Sensitivity of Limited-Area Hybrid Variational-Ensemble Analyses and Forecasts to Ensemble Perturbation Resolution

CRAIG S. SCHWARTZ, ZHIQUAN LIU, AND XIANG-YU HUANG

National Center for Atmospheric Research, Boulder, Colorado*

(Manuscript received 7 August 2014, in final form 21 May 2015)

ABSTRACT

Dual-resolution (DR) hybrid variational-ensemble analysis capability was implemented within the community Weather Research and Forecasting (WRF) Model data assimilation (DA) system, which is designed for limited-area applications. The DR hybrid system combines a high-resolution (HR) background, flow-dependent background error covariances (BECs) derived from a *low-resolution* ensemble, and observations to produce a deterministic HR analysis. As DR systems do not require HR ensembles, they are computationally cheaper than single-resolution (SR) hybrid configurations, where the background and ensemble have equal resolutions.

Single-observation tests were performed to document some characteristics of limited-area DR hybrid analyses. Additionally, the DR hybrid system was evaluated within a continuously cycling framework, where new DR hybrid analyses were produced every 6 h over ~3.5 weeks. In the DR configuration presented here, the deterministic backgrounds and analyses had 15-km horizontal grid spacing, but the 32-member WRF Model–based ensembles providing flow-dependent BECs for the hybrid had 45-km horizontal grid spacing. The DR hybrid analyses initialized 72-h WRF Model forecasts that were compared to forecasts initialized by an SR hybrid system where both the ensemble and background had 15-km horizontal grid spacing. The SR and DR hybrid systems were coupled to an ensemble adjustment Kalman filter that updated ensembles each DA cycle.

On average, forecasts initialized from 15-km DR and SR hybrid analyses were not statistically significantly different, although tropical cyclone track forecast errors favored the SR-initialized forecasts. Although additional studies over longer time periods and at finer grid spacing are needed to further understand sensitivity to ensemble perturbation resolution, these results suggest users should carefully consider whether SR hybrid systems are worth the extra cost.

1. Introduction

Ensemble-based data assimilation (DA) methods, such as the ensemble Kalman filter (EnKF; Evensen 1994; Burgers et al. 1998; Houtekamer and Mitchell 1998), have become popular alternatives to traditional variational DA approaches. EnKFs use short-term ensemble forecasts to calculate flow-dependent, multivariate background error covariances (BECs), contrasting the static, isotropic BECs typically employed in three-dimensional variational data assimilation (3DVAR; e.g., Barker et al. 2004).

* The National Center for Atmospheric Research is sponsored by the National Science Foundation.

Corresponding author address: Craig Schwartz, NCAR, 3090 Center Green Dr., Boulder, CO 80301.
E-mail: schwartz@ucar.edu

DOI: 10.1175/MWR-D-14-00259.1

Flow-dependent BECs can also be incorporated within a variational framework in a ''hybrid'' variational-ensemble DA algorithm (e.g., Hamill and Snyder 2000; Lorenc 2003; Buehner 2005; Wang et al. 2008a; Zhang et al. 2009; Wang 2010; Clayton et al. 2012; Kuhl et al. 2013). Moreover, hybrid paradigms permit flexibility regarding how much the total BECs are weighted toward ensemble and static (i.e., 3DVAR) contributions. Although hybrid analyses are deterministic, since an ensemble is required to provide flow-dependent BECs, hybrid systems are often coupled with EnKFs that update the ensemble each DA cycle (e.g., Wang et al. 2008a,b; Hamill et al. 2011; Wang 2011; Zhang and Zhang 2012; Gao et al. 2013; Schwartz et al. 2013; Wang et al. 2013; Zhang et al. 2013; Pan et al. 2014; Schwartz and Liu 2014).

The hybrid method has shown great promise for initializing numerical weather prediction (NWP) model forecasts. It has been demonstrated that hybrid approaches typically initialize comparable or better forecasts than purely variational methods that do not incorporate ensemble BECs and can outperform forecasts initialized by stand-alone EnKFs (e.g., Wang et al. 2008b; Buehner et al. 2010; Hamill et al. 2011; Wang 2011; Li et al. 2012; Zhang and Zhang 2012; Wang et al. 2013; Zhang et al. 2013; Schwartz et al. 2013; Pan et al. 2014; Poterjoy and Zhang 2014; Schwartz and Liu 2014; Li et al. 2015; Xu et al. 2015). Additionally, the hybrid technique can be easily implemented in preexisting variational DA systems and may produce results similar to those of EnKFs, but with a smaller ensemble (e.g., Wang et al. 2007a; Zhang et al. 2013; Pan et al. 2014). Moreover, as the hybrid employs model-space covariance localization, assimilation of nonlocal observations, such as satellite radiances, may be more effective within hybrid frameworks than within EnKFs that use observation-space localization (Campbell et al. 2010). Given these attractive features and successful hybrid-initialized forecasts, the National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS) model is now initialized with a hybrid-3DVAR system (Wang et al. 2013) and the Met Office uses a four-dimensional variational data assimilation (4DVAR; e.g., Courtier et al. 1994) hybrid system to initialize their global model (Clayton et al. 2012).

Many studies have described limited-area hybrid systems that employ "single resolution" (SR) configurations, where the ensemble providing flow-dependent BECs has the same resolution as the deterministic background and analysis (e.g., Wang et al. 2008b; Wang 2011; Li et al. 2012; Zhang and Zhang 2012; Zhang et al. 2013; Schwartz et al. 2013; Schwartz and Liu 2014; Pan et al. 2014). However, a "dual resolution" (DR) hybrid analysis can be produced that combines a high-resolution (HR) background with a low-resolution (LR) ensemble to produce a HR analysis, obviating the need for a costly HR ensemble.[1] Given the savings afforded by DR hybrid systems—and out of practical necessity—several global hybrid DA configurations have employed DR approaches (e.g., Buehner et al. 2010; Hamill et al. 2011; Clayton et al. 2012; Kuhl et al. 2013), including the operational NCEP GFS hybrid-3DVAR system [as noted in Wang et al. (2013)].

However, perhaps because of the expense of producing HR ensembles, the relative quality of forecasts initialized from DR and SR hybrid analysis–forecast systems has not been thoroughly documented for either global or regional real-data applications. Yet, the performance of SR versus DR hybrid systems has both important practical and scientific consequences, and it is important to consider whether ensemble resolution matters for hybrid DA purposes to enable wise decisions about allocation of computational resources of future operational hybrid DA systems. As the most expensive component of ensemble DA involves advancing an ensemble of forecasts between analyses, if hybrid analyses incorporating flow-dependent BECs provided by an LR ensemble can initialize forecasts with comparable quality as those initialized by hybrid analyses that ingest HR perturbations, considerable computational savings can be realized. Conversely, if increasing the ensemble resolution improves hybrid analyses and subsequent forecasts, increasing ensemble resolution may be justified. Moreover, on a deeper level, the sensitivity of ensemble covariance structures to resolution and how these multivariate BECs interact with hybrid algorithms are interesting scientific questions that may have meaningful implications.

In recognition of these considerations, this paper investigates the performance of limited-area DR and SR hybrid systems. We primarily focus on practical aspects regarding the sensitivity of hybrid analyses and forecasts to the resolution of ensemble perturbations, while delving into the complexities of ensemble correlation structures requires further work. Specifically, we describe the implementation of a DR hybrid analysis system within the community Weather Research and Forecasting (WRF; Skamarock et al. 2008) Model DA system (WRFDA; Barker et al. 2012) that is designed for limited-area modeling applications. Single-observation tests are performed to understand basic properties of the DR analyses. Additionally, we assimilate real observations with the newly developed DR hybrid system by combining 15-km backgrounds and 45-km ensembles in a continuously cycling configuration over a ~3.5-week period. The DR analyses initialized 72-h WRF Model forecasts. Similarly configured 15-km SR hybrid analyses and forecasts were also generated and compared to those produced by the DR system. The DR and SR hybrid systems were coupled to an ensemble adjustment Kalman filter (EAKF; Anderson 2001, 2003) from the Data Assimilation Research Testbed (DART; Anderson et al. 2009) software that updated the ensemble each DA cycle. This work also extends that of Schwartz et al. (2013, hereafter S13), who examined

[1] We note that use of multiple resolutions within DA systems is not confined to hybrid methods. Multiple resolutions are commonly employed in incremental 4DVAR (Courtier et al. 1994) systems, where an HR nonlinear model is used to calculate innovations based on an HR guess field and to define a trajectory about which LR tangent linear and adjoint models are formulated. Moreover, Gao and Xue (2008) and Rainwater and Hunt (2013) discussed the merits of DR nonhybrid ensemble DA systems within idealized frameworks.

45-km 3DVAR and SR hybrid analysis–forecast systems over the same region and time period, and, to our knowledge, represents the first time limited-area DR hybrid analyses assimilating real observations have been produced in conjunction with a limited-area ensemble.

Section 2 describes the DR hybrid algorithm, while section 3 details the WRF Model configurations and DA settings. The experimental design is presented in section 4, and section 5 briefly discusses the observations. Results regarding single-observation experiments are described in section 6, section 7 examines analyses and forecasts produced by continuously cycling DR and SR hybrid systems that assimilated real observations, and we conclude in section 8.

2. The WRFDA dual-resolution hybrid algorithm

WRFDA's hybrid formulation is described by Wang et al. (2008a). The algorithm incorporates BECs from an N-member ensemble into a variational cost function using the extended control variable approach (Lorenc 2003; Wang et al. 2008a), where the total n-dimensional analysis increment vector ($\delta\mathbf{x}$) is written as

$$\delta\mathbf{x} = \mathbf{x}_1 + \sum_{i=1}^{N} \mathbf{a}_i \circ \mathbf{x}'_i. \tag{1}$$

In Eq. (1), $\mathbf{x}_1$ is the n-dimensional analysis increment vector associated with the static BECs (i.e., 3DVAR) and the second term on the right-hand side (rhs) is the increment associated with the ensemble BECs. Vector $\mathbf{x}'_i$ is the perturbation of the ith prior (before assimilation) ensemble member about the prior ensemble mean normalized by $(N-1)^{1/2}$, vector $\mathbf{a}_i$ is an extended control variable for each ensemble member (Lorenc 2003) that determines weighting for the ensemble perturbations, and the symbol $\circ$ denotes a Schur product (element by element multiplication). The fields contained in $\delta\mathbf{x}$ and $\mathbf{x}_1$ were model variables (e.g., wind, temperature, water vapor mixing ratio, and surface pressure). However, during variational minimization, the static portion of the analysis increment (i.e., $\mathbf{x}_1$) was transformed into control variables ($\mathbf{v}$)—streamfunction, pseudo–relative humidity, and unbalanced velocity potential, temperature, and surface pressure—by the relationship $\mathbf{x}_1 = \mathbf{U}\mathbf{v}$, where $\mathbf{U}$ is a transformation matrix. This procedure, also called preconditioning, is quite common (e.g., Barker et al. 2004; Wang et al. 2007b, 2008a; Wang 2010; Clayton et al. 2012, and many others), and delving into the details of preconditioning $\mathbf{x}_1$ is unnecessary to understand the DR hybrid algorithm.

Each $\mathbf{x}'_i$ is a vector of length n_l, where $n_l \le n$, and is composed of model variables that were not transformed into control variable space (note: in term n_l, the subscript l is not a free index). Necessarily, each $\mathbf{a}_i$ is also a vector of length n_l and applied in model space. In an SR hybrid system, $n_l = n$ and the ensemble and background are at identical resolutions. But, in a DR hybrid system, $n_l < n$, meaning the ensemble is at coarser resolution than the background. Therefore, DR hybrid analyses have fewer extended control variables (i.e., $\mathbf{a}_i$) than SR hybrid analyses.

Following Wang (2010), we define $n_l \times n_l$ matrix $\mathbf{d}_i = \mathrm{diag}(\mathbf{x}'_i)$, where diag is an operator that converts vector $\mathbf{x}'_i$ into diagonal matrix $\mathbf{d}_i$, whose pth diagonal element is the pth element of $\mathbf{x}'_i$. Further, let $\mathbf{D}$ be the $n_l \times (Nn_l)$ matrix defined as $\mathbf{D} = [\mathbf{d}_1\, \mathbf{d}_2\, \mathbf{d}_3 \cdots \mathbf{d}_N]$, and concatenate each $\mathbf{a}_i$ to form vector $\mathbf{a}$ of length (Nn_l):

$$\mathbf{a} = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \\ \vdots \\ \mathbf{a}_N \end{bmatrix}. \tag{2}$$

Then,

$$\delta\mathbf{x} = \mathbf{x}_1 + \mathbf{D}\mathbf{a}. \tag{3}$$

Equations (1) and (3) are identical, but Eq. (3) is simpler because it does not contain summations or Schur products. When the ensemble and background are at the same resolution (SR hybrid), Eq. (3) is valid since $n_l = n$ and both terms on the rhs of Eq. (3) are n-dimensional vectors. However, if $n_l < n$, as in a DR application, Eq. (3) is *invalid* since the two terms on the rhs of Eq. (3) are vectors of different lengths. Thus, for DR applications, interpolation of one term is needed. Since we wish to produce HR analyses, we introduce an interpolation operator $\mathbf{L}$ to interpolate the quantity $\mathbf{D}\mathbf{a}$ from LR to HR space.

Strictly, $\mathbf{L}$ is an $n \times n_l$ matrix, where each row of $\mathbf{L}$ specifies how a single HR grid point is related to each LR grid point. While, theoretically, $\mathbf{L}$ could be any interpolation method, we defined $\mathbf{L}$ as the same bilinear interpolation operator used to interpolate the model state to observation locations to use existing WRFDA code.

Introducing $\mathbf{L}$ into Eq. (3) gives

$$\delta\mathbf{x} = \mathbf{x}_1 + \mathbf{L}\mathbf{D}\mathbf{a}. \tag{4}$$

For an SR application ($n_l = n$), $\mathbf{L} = \mathbf{I}$, the identity matrix, and Eq. (3) is recovered. Thus, Eq. (4) is a general expression for the total increment since it is valid even if $n \ne n_l$.

The corresponding cost function (J) that is minimized with respect to $\mathbf{x}_1$ and $\mathbf{a}$ to obtain the hybrid analysis increment is

$$J(\mathbf{x}_1,\mathbf{a}) = \beta_1\frac{1}{2}(\mathbf{x}_1)^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x}_1) + \beta_2\frac{1}{2}(\mathbf{a})^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{a} + \frac{1}{2}(\mathbf{H}\delta\mathbf{x}-\mathbf{y}')^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x}-\mathbf{y}'), \quad (5)$$

where $\delta\mathbf{x}$ is given by Eq. (4). In Eq. (5), $\mathbf{y}'$ is the innovation vector; $\mathbf{B}$ and $\mathbf{R}$ are the background and observation error covariance matrices, respectively; matrix $\mathbf{H}$ is the linearized "observation operator" that interpolates gridpoint values to observation locations and transforms model-predicted variables to observed quantities; and $\mathbf{A}$ is an $(Nn_l) \times (Nn_l)$ block diagonal matrix that controls the spatial correlation of $\mathbf{a}$, effectively performing localization of the ensemble BECs (Wang et al. 2007b). Note that $\mathbf{A}$ is in the ensemble space, while $\mathbf{B}$ is in the space of the background. Moreover, for a sufficiently large ensemble, $\mathbf{A}$ is typically defined with localization length scales substantially larger than the horizontal grid spacing, which constrains $\mathbf{a}$ to be spatially smooth (e.g., Wang 2010) and motivates the potential for successful DR hybrid systems. The terms β_1 and β_2 determine how much weight is given to the ensemble and static BECs and are constrained such that

$$\frac{1}{\beta_1}+\frac{1}{\beta_2}=1. \quad (6)$$

Following Wang (2010), Eq. (5) is minimized by taking its gradient with respect to $\mathbf{x}_1$ and $\mathbf{a}$ and equating with zero, which yields

$$\nabla_{\mathbf{x}_1}J=\beta_1\mathbf{B}^{-1}\mathbf{x}_1+\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x}-\mathbf{y}')=0 \quad (7)$$

and

$$\nabla_{\mathbf{a}}J=\beta_2\mathbf{A}^{-1}\mathbf{a}+\mathbf{D}^{\mathrm{T}}\mathbf{L}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x}-\mathbf{y}')=0. \quad (8)$$

In Eq. (8), $\mathbf{L}^{\mathrm{T}}$ is the adjoint of $\mathbf{L}$, which transforms $\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x} - \mathbf{y}')$ from HR to LR space. Within the context of variational minimization, for DR hybrid applications, each iteration, $\mathbf{L}^{\mathrm{T}}$, is applied to $\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x} - \mathbf{y}')$ and $\mathbf{L}$ is applied to $\mathbf{Da}$. It is unclear how much representativeness error is introduced by interpolating quantities from LR to HR (and vice versa) each iteration, although representativeness errors should increase as the ratio of LR to HR horizontal grid spacing increases. However, since the interpolated quantities are the ensemble contribution to the increment ($\mathbf{Da}$) and the adjoint vector $[\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x} - \mathbf{y}')]$, which are spatially smooth compared to the deterministic hybrid background, these representativeness errors may be somewhat diminished.

3. WRF Model and data assimilation configurations

The WRF Model and DA configurations were very similar to those in S13. Thus, generally brief descriptions follow.

a. Forecast model

Weather forecasts were produced by version 3.3.1 of the nonhydrostatic Advanced Research core of the WRF (Skamarock et al. 2008) Model. All experiments ran over a one-way-nested computational domain encompassing the western Pacific Ocean and eastern Asia (Fig. 1). The horizontal grid spacing was 45 km (222 × 128 grid points) in the outer domain and 15 km (316 × 274 grid boxes) in the inner nest. While testing at higher resolution is desirable, these resolutions were chosen because of limited computational resources. Given these resolutions, we focused on meso-α- to synoptic-scale weather patterns and features and caution that results regarding the relative performance of DR and SR hybrid systems may differ for finer-scale phenomena or modeling systems at higher resolution.

Both domains were configured with 45 vertical levels and a 30-hPa top. The time step was 180 s in the 45-km domain and 60 s in the 15-km nest. GFS forecasts provided lateral boundary condition (LBC) forcing for the 45-km domain every 6 h and the 45-km domain provided LBCs for the 15-km nest. The same physical parameterizations as in S13 were used in both domains and are listed in Table 1.

b. EAKF and hybrid data assimilation settings

The hybrid uses an ensemble of short-term forecasts to incorporate flow-dependent BECs in the variational cost function [i.e., Eq. (5)] and the ensemble needs to be updated when new observations are available. The EAKF from the DART was used to update a 32-member WRF Model–based ensemble. To reduce spurious correlations due to sampling error, localization forced EAKF analysis increments to zero ~1280 km from an observation in the horizontal and ~10 km in the vertical. Adaptive inflation (Anderson 2009) was applied immediately before prior model-simulated observations were computed to maintain ensemble spread. A stochastic kinetic energy backscatter scheme (SKEBS; Shutts 2005; Berner et al. 2011) was applied during the ensemble of WRF Model advances between each EAKF analysis to further preserve spread. SKEBS

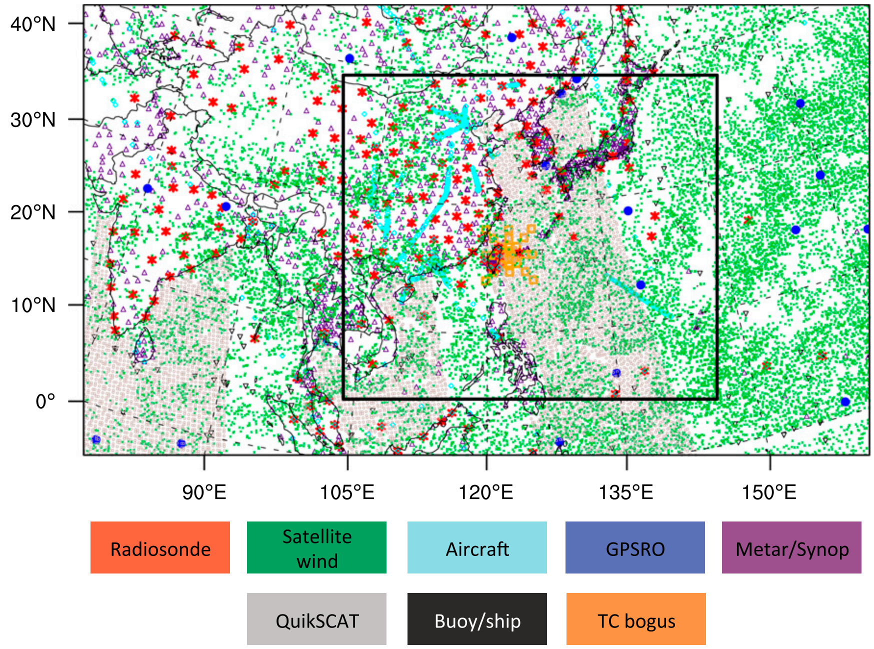

FIG. 1. Computational domain overlaid with observations available for assimilation during the 0000 UTC 13 Sep analysis. The inner box represents the bounds of the 15-km domain, which was nested within the 45-km domain.

parameters were identical for each domain (Table 2), which may be suboptimal, as different horizontal grid lengths may necessitate different SKEBS settings for best performance (e.g., Sanchez et al. 2014). Nonetheless, both the 45- and 15-km ensemble spread–skill relationships were reasonable (section 7a), suggesting that our tunings had the intended effects of engendering appropriate ensemble spread. Furthermore, in experiments without application of SKEBS during ensemble model advances, the ensemble spread was clearly insufficient.

Localization was also applied in the hybrid to limit the spatial extent of the ensemble contribution to the analysis increments. Horizontal localization of approximately the same length scale in DART was applied in the hybrid (i.e., ~1280 km). Vertical localization length scales in the hybrid increased with height (see S13 for more information). For consistency with the EAKF, the prior ensemble perturbations were inflated before they were ingested into hybrid analyses using the same algorithm DART employed to inflate prior ensembles.

Static 45- and 15-km BECs used in the hybrid algorithm were constructed using the National Meteorological Center (NMC, now known as NCEP) method (Parrish and Derber 1992) from WRF Model forecasts produced over this domain for multiple months and used operationally by the Central Weather Bureau (CWB) of Taiwan, as described by S13. Three outer loops (Courtier et al. 1994) were used in the hybrid minimization. As in S13, hybrid BECs were weighted 75% toward the ensemble contribution and 25% toward the static (i.e., 3DVAR) component. We also weighted the BECs equally between the ensemble and static contributions and achieved similar results. Limited sensitivity to BEC weightings in SR hybrid configurations has also been noted elsewhere (e.g., Wang 2011; Wang et al.

TABLE 1. Physical parameterizations used in both WRF domains.

Physical parameterization	WRF option	Reference
Microphysics	Goddard	Tao and Simpson (1993); Tao et al. (2003)
Longwave radiation	Rapid Radiative Transfer Model	Mlawer et al. (1997)
Shortwave radiation	Goddard	Chou and Suarez (1994)
Planetary boundary layer	Yonsei University	Hong et al. (2006)
Land surface model	Noah	Chen and Dudhia (2001)
Cumulus parameterization	Kain–Fritsch with modified trigger function	Kain and Fritsch (1990, 1993); Kain (2004); Ma and Tan (2009)

TABLE 2. SKEBS parameters.

Parameter	Value
Backscatter dissipation rate for streamfunction	$6 \times 10^{-6}\,\mathrm{m^2\,s^{-3}}$
Backscatter dissipation rate for temperature	$5 \times 10^{-7}\,\mathrm{m^2\,s^{-3}}$
Decorrelation time	~30 min
Power law for streamfunction perturbations	−1.83
Power law for temperature perturbations	−1.83

2013; Zhang et al. 2013), but Wang et al. (2013) stated that in preliminary testing, forecasts were improved in a global DR hybrid-3DVAR system when the total BECs were weighted equally between the static and ensemble contributions compared to when ensemble BECs provided the total BECs (i.e., no static contribution).

4. Experimental design

Three hybrid experiments were designed to investigate the performance of limited-area DR and SR hybrid analysis–forecast systems. All experiments began at 0000 UTC 4 September by interpolating the deterministic 0.5° × 0.5° NCEP GFS analysis onto the nested computational domain (Fig. 1). The initial 45-km ensemble was constructed at this time by taking Gaussian random draws with zero mean and static BECs (Barker 2005; Torn et al. 2006) and adding them to the GFS field. LBCs for the ensemble system were perturbed similarly. The initial 15-km ensemble was produced by downscaling the perturbed 45-km fields onto the 15-km grid, similar to Ha and Snyder (2014).

The deterministic and ensemble fields produced at 0000 UTC 4 September initialized 6-h WRF Model forecasts, which served as backgrounds for the first hybrid and EAKF analyses at 0600 UTC 4 September. Thereafter, the EAKF and hybrid configurations cycled continuously until 0000 UTC 28 September, and a new analysis was produced every 6 h. The background for DA was always the previous cycle's 6-h forecast. Nested 45- and 15-km 72-h WRF Model forecasts were initialized every 6 h from hybrid analyses between 1800 UTC 8 September and 0000 UTC 28 September (inclusive; 78 total forecasts). Identical to S13, digital filter initialization (DFI; Lynch and Huang 1992; Huang and Lynch 1993) using a twice-DFI scheme and the Dolph filter (Lynch 1997) with a 2-h backward integration was applied to all 72-h forecasts, but not during the 6-h cycling between analyses. S13 examined this same period and employed an identical experimental design, but they only produced 45-km SR hybrid analyses. Thus, while 45-km analyses and forecasts were necessarily produced here because of WRF nesting, we focus exclusively on the 15-km forecasts.

When 15-km EAKF analyses were required, the EAKF produced separate, independent 45- and 15-km analyses. The 45- and 15-km prior ensembles produced by cycling EAKF-WRF systems were used as input into hybrid analyses. Like the EAKF, all hybrid experiments produced separate, independent 45- and 15-km analyses. The three hybrid experiments differed in the resolution of the ensemble perturbations ingested by the 15-km hybrid analyses (which determined whether 15-km EAKF analyses and ensemble forecasts were needed) and whether the EAKF analysis ensemble was recentered about the hybrid analysis (e.g., Zhang et al. 2013; Wang et al. 2013; Pan et al. 2014):

1) "Hybrid_SR"—separate, independent SR 45- and 15-km hybrid analyses were produced each DA cycle. The 45-km hybrid analyses incorporated BECs from the cycling 45-km EAKF-WRF ensemble system, while the 15-km hybrid analyses used BECs from the cycling 15-km EAKF-WRF ensemble system. Since 15-km ensembles were needed for the 15-km SR hybrid, each ensemble member was advanced between analysis cycles with the 15-km nest embedded within the 45-km domain. EAKF analysis ensembles were not recentered about hybrid analyses. Because of the necessity of 15-km ensembles, this experiment was the most computationally expensive. This experiment's procedure is illustrated in Fig. 2.
2) "Hybrid_DR_1way"—45-km hybrid analyses were produced as in Hybrid_SR, but ensemble BECs for 15-km hybrid analyses were provided by 45-km prior ensembles. Thus, the same 45-km ensembles provided BECs for 45-km SR hybrid analyses and 15-km DR hybrid analyses. Since 15-km ensembles were *not* required, the EAKF-WRF ensemble system performed solely 45-km analyses, allowing for the *removal* of the 15-km nest during the ensemble of WRF Model advances between EAKF analyses, enabling considerable savings compared to Hybrid_SR. EAKF analysis ensembles were not recentered about hybrid analyses. Omission of the recentering step in Fig. 3 yields this experiment's methodology.
3) "Hybrid_DR_2way"—identical to Hybrid_DR_1way, except the 45-km EAKF analysis ensembles were recentered about hybrid analyses. Again, 15-km ensembles were not needed, so the EAKF-WRF ensemble system ran solely at 45-km horizontal grid spacing. To perform recentering, first, the 15-km hybrid analyses were upscaled to 45 km and replaced the 45-km hybrid analyses over the 45-km geographic region collocated with the 15-km grid. Then, each 45-km EAKF analysis ensemble member was recentered about the 45-km hybrid analysis that contained the

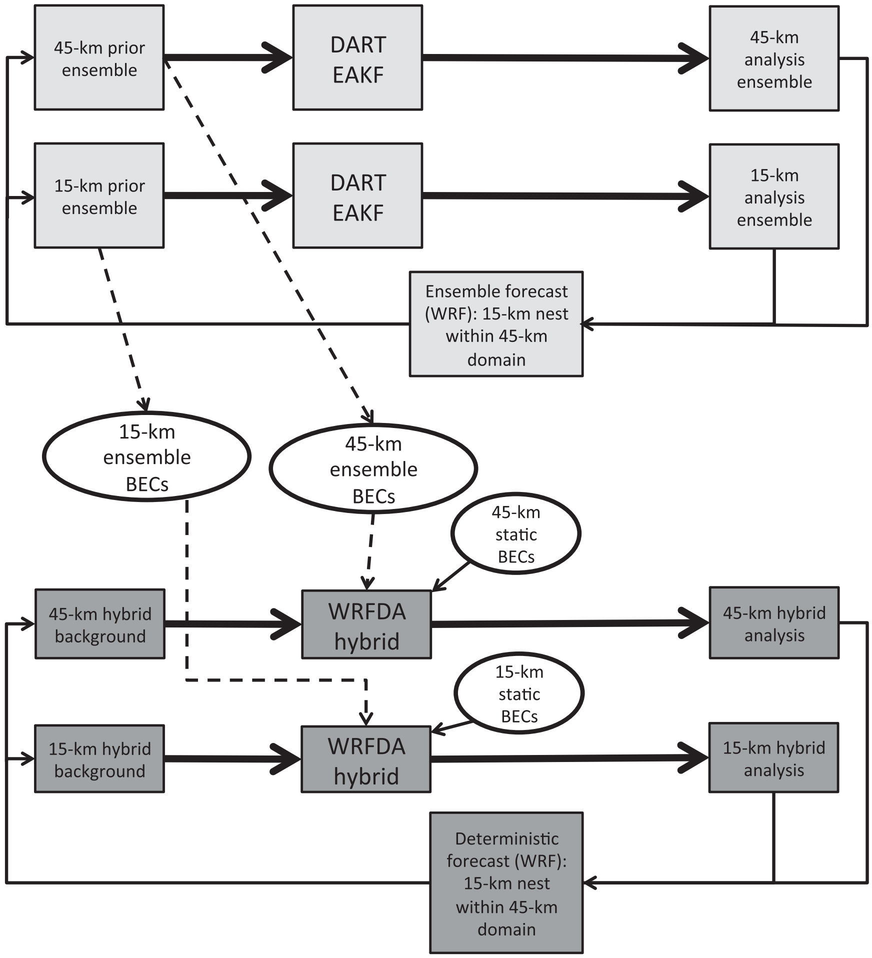

FIG. 2. Flow chart describing a cycling EAKF and single-resolution hybrid system where separate, independent 45- and 15-km EAKF and hybrid analyses are performed.

upscaled 15-km hybrid analysis information. Figure 3 exactly depicts this experiment's procedure. The cost of recentering was negligible, and this experiment had a similar cost as Hybrid_DR_1way.

A fourth experiment ("3DVAR") configured exactly as Hybrid_DR_1way was also performed, except pure 3DVAR analyses were produced in both the 45- and 15-km domains. This experiment was a control for and considerably cheaper than the hybrid experiments because it did not require an ensemble. However, as SR WRFDA-hybrid analyses have been shown to regularly initialize better forecasts than WRFDA-3DVAR analyses (e.g., Wang et al. 2008b; Wang 2011; Barker et al. 2012; Li et al. 2012; Zhang and Zhang 2012; S13; Zhang et al. 2013; Poterjoy and Zhang 2014; Li et al. 2015; Xu et al. 2015), this work does not emphasize the relative performance of the 3DVAR and SR hybrid experiments and focuses on comparing the DR and SR hybrid systems. Nonetheless, it is important to consider whether Hybrid_DR_1way can initialize 15-km forecasts that outperform those initialized by a pure 3DVAR system to justify its much greater cost.[2]

Comparison of Hybrid_SR with Hybrid_DR_1way assesses the sensitivity to the resolution of the ensemble

[2] We recognize that different 15-km Hybrid_DR_1way and 3DVAR forecasts cannot be solely attributed to differing 15-km analysis systems because the 45-km forecasts providing LBCs for the 15-km domain differed. To quantify the impact of 45-km LBCs on 15-km forecasts, an auxiliary experiment was performed where hybrid analyses were produced in the 45-km domain (as in Hybrid_DR_1way) but 3DVAR analyses were produced in the 15-km domain. This mixed hybrid–3DVAR experiment cost substantially more than the pure 3DVAR experiment because 45-km ensembles were required, yet initialized similar quality 15-km forecasts as the pure 3DVAR experiment. Therefore, the impact of the 45-km LBCs was small, and we attribute 15-km forecast differences between the pure 3DVAR and Hybrid_DR_1way experiments to the 15-km analysis algorithms.

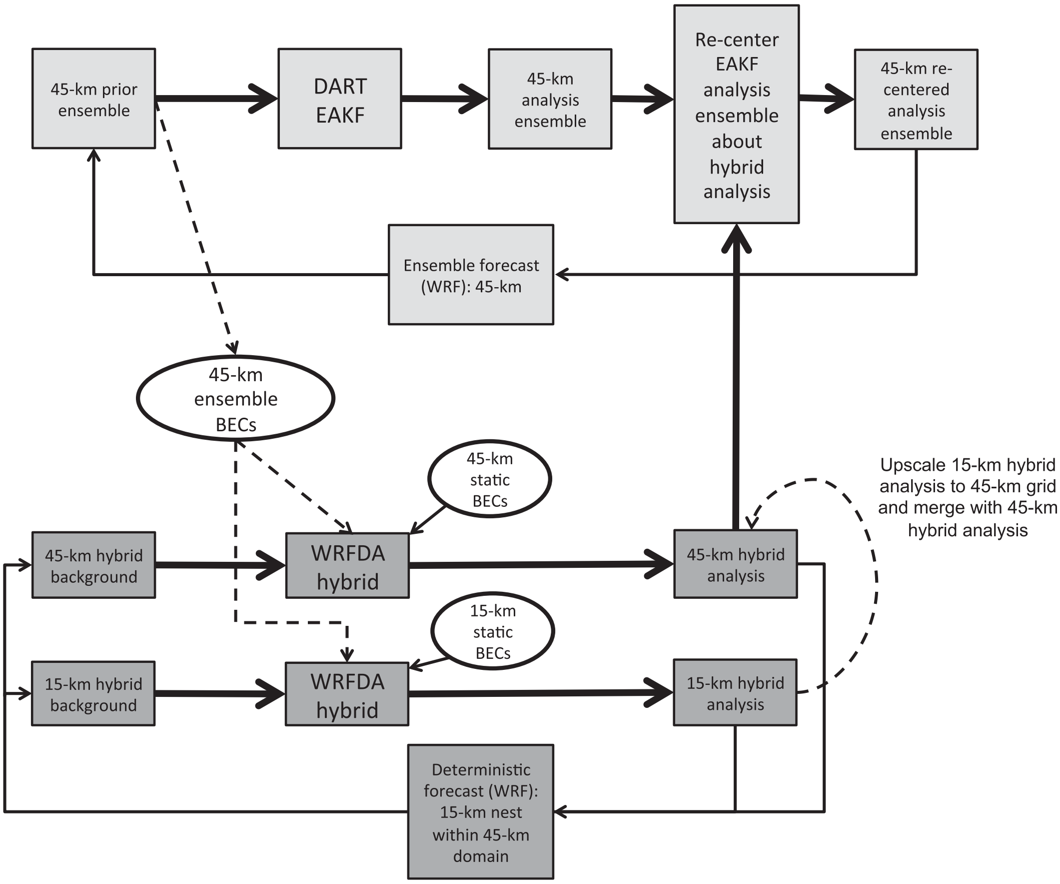

FIG. 3. Flow chart describing a cycling EAKF and DR hybrid system where the EAKF analysis ensemble is recentered about the hybrid analysis.

perturbations, while comparing Hybrid_DR_1way with Hybrid_DR_2way isolates whether recentering benefits DR hybrid systems. Wang et al. (2013) and Pan et al. (2014) noted little practical difference between SR hybrid systems with and without recentering steps. Additionally, S13 performed 45-km SR hybrid analyses for this period and domain and noted little sensitivity to whether recentering occurred, so, here, SR analyses with EAKF recentering were not performed.

Results from these experiments are presented in section 7.

5. Observations

As in S13, the WRFDA and EAKF systems assimilated different observations, as summarized in Table 3. Because of this difference, we did not compare forecasts initialized from EAKF mean and hybrid analyses and solely used the EAKF system to produce ensembles for hybrid DA purposes. Observations taken within ±3 h of each analysis time were assimilated and observations were assumed to be valid at the analysis time. A typical distribution of observations available for assimilation at 0000 UTC is shown in Fig. 1. At this time, bogus tropical cyclone (TC) observations produced as in Hsiao et al. (2010) were distributed around Typhoon Sinlaku, and a similar spatial distribution of TC bogus observations was used for other TCs. Analyses in both domains only assimilated observations located within their bounds, meaning the 15-km analyses assimilated fewer observations than the 45-km analyses.

All observations were subject to various forms of quality control, as in S13. Observations above the model top were excluded from assimilation and at stations where multiple observations were received during the ±3-h time window, only the observation nearest the analysis time was kept. Additionally, "outlier checks" were applied. In the hybrid, an observation was not assimilated if its innovation exceeded $5\sigma_o$, where σ_o is the observation error standard deviation. As in S13, a different outlier check was applied in DART compared to that in the hybrid to account for

TABLE 3. Assimilated meteorological observations in the WRFDA-hybrid and DART systems. See Schwartz et al. (2013) for more information.

Observing platform	Observation type assimilated in WRFDA-hybrid	Observation type assimilated in DART	Notes
Radiosonde	Surface pressure	Surface pressure	
	Temperature	Temperature	
	Specific humidity	Specific humidity	
	Wind	Wind	
Aircraft	Temperature	Temperature	DART: superobbed into 100 km × 100 km × 25 hPa boxes
	Wind	Wind	
Global positioning system radio occultation (GPSRO)	Refractivity	Refractivity	
Satellite-tracked winds	Wind	Wind	DART: assimilated over water only
			DART: superobbed in 100 km × 100 km × 25 hPa boxes
QuikScat	Wind	Not assimilated	WRFDA-hybrid: assimilated over water only
Ship and buoy	Surface pressure	Surface pressure	
	Temperature	Temperature	
	Specific humidity	Specific humidity	
	Wind	Wind	
SYNOP and METAR	Surface pressure	Surface pressure	
	Temperature		
	Specific humidity		
	Wind		
Bogus	Temperature		DART: only assimilated at 700 hPa
	Specific humidity	Relative humidity	
	Wind	Wind	

ensemble spread. Specifically, the EAKF did not assimilate an observation if the ensemble mean innovation was greater than 3 times the square root of the sum of σ_o^2 and σ_f^2, where σ_f^2 is the ensemble variance of the simulated observation.

6. Results: Single-observation experiments

To understand hybrid analysis sensitivity to the resolution of ensemble perturbations, two separate sets of hybrid analyses were performed where solely a single observation was assimilated. The two sets differed by the location of the observation—one was placed within a strong typhoon and the other in westerly flow. Within each set, SR and DR hybrid analyses were performed that differed by the resolution of the ensemble perturbations. The SR analyses used the 15-km ensemble produced in Hybrid_SR to provide BECs whereas the DR analyses used BECs provided by the 45-km ensemble produced in Hybrid_DR_1way. To ensure that analysis differences were solely attributable to the different ensembles, the background for *all* single-observation experiments was the 15-km Hybrid_DR_1way background valid at 0000 UTC 12 September. As in the real data experiments (section 7), the ensemble (static) BECs contributed 75% (25%) to the total BECs.

a. *Single observation in typhoon core*

A single 500-hPa temperature observation placed near the center of Typhoon Sinlaku that was 2 K colder than the background (i.e., innovation of −2 K) with an error standard deviation of 1 K was assimilated. There were many differences between the 15-km SR and DR increments. For example, the SR hybrid 500-hPa potential temperature (θ) increments (Fig. 4a) were more negative near the observation than the DR hybrid analysis increments (Fig. 4b), indicating the SR analysis more closely fit the observation. Additionally, while both increments were positive west of the observation, northeast of Taiwan, the DR increments were slightly negative or neutral while the SR increments were positive. Furthermore, the DR analysis had a greater area of negative increments north and east of the observation. Everywhere, the SR increments had more finescale detail than the DR increments, and the circulation around Sinlaku was more prominent in the SR increments.

Similarly, near the observation location, the 15-km 500-hPa water vapor mixing ratio increments (Figs. 5a,b) were larger in the SR analysis. While the DR and SR moisture increments were broadly similar west of ~123°E, there were substantial differences near and east of the observation. Specifically, the DR increments were more

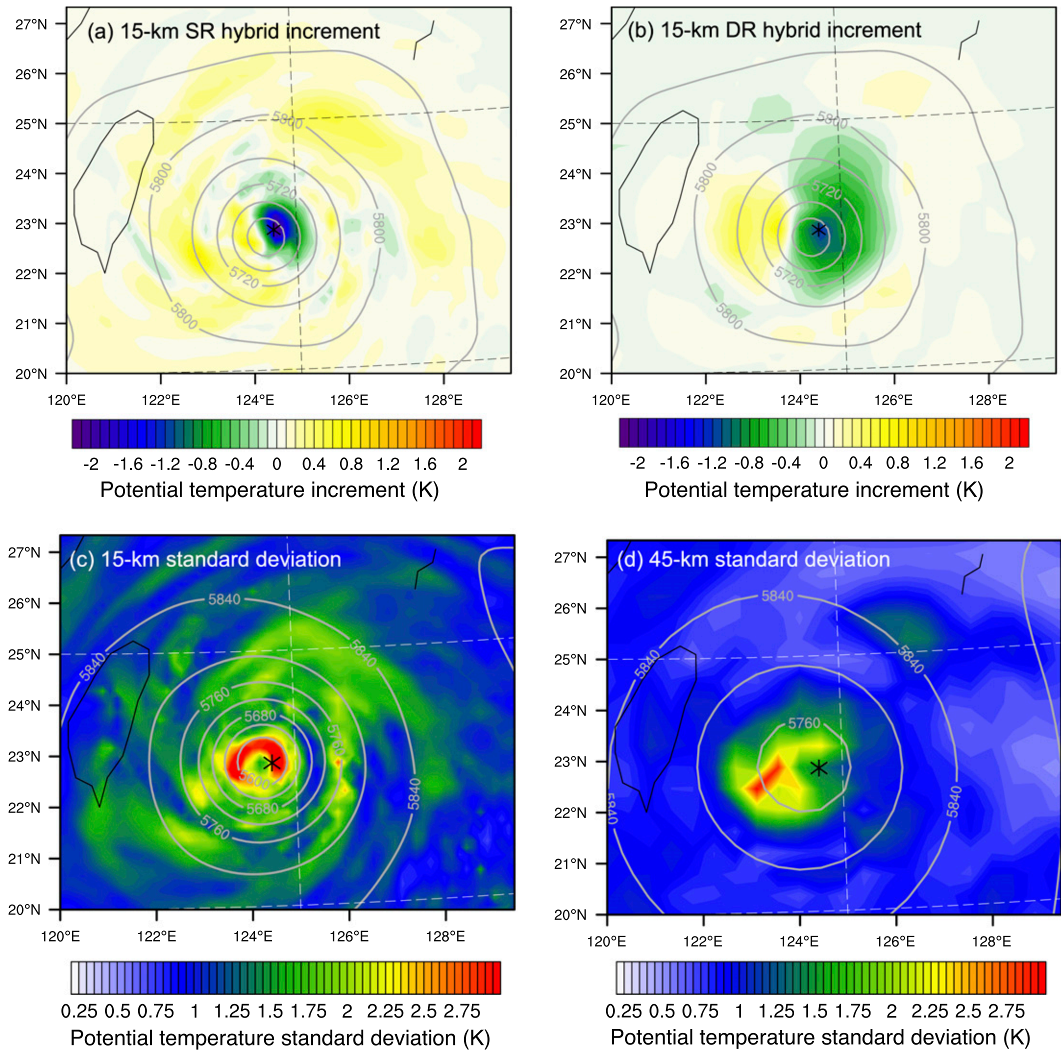

FIG. 4. The 15-km, 500-hPa potential temperature analysis increments at 0000 UTC 12 Sep for (a) SR and (b) DR analyses that assimilated a single observation at the location indicated by asterisks. The background 500-hPa height (m; contoured every 40 m) is overlaid. The 500-hPa potential temperature (c) 15- and (d) 45-km prior ensemble standard deviations (after inflation) at 0000 UTC 12 Sep overlaid with the ensemble mean prior 500-hPa height (m; contoured every 40 m). The asterisks in (c) and (d) mark the locations of the single assimilated observation that produced increments in (a) and (b). Note that the height fields in (a),(b) differ from those in (c),(d) because the heights in (a),(b) were from the deterministic background while those in (c),(d) were from the ensemble mean.

negative immediately west of the observation, and the SR and DR increments had opposite signs at many locations east of ~125°E. Both increments clearly captured the circulation around the typhoon, illustrating the incorporation of flow-dependent BECs in the hybrid, but the SR increments featured more banded structures and greater detail than did the DR increments.

Those disparities between the SR and DR hybrid increments can largely be explained by differences regarding the 45- and 15-km ensembles that provided the BECs for the analyses. Figures 4c and 4d show the 15- and 45-km ensemble standard deviations of 500-hPa θ (after inflation[3]) at 0000 UTC 12 September overlaid with the ensemble mean 500-hPa height. The 15-km ensemble had a stronger TC than the 45-km ensemble, consistent with the expectation that HR models can better resolve strong TCs than LR models (e.g., Xue et al. 2013). In most areas, the 15-km ensemble had

[3] Here and throughout the paper, qualitatively, examination of the 15- and 45-km standard deviations *before* inflation yielded identical conclusions compared to assessing the spreads after inflation.

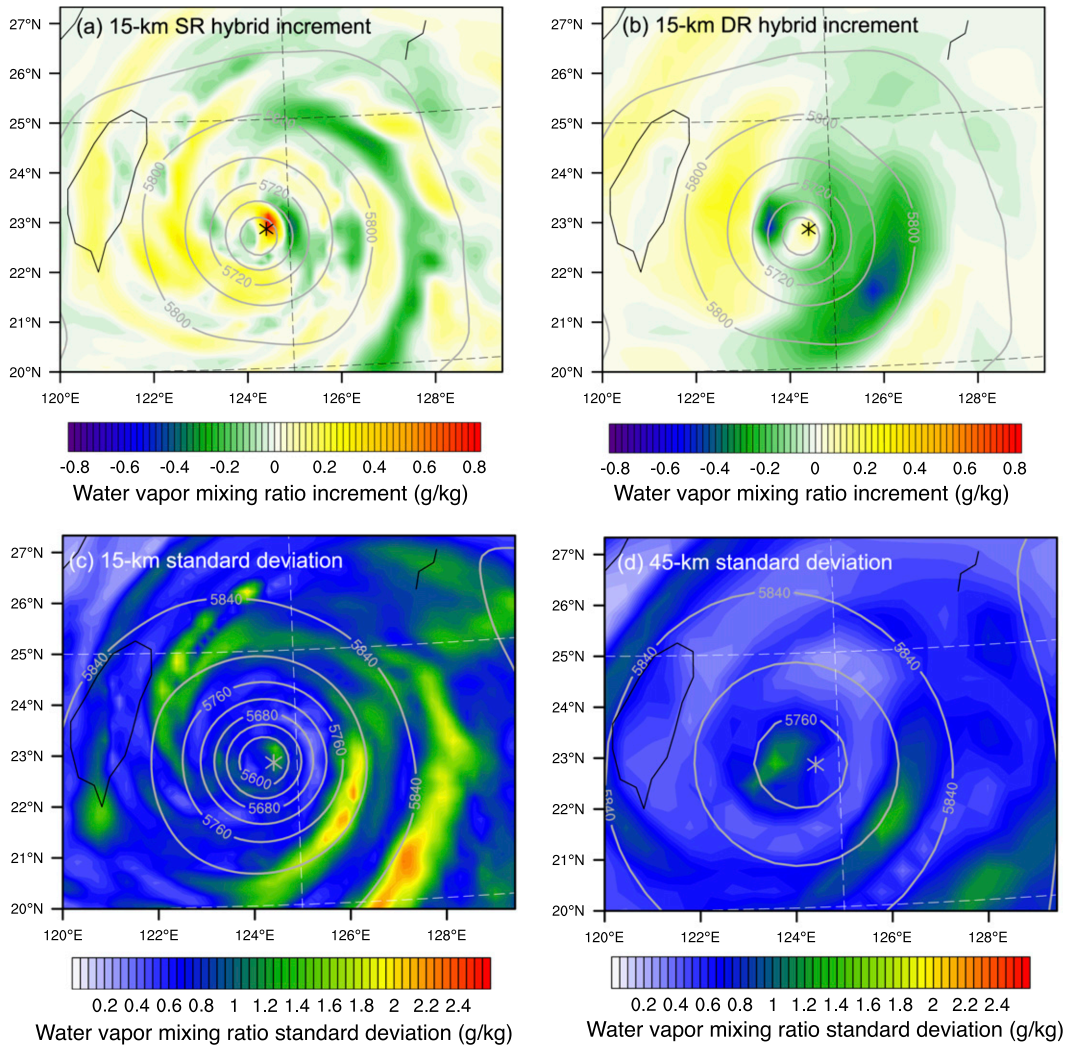

FIG. 5. As in Fig. 4, but for 500-hPa water vapor mixing ratio.

larger θ spread than the 45-km ensemble, which permitted the SR analysis greater ability to adjust toward the observation than the DR analysis. The 15-km ensemble θ spread was organized into bands associated with the TC, while the 45-km ensemble θ spread had less-coherent spiraling structures. However, the 45-km 500-hPa ensemble water vapor mixing ratio spread (after inflation) more clearly reflected the TC, but the 15-km spread again had more banding and generally larger standard deviations (Figs. 5c,d). Overall, the SR and DR increments usually reflected the ensemble spreads, as the largest increments often corresponded to those regions where ensemble spread was greatest.

b. *Single observation in midlatitude westerly flow*

The second set of single-observation experiments assimilated a 500-hPa temperature observation placed at 35°N, 120°E, on the southern periphery of the jet stream. Again, the observation error standard deviation and innovation were 1 and −2 K, respectively. For this case, the SR and DR 500-hPa θ increments were remarkably similar (Figs. 6a,b), although the SR increments again had finer structures. Furthermore, the 500-hPa 45- and 15-km θ spreads (after inflation) over this region were broadly similar (Figs. 6c,d) and small compared to spread near the TC core. Thus, the increments were smaller than those near the TC core. For other meteorological variables and vertical levels, the DR and SR increments were also very similar (not shown).

c. *Discussion*

The extent of the differences between the SR and DR hybrid analysis increments depended on the nature of the flow. These single-observation tests suggest that DR and

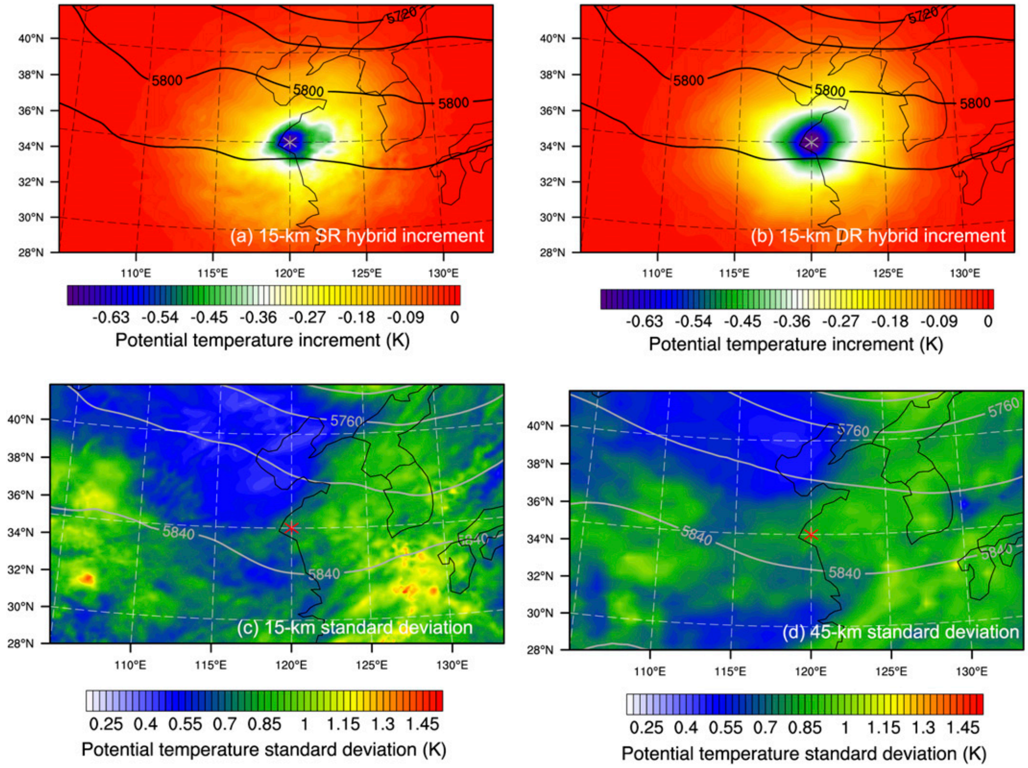

FIG. 6. As in Fig. 4, but increments were engendered by assimilation of a different observation, whose location is indicated by the asterisks.

SR hybrid analyses will potentially be most disparate around small-scale features that HR ensembles can better resolve than LR ensembles. In these cases, HR ensembles can be expected to better represent uncertainty, which should lead to more spread compared to LR ensembles. Conversely, in regimes where synoptic-scale flow dominates, HR and LR ensembles are more likely to resolve features similarly, and these single-observation tests suggest that for large-scale patterns, SR and HR hybrid analyses may be quite similar.

The next section objectively verifies the analyses and forecasts produced by the SR and DR hybrid systems that assimilated real observations.

7. Results: Real-data experiments

Model output was compared to TC track, radiosonde, and dropwindsonde observations. Aspects of the ensemble forecasts were also examined since they are important inputs to the hybrid. The first ~5 days of the simulations were excluded from all verification statistics to allow ample time for the ensemble to ''spin up'' from the initial, randomly generated ensemble.

Statistical significance was assessed by a bootstrap resampling technique (Wilks 2006). For each experiment and 1000 iterations, random samples (with replacement) were drawn from the distribution of daily error statistics and aggregate error statistics were computed from the daily resamples. The 90% confidence interval (CI) was estimated from the distribution of the resampled aggregate statistics. If the bounds of two experiments' CIs did not overlap, then the two experiments had statistically significant (SS) differences at the 95% level.

Bootstrap CIs were also computed based upon pairwise differences of two experiments' errors (e.g., Hamill 1999; Davis et al. 2010; Schwartz and Liu 2014), which, assuming the two distributions have similar variances, is more robust than and yields larger significance levels compared to bootstrapping distributions separately. However, for metrics other than TC track errors, bootstrap CIs based on pairwise differences yielded identical conclusions as when CIs were computed for individual distributions. Thus, for ease of presentation, aside from TC track errors, we present bootstrap CIs based on unpaired resamples.

a. Ensemble performance

A high quality prior ensemble is instrumental in performing successful hybrid analyses. In a well-calibrated

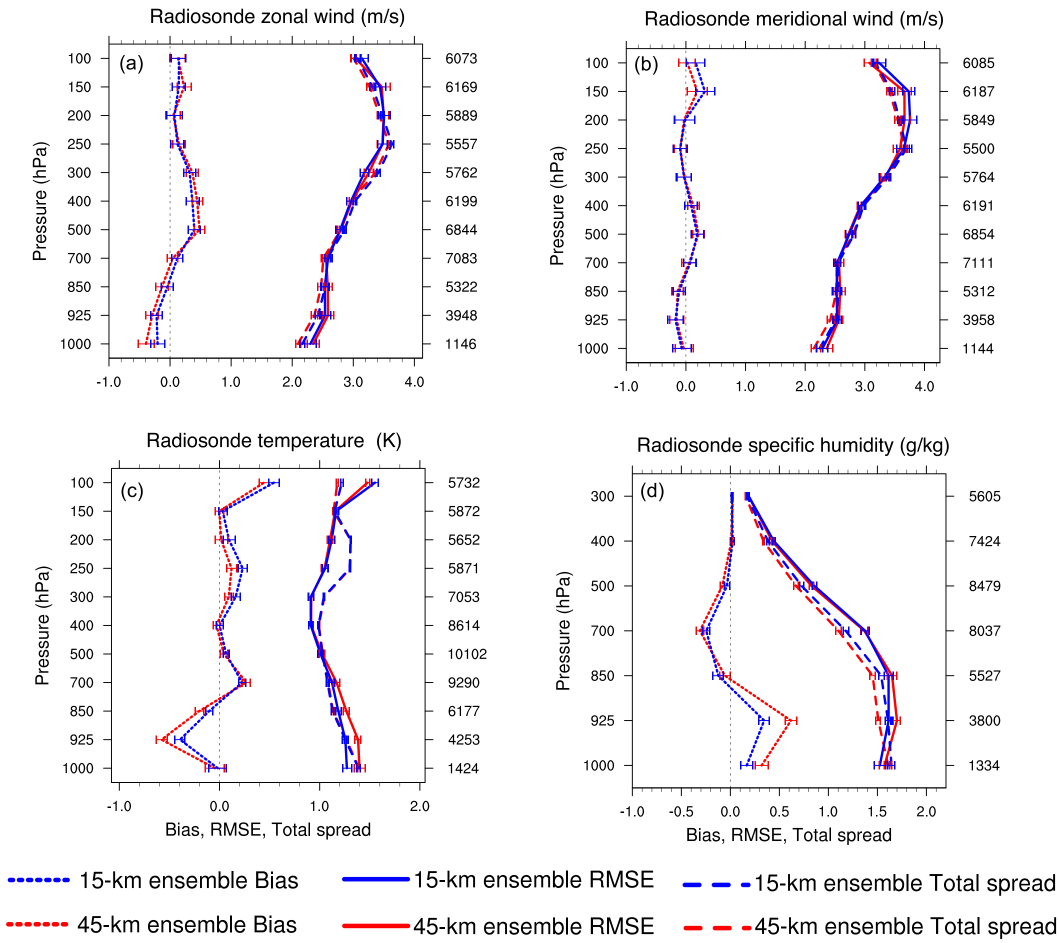

FIG. 7. Average prior total spread, ensemble mean RMSE, and ensemble mean bias (after inflation) of radiosonde (a) zonal wind ($m\,s^{-1}$), (b) meridional wind ($m\,s^{-1}$), (c) temperature (K), and (d) specific humidity ($g\,kg^{-1}$) between 1800 UTC 8 Sep and 0000 UTC 28 Sep. The sample size at each pressure level is shown at the right of each panel. Error bars denote bounds of 90% confidence intervals.

EnKF analysis–forecast system, when compared to observations, the prior ensemble mean root-mean-square error (RMSE) will equal the prior "total spread," defined as the square root of the sum of the observation error variance and prior ensemble variance of the simulated observations (Houtekamer et al. 2005). Therefore, the ratio of the prior total spread to the prior ensemble mean RMSE, called the "consistency ratio" (CR; Dowell and Wicker 2009), should equal 1 in a well-calibrated system. CRs < 1 indicate insufficient ensemble spread.

To enable comparison between the 45- and 15-km prior ensembles, verification occurred against a dataset composed solely of radiosonde observations assimilated by both the 15- and 45-km EAKFs. The 15- and 45-km ensembles were produced in Hybrid_SR and Hybrid_DR_1way, respectively. The prior RMSE, total spread, and ensemble mean additive bias (after inflation) aggregated between 1800 UTC 8 September and 0000 UTC 28 September are shown in Fig. 7 for radiosonde observations. Both ensembles had comparable wind biases and RMSEs (Figs. 7a,b), and the total spread agreed well with the RMSEs at most levels. The 45-km ensemble had statistically significantly poorer temperature biases and RMSEs (Fig. 7c) than the 15-km ensemble at 850 and 925 hPa but performed comparably to or better than the 15-km ensemble at higher levels. For temperature observations, both ensembles had similar total spread that was greater than the corresponding RMSEs between ~400 and 200 hPa. Regarding specific humidity, at 500, 700, and 850 hPa, both ensembles had comparable RMSEs and dry biases (Fig. 7d). However, at and below 925 hPa, the 15-km ensemble had lower RMSEs than the 45-km ensemble and there were moist biases, although the 15-km ensemble bias was statistically

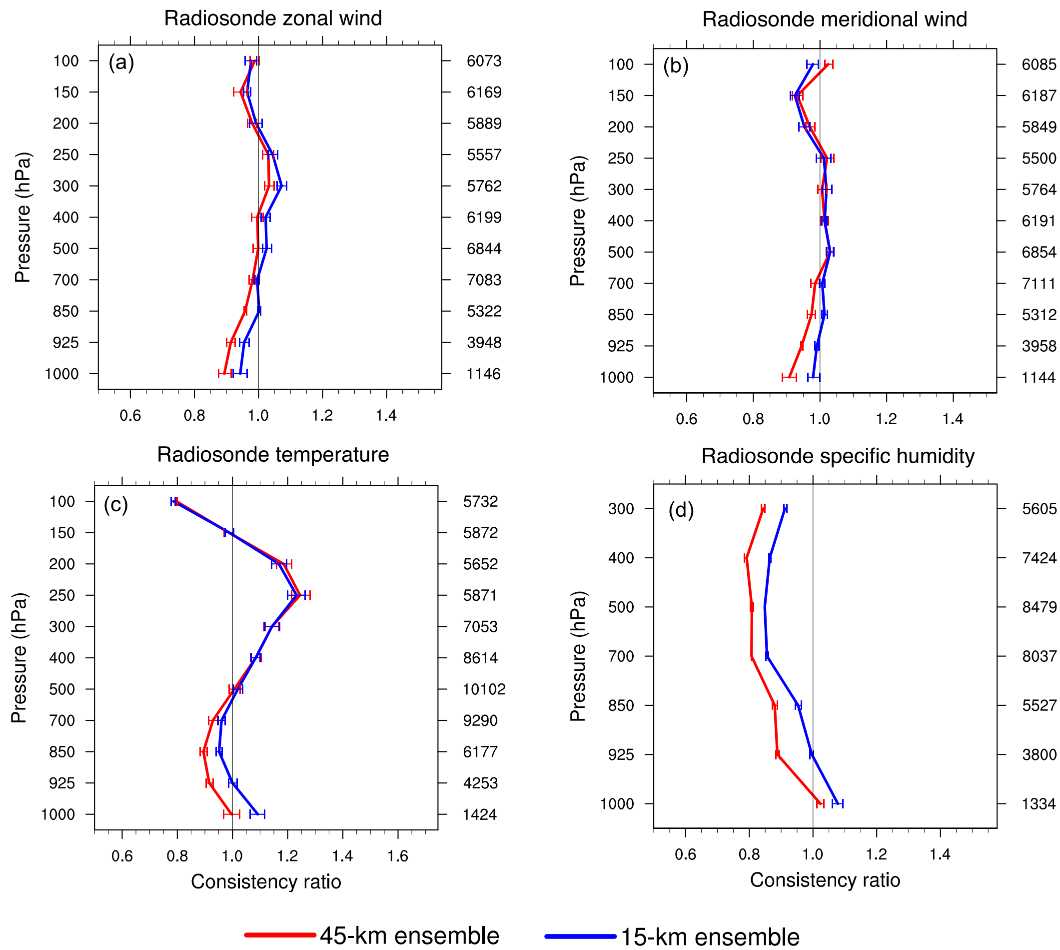

FIG. 8. As in Fig. 7, but for consistency ratios.

significantly smaller. Throughout the column, the 15-km ensemble had more moisture spread than the 45-km ensemble, but both ensembles had insufficient spread at most levels.

Both ensembles had CRs near 1 at most levels for wind (Figs. 8a,b), with the 15-km ensemble performing significantly better below 700 hPa. For temperature observations (Fig. 8c), at and above 500 hPa, the 45- and 15-km ensembles had comparable CRs, but below 500 hPa the 15-km ensemble had CRs closer to 1 than the 45-km ensemble, except near 1000 hPa. Similarly, 45-km CRs for specific humidity were significantly closer to 1 than the 15-km ensemble near 1000 hPa (Fig. 8d), but at all other levels, the 15-km CRs for moisture were statistically significantly nearer to 1 than the 45-km CRs.

It is also interesting to examine the spatial distribution of the 45- and 15-km ensemble spreads. The average prior ensemble standard deviation (after inflation) of 500-hPa wind speed between 1800 UTC 8 September and 0000 UTC 28 September (Figs. 9a,b) was smallest over eastern China, where observations were plentiful, and portions of the Pacific Ocean, where there was little uncertainty about the location of the subtropical high pressure system. The 15-km ensemble had slightly higher spread in most areas. Similar patterns were evident with the mean 500-hPa potential temperature spread (Figs. 9c,d). A local spread maximum was evident in both 500-hPa wind and potential temperature southeast of Taiwan, where three TCs moved, reflecting the uncertainty of TC prediction.

Consistent with Fig. 9, the 15-km ensemble typically had more spread than the 45-km ensemble throughout the column, as evidenced by the domain average prior ensemble standard deviations (after inflation) between 1800 UTC 8 September and 0000 UTC 28 September (Fig. 10). The 45-km statistics were computed solely over the portion of the 45-km domain collocated with the 15-km nest. At most levels for wind and water vapor mixing ratio (Figs. 10a,b,d), the 15-km ensemble had greater spread than the 45-km ensemble, but the 15-km

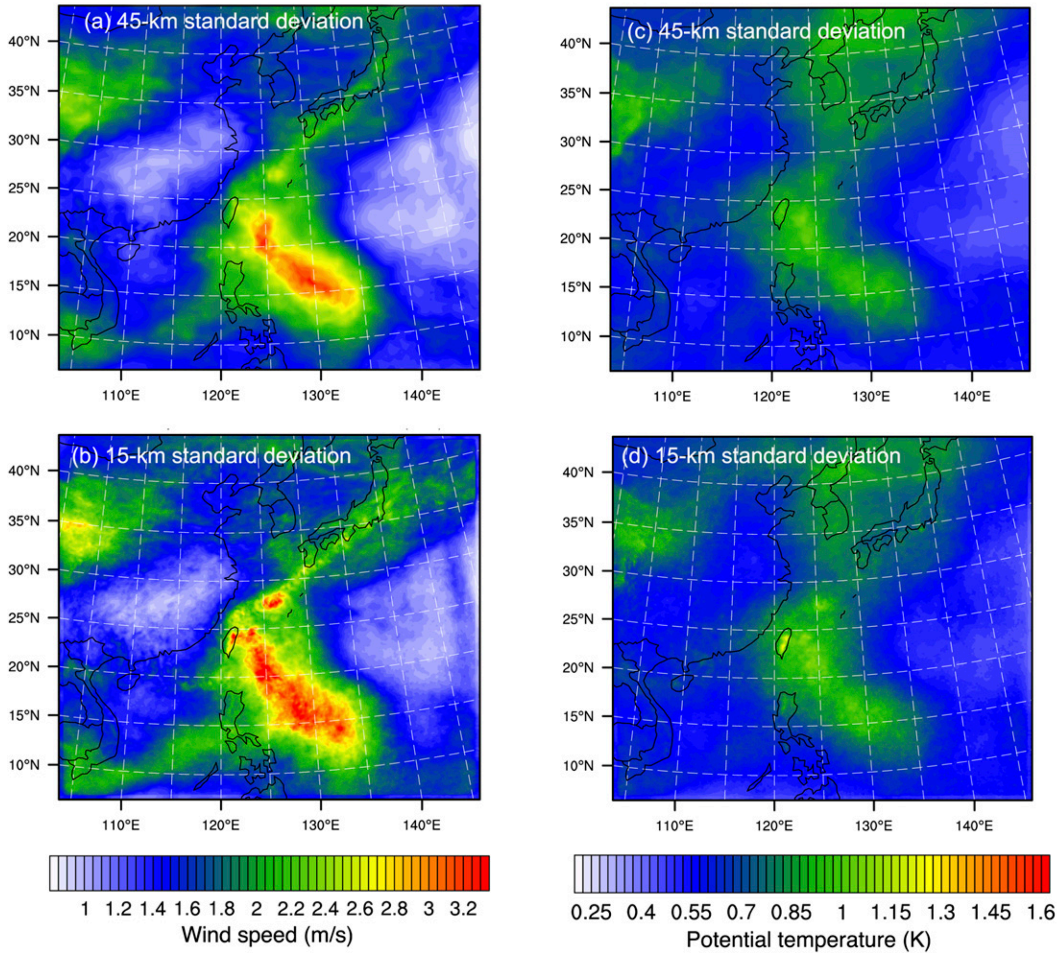

FIG. 9. Average prior ensemble standard deviation (after inflation) of 500-hPa (a),(b) wind speed (m s^{-1}) and (c),(d) potential temperature (K) between 1800 UTC 8 Sep and 0000 UTC 28 Sep for the (a),(c) 45- and (b),(d) 15-km ensembles.

ensemble spread was typically at most 10% greater than the 45-km ensemble spread. Differences between the 15- and 45-km ensemble potential temperature spread (Fig. 10c) were small compared to those for other variables except above model level 40.

Overall, both the 15- and 45-km ensembles were reasonably well calibrated, as CRs were typically within 10% of 1 for most levels and variables. The 15-km ensemble CRs were usually comparable to or better than the 45-km CRs, and the 15-km ensemble performed notably better than the 45-km ensemble below ~700 hPa, particularly for temperature and moisture. Additionally, the 15-km ensemble had greater spread than the 45-km ensemble, which is sensible, since errors on HR grids grow faster than those on LR grids (e.g., Lorenz 1969). Yet, the differences in spread were usually small, and the next subsection assesses how these different ensemble spreads impacted the DR and SR hybrid analysis systems.

b. Mean hybrid background and analysis characteristics

Fits to observations were aggregated over each 15-km hybrid background (6-h forecasts) and analysis between 1800 UTC 8 September and 0000 UTC 28 September (78 total). All backgrounds had similar aggregate fits to radiosonde observations at most levels (Fig. 11), which suggests all hybrid systems had similar quality. Additionally, there were no SS differences regarding aggregate analysis fits to radiosonde observations (not shown). However, Hybrid_SR analysis root-mean-square fits compared to radiosondes were smaller than those of Hybrid_DR_1way for wind and specific humidity, which is consistent with the 15-km ensemble having slightly more

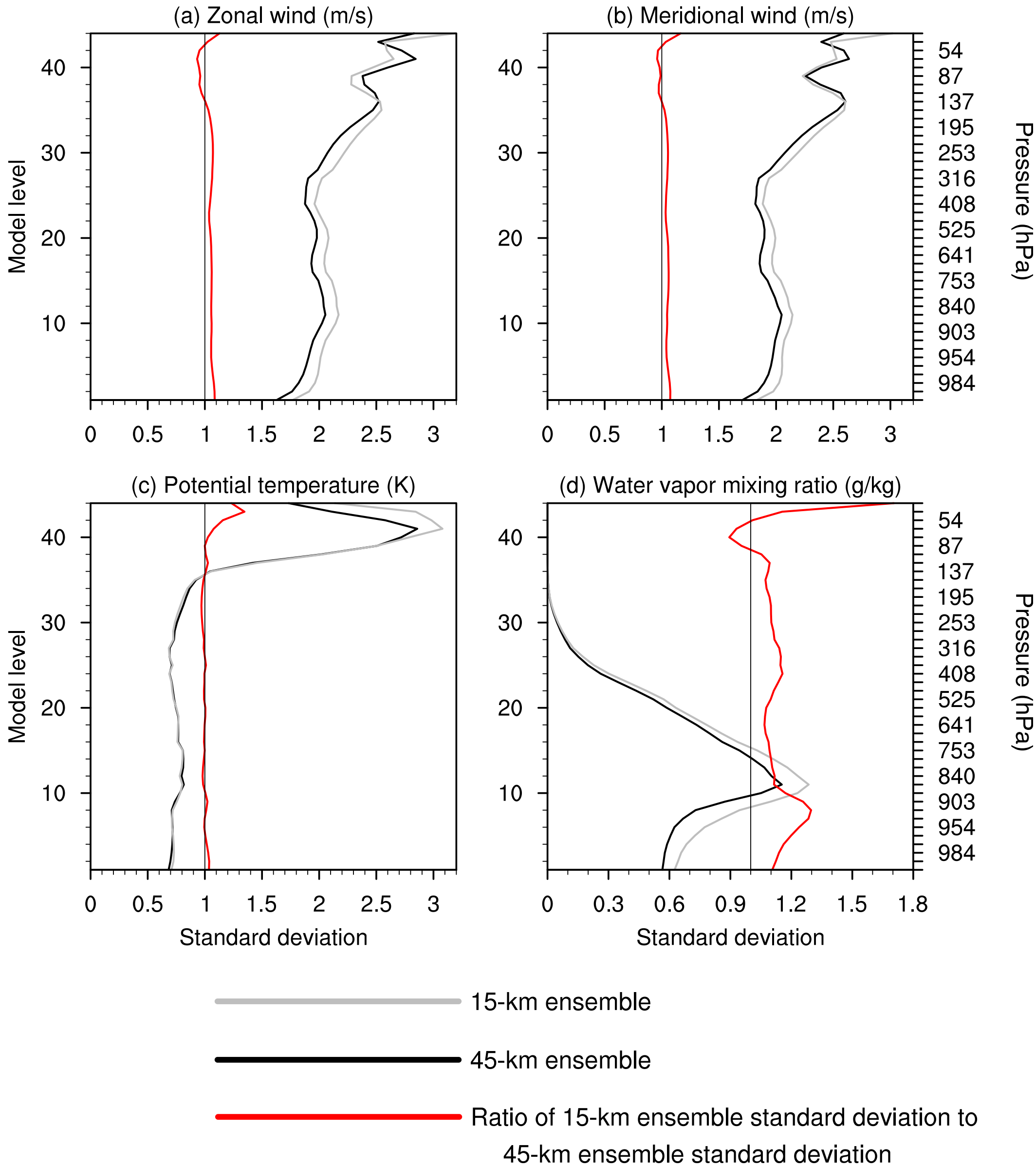

FIG. 10. Domain average prior ensemble standard deviations (after inflation) between 1800 UTC 8 Sep and 0000 UTC 28 Sep for (a) zonal wind ($m\,s^{-1}$), (b) meridional wind ($m\,s^{-1}$), (c) potential temperature (K), and (d) water vapor mixing ratio ($g\,kg^{-1}$). The approximate pressures (hPa) of selected model levels are shown along the right axes of (b) and (d). The 45-km statistics were computed solely over the portion of the 45-km domain collocated with the 15-km domain.

spread than the 45-km ensemble for these variables (e.g., Fig. 10).

The mean 15-km Hybrid_DR_1way and Hybrid_SR 500 hPa potential temperature (Figs. 12a,b) and 700-hPa water vapor mixing ratio (Figs. 12c,d) analysis increments between 1800 UTC 8 September and 0000 UTC 28 September were very similar, although the Hybrid_SR patterns were less smooth. Furthermore, the mean Hybrid_DR_1way and Hybrid_SR 500 and 700-hPa heights (overlaid in Fig. 12) were remarkably similar. The corresponding Hybrid_DR_2way increments and heights were also similar to those of Hybrid_DR_1way and Hybrid_SR (not shown). Despite the 15-km Hybrid_SR analyses sometimes fitting observations slightly closer than the other analyses, the mean increments and prior fits to observations suggest that the three 15-km hybrid DA systems performed similarly, on average. We now assess

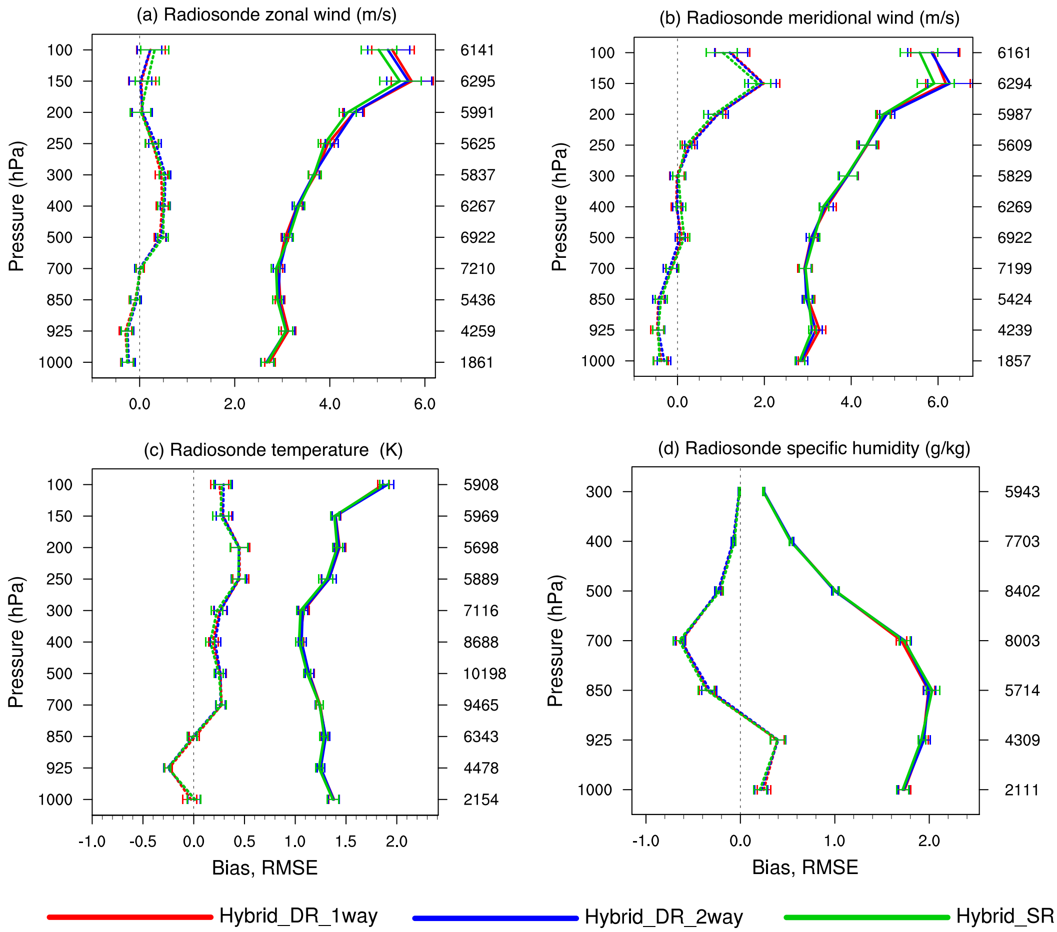

FIG. 11. RMSE (solid lines) and bias (dashed lines) for verification vs radiosonde (a) zonal wind (m s^{-1}), (b) meridional wind (m s^{-1}), (c) temperature (K), and (d) specific humidity (g kg^{-1}) observations aggregated over all 15-km backgrounds (6-h forecasts) between 1800 UTC 8 Sep and 0000 UTC 28 Sep. The sample size at each level is denoted to the right of each panel. Error bars denote bounds of 90% confidence intervals.

whether these similar analyses translated into comparable forecasts.

c. Forecast verification

1) TROPICAL CYCLONE TRACK FORECASTS

TC track forecasts were verified as in S13 using ''best track'' positions from the CWB as ''truth.'' TC positions were diagnosed using a DART forward operator that locates TCs using 800-hPa circulation (e.g., Cavallo et al. 2013). Track error statistics for each storm were computed from multiple WRF Model forecasts spanning the lifetime of each TC (Table 4), and the track of each TC is shown in Fig. 13a. Homogeneous quantitative comparisons were produced based on TCs that all experiments successfully predicted. We note that physical processes too small to be resolved with 15-km horizontal grid spacing impact TC motion, and, thus, results regarding the relative performance of higher-resolution SR and DR hybrid systems for TC track forecasts may differ from ours.

Figure 13b shows mean absolute track errors and sample sizes at each forecast hour averaged over all three TCs. Horizontal lines are ''zero'' lines for 90% bootstrap CIs based upon pairwise differences of two experiments' errors, and if the CI did not include zero, then differences between the experiments' errors were statistically significantly different at the 95% level. The Hybrid_DR_1way and Hybrid_DR_2way track errors were very similar, but track errors from Hybrid_SR were smallest and statistically significantly better at the 95% level compared to Hybrid_DR_1way at four forecast times. All TC track forecasts initialized by hybrid analyses had smaller errors than the corresponding 3DVAR-initialized forecasts, with Hybrid_DR_1way

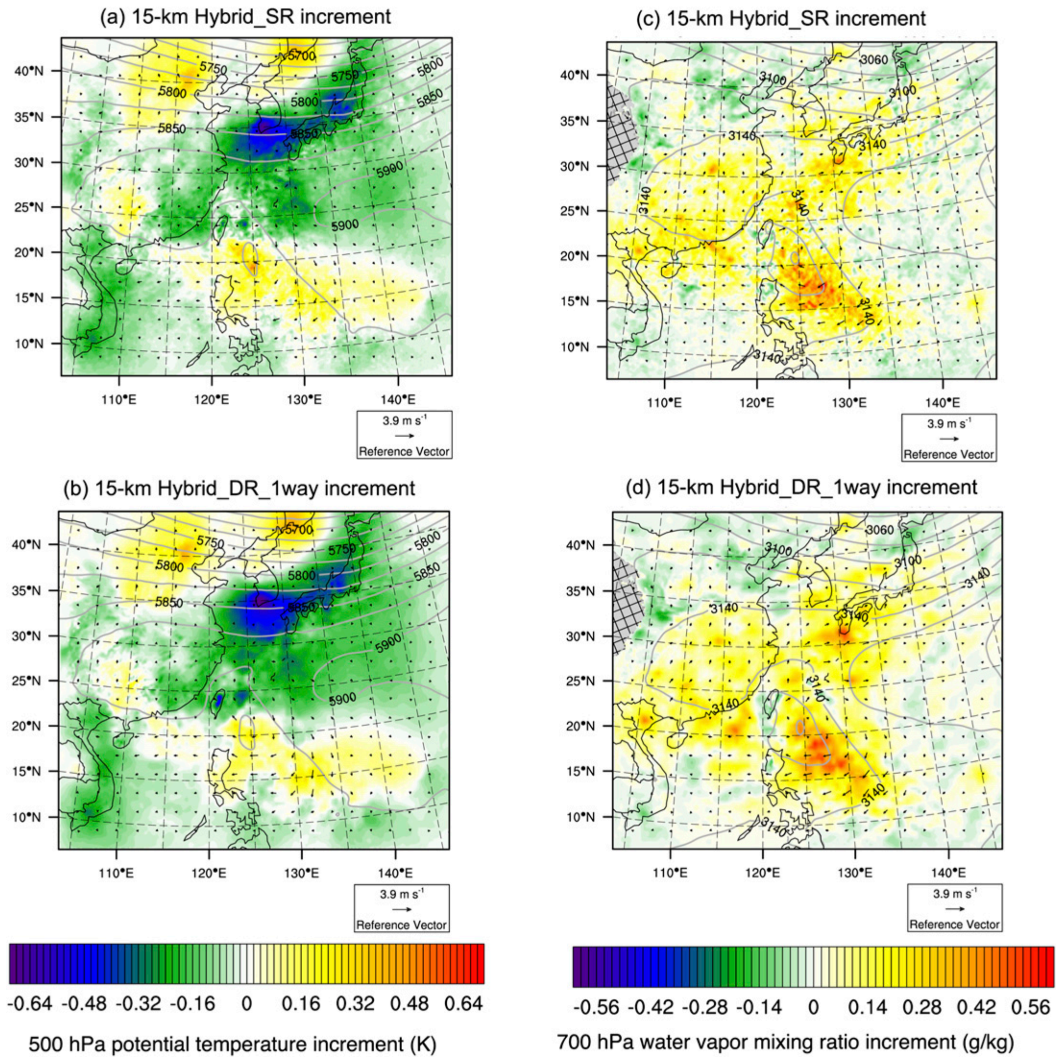

FIG. 12. The 15-km, 500-hPa potential temperature analysis increments (K), wind vector increments (arrows), and mean background 500-hPa height (m) averaged between 1800 UTC 8 Sep and 0000 UTC 28 Sep for (a) Hybrid_SR and (b) Hybrid_DR_1way. (c),(d) As in (a),(b), but for 700-hPa water vapor mixing ratio increments (g kg^{-1}), wind vector increments, and mean background height. Hatching in (c) and (d) indicates those areas where the 700-hPa surface was underground. Heights are contoured every 25 m in (a),(b) and every 20 m in (c),(d).

having statistically significantly smaller errors at the 95% level at eight times.

2) VERIFICATION VERSUS RADIOSONDE OBSERVATIONS

The 15-km model output was also verified against radiosonde observations at several forecast times. Statistics were aggregated over 78 forecasts initialized every 6 h between 1800 UTC 8 September and 0000 UTC 28 September.

At 24 h, all hybrid experiments had similar RMSEs and biases compared to radiosonde observations (Fig. 14). The 3DVAR RMSEs were often slightly larger than the hybrid experiments' RMSEs, but the differences were not SS at the 95% level. Similar patterns with no SS differences were also noted both for verification versus other observation types (e.g., aircraft observations) and at later forecast times (not shown).

TABLE 4. The beginning and ending times that were verified for each TC.

Storm	Beginning time	Ending time
Sinlaku	1800 UTC 8 Sep	0600 UTC 20 Sep
Hagupit	1200 UTC 19 Sep	1800 UTC 24 Sep
Jangmi	1200 UTC 24 Sep	0000 UTC 1 Oct

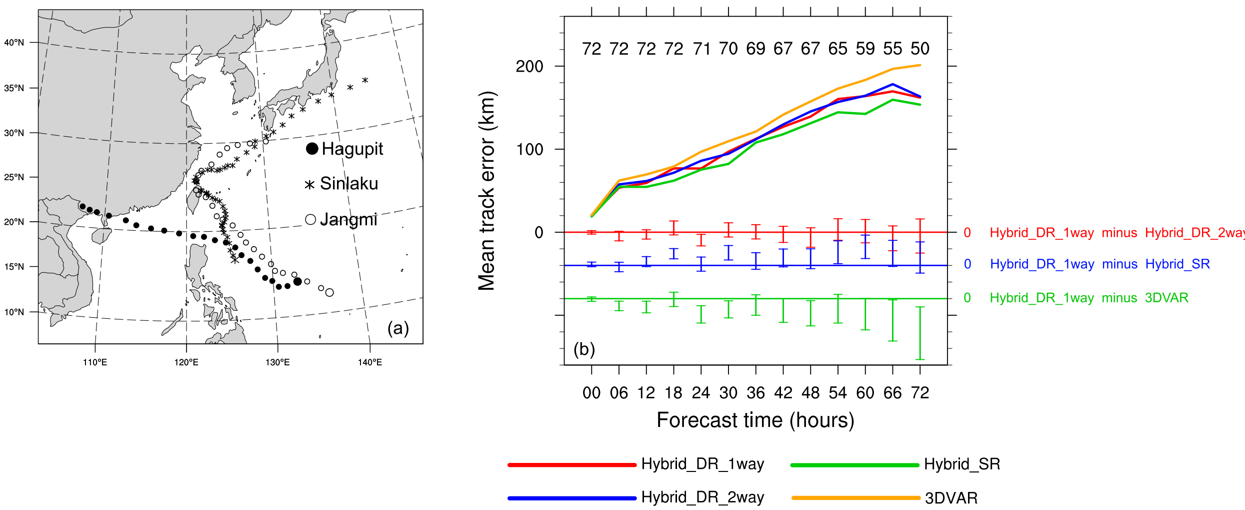

FIG. 13. (a) Best track positions of TCs Sinlaku, Hagupit, and Jangmi. Locations are plotted every 6 h. See Table 4 for the starting and ending times of each storm. (b) Mean 0–72-h absolute track errors (km) averaged over the three TCs. The sample size at each forecast hour is denoted along the top axis. Horizontal lines are "zero lines" for 90% bootstrap CIs based upon track error differences between pairs of experiments. The differences between two experiments were statistically significant at the 95% level if the bounds of the 90% CI did not include zero.

3) Verification versus dropwindsonde observations

As in S13, forecasts were also compared to dropwindsonde observations taken during The Observing System Research and Predictability Experiment (THORPEX) Pacific Asian Regional Campaign (T-PARC; Elsberry and Harr 2008; Wang et al. 2010). These observations were not assimilated and provide an independent dataset for model validation. Most dropwindsondes sampled meso-α-scale environments surrounding the TCs (see Fig. 12 in S13 for dropwindsonde locations).

At 24 h (Fig. 15), consistent with verification versus radiosondes, there were no SS differences at the 95% level between the three hybrid experiments. Of the hybrid experiments, Hybrid_SR had the highest RMSEs for zonal wind below 500 hPa (Fig. 15a) and specific humidity between 700 and 850 hPa (Fig. 15d). However, the 3DVAR RMSEs were usually largest, particularly for wind. At later forecast times, Hybrid_SR and Hybrid_DR_1way wind RMSEs were usually more similar, and, again, there were no SS differences between the experiments at the 95th percentile (not shown).

4) Discussion

A general lack of SS differences (at the 95th percentile) between the various hybrid experiments was consistent across the different verification metrics. However, while all verification scores indicated hybrid experiments usually performed better than the 3DVAR experiment, the statistical significance between the 3DVAR and hybrid results differed depending on the verifying observations.

Verification against radiosondes primarily measured forecast accuracy over land (Fig. 1). It is unsurprising that the 3DVAR and hybrid forecasts were most similar over land, as many land-based observations were present to constrain analyses, lessening the importance of the BECs. Conversely, comparing forecasts to TC center locations and dropwindsondes measured forecast performance over the sea, where observations were relatively sparse. Therefore, the BECs assumed more importance in maritime regions, and, indeed, the hybrid and 3DVAR experiments differed more over oceanic areas, with aggregate TC track error and dropwindsonde statistics clearly favoring the hybrid experiments (Figs. 13b and 15).

However, while aggregate 3DVAR TC track errors were statistically significantly worse than aggregate hybrid TC track errors at the 95% level for eight forecast times (Fig. 13b), because of wide CIs, there were no SS differences at the 95th percentile between the 3DVAR and hybrid experiments for dropwindsonde verification, despite generally superior hybrid statistics (Fig. 15). Although temporal correlation of forecast errors may have contributed to ambiguity regarding statistical significance, wide CIs suggested considerable forecast accuracy variability in near-TC environments, which was confirmed by examining the distributions of daily biases and RMSEs compared to dropwindsondes (not shown). That forecast goodness varied considerably around TCs seems reasonable, as forecasts in these regions are sensitive to even small errors regarding not only TC track, but also intensity and structure, because of sharp gradients associated with TCs.

Thus, regarding dropwindsonde verification, we suggest the absence of statistical significance at the 95th

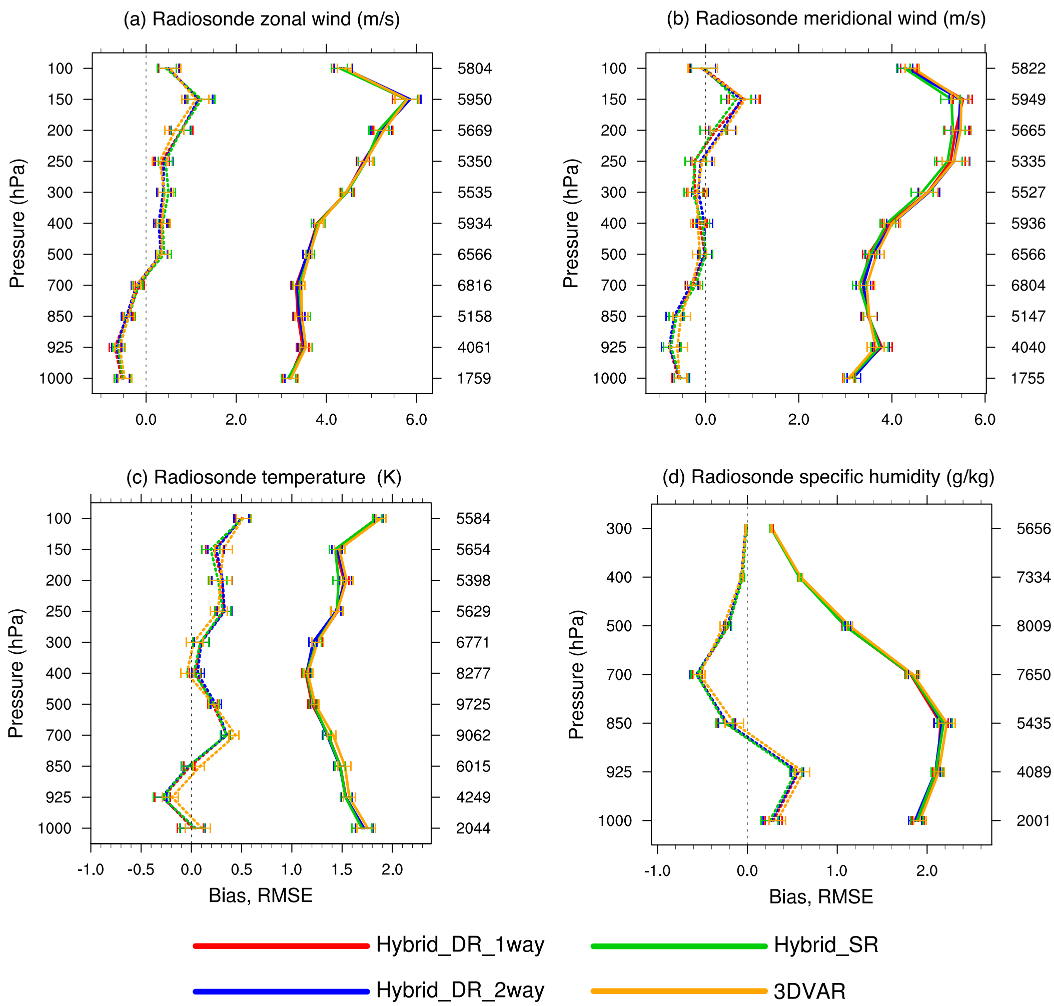

FIG. 14. Average RMSE (solid lines) and bias (dashed lines) vs radiosonde (a) zonal wind (m s^{-1}), (b) meridional wind (m s^{-1}), (c) temperature (K), and (d) specific humidity observations averaged over all 24-h, 15-km forecasts. The sample size at each level is denoted to the right of each panel. Error bars denote the bounds of the 90% confidence intervals.

percentile be interpreted as an indication of wide variability and not as evidence that the 3DVAR and hybrid experiments performed similarly. Moreover, despite the lack of statistical significance at the 95th percentile, dropwindsonde verification statistics complemented TC track error statistics by indicating the hybrid experiments collectively performed best over ocean. Overall, our results are consistent with previous studies showing that flow-dependent BECs often provide the greatest benefit compared to static BECs over regions with relatively few observations (e.g., Hamill and Snyder 2000; Whitaker et al. 2008; Buehner et al. 2010; Kleist and Ide 2015).

8. Summary and conclusions

DR hybrid analysis capability was implemented within the community WRFDA system. The DR hybrid combines observations, a HR background, and an LR ensemble to produce a deterministic HR analysis, permitting considerable computational savings compared to an SR hybrid configuration. DR and SR experiments were performed that produced new hybrid analyses every 6 h within a continuously cycling framework over a ~3.5-week period and initialized 72-h WRF Model forecasts. Both the DR and SR hybrid systems ingested flow-dependent BECs provided by 32-member ensembles that were updated by an EAKF, and different DR configurations examined whether it was preferable to recenter EAKF analysis ensembles about DR hybrid analyses. The DR system combined 15-km backgrounds with 45-km ensembles, while the SR system combined backgrounds and ensembles with equal, 15-km horizontal grid lengths.

On average, 15-km prior ensembles had slightly more spread than 45-km prior ensembles. However, the mean

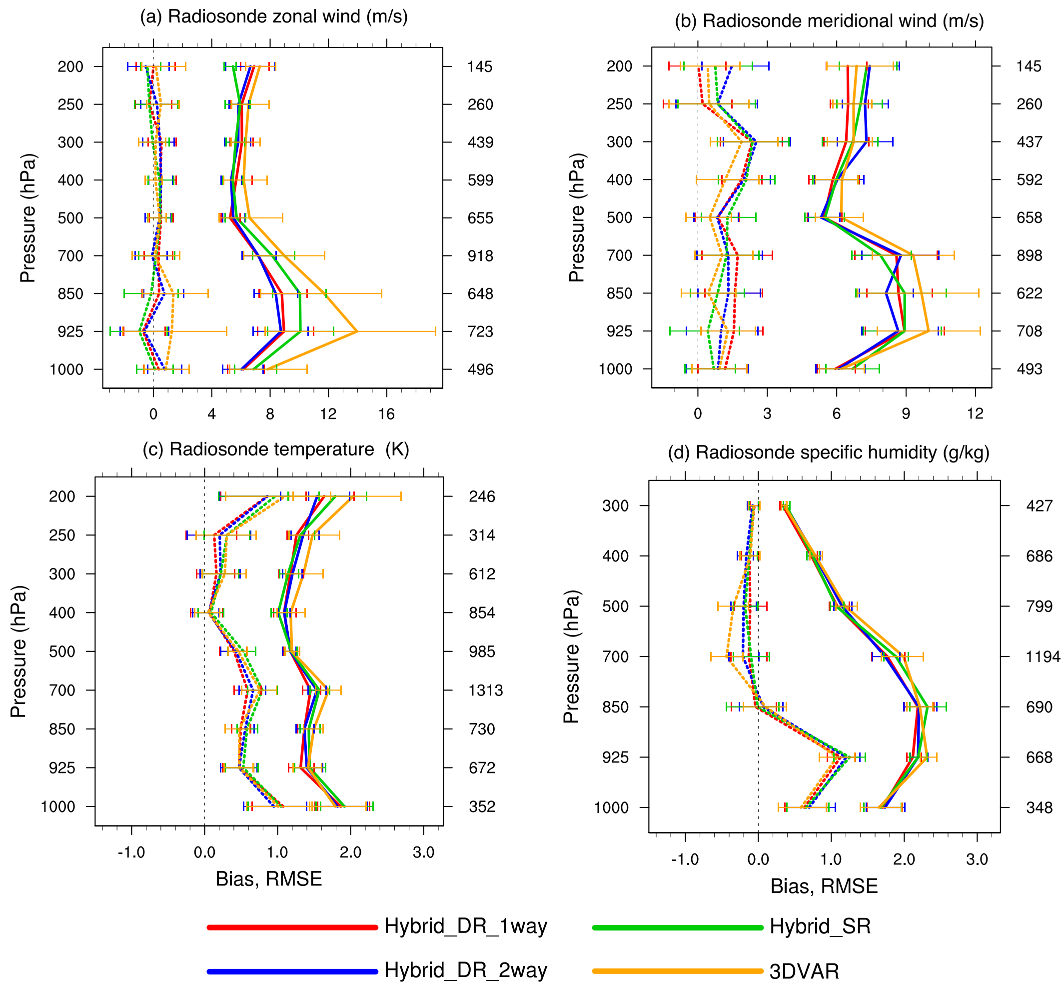

FIG. 15. As in Fig. 14, but for 24-h forecast verification vs dropwindsonde observations.

15-km SR and DR hybrid analysis increments and prior fits to radiosonde observations were very similar. Overall, 15-km forecasts of wind, temperature, and moisture initialized by 15-km DR and SR hybrid analyses were comparable and not statistically significantly different. However, Hybrid_SR TC track forecast errors were clearly—but only sometimes statistically significantly—smaller than those of the DR experiments. That the most consistent difference between the SR and DR configurations involved TC forecasts agrees with the expectation that HR and LR ensembles will be most different around small-scale features (e.g., section 6). Recentering EAKF analysis ensembles about DR hybrid analyses only had a small impact, commensurate with previous studies (e.g., Clayton et al. 2012; S13; Wang et al. 2013; Pan et al. 2014). As the recentering procedure simply shifts the ensemble perturbations without changing their amplitudes, this small impact was not surprising. The various hybrid-initialized 15-km forecasts improved upon those initialized by 15-km 3DVAR analyses, particularly for TC track, where many differences were SS at the 95% level.

These collective results suggest that DR hybrid analyses can often initialize similar quality forecasts as SR hybrid analyses, although SR systems may be preferable for forecasting smaller-scale features, including TCs. However, we only examined a ~3.5-week period and thus encourage further experimentation with DR and SR hybrid systems over longer time periods to further understand how ensemble perturbation resolution impacts analyses and forecasts.

Practically, users should carefully consider whether any gains in forecast skill from SR systems are worth the added computational cost. For our experiments, the 15-km DR analyses completed ~3 times faster than the 15-km SR analyses because the DR hybrid had fewer extended control variables. Additionally, during the ensemble of WRF Model advances between EAKF analyses, the DR configuration realized approximately a sixfold savings compared to the 15-km SR hybrid because the 15-km nest was removed in the DR configuration for each ensemble member (e.g., Fig. 3). Moreover, the 15-km SR hybrid necessitated ~4 times more disk space than the 15-km DR hybrid, as the 15-km SR hybrid required storage of 15-km perturbations, whereas the 15-km DR hybrid solely needed 45-km perturbations. These savings could be utilized for many purposes, including increasing the ensemble size or resolution of the deterministic background.

Here, the HR horizontal grid spacing was 3 times finer than the LR horizontal grid length. As the ratio of LR to HR horizontal grid spacing increases, so do the computational savings, but a larger grid ratio may translate into greater differences between SR and DR hybrid analysis–forecast systems than are documented here.

Additionally, an important question regards the utility of DR hybrid systems at increased resolution, particularly when the background is at sufficiently fine resolution that convective parameterization (CP) can be removed but the ensemble resolution is coarse enough that CP is required. In such a configuration, the CP scheme may engender very different bias characteristics (e.g., Romine et al. 2013) in the prior ensemble compared to those of the convection-allowing background. It is unclear how much of an impact this disparity may have, but this topic demands investigation as NWP models continue their progression toward higher resolution.

Acknowledgments. The Taiwan Central Weather Bureau (CWB) partially funded this work. Three anonymous reviewers provided constructive comments that improved this paper. We also thank the Editor, Dr. Altuğ Aksoy, for his suggestions.

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.

——, 2003: A local least squares framework for ensemble filtering. *Mon. Wea. Rev.,* **131,** 634–642, doi:10.1175/1520-0493(2003)131<0634:ALLSFF>2.0.CO;2.

——, 2009: Spatially and temporally varying adaptive covariance inflation for ensemble filters. *Tellus,* **61A,** 72–83, doi:10.1111/j.1600-0870.2008.00361.x.

——, T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Arellano, 2009: The Data Assimilation Research Testbed: A community facility. *Bull. Amer. Meteor. Soc.,* **90,** 1283–1296, doi:10.1175/2009BAMS2618.1.

Barker, D. M., 2005: Southern high-latitude ensemble data assimilation in the Antarctic Mesoscale Prediction System. *Mon. Wea. Rev.,* **133,** 3431–3449, doi:10.1175/MWR3042.1.

——, W. Huang, Y.-R. Guo, A. Bourgeois, and X. N. Xio, 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.

——, and Coauthors, 2012: The Weather Research and Forecasting Model's Community Variational/Ensemble Data Assimilation System: WRFDA. *Bull. Amer. Meteor. Soc.,* **93,** 831–843, doi:10.1175/BAMS-D-11-00167.1.

Berner, J., S.-Y. Ha, J. P. Hacker, A. Fournier, and C. Snyder, 2011: Model uncertainty in a mesoscale ensemble prediction system: Stochastic versus multiphysics representations. *Mon. Wea. Rev.,* **139,** 1972–1995, doi:10.1175/2010MWR3595.1.

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.

——, P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, 2010: 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.

Cavallo, S. M., R. D. Torn, C. Snyder, C. Davis, W. Wang, and J. Done, 2013: Evaluation of the Advanced Hurricane WRF data assimilation system for the 2009 Atlantic hurricane season. *Mon. Wea. Rev.,* **141,** 523–541, doi:10.1175/MWR-D-12-00139.1.

Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model description and implementation. *Mon. Wea. Rev.,* **129,** 569–585, doi:10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2.

Chou, M.-D., and M. J. Suarez, 1994: An efficient thermal infrared radiation parameterization for use in general circulation models. NASA Tech. Memo. 104606, Vol. 3, 85 pp.

Clayton, A. M., A. C. Lorenc, and D. M. Barker, 2012: Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office. *Quart. J. Roy. Meteor. Soc.,* **139,** 1445–1461, doi:10.1002/qj.2054.

Courtier, P., J.-N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.,* **120,** 1367–1387, doi:10.1002/qj.49712051912.

Davis, C., W. Wang, J. Dudhia, and R. Torn, 2010: Does increased horizontal resolution improve hurricane wind forecasts? *Wea. Forecasting,* **25,** 1826–1841, doi:10.1175/2010WAF2222423.1.

Dowell, D. C., and L. J. Wicker, 2009: Additive noise for storm-scale ensemble data assimilation. *J. Atmos. Oceanic Technol.,* **26,** 911–927, doi:10.1175/2008JTECHA1156.1.

Elsberry, R. L., and P. A. Harr, 2008: Tropical cyclone structure (TCS08) field experiment science basis, observational platforms, and strategy. *Asia-Pac. J. Atmos. Sci.,* **44,** 209–231.

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.

Gao, J., and M. Xue, 2008: An efficient dual-resolution approach for ensemble data assimilation and tests with assimilated Doppler radar data. *Mon. Wea. Rev.,* **136,** 945–963, doi:10.1175/2007MWR2120.1.

——, ——, and D. J. Stensrud, 2013: The development of a hybrid EnKF-3DVAR algorithm for storm-scale data assimilation. *Adv. Meteor.,* **2013,** 512656, doi:10.1155/2013/512656.

Ha, S.-Y., and C. Snyder, 2014: Influence of surface observations in mesoscale data assimilation using an ensemble Kalman filter. *Mon. Wea. Rev.,* **142,** 1489–1508, doi:10.1175/MWR-D-13-00108.1.

Hamill, T. M., 1999: Hypothesis tests for evaluating numerical precipitation forecasts. *Wea. Forecasting,* **14,** 155–167, doi:10.1175/1520-0434(1999)014<0155:HTFENP>2.0.CO;2.

——, 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:AHEKFV>2.0.CO;2.

——, J. S. Whitaker, D. T. Kleist, M. Fiorino, and S. G. Benjamin, 2011: 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.

Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.,* **134,** 2318–2341, doi:10.1175/MWR3199.1.

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.

Hsiao, L. F., C. S. Liou, T. C. Yeh, Y. R. Guo, D. S. Chen, K. N. Huang, C. T. Terng, and J. H. Chen, 2010: A vortex relocation scheme for tropical cyclone initialization in Advanced Research WRF. *Mon. Wea. Rev.,* **138,** 3298–3315, doi:10.1175/2010MWR3275.1.

Huang, X.-Y., and P. Lynch, 1993: Diabatic digital filter initialization: Application to the HIRLAM model. *Mon. Wea. Rev.,* **121,** 589–603, doi:10.1175/1520-0493(1993)121<0589:DDFIAT>2.0.CO;2.

Kain, J. S., 2004: The Kain–Fritsch convective parameterization: An update. *J. Appl. Meteor.,* **43,** 170–181, doi:10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2.

——, and J. M. Fritsch, 1990: A one-dimensional entraining/detraining plume model and its application in convective parameterization. *J. Atmos. Sci.,* **47,** 2784–2802, doi:10.1175/1520-0469(1990)047<2784:AODEPM>2.0.CO;2.

——, and ——, 1993: Convective parameterization for mesoscale models: The Kain–Fritsch scheme. *The Representation of Cumulus Convection in Numerical Models, Meteor. Monogr.,* No. 24, Amer. Meteor. Soc., 165–170.

Kleist, D. T., and K. Ide, 2015: An OSSE-based evaluation of hybrid variational–ensemble data assimilation for the NCEP GFS. Part I: System description and 3D-hybrid results. *Mon. Wea. Rev.,* **143,** 433–451, doi:10.1175/MWR-D-13-00351.1.

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, doi:10.1175/MWR-D-12-00182.1.

Li, X., J. Ming, M. Xue, Y. Wang, and K. Zhao, 2015: Implementation of a dynamic equation constraint based on the steady state momentum equations within the WRF hybrid ensemble-3DVar data assimilation system and test with radar T-TREC wind assimilation for Tropical Cyclone Chanthu (2010). *J. Geophys. Res. Atmos.,* **120,** 4017–4039, doi:10.1002/2014JD022706.

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.

Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP—A comparison with 4D-VAR. *Quart. J. Roy. Meteor. Soc.,* **129,** 3183–3203, doi:10.1256/qj.02.132.

Lorenz, E. N., 1969: The predictability of a flow which possesses many scales of motion. *Tellus,* **21,** 289–307, doi:10.1111/j.2153-3490.1969.tb00444.x.

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.

Ma, L.-M., and Z.-M. Tan, 2009: Improving the behavior of the cumulus parameterization for tropical cyclone prediction: Convection trigger. *Atmos. Res.,* **92,** 190–211, doi:10.1016/j.atmosres.2008.09.022.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmosphere: RRTM, a validated correlated-*k* model for the long-wave. *J. Geophys. Res.,* **102,** 16 663–16 682, doi:10.1029/97JD00237.

Pan, Y., K. Zhu, M. Xue, X. 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, doi:10.1175/MWR-D-13-00242.1.

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.

Poterjoy, J., and F. Zhang, 2014: Intercomparison and coupling of ensemble and four-dimensional variational data assimilation methods for the analysis and forecasting of Hurricane Karl (2010). *Mon. Wea. Rev.,* **142,** 3347–3364, doi:10.1175/MWR-D-13-00394.1.

Rainwater, S., and B. Hunt, 2013: Mixed resolution ensemble data assimilation. *Mon. Wea. Rev.,* **141,** 3007–3021, doi:10.1175/MWR-D-12-00234.1.

Romine, G., C. S. Schwartz, C. Snyder, J. Anderson, and M. Weisman, 2013: Model bias in a continuously cycled assimilation system and its influence on convection-permitting forecasts. *Mon. Wea. Rev.,* **141,** 1263–1284, doi:10.1175/MWR-D-12-00112.1.

Sanchez, C., K. D. Williams, G. Shutts, and M. Collins, 2014: Impact of a stochastic kinetic energy backscatter scheme across time-scales and resolutions. *Quart. J. Roy. Meteor. Soc.,* **140B,** 2625–2637, doi:10.1002/qj.2328.

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.

——, ——, X.-Y. Huang, Y.-H. Kuo, and C.-T. Fong, 2013: Comparing limited-area 3DVAR and hybrid variational-ensemble data assimilation methods for typhoon track forecasts: Sensitivity to outer loops and vortex relocation. *Mon. Wea. Rev.,* **141,** 4350–4372, doi:10.1175/MWR-D-13-00028.1.

Shutts, G., 2005: A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quart. J. Roy. Meteor. Soc.,* **131,** 3079–3102, doi:10.1256/qj.04.106.

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech Note NCAR/TN–475+STR, 113 pp. [Available online at http://www2.mmm.ucar.edu/wrf/users/docs/arw_v3.pdf.]

Tao, W.-K., and J. Simpson, 1993: The Goddard cumulus ensemble model. Part I: Model description. *Terr. Atmos. Oceanic Sci.,* **4,** 35–72.

——, and Coauthors, 2003: Microphysics, radiation and surface processes in the Goddard Cumulus Ensemble (GCE) model. *Meteor. Atmos. Phys.,* **82,** 97–137, doi:10.1007/s00703-001-0594-7.

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

Wang, J., and Coauthors, 2010: Water vapor variability and comparisons in the subtropical Pacific from The Observing System Research and Predictability Experiment-Pacific Asian Regional Campaign (T-PARC) Driftsonde, Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC), and reanalyses. *J. Geophys. Res.,* **115,** D21108, doi:10.1029/2010JD014494.

Wang, X., 2010: Incorporating ensemble covariance in the gridpoint statistical interpolation (GSI) 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.

——, T. M. Hamill, J. S. Whitaker, and C. H. Bishop, 2007a: A comparison of hybrid ensemble transform Kalman filter–OI and ensemble square-root filter analysis schemes. *Mon. Wea. Rev.,* **135,** 1055–1076, doi:10.1175/MWR3307.1.

——, C. Snyder, and T. M. Hamill, 2007b: On the theoretical equivalence of differently proposed ensemble—3DVARhybrid analysis schemes. *Mon. Wea. Rev.,* **135,** 222–227, doi:10.1175/MWR3282.1.

——, D. 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 experiments. *Mon. Wea. Rev.,* **136,** 5132–5147, doi:10.1175/2008MWR2445.1.

——, D. F. Parrish, D. T. Kleist, and J. S. 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., T. M. Hamill, 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.

Wilks, D., 2006: *Statistical Methods in the Atmospheric Sciences: An Introduction.* 2nd ed. Academic Press, 627 pp.

Xu, D., X.-Y. Huang, H. Wang, A. P. Mizzi, and J. Min, 2015: Impact of assimilating radiances with the WRFDA ETKF/3DVAR hybrid system on prediction of two typhoons in 2012. *J. Meteor. Res.,* **29,** 28–40, doi:10.1007/s13351-014-4053-z.

Xue, M., J. Schleif, F. Kong, K. W. Thomas, Y. Wang, and K. 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, doi:10.1175/WAF-D-12-00063.1.

Zhang, F., M. Zhang, and J. A. Hansen, 2009: Coupling ensemble Kalman filter with four-dimensional variational data assimilation. *Adv. Atmos. Sci.,* **26,** 1–8, doi:10.1007/s00376-009-0001-8.

——, ——, 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.