doi:10.1175/2009MWR2923.1.
- , 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, doi: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, doi:10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2.
- , —, X. Wei, Y. Song, and Z. Toth, 2008: Ensemble data assimilation with the NCEP Global Forecast System. *Mon. Wea. Rev.*, **136**, 463–482, doi:10.1175/2007MWR2018.1.
- , D. T. Kleist, X. Wang, and T. Hamill, 2011: Tests of a hybrid variational-ensemble global data assimilation system for hurricane prediction. *24th Conf. on Weather Forecasting/20th Conf. on Numerical Weather Prediction*, Seattle, WA, Amer. Meteor. Soc., J16.2. [Available online at <https://ams.confex.com/ams/91Annual/webprogram/Paper178847.html>.]
- 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, doi:10.1175/1520-0493(2002)130<2905:TDVAWS>2.0.CO;2.
- Xue, M., Y. Jung, and G. Zhang, 2010: State estimation of convective storms with a two-moment microphysics scheme and an ensemble Kalman filter: Experiments with simulated radar data. *Quart. J. Roy. Meteor. Soc.*, **136**, 685–700, doi:10.1002/qj.593.
- , J. Schreif, F. Kong, K. K. Thomas, Y. Wang, and K. Zhu, 2013: Track and intensity forecasting of hurricanes: Impact of cloud-resolving resolution and ensemble Kalman filter data assimilation on 2010 Atlantic season forecasts. *Wea. Forecasting*, **28**, 1366–1384, doi:10.1175/WAF-D-12-00063.1.
- Zhang, F., 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, doi:10.1175/MWR-D-12-00075.1.
- Zhang, M., and F. 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, doi:10.1175/MWR-D-11-00023.1.
- Zhu, K., Y. Pan, M. Xue, X. 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, doi:10.1175/MWR-D-13-00039.1.