Comparison of Hybrid Four-Dimensional Data Assimilation Methods with and without the Tangent Linear and Adjoint Models for Predicting the Life Cycle of Hurricane Karl (2010)

JONATHAN POTERJOY

Advanced Study Program, National Center for Atmospheric Research,* Boulder, Colorado, and Department of Meteorology, The Pennsylvania State University, University Park, Pennsylvania

FUQING ZHANG

Department of Meteorology, The Pennsylvania State University, University Park, Pennsylvania

ABSTRACT

Two four-dimensional hybrid data assimilation systems based on the Weather Research and Forecasting (WRF) Model are applied over the life cycle of Hurricane Karl (2010). One method uses a mix of ensemble- and climate-based background error covariance in a four-dimensional variational data assimilation (4DVar) system that uses an adjoint model to assimilate observations over a time window (denoted E4DVar). The second method approximates the function of linearized models in 4DVar with perturbations generated from an ensemble forecast using the full nonlinear model (denoted 4DEnVar). Ensemble perturbations in 4DEnVar provide a four-dimensional covariance, which is combined with a static climate-based covariance for performing the data assimilation. In cycling data assimilation experiments, analyses produced by both methods provide more accurate intensity forecasts than E3DVar, owing mostly to a better representation of moisture near the developing storm. Despite providing a computationally efficient alternative to E4DVar, predictions made from 4DEnVar analyses are less accurate than E4DVar for the tested case study. Numerical experiments using identical background error statistics in both schemes reveal differences in the mesoscale structure of the developing storm, which are suspected to be responsible for this result. In particular, 4DEnVar analyses contain a less intense inner-core circulation and lower column relative humidity than E4DVar at analysis times closest to Karl's genesis, which lead to a persistent slow bias in intensifying the storm. These results suggest errors introduced in the linearization of the model for E4DVar may be less severe than errors introduced by the localization of time covariances. This study provides the first comparison of hybrid E4DVar and 4DEnVar for a tropical cyclone application.

1. Introduction

Data assimilation approaches that combine strategies from variational and ensemble Kalman filtering (EnKF) methods are applied regularly for operational weather forecasting and research. Applications include global numerical weather prediction (e.g., Buehner 2005; Buehner et al. 2010a,b; Bishop and Hodyss 2011; Clayton et al. 2013; Kuhl et al. 2013; Wang et al. 2013; Wang and Lei 2014; Lorenc et al. 2015) as well as regional analysis and forecasting (e.g., Liu et al. 2009; Zhang and Zhang 2012; F. Zhang et al. 2013; Liu and Xiao 2013; Poterjoy and Zhang 2014a, hereafter PZ14a; Gustafsson and Bojarova 2014). One common result from these studies is that variational data assimilation methods benefit from using either an ensemble-based estimate of prior error covariance in place of the climate-based static covariance, or from a hybrid covariance that combines ensemble and static background errors. This finding has been verified by systematic comparisons of hybrid and nonhybrid data assimilation systems as well (e.g., F. Zhang et al. 2013).

Among the ensemble–variational data assimilation strategies listed above are four-dimensional (4D) methods, which seek the maximum likelihood model

---

Corresponding author address: Dr. Jonathan Poterjoy, NCAR, P.O. Box 3000, Boulder, CO 80307.
E-mail: poterjoy@ucar.edu

solution, given a prior estimate of the model state or background, a set of observations over a time window, and the uncertainty of both quantities. The first operational use of 4D data assimilation for weather prediction came in the form of incremental four-dimensional variational data assimilation (4DVar; Courtier et al. 1994; Rabier et al. 2000), which applies tangent linear and adjoint versions of the forecast model to propagate information in time to assimilate observations displaced over a window. This approach uses a model forecast from the previous analysis as the background state, and approximates the background uncertainty using a climate-based error covariance (Parrish and Derber 1992). The tangent linear model propagates increments, or differences between the control variable and a background state, along a nonlinear trajectory to future times, while the adjoint model propagates sensitivity gradients backward from observation times to the beginning of the window. The linear model operators evolve the background error covariance implicitly in time, thus providing a source of flow-dependent information for the state estimation (Lorenc 2003a).

When ensemble information is used in 4D data assimilation systems, a choice emerges regarding how to model the temporal covariances required for performing the state estimation. The EnKF, first proposed by Evensen (1994), presents a means of propagating error covariances forward in time without linearized model operators. EnKFs rely on an ensemble of model forecasts to approximate the prior mean and error covariance, which are updated each time observations are available. Algorithms that expand the EnKF to include a time dimension include the ensemble Kalman smoother (Evensen and van Leeuwen 2000) and 4DEnKF (Hunt et al. 2004). Likewise, ensemble 3DVar data assimilation systems have been conceived by Lorenc (2003b), Zupanski (2005), and Buehner (2005). Extending these methods to 4D, Liu et al. (2008) introduced the 4D-ensemble-variational (4DEnVar) technique, which uses an ensemble of model trajectories to find the maximum likelihood solution. 4DEnVar contains a number of qualities that make it a practical alternative to 4DVar as a data assimilation method for atmospheric models. Minimizing the 4DEnVar cost function does not require the coding of tangent linear and adjoint models, which is appealing for operational centers and research institutions that do not currently maintain a 4DVar system. An equally important factor is the scalability of the algorithm on massively parallel computing platforms; unlike 4DVar, 4DEnVar does not require serial integrations of a tangent linear model and its adjoint during each iteration of the cost function minimization.

Despite the advantages 4DEnVar presents for implementation, methods that use ensemble covariance in 4DVar are found to be more accurate for global numerical weather prediction models (Buehner et al. 2010b; Lorenc et al. 2015). One challenge for 4DEnVar comes from the localization of 4D ensemble covariances; this problem is avoided in the 4DVar framework because a full-rank localized covariance is evolved implicitly in time by linearized models.[1] The lack of linear model operators in 4DEnVar also complicates the process of evolving errors from the climate-based static error covariance when a hybrid background error covariance is used. How to model this covariance component in an efficient manner remains an open research question at the time of writing this manuscript. Using a hybrid covariance in 4DVar, Poterjoy and Zhang (2015) find that the ability of the tangent linear and adjoint model to evolve temporal information from a static covariance component is most advantageous for treating model errors not accounted for by the ensemble. The static covariance improves the analysis in a manner similar to additive covariance inflation in ensemble filters: It provides additional orthogonal directions for errors to grow in the assimilation window, but with the added advantage of remaining full rank. Nevertheless, 4DEnVar may contain benefits for certain weather regimes that have not been realized in past global modeling applications. For example, the linearized model in 4DVar often contains simplified physics to avoid strong nonlinearities in parameterization schemes (Sun and Crook 1997; Honda et al. 2005; Huang et al. 2009), which is avoided in 4DEnVar. Evidence of 4DEnVar producing more accurate analyses than ensemble implementations of 4DVar for regional-scale modeling have been demonstrated by Gustafsson and Bojarova (2014). At the same time, the 4D ensemble introduces nonlinearities in the temporal evolution of covariances, possibly leading to a poorly conditioned minimization problem (J. Sun 2014, personal communication).

In a recent study by Poterjoy and Zhang (2014b), the authors apply a coupled EnKF–4DVar (E4DVar) data assimilation system to assimilate routinely collected observations and field measurements taken during the development and intensification of Hurricane Karl (2010). Measurements assimilated during this study were collected during the Pre-Depression Investigation of Cloud-systems in the Tropics (PREDICT) field

---

[1] See section 3 of Poterjoy and Zhang (2015) for a discussion on the different localization strategies used by 4DEnVar and E4DVar.

campaign. Poterjoy and Zhang (2014b) provide a detailed introduction into this application by discussing the spatial and temporal frequency of observations, along with their impact on synoptic and mesoscale processes leading to Karl's genesis and intensification in the Weather Research and Forecasting (WRF) Model. Poterjoy and Zhang (2014b) also highlight the important role of mesoscale processes in the evolving weather event, which motivated PZ14a to test a newly developed E4DVar system for generating analyses of Karl's entire life cycle.

The E4DVar method used in PZ14a incorporates a hybrid ensemble/climate-based background error covariance in a 4DVar system developed for the WRF Model (Zhang et al. 2014). Each E4DVar data assimilation cycle includes a hybrid update of the ensemble mean state using 4DVar, and an EnKF update of the ensemble perturbations. The authors find E4DVar to produce a more accurate representation of wind, temperature, and moisture near the pre-Karl disturbance than similarly configured 4DVar and EnKF data assimilation systems, which leads to better predictions for the timing and location of genesis and rapid intensification. In the current study, the authors apply a coupled EnKF–4DEnVar system for the same application to examine the effectiveness of a 4D-ensemble data assimilation system that does not require linearized models. Results are compared with a coupled EnKF–3DVar (E3DVar) data assimilation system and the E4DVar method presented in PZ14a. The purpose of this study is to examine E3DVar, E4DVar, and 4DEnVar analyses for an application where theoretical differences highlighted in Poterjoy and Zhang (2015) may have a large impact on the modeling of 4D covariances by E4DVar and 4DEnVar. For this reason, the authors focus on a single well-observed weather event. The goal of this evaluation is to supplement future and existing studies that provide more rigorous statistical comparisons of adjoint- and ensemble-based 4D data assimilation methods.

The organization of the manuscript is as follows. We describe the model, data assimilation systems, and cycling data assimilation setup in section 2. In section 3 we present a set of single-observation experiments using ensemble- and climate-based background error covariance for observations displaced at multiple times in the assimilation window. Sections 4 and 5 contain results from cycling and noncycling data assimilation experiments, respectively, where noncycling refers to the use of the same background at each analysis time. We provide our conclusions and a discussion in section 6.

2. Experiment setup

a. WRF Model setup and case study

This study uses version 3.4.1 of the Advanced Research WRF Model (Skamarock et al. 2008). The model domain for ensemble forecasts has 251×226 horizontal grid points with a spacing of 13.5 km (see Fig. 1 of PZ14a) and 35 vertical levels with a 5-mb (1 mb = 1 hPa) upper boundary. Deterministic forecasts initialized from each analysis use a 4.5-km spacing two-way nested domain of size 253×253 that follows the storm with preset moves. We use the following physical parameterization schemes for unresolved processes in the model: Monin–Obukhov similarity (Monin and Obukhov 1954) for the surface layer, the Yonsei University planetary boundary layer scheme (Noh et al. 2003), five-layer thermal diffusion for surface layer physics, the Rapid Radiative Transfer Model (RRTM; Mlawer et al. 1997) and Dudhia (Dudhia 1989) radiation schemes, and the WRF single-moment 6-class microphysics scheme (Hong et al. 2004). Both domains represent cumulus convection explicitly, which was found in offline sensitivity experiments to provide more accurate temperature and moisture fields than simulations using parameterized convection for this model setup and case study.

Each data assimilation method tested during this study uses ensemble forecast data to estimate background errors. The initial ensemble of model states is formed at 1800 UTC 7 September by adding random perturbations generated from a month-long estimated climatological background error covariance to the National Centers for Environmental Prediction (NCEP) Global Data Assimilation System (GDAS) analysis at this time. Following a 12-h ensemble forecast, the first data assimilation cycle occurs at 0600 UTC 8 September, and observations are assimilated every 6 h. The data assimilation systems are cycled over a period that includes the days leading up to Karl's development at 1200 UTC 14 September, as well as its initial intensification on 15 September. Deterministic forecasts from each cycle are run to 0000 UTC 18 September, which spans the entire life cycle of Karl. Observations assimilated during these experiments include regularly collected surface and upper-air data from the National Oceanic and Atmospheric Administration (NOAA) Meteorological Assimilation Data Ingest System (MADIS) and dropsonde measurements taken during the PREDICT field campaign (Montgomery et al. 2012). Observations from the National Aeronautics and Space Administration (NASA) Genesis and Rapid Intensification Processes (GRIP) campaign (Braun et al. 2013) are combined with MADIS and PREDICT soundings to verify deterministic forecasts from each method. See

Poterjoy and Zhang (2014b) and Torn and Cook (2013) for details regarding the impact of the PREDICT field observations on analyses and forecasts for this event.

b. Hybrid 3DVar, 4DVar, and 4DEnVar

For the variational component of the data assimilation, we use version 3.4.1 of the WRF data assimilation (WRFDA) package (Huang et al. 2009; Barker et al. 2012). For a given observation time t, the variational analysis comes from minimizing a cost function that measures the misfit of a control variable to the background state $\mathbf{x}_t^b$ and observations $\mathbf{y}_t$. The minimization is carried out with respect to increments $\delta\mathbf{x}_t$ from $\mathbf{x}_t^b$ (Courtier et al. 1994), and the cost function is expressed as the sum of background (J_b) and observation (J_o) terms. The incremental 3DVar cost function depends on quantities valid at the data assimilation time, represented by time 0:

$$
\begin{aligned}
J(\delta\mathbf{x}_0) &= J_b(\delta\mathbf{x}_0) + J_o(\delta\mathbf{x}_0) \\
&= \frac{1}{2}\delta\mathbf{x}_0^{\mathrm{T}}\mathbf{B}^{-1}\delta\mathbf{x}_0 \\
&\quad + \frac{1}{2}(\mathbf{H}_0\delta\mathbf{x}_0 - \mathbf{d}_0)^{\mathrm{T}}\mathbf{R}_0^{-1}(\mathbf{H}_0\delta\mathbf{x}_0 - \mathbf{d}_0),
\end{aligned} \tag{1}
$$

where $\mathbf{B}$ and $\mathbf{R}_0$ are the background and observation error covariance matrices, respectively. The quantity $\mathbf{d}_0$ is the innovation vector, which is given by the difference between $\mathbf{y}_0$ and the background projected into observation space by the observation operator $H_0(\mathbf{x}_0^b)$. Likewise, increments are projected into observation space using the tangent linear version of H_0, denoted by $\mathbf{H}_0$. Equation (1) is extended to 4DVar for observations displaced temporally between $t = 0$ and $t = \tau$:

$$
\begin{aligned}
J(\delta\mathbf{x}_0) &= J_b(\delta\mathbf{x}_0) + J_o(\delta\mathbf{x}_0) \\
&= \frac{1}{2}\delta\mathbf{x}_0^{\mathrm{T}}\mathbf{B}^{-1}\delta\mathbf{x}_0 \\
&\quad + \frac{1}{2}\sum_{t=0}^{\tau}(\mathbf{H}_t\mathbf{M}_t\delta\mathbf{x}_0 - \mathbf{d}_t)^{\mathrm{T}}\mathbf{R}_t^{-1}(\mathbf{H}_t\mathbf{M}_t\delta\mathbf{x}_0 - \mathbf{d}_t)
\end{aligned} \tag{2}
$$

In (2), the control variable is the vector of increments at the beginning of an assimilation window $\delta\mathbf{x}_0$, which relates to increments at future times through the tangent linear forecast model $\mathbf{M}_t$.[2] The time displacement of observations is also reflected in the 4DVar innovation vector, which is given by

$$
\mathbf{d}_t = \mathbf{y}_t - H_t[M_t(\mathbf{x}_0^b)], \tag{3}
$$

where the background state at $t = 0$ is propagated forward to the time of observations using the nonlinear forecast model M_t. See X. Zhang et al. (2013) and Zhang et al. (2014) for detailed descriptions of the WRF 4DVar data assimilation system and the WRF tangent linear model, respectively.

To include a hybrid covariance in the variational cost function, $\delta\mathbf{x}_0$ is separated into two terms:

$$
\delta\mathbf{x}_0 = \delta\mathbf{x}_0^c + \frac{1}{\sqrt{N-1}}\sum_{n=1}^{N}(\mathbf{a}_n \circ \mathbf{x}_{n,0}'^f). \tag{4}
$$

The first and second terms of (4) are the climatological and ensemble contribution to the analysis increment, respectively (Wang et al. 2008b). Each $\mathbf{a}_n \circ \mathbf{x}_{n,0}'^f$ in the second term is a Schur product between the vector $\mathbf{a}_n$ and the nth ensemble perturbation, where $\mathbf{x}_{n,0}'^f \equiv \mathbf{x}_{n,0}^f - \overline{\mathbf{x}}_0^f$. Following Lorenc (2003b), the ensemble information is introduced in the variational cost function using an additional term for the ensemble background information (J_e):

$$
\begin{aligned}
J(\delta\mathbf{x}_0^c, \alpha) &= \beta_c J_b(\delta\mathbf{x}_0^c) + \beta_e J_e(\alpha) + J_o(\delta\mathbf{x}_0^c, \alpha) \\
&= \beta_c \frac{1}{2}\delta\mathbf{x}_0^{c\mathrm{T}}\mathbf{B}^{-1}\delta\mathbf{x}_0^c + \beta_e\frac{1}{2}\alpha^{\mathrm{T}}\mathbf{A}^{-1}\alpha \\
&\quad + \frac{1}{2}\sum_{t=0}^{\tau}(\mathbf{H}_t\mathbf{M}_t\delta\mathbf{x}_0 - \mathbf{d}_t)^{\mathrm{T}}\mathbf{R}_t^{-1}(\mathbf{H}_t\mathbf{M}_t\delta\mathbf{x}_0 - \mathbf{d}_t)
\end{aligned} \tag{5}
$$

The α control variables in the second term on the right-hand side of (5) store the weighting coefficients ($\mathbf{a}_n$) for each ensemble perturbation, which are constrained by $\mathbf{A}$; matrix $\mathbf{A}$ is block diagonal, and holds the correlations used to localize the ensemble covariance. The climate and ensemble background cost terms are weighted by the scalar coefficients β_c and β_e to determine the contribution of the two background error covariances during the minimization. The quantities β_c and β_e are chosen to satisfy $(1/\beta_c) + (1/\beta_e) = 1$, which guarantees the weights assigned to the climate- and ensemble-based covariance sum to unity[3] (Wang et al. 2008b). Equation (5) provides the cost function used in E4DVar experiments for introducing ensemble information in 4DVar, which reduces to E3DVar for $\tau = 0$.

Using (4), the $\mathbf{H}_t\mathbf{M}_t\delta\mathbf{x}_0$ term in (5), which projects $\delta\mathbf{x}_0$ into observation space at future times, expands to

---

[2] In practice, the cost function is preconditioned by replacing $\delta\mathbf{x}_0$ with $\mathbf{U}\mathbf{v}$, where $\mathbf{U}$ is a square root of the background error covariance and $\mathbf{v}$ is the new control variable. For simplicity, this substitution is not shown in the manuscript.

[3] We enforce this constraint for simplicity only. Bishop and Satterfield (2013) show analytically that the optimal weighting coefficients do not have to follow this constraint.

$$\mathbf{H}_t\mathbf{M}_t\delta\mathbf{x}_0 = \mathbf{H}_t\mathbf{M}_t\delta\mathbf{x}_0^c + \mathbf{H}_t\frac{1}{\sqrt{N-1}}\sum_{n=1}^{N}(\mathbf{a}_n \circ \mathbf{M}_t x_{n,0}'^f). \quad (6)$$

To avoid the use of a tangent linear model, (6) can be approximated as

$$\mathbf{H}_t\mathbf{M}_t\delta\mathbf{x}_0 \approx \mathbf{H}_t\delta\mathbf{x}_0^c + \mathbf{H}_t\frac{1}{\sqrt{N-1}}$$
$$\times \sum_{n=1}^{N}\{\mathbf{a}_n \circ [M_t(\mathbf{x}_{n,0}^f) - \overline{M_t(\mathbf{x}_0^f)}]\}, \quad (7)$$

where $\mathbf{M}_t$ is neglected from the climate contribution (first term on the right-hand side), and the propagation of ensemble perturbations by $\mathbf{M}_t$ (second term on the right-hand side) is replaced with perturbations from a nonlinear ensemble forecast. The assumptions made in the first and second terms in (7) are equivalent to those made by 3DVar first guess at appropriate time (FGAT; Fisher and Andersson 2001) and 4DEnVar (Liu et al. 2008, 2009), respectively. The hybrid 4DEnVar system applied in this study minimizes (5) using the approximations made in (7) to incorporate static and ensemble covariance in the cost function. Unlike Liu et al. (2008, 2009), ensemble perturbations are stored in model space and projected into observation space using the tangent linear observation operator.

c. Coupled data assimilation

The three data assimilation methods tested in this study require the execution of both EnKF and variational data assimilation systems—which can be performed simultaneously. From an ensemble forecast, the members are separated into mean and perturbations, which are introduced into the variational system as the background state $\mathbf{x}_0^b$ and perturbation vectors $\mathbf{x}_{n,0}'^f$, respectively. To estimate the analysis state at the center of the window for E4DVar and 4DEnVar, we propagate the sum of analysis increments and background state forward from the beginning of the window.[4] Assuming all observations displaced through the assimilation window are valid at the analysis time (center of the window), the EnKF estimates the posterior perturbations, which are added to the maximum likelihood solution found during the variational analysis. The posterior ensemble is then used to initialize

the ensemble forecast for the next data assimilation cycle. This approach results in a two-way coupling of ensemble and variational data assimilation methods (Zhang et al. 2009) for E3DVar, E4DVar, and 4DEnVar.

d. Data assimilation configuration

For the 4D data assimilation experiments, observations are binned into 1-h intervals and assimilated over a 6-h period that spans ± 3 h from the analysis time. Likewise, the E3DVar experiment assimilates all observations over the same period, but assumes all data are valid at the same time. We also applied E3DVar with 3-h assimilation cycles (and a 3-h window), but results from this experiment are omitted from the manuscript because no significant improvements were found over the 6-h configuration. All three methods use the same climatological background error covariance, which is estimated over the previous month using 24- and 12-h forecast differences (Parrish and Derber 1992). We use the default covariance amplitudes and length scales for the climatological errors, since tuning these parameters yielded no significant change in performance. The ensemble contains 60 members, and its covariance is localized in the variational and EnKF components using a 900-km horizontal radius of influence and 34 vertical levels for vertical localization. Here, the variational methods perform a model-space Gaussian localization using a recursive filter (Wang et al. 2008a), while the EnKF applies a Gaspari and Cohn (1999) localization function directly in the Kalman gain matrix when assimilating each observation.

Ensemble spread between data assimilation cycles is maintained by inflating the posterior ensemble perturbations after each analysis using the "covariance relaxation" technique described in Zhang et al. (2004). This method uses 80% of the prior perturbations and 20% of the posterior perturbations to update the posterior ensemble members. The data assimilation sensitivity to localization and inflation parameters was explored in offline experiments before deciding on the current configuration. Our choice of weighting coefficients for the climate and ensemble background cost terms come from results of sensitivity tests performed in previous regional modeling studies (Zhang and Zhang 2012; F. Zhang et al. 2013). The specified weights translate into a hybrid configuration that uses 20% of the static covariance and 80% of the ensemble covariance. Extensive tests of hybrid E4DVar and 4DEnVar in Poterjoy and Zhang (2015) using a low-dimensional dynamical system show the methods having a similar sensitivity to covariance localization and inflation, window length, ensemble size, and

---

[4] Another option is to use increments from the middle of the assimilation window. Section 4f of Poterjoy and Zhang (2015) provides an explanation as to why the approach taken in the current study is more accurate than the alternative method when observations are spaced equally in time through the assimilation window.
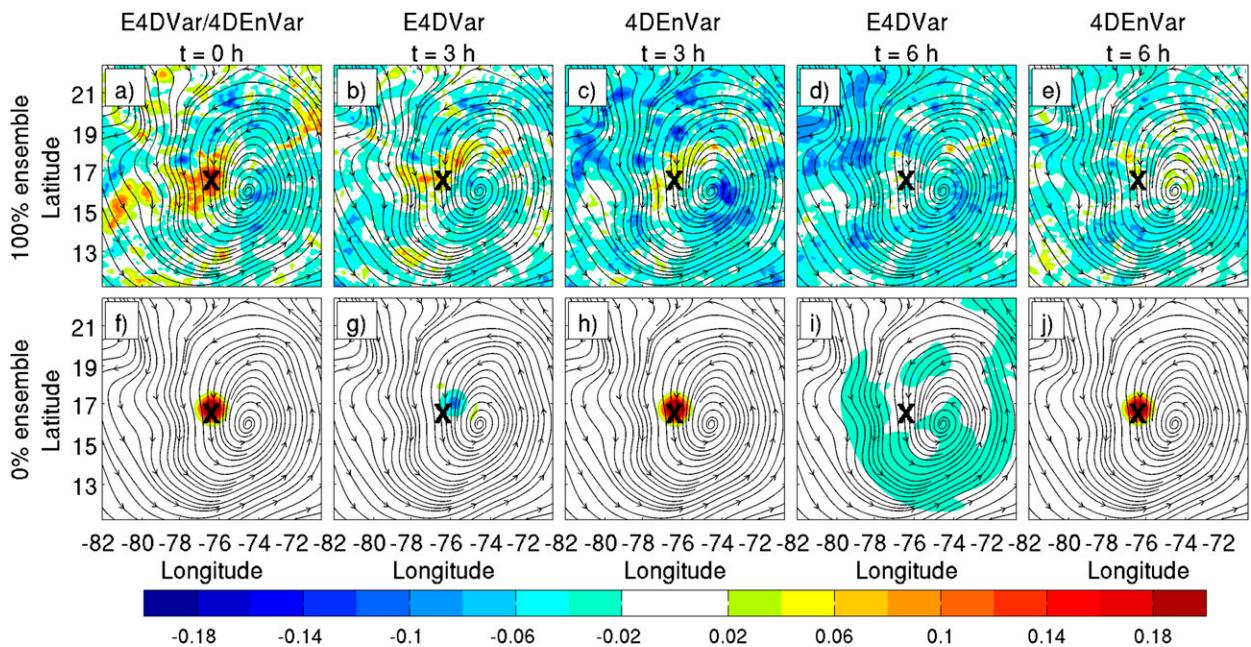
FIG. 1. The 850-mb θ analysis increments after assimilating a single 850-mb temperature observation located at the $\times$ in the domain. The increments are generated using (top) 100% and (bottom) 0% ensemble background error covariances. Each panel contains increments generated using either E4DVar or 4DEnVar for an observation at (left to right) $t = 0$, $t = 3$, and $t = 6$ h in the assimilation window, as indicated by the labels above each column. System-relative streamlines at 850 mb are also overlaid for reference.

hybrid weighting coefficients. As a result, the optimal set of parameters for the two methods were found to be near identical for almost all model and observing system scenarios tested. While the parameter sensitivity of both methods may not be as comparable for complex applications, like tropical cyclogenesis in WRF, this result provides a basis for using the same parameter values for E4DVar and 4DEnVar, which is done in the current study. Maintaining the same parameter values between experiments also allows for a simpler interpretation of results. As applied in this study, the only difference between these two methods is the modeling of covariances via the tangent linear and adjoint operators or an ensemble of model trajectories.

To reduce the computational cost of running the tangent linear and adjoint models for the 4DVar component of E4DVar, we perform the analyses using a reduced 40.5-km resolution to minimize the cost function (Zhang et al. 2014). We also generate the background trajectory for calculating innovations and linearizing the model and observation operators only once for each cycle (i.e., we do not run multiple outer loops). Because the 4DEnVar and E3DVar analysis steps require less computing time, we make use of the full-resolution 13.5-km grid for the minimization in these experiments.

3. Single-observation experiments

E4DVar and 4DEnVar differ primarily in how each method models background error covariance in the assimilation window. In this section, we explore these differences by assimilating synthetic observations and comparing the resulting analysis increments. We apply E4DVar and 4DEnVar using ensemble and static covariance to assimilate single observations displaced at three different times in the assimilation window. The comparison is performed for the 1200 UTC 13 September cycling time using the background (ensemble forecast) from the cycling E4DVar experiment. This cycle occurs 24 h prior to genesis, during a time in which the low-level mesoscale vortex undergoes a large increase in circulation (see discussion in section 4a). We create 850-mb temperature observations at 0, 3, and 6 h from the analysis time by taking temperature values from a 6-h forecast from the beginning of the assimilation window and adding a 1-K error. We then assimilate the observations separately using E4DVar and 4DEnVar (Fig. 1) to show potential temperature (θ) increments at the beginning of the assimilation window associated with observations at times $t = 0$, $t = 3$, and $t = 6$ h. These experiments are performed using 100% and 0% contributions of ensemble covariance to demonstrate how each method estimates

time covariances from the ensemble and static covariance components.

E4DVar and 4DEnVar provide identical analysis increments for observations at the beginning of the assimilation window ($t = 0$ h, Figs. 1a,f), but clear differences emerge for observations placed at future times. The θ increments change sign over the window, becoming negative for observations at $t = 3$ h and $t = 6$ h. We suspect that uncertainty in vortex position contributes the most to this sign change: negative increments produced from assimilating observations later in the assimilation window contain increasingly larger scales, which serve to correct track error inferred from future observations. Nevertheless, intensity errors are also large at this stage of the storm, which further complicates the interpretation of these experiments. The increments in the top row of Fig. 1, are produced using ensemble covariance only. In this case, the differences between increments for $t > 0$ h result from the following: 1) the unique localization strategies applied by the methods—E4DVar localizes the ensemble covariance at the beginning of the window, then propagates information forward using linearized models, whereas 4DEnVar performs the localization after estimating time covariances; and 2) simplifications introduced in the WRF tangent linear and adjoint models used for E4DVar. For the application chosen to perform these comparisons, the effects of using a different localization strategy are accelerated by the rapid development of the storm over the assimilation window. The sensitivity of E4DVar and 4DEnVar to forecast error growth was explored previously in Poterjoy and Zhang (2015) using the 40-variable Lorenz (1996) system with varying degrees of model forcing. Results from this study show that faster forecast error growth in the assimilation window leads to a faster degradation of 4DEnVar analyses, because of the localization strategy used to remove spurious temporal correlations. This finding explains some of the differences seen in the 100% ensemble covariance cases plotted in the top row of Fig. 1. Likewise, the elevated forecast error growth rate for this application may also accelerate the effects of errors introduced by imperfect linearized models in 4DVar, which may degrade the accuracy of E4DVar increments in this example.

In the other limiting case, where 0% of the ensemble covariance is used, increments plotted in the bottom row of Fig. 1 reflect the contribution of the static background error covariance. The E4DVar increments in this row (Figs. 1f,g,i) are identical to traditional implementations of 4DVar using a climate-based background covariance, while the 4DEnVar increments (Figs. 1f,h,j) are identical to 3DVar-FGAT. As the observation is displaced

further in time, the structure of the E4DVar increments becomes closer to the 100% ensemble covariance case (i.e., a large area of negative θ increments forms around the developing storm). The result follows from the tangent linear model and its adjoint having more time to evolve flow-dependent information over the assimilation window. Comparisons using hybrid covariance in the two systems are omitted from Fig. 1, because the resulting increments are a linear combination of the 100% and 0% ensemble cases described in this section [cf. (4)].

4. Results from cycling experiments

a. Analysis storm structure

In this section, we compare E3DVar, E4DVar, and 4DEnVar analyses over a weeklong cycling period between 0600 UTC 8 September and 0000 UTC 15 September, which spans the development and intensification of Karl. For this comparison, we focus on the kinematic and thermodynamic characteristics of analyses in the vicinity of the storm. As described in PZ14a, we estimate the pregenesis storm center by averaging the 950- and 700-mb circulation centers, which are designated as the positions that maximize the azimuthal mean winds within a 3° radius at each level. If a tropical storm is present in the simulation (10-m winds exceeding 18 m s^{-1}), we perform a Barnes's analysis (Barnes 1964) using the 10-m, 850-mb, and 700-mb vorticity fields to locate the storm center. Figures 2 and 3 summarize the evolution of the pregenesis disturbance in each set of analyses by comparing the time series of area-averaged quantities from 1200 UTC 9 September to 0000 UTC 15 September, following a short spinup period between 0600 UTC 8 September and 0600 UTC 9 September.

Figures 2a,b compare 950- and 500-mb relative vorticity (ζ) values averaged within 3° of the storm center. In all three experiments, the evolution of the low- and midlevel circulation at the meso-α scale is qualitatively consistent with the synopsis provided by Stewart (2010), likely because of the large spatial and temporal extent of in situ observations assimilated. The low-level circulation increases on 11 September, followed by a decrease between 11 and 13 September before strengthening until Karl intensified into a tropical storm at 1800 UTC 14 September (indicated by the vertical dashed line in Fig. 2). Meanwhile, the strength of the midlevel vortex increases steadily throughout the entire cycling period (Fig. 2b). The development of the 950-mb vortex between 13 and 15 September in the analyses coincides with the alignment of the low- and midlevel circulation
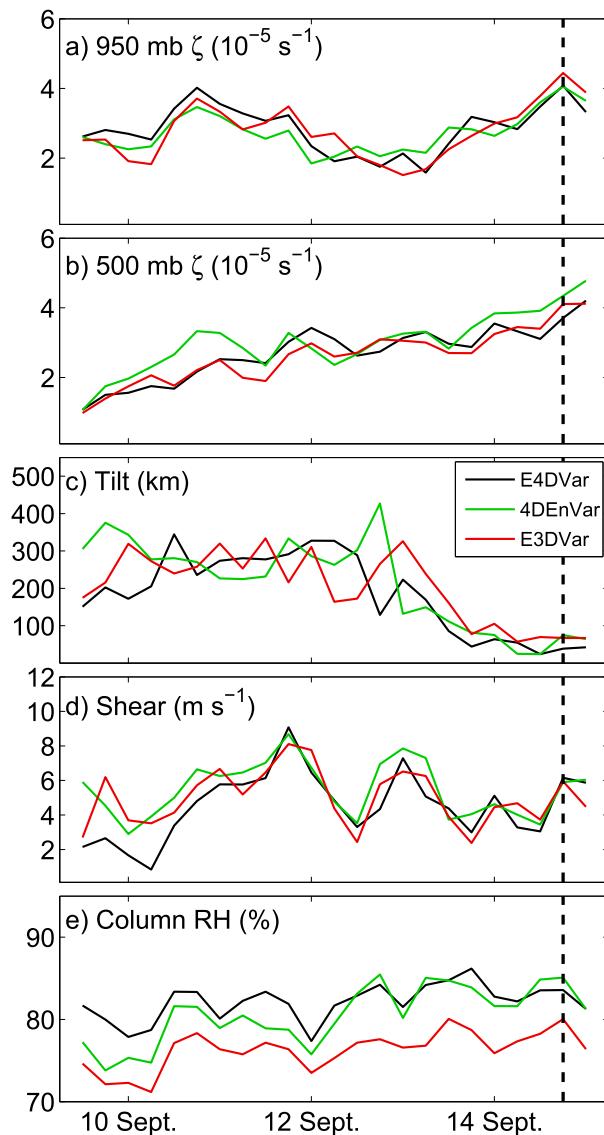
FIG. 2. The 3° average (a) 950- and (b) 500-mb relative vorticity, (c) 950–500-mb tilt and (d) vertical shear, and (e) 950–500-mb column relative humidity plotted every 6 h for E4DVar (black), 4DEnVar (green), and E3DVar (red) analyses between 1200 UTC 9 Sep and 0000 UTC 15 Sep. The vertical black dashed line indicates the time when Karl intensified into a tropical storm at 1800 UTC 14 Sep.

FIG. 3. As in Fig. 2, but for 1°-averaged quantities.

centers; this alignment is indicated by the time series of the 950–500-mb tilt magnitude plotted in Fig. 2c. The vortex alignment in the analyses leads to a reduction in the local 950–500-mb vertical wind shear (Fig. 2d), which we estimate from winds within 3° of the storm center. The plotted shear magnitude takes into account the environmental shear, as well as the effects of asymmetries induced by the tilted vortex. Mechanisms that can contribute to the alignment of a vortex in vertical wind shear are discussed in Schecter et al. (2002), Nolan

and McGauley (2012), Rappin and Nolan (2012), and Zhang and Tao (2013). Poterjoy and Zhang (2014b) found vortex alignment to be an important factor in determining when tropical cyclogenesis occurs in simulations initialized from EnKF and 4DVar analyses of this event, and noted the importance of dropsondes near the pregenesis disturbance in capturing the low-level development of the vortex with time.

With the exception of a faster vortex alignment in the E4DVar analyses, the three data assimilation methods yield qualitatively similar kinematic structure for meso-α-scale features of the disturbance preceding genesis. One notable distinction between analyses produced from the 4D data assimilation methods and E3DVar is the amount of near-saturated air surrounding the

developing storm. Figure 2e compares the time series of spatially averaged 950–500-mb column relative humidity (CRH), which is defined as the ratio of vertically integrated water vapor to vertically integrated saturation water vapor. Over the cycling period examined here, the E3DVar analyses remain 2%–8% (RH) drier than E4DVar and 4DEnVar, which inhibits the generation of convectively induced vorticity anomalies in the pregenesis disturbance, and decreases the likelihood of genesis (Hendricks et al. 2004). The relatively low values of CRH near the developing storm in E3DVar analyses are comparable to EnKF and 4DVar analyses discussed in Poterjoy and Zhang (2014b) for Karl (see their Fig. 3). In addition, verifications of deterministic forecasts in section 4b show that E4DVar and 4DEnVar improve the representation of thermodynamic variables over E3DVar. These results suggest that the hybrid 4D data assimilation methods capture moist processes related to Karl's genesis more accurately than the ensemble filters and non-ensemble methods tested for this case study (PZ14a; Poterjoy and Zhang 2014b).

Figure 3 compares the same quantities plotted in Fig. 2, except values are averaged within 1° of the storm center. Here, distinctions emerge with regard to the inner core of the storm in each set of analyses. The 13–15 September assimilation cycles reveal a gradual intensification of the midlevel vortex in the E4DVar and 4DEnVar data assimilation experiments (Fig. 3b); the E3DVar analyses contain a similar increase in circulation at these times, except the vortex remains much weaker. Following the increase in circulation in the middle troposphere, the low-level vortex intensifies rapidly between 14 September and the end of the cycling period on 15 September. The intensification of the inner-core midlevel vortex in E4DVar and 4DEnVar analyses follows a 24-h period of elevated CRH values. These magnitudes agree with the 80% average relative humidity threshold found in WRF simulations by Nolan (2007) to be a precursor for genesis. Furthermore, the cycle at which the initial intensification occurs in each experiment is consistent with the rate at which the atmosphere approaches saturation in the volume used to perform these calculations. The E4DVar experiment produces elevated values of CRH earlier in the cycling period and develops a mesoscale vortex faster than E3DVar and 4DEnVar in the analyses.

b. Deterministic forecasts

Here we compare deterministic forecasts initialized with the analyses described in section 4a. The prediction accuracy in each experiment is quantified by comparing forecasts with routine soundings and 234 total dropsonde measurements collected during PREDICT and

GRIP. We again focus on meteorological fields in the vicinity of the storm by calculating forecast root-mean-square differences (RMSDs) to observations within 800 km of the National Hurricane Center (NHC) best-track center.[5] This verification is performed for zonal and meridional wind (u and v), temperature (T), and water vapor mixing ratio (q_v). We average the RMSDs over the volume that spans the verification region between 950 and 200 mb, and temporally over 28 cycles between 0600 UTC 8 September and 0000 UTC 15 September. Results are plotted for forecast times between 0 and 72 h in Fig. 4, with vertical bars that indicate the 99% confidence interval, assuming the samples of RMSDs are uncorrelated in time and distributed normally. Figure 5 provides a more detailed overview of forecast errors for 24- and 72-h lead times by showing forecast bias and RMSDs averaged over 9 vertical columns between 1000 and 200 mb (observations within 50 mb of each level are used). This figure also indicates the number of verifying observations for each level on the right-hand side of each plot.

The 0-h verification in Fig. 4 measures the fit of observations to the analysis. Because the verifying observations include data assimilated by all three schemes (i.e., measurements from MADIS and PREDICT), a component of the mean RMSDs at the initial time comes from a model state that is dependent on the verifying observations. For this reason, we omit an interpretation of the analysis accuracy using the 0-h RMSDs. The time series of errors beginning from 0 h, however, provides some evidence regarding imbalance in the initial conditions for the two 4D methods, which require a 3-h integration from the beginning of the assimilation window. For example, the T RMSDs decrease immediately after 0 h in the 4DEnVar experiments. Likewise, 4DEnVar produces the largest 0-h RMSDs for u, v, and q, which we verified to be caused in part by elevated gravity wave activity in the 3-h forecasts to the analysis time (not shown). Previous studies using 4DEnVar in global models have encountered similar issues related to balance because of the use of a discontinuous set of data to characterize the 4D prior statistics (e.g., Kleist and Ide 2015). Lorenc et al. (2015) and Buehner et al. (2015) both found benefits using four-dimensional incremental analysis update (4DIAU) to control noise during initialization. The E4DVar system used in this study uses a weak digital filter constraint 4DVar that penalizes high-frequency oscillations in the

---

[5] We use an 800-km radius around the best-track center for this analysis because this region includes nearly all the PREDICT and GRIP dropsondes at each verification time.
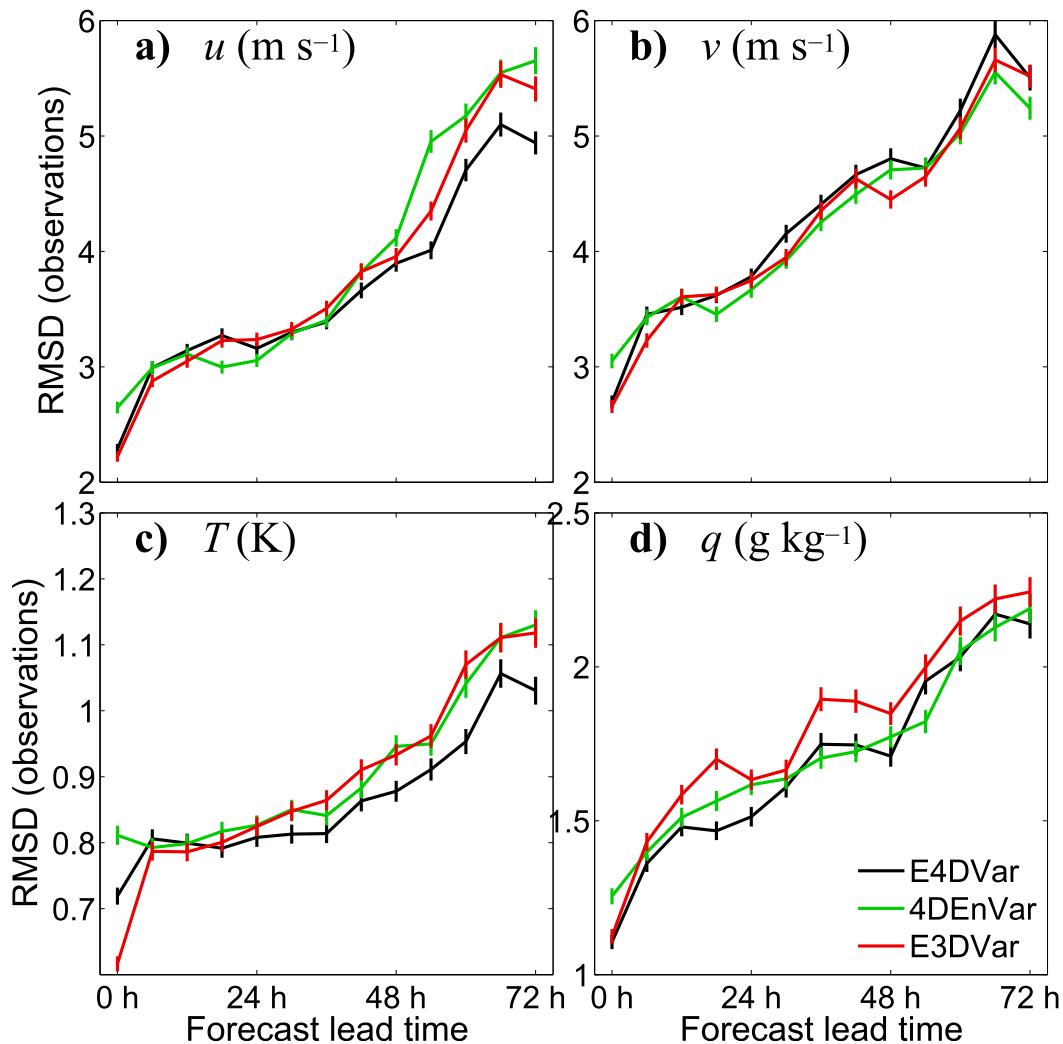
FIG. 4. Deterministic forecast RMSDs for E4DVar (black), 4DEnVar (green), and E3DVar (red) experiments. Vertical lines (whiskers) indicate 99% confidence interval for the RMSDs: (a) u, (b) v, (c) T, and (d) q.

cost function, but no such mechanism is available in the WRF-4DEnVar system used in this study. PZ14a used the digital filter constraint in their comparison of 4DVar and E4DVar because of minor benefits exhibited by 4DVar with the additional cost function term. This benefit occurs mostly because the increments produced during inner-loop iterations depend largely on a climatological covariance estimate, which represents background errors rather poorly for tropical cyclone applications. These increments result in an elevated level of imbalance, which can be dealt with using the digital filter. Because this problem is partially alleviated by the improved covariance estimates in E4DVar, the role of this filter becomes less important. Offline tests using E4DVar without the digital filter constraint show no significant benefit of this additional cost function term, so we use the same E4DVar experiment in PZ14a

and the current paper to maintain consistency between the two studies. While repeating the E4DVar experiment without the digital filter would provide a more consistent comparison between the methods, we have no reason to suspect the digital filter changes the major results of this study.

We use short-term forecasts after 0 h to verify the effectiveness of the data assimilation methods in producing accurate predictions of state variables in the region surrounding the disturbance. From this verification, the 4DEnVar and E3DVar experiments have comparable predictive skill for u, v, and T out to 72 h. Forecasts from E4DVar analyses contain similar v RMSDs over the same forecast range, but have significantly lower u and T RMSDs than E3DVar and 4DEnVar for lead times greater than 48 h. The lower u and T RMSDs in E4DVar forecasts are found mostly in the upper model
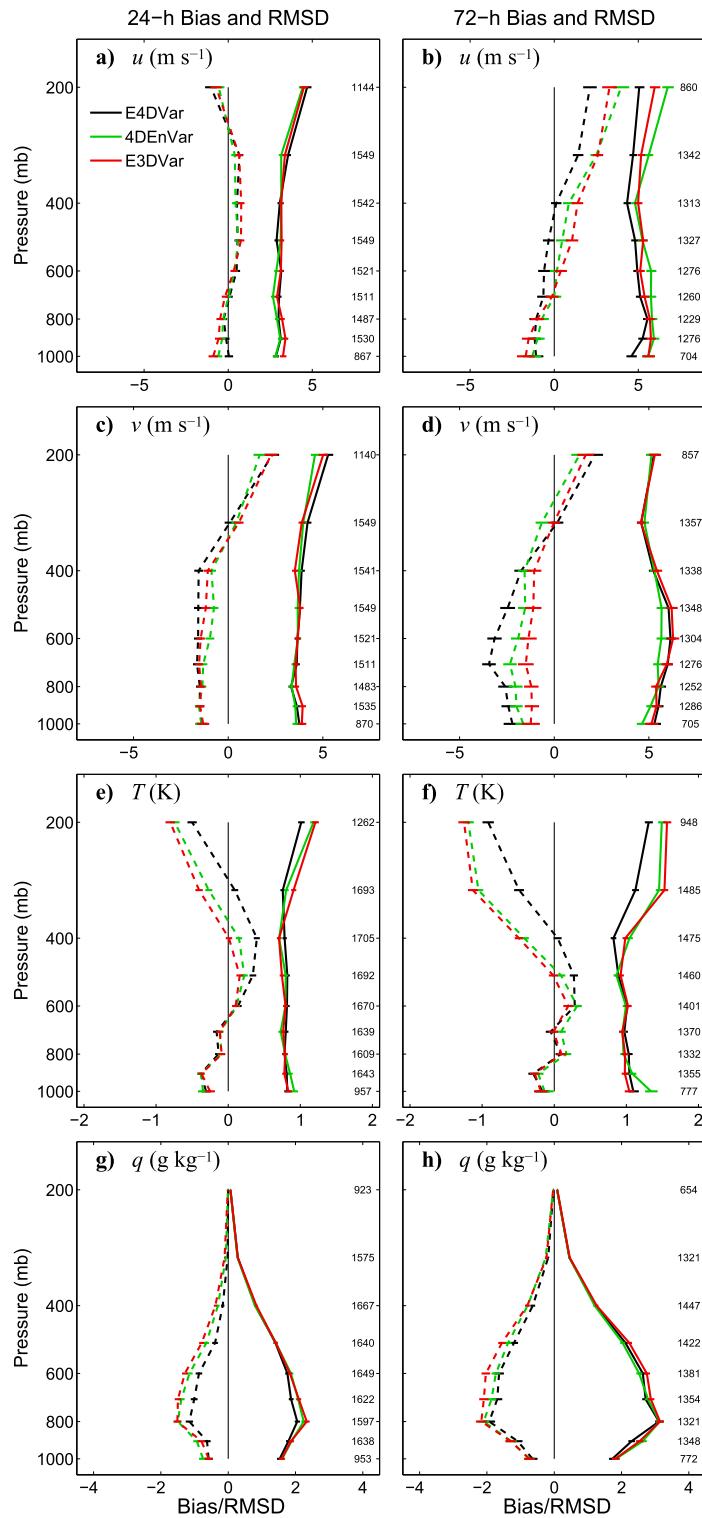
FIG. 5. Vertical profiles of (left) 24-h and (right) 72-h deterministic fore-cast bias (dashed lines) and RMSD (solid lines) for E4DVar (black), 4DEnVar (green), and E3DVar (red) experiments: (a),(b) *u*; (c),(d) *v*; (e),(f) *T*; and (g),(h) *q*. Horizontal lines (whiskers) indicate 99% confidence interval for the metrics; the numbers along the right-hand side of each plot are the number of verifying observations.
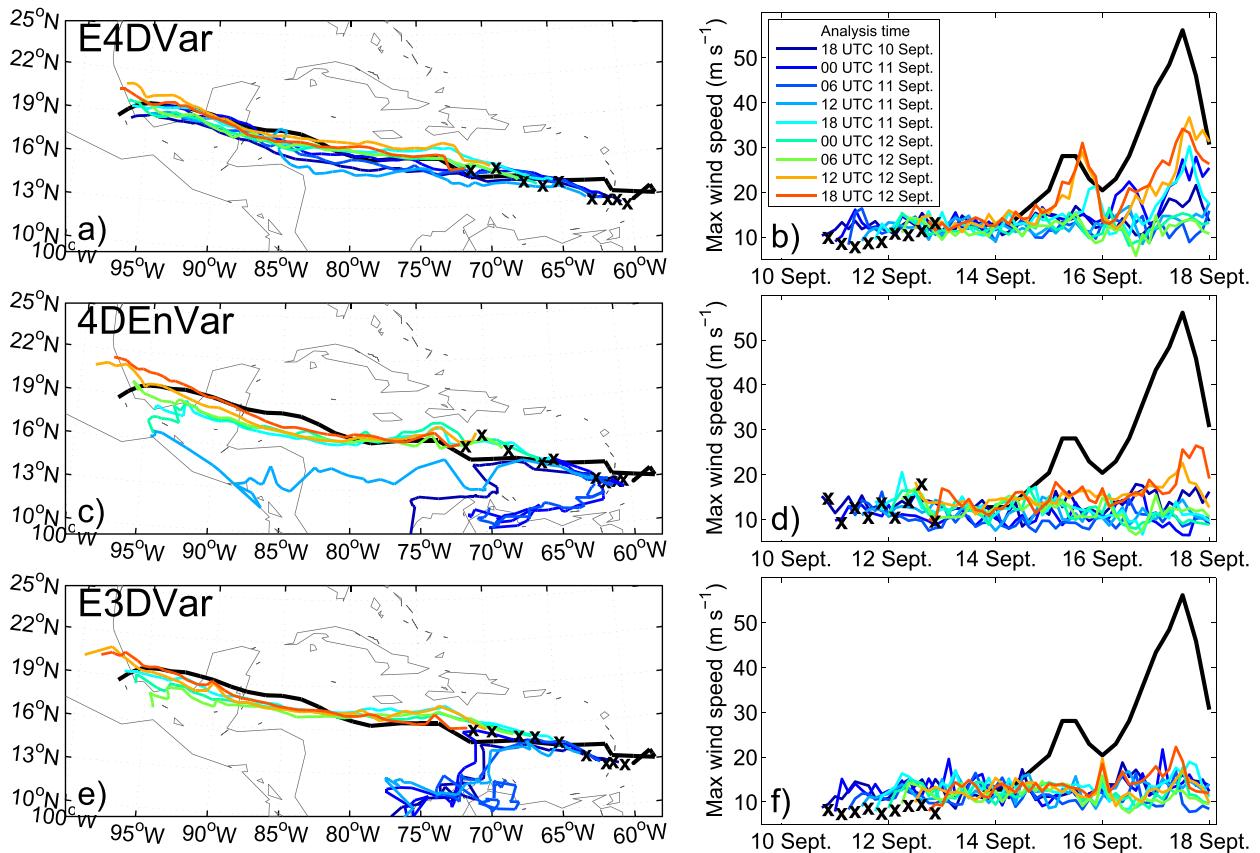
FIG. 6. Deterministic (left) track and (right) intensity forecasts from (top) E4DVar, (middle) 4DEnVar, and (bottom) E3DVar experiments between 1800 UTC 10 Sep and 1800 UTC 12 Sep. Forecasts are colored according to initialization time and NHC best-track data are plotted in black.

levels (Fig. 5), and coincide with lower biases than E3DVar and 4DEnVar. On the other hand, E4DVar produces the largest bias for v in low levels, despite having RMSDs that are comparable to E3DVar and 4DEnVar for most model levels. Both 4D data assimilation methods produce a more accurate representation of moisture than E3DVar near the disturbance, which is reflected in the significantly lower RMSDs for q_v starting from 12 h. Figures 5g,h indicate that most of the forecast error in these experiments is due to a negative bias, which is smaller for the E4DVar and 4DEnVar cases at most levels. This result suggests that the elevated values of CRH found near the disturbance in E4DVar and 4DEnVar analyses, compared to E3DVar in Figs. 2e and 3e, provide a better depiction of Karl's pregenesis relative humidity field.

We also verify the predicted track and intensity of the developing cyclone in deterministic forecasts. Here, the performance of the data assimilation techniques depends on how well each method captures features in analyses that produce accurate predictions of the storm track before and after genesis, as well as the correct

timing of genesis and rapid intensification in the simulations. Figures 6 and 7 show deterministic track and intensity (maximum 10-m wind speed) forecasts, initialized every 6 h between 1800 UTC 12 September and 0000 UTC 15 September. Simulations are run through the entire life cycle of Karl, thus providing insight into the long-range forecast performance of the three data assimilation methods. Black lines plotted in Figs. 6 and 7 represent the observed track and intensity, which are taken from the NHC best-track dataset; pregenesis storm positions come from the wave-tracking product described in Wang et al. (2012). Table 1 also provides the mean RMSEs, calculated from all deterministic forecasts plotted in these figures at 0-, 6-, 12-, 24-, 48-, and 96-h forecast lead times.

For forecasts initialized 3–4 days before genesis (Fig. 6), the E4DVar forecasts propagate the pregenesis wave westward, and produce accurate genesis predictions as early as 1200 UTC 12 September, 54 h before Karl became a tropical storm. Forecasts from 4DEnVar and E3DVar analyses fail to capture the correct translation speed and direction of the wave until 1800 UTC
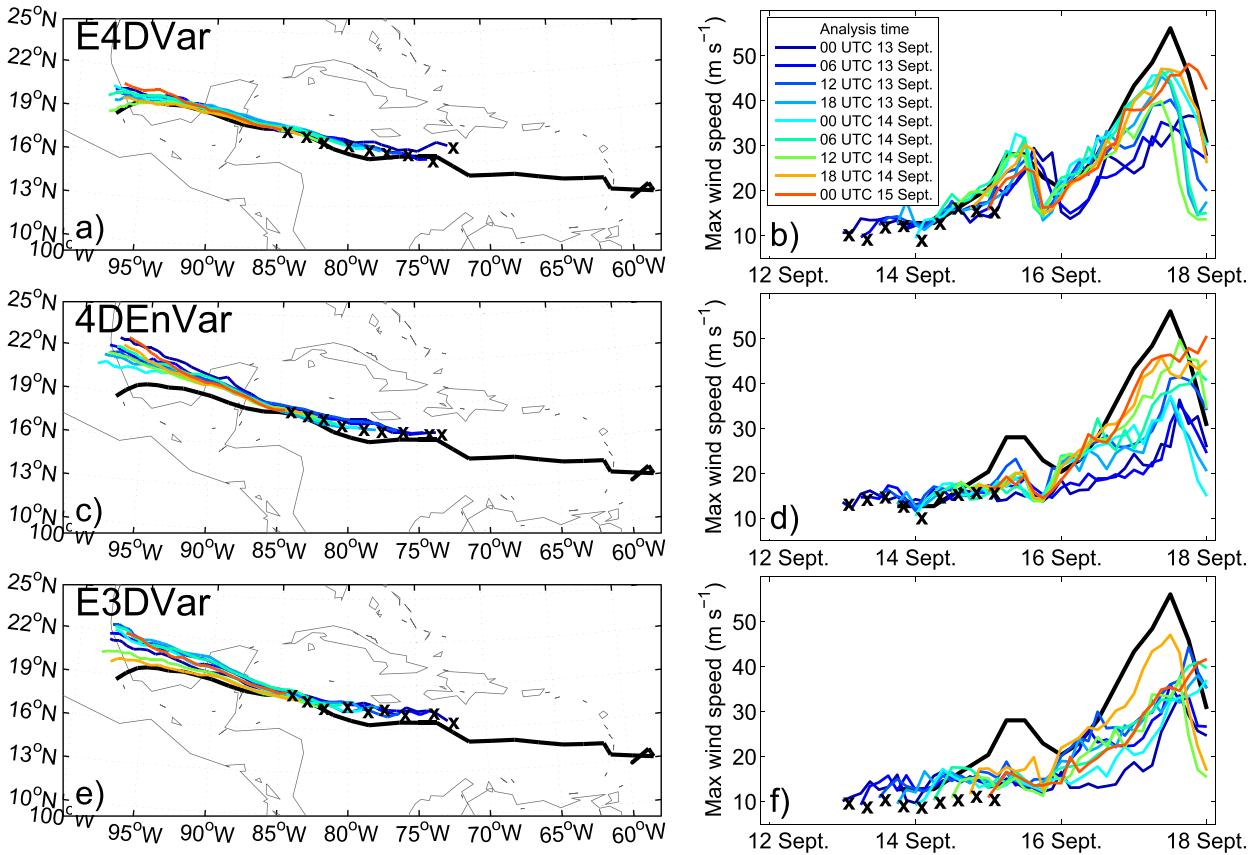
FIG. 7. As in Fig. 6, but for forecasts initialized between 0000 UTC 13 Sep and 0000 UTC 15 Sep.

11 September, which coincides with the second PREDICT flight mission. Likewise, none of the 4DEnVar and E3DVar forecasts prior to 13 September capture Karl's genesis before making landfall on the Yucatan Peninsula. Figure 7 compares the track and intensity forecasts from the three experiments within two days of Karl becoming a tropical storm. During this period, all three experiments produce a persistent right-of-track forecast bias, which also occurred in operational hurricane model forecasts for this event (Stewart 2010). The variability in Karl's track between data assimilation cycles after 1800 UTC 11 September also remains relatively low, which is consistent with operational

Global Forecast System forecasts—used in this study for lateral boundary conditions. Nevertheless, the E4DVar forecast errors are consistently smaller than forecast errors from 4DEnVar and E3DVar analyses. In addition to providing the most accurate track forecasts, the E4DVar analyses produce superior forecasts for the timing of Karl's development, which leads to better predictions of the storm's rapid intensification over the Bay of Campeche. The faster development of Karl in E4DVar analyses—compared to 4DEnVar and E3DVar—is demonstrated by the 1° averaged vorticity comparisons in Figs. 3a,b. In the 4DEnVar and E3DVar simulations, Karl's development is delayed

TABLE 1. Forecast track and intensity (maximum 10-m wind) errors averaged over all deterministic forecasts plotted in Figs. 6 and 7.

		Track RMSE (km)					Intensity RMSE (m s^{-1})		
		E4DVar	4DEnVar	E3DVar			E4DVar	4DEnVar	E3DVar
Lead time (h)	0	131.25	119.54	116.48	Lead time (h)	0	1.83	1.68	5.46
	6	141.56	197.89	136.08		6	1.88	2.19	4.81
	12	153.11	206.89	148.78		12	1.63	4.86	6.28
	24	205.00	258.58	176.33		24	2.95	6.24	8.01
	48	203.86	470.06	272.77		48	4.56	5.08	8.14
	96	286.75	994.44	800.42		96	14.02	15.74	15.11

until the weather system crosses the Yucatan Peninsula, which leads to a persistent low bias in the intensity forecasts (Fig. 7).

5. Results from noncycling experiments

For the experiments analyzed in section 4, the mean and covariance from a given data assimilation cycle is propagated forward to the next cycle and used to define the new prior error distribution. The cycling causes errors to accumulate between analyses because of deficiencies in how each method estimates posterior quantities. One negative outcome of the cumulative errors is that each method may respond differently to random noise introduced by sampling errors in the prior ensemble statistics and observation errors, thus affecting the outcome of the results presented in section 4. The impact of cumulative errors during cycling can be especially large for tropical cyclogenesis cases, where the underlying dynamics are very sensitive to small-scale initial condition errors in kinematic and thermodynamic variables such as vorticity and humidity (e.g., Sippel and Zhang 2008, 2010; Zhang and Tao 2013; Munsell et al. 2013). To remove differences due to cumulative errors introduced by the data assimilation system and observation noise, we perform a new set of analyses and forecasts using the same prior ensemble in the three data assimilation systems. For these experiments, each method uses forecast members from the cycling E4DVar experiment discussed in section 4. We generate new analyses and deterministic forecasts for 10 cycles between 1800 UTC 12 September and 0000 UTC 15 September, which covers 48 h before Karl's genesis, as well as its initial intensification on 15 September. Results from these experiments depend exclusively on each method's unique data assimilation strategy, which allows for an investigation into possible systematic differences in the three methods.

We also performed experiments using a hybrid configuration of 4DEnVar that does not include the climate-based background error covariance **B** in the hybrid cost function (denoted 4DEnVar-ep). Instead, we introduce **B** by replacing the ensemble forecast perturbations at the beginning of the window with hybrid perturbations:

$$\mathbf{x}_{n,0}^{\prime f} \leftarrow \frac{1}{\beta^c}\mathbf{x}_{n,0}^{\prime s} + \frac{1}{\beta^e}\mathbf{x}_{n,0}^{\prime f}, \qquad (8)$$

where each $\mathbf{x}_{n,0}^{\prime s}$ is drawn randomly from $N(0, \mathbf{B})$. This approach allows 4DEnVar to evolve perturbations from the static contribution of the background error covariance in the assimilation window to approximate the

time evolution of **B**. Poterjoy and Zhang (2015) demonstrate advantages of this application of 4DEnVar for cases when substantial model error is present in the forecast system. Representing **B** with hybrid perturbations has a similar effect as additive inflation because it allows forecast errors to grow in new orthogonal directions during the assimilation window. One disadvantage, however, is that the method introduces additional sampling error during the data assimilation because of the replacement of a full rank **B** with an ensemble representation.

Deterministic track and intensity forecasts initialized from each of the four hybrid analyses are plotted in Fig. 8, along with forecasts from the background mean state used to perform each analysis (denoted control throughout this section). When no data are assimilated to improve the background state, the forecasts generate persistently accurate track forecasts, but contain a slow bias in the intensification of Karl (Figs. 8a,b). Forecasts from 4DEnVar and E3DVar analyses in these experiments show improvements over the cycling cases discussed in section 4, which include the removal of a northward track bias and the prediction of Karl's genesis before making landfall on the Yucatan Peninsula (cf. Figs. 7 and 8). Nevertheless, these forecasts contain a slow bias for the timing of genesis that is similar to the delayed intensification found in the control simulations (Fig. 8b). The 4DEnVar-ep case yields marginal improvements over 4DEnVar, including more accurate intensity predictions in forecasts generated at earlier analysis times, but the advantages become negligible for analyses close to the genesis time. The E4DVar analyses produce a larger reduction in the intensity bias over the control, demonstrating additional skill over the other methods in forecasting Karl's evolution. Furthermore, the E3DVar and 4DEnVar analyses produce a larger amount of variability in track forecasts between data assimilation cycles compared to the E4DVar and control simulations. The increased variability in track is partly due to the higher magnitude and variance of analysis increments generated over these cycles (see discussion at the end of this section).

To summarize differences in the four sets of analyses, we again compare time series of 1°-averaged 950- and 500-mb ζ, vortex tilt, shear, and 950–500-mb CRH near the storm center. These quantities are plotted for each assimilation cycle between 1800 UTC 12 September and 0000 UTC 15 September in Fig. 9, along with values from the control state, or prior mean used to perform the data assimilation. The time series show the low- and midlevel circulation in the 4DEnVar and E4DVar analyses being comparable to the prior mean for most cycles of the experiment. 4DEnVar-ep tends to produce a

FIG. 8. Deterministic (left) track and (right) intensity forecasts from (a),(b) control; (c),(d) E4DVar; (e),(f) 4DEnVar; (g),(h) 4DEnVar-ep; and (i),(j) E3DVar analyses using same background between 1800 UTC 12 Sep and 0000 UTC 15 Sep. Forecasts are colored according to initialization time and NHC best-track data are plotted in black.

FIG. 9. As in Fig. 3, but for experiments using the same background.

results is the microphysics schemes used by the non-linear and tangent linear models: 4DEnVar uses the nonlinear WRF Model with single-moment 6-class microphysics to approximate covariances, while E4DVar uses a simplified large-scale condensation microphysics scheme in the tangent linear and adjoint models. 4DEnVar-ep also produces conditions that are more favorable for genesis than 4DEnVar, including a decrease in vortex tilt and increase in mean ζ and CRH near the storm center at some cycles. Nevertheless, the column of air in the vicinity of the 4DEnVar-ep vortex remains consistently less saturated than the E4DVar case, which is one reason why Karl's intensification rate in forecasts contains a slow bias that is similar to the 4DEnVar experiment at most cycles.

The time series of analysis quantities in Fig. 9 show differing magnitudes for analysis increments generated by each method during the 10 cycles examined in this section. To compare the evolution of these increments with time, Fig. 10 shows the mean squared difference between the control simulations and forecasts generated from each analysis; values are plotted for the 850-mb ζ (left column) and CRH (right column) from 0 to 72 h. The increments are separated according to wavelength (L) using high- and low-pass filters, then binned to reflect increments at the system scale ($L > 150$ km), intermediate scale ($50 < L < 150$ km), and convective scale ($L < 50$ km). The mean squared increments are averaged spatially over a 2000×2000 km^2 region near the storm location, and then averaged over the 10 forecasts in this sample (solid lines in Fig. 10). Likewise, dashed lines in Fig. 10 show the standard deviation of the mean squared increments.

On average, E4DVar produces the smallest ζ analysis increments for all three sets of filtered data. At the system scale, the average squared ζ differences between E4DVar and control simulations is less than half the values found in the other data assimilation experiments at the analysis time. The E4DVar increments grow at a faster rate, and approach the same magnitude as increments from the other experiments by 36 h (Fig. 10a). A similar result occurs for the intermediate- and convective-scale ζ field (Figs. 10b,c), except the rate at which the increments reach saturation is increased for the smaller scales. The growth of CRH increments with time is consistent with the ζ increments for the intermediate and convective scales because of the relationship between moist processes and the intensification of the storm. The system-scale CRH increments, however, have negative or zero growth during the first 6 h of the 4DEnVar-ep and E3DVar forecasts, suggesting that portions of the moisture and temperature increments made by these data assimilation

stronger low-level vortex between 13 and 14 September, which likely causes the faster intensification of Karl in these cycles compared to 4DEnVar. Likewise, E3DVar produces the weakest 950- and 500-mb circulation from 14 September onward, which is consistent with analyses from the cycling data assimilation experiments (see Figs. 3a,b). Another result that is consistent with the cycling experiments is the persistently higher values of CRH in the E4DVar analyses compared to analyses produced by 4DEnVar and E3DVar (Fig. 9e). These experiments confirm that the higher values of saturation found in the E4DVar analyses, compared to the other experiments, come mostly from the use of a tangent linear model and its adjoint to model temporal co-variances. Another factor leading to differences in these

FIG. 10. Mean squared (left) 950-mb vorticity and (right) 950–500-mb CRH increments averaged every 3 h from 0 to 72 h in the forecast period for (solid black) E4DVar, (solid green) 4DEnVar, and (solid blue) 4DEnVar-ep with standard deviations (dashed lines). Increments are separated by wavelength according to (a),(b) $L > 150$ km; (c),(d) 50 km $< L < 150$ km; and (e),(f) $L < 50$ km.

methods do not persist through the forecast period. Both of these experiments also exhibit higher variance in the mean squared CRH increments at the initial time, suggesting a possible overfitting of the observations. E3DVar produces larger analysis increments because the assimilated observations are binned over a longer time window, thus leading to more substantial innovations. Likewise, 4DEnVar-ep produces larger analysis increments than the standard cost function form of 4DEnVar because the static component of the background error covariance is able to grow over the assimilation window via the hybrid ensemble perturbations. The increase in background error variance in 4DEnVar-ep leads to larger analysis increments from this method. The E4DVar and 4DEnVar experiments yield system-scale analysis increments to CRH that have a similar magnitude, and exhibit a log-linear

increase with time. Nevertheless, the E4DVar increments (for all wavelengths) approach a larger mean and variance by 24 h in the forecast period because of the faster intensification rate of the storm in these simulations.

6. Conclusions

This study uses three ensemble-variational hybrid data assimilation methods to generate analyses and forecasts for Hurricane Karl, an Atlantic storm targeted during the PREDICT and GRIP field campaigns. Three variational schemes are coupled with an EnKF data assimilation system, which generates posterior ensemble perturbations that are added to the hybrid variational analysis during each data assimilation cycle. The methods include hybrid E3DVar and E4DVar, as well as a hybrid 4DEnVar scheme that combines

3DVar-FGAT with temporal covariances estimated from an ensemble of model trajectories. Hybrid E4DVar and 4DEnVar apply two different strategies for assimilating observations over a time window: E4DVar uses linearized models to propagate implicitly in time a climate-based covariance and a localized ensemble covariance from the beginning of the assimilation window, while 4DEnVar uses a static climate-based covariance and localized covariances from a previously run ensemble forecast through the assimilation window. E4DVar has the advantage of propagating a full-rank hybrid background error covariance in time, while 4DEnVar benefits from using the full nonlinear model to estimate temporal covariances and is more computationally efficient to run on large parallel computing platforms. To the best of our knowledge, this study presents the first comparison of these two ensemble 4D data assimilation techniques for a tropical cyclone application.

The data assimilation systems are cycled over a weeklong period to assimilate routinely collected observations and dropsonde measurements from the PREDICT field campaign. The cycling period spans 4 days leading up to Karl's genesis, as well as its initial intensification, and deterministic forecasts from each hybrid analysis are run over the entire life cycle of the storm. The meso-α-scale vortex in the E3DVar, E4DVar and 4DEnVar analyses contains qualitatively similar kinematic structure throughout the cycling period, including comparable low- and midlevel circulation strength and tilt. The two 4D data assimilation systems, however, produce values of column saturation that are more favorable for genesis. Elevated relative humidity values found near the developing storm in E4DVar and 4DEnVar analyses coincide with a faster strengthening of the mid- and low-level mesoscale vortex in the analyses. The circulation region near the pregenesis disturbance approaches saturation faster in the E4DVar experiment than in the 4DEnVar experiment, which coincides with earlier genesis times in forecasts. This factor leads to more accurate predictions for Karl's formation before making its first landfall over the Yucatan Peninsula.

To remove the effects of errors that accumulate during successive data assimilation cycles, the three methods are compared using the same background ensemble in the variational analyses. These experiments are performed over a period that spans the pregenesis phase of Karl's development and initial intensification, revealing systematic differences in the mesoscale structure of the storm. Analyses produced from these experiments show a stronger circulation in the low-level (950 mb) vortex in the E3DVar and 4DEnVar cases, compared to E4DVar. Nevertheless, E4DVar analyses contain the highest values of column relative humidity in a deep layer near the disturbance center, which contributes to faster rates of intensification in forecasts. E4DVar also produces systematically smaller analysis increments than E3DVar and 4DEnVar, which grow faster in time during the first 0–12 h of the forecast period. This result suggests E4DVar targets different dynamical instabilities during 4D data assimilation than 4DEnVar, thus reflecting differences in how covariance is modeled by the two methods.

The current study highlights several differences between hybrid E4DVar and 4DEnVar for a weather application where mesoscale processes play an important role in the system's development. Additional data assimilation experiments and a more idealized model setup are necessary to fully explore some of the findings in this study. In particular, the modeling of covariances related to moist processes appears to be a major difference between E4DVar and 4DEnVar. We hypothesize that E4DVar may provide better estimates of temperature and moisture covariances because of its superior localization strategy. E4DVar may also benefit from using a simple linearized microphysics scheme, which improves the conditioning of the minimization problem. This study motivates further testing of E4DVar and 4DEnVar using a diverse set of case studies to gain a more comprehensive understanding of the advantages and deficiencies of these methods for mesoscale data assimilation.

REFERENCES

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

Barnes, S. L., 1964: A technique for maximizing details in numerical weather map analysis. *J. Appl. Meteor.*, **3**, 396–409, doi:10.1175/1520-0450(1964)003<0396:ATFMDI>2.0.CO;2.

Bishop, C. H., and D. Hodyss, 2011: Adaptive ensemble covariance localization in ensemble 4D-VAR state estimation. *Mon. Wea. Rev.*, **139**, 1241–1255, doi:10.1175/2010MWR3403.1.

——, and E. A. Satterfield, 2013: Hidden error variance theory. Part I: Exposition and analytic model. *Mon. Wea. Rev.*, **141**, 1454–1468, doi:10.1175/MWR-D-12-00118.1.

Braun, S. A., and Coauthors, 2013: NASA's Genesis and Rapid Intensification Processes (GRIP) field experiment. *Bull. Amer. Meteor. Soc.*, **94**, 345–363, doi:10.1175/BAMS-D-11-00232.1.

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. 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.

——, and Coauthors, 2015: Implementation of deterministic weather forecasting systems based on ensemble–variational data assimilation at Environment Canada. Part I: The global system. *Mon. Wea. Rev.*, **143**, 2532–2559, doi:10.1175/MWR-D-14-00354.1.

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

Courtier, P., J.-N. Thepáut, 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.

Dudhia, J., 1989: Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.*, **46**, 3077–3107, doi:10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2.

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.

——, and P. J. van Leeuwen, 2000: An ensemble Kalman smoother for nonlinear dynamics. *Mon. Wea. Rev.*, **128**, 1852–1867, doi:10.1175/1520-0493(2000)128<1852:AEKSFN>2.0.CO;2.

Fisher, M., and E. Andersson, 2001: Developments in 4D-Var and Kalman filtering. Tech. Memo. 347, ECMWF, 36 pp.

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.

Gustafsson, N., and J. Bojarova, 2014: Four-dimensional ensemble variational (4D-En-Var) data assimilation for the High Resolution Limited Area Model. *Nonlinear Processes Geophys.*, **21**, 745–762, doi:10.5194/npg-21-745-2014.

Hendricks, E. A., M. T. Montgomery, and C. A. Davis, 2004: The role of vortical hot towers in the formation of Tropical Cyclone Diane (1984). *J. Atmos. Sci.*, **61**, 1209–1232, doi:10.1175/1520-0469(2004)061<1209:TROVHT>2.0.CO;2.

Honda, Y., M. Nishijima, K. Koizumi, Y. Ohta, K. Tamiya, T. Kawabata, and T. Tsuyuki, 2005: A pre-operational variational data assimilation system for a nonhydrostatic model at Japan Meteorological Agency: Formulation and preliminary results. *Quart. J. Roy. Meteor. Soc.*, **131**, 3465–3475, doi:10.1256/qj.05.132.

Hong, S.-Y., J. Dudhia, and S.-H. Chen, 2004: A revised approach to ice-microphysical processes for the bulk parameterization of cloud and precipitation. *Mon. Wea. Rev.*, **132**, 103–120, doi:10.1175/1520-0493(2004)132<0103:ARATIM>2.0.CO;2.

Huang, X., and Coauthors, 2009: Four-dimensional variational data assimilation for WRF: Formulation and preliminary results. *Mon. Wea. Rev.*, **137**, 299–314, doi:10.1175/2008MWR2577.1.

Hunt, B. R., and Coauthors, 2004: Four-dimensional ensemble Kalman filtering. *Tellus*, **56A**, 273–277, doi:10.1111/j.1600-0870.2004.00066.x.

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

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.

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.

——, ——, and ——, 2009: An ensemble-based four-dimensional variational data assimilation scheme. Part II: Observing system simulation experiments with advanced research WRF (ARW). *Mon. Wea. Rev.*, **137**, 1687–1704, doi:10.1175/2008MWR2699.1.

Lorenc, A. C., 2003a: Modeling of error covariances by 4D-Var data assimilation. *Quart. J. Roy. Meteor. Soc.*, **129**, 3167–3182, doi:10.1256/qj.02.131.

——, 2003b: 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.

——, N. E. Howler, A. M. Clayton, S. R. Pring, and D. Fairbairn, 2015: Comparison of hybrid-4DEnVar and hybrid-4DVar data assimilation methods for global NWP. *Mon. Wea. Rev.*, **143**, 212–229, doi:10.1175/MWR-D-14-00195.1.

Lorenz, E. N., 1996: Predictability: A problem partly solved. *Proc. Seminar on Predictability*, Vol. 1, ECMWF, Reading, United Kingdom, 1–18.

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

Monin, A. S., and A. M. Obukhov, 1954: Basic laws of turbulent mixing in the surface layer of the atmosphere. *Tr. Akad. Nauk SSSR Geophiz. Inst.*, **24**, 163–187.

Montgomery, M. T., and Coauthors, 2012: The Pre-Depression Investigation of Cloud-Systems in the Tropics (PREDICT) experiment: Scientific basis, new analysis tools, and some first results. *Bull. Amer. Meteor. Soc.*, **93**, 153–172, doi:10.1175/BAMS-D-11-00046.1.

Munsell, E. B., F. Zhang, and D. P. Stern, 2013: Predictability and dynamics of a nonintensifying tropical storm: Erica (2009). *J. Atmos. Sci.*, **70**, 2505–2524, doi:10.1175/JAS-D-12-0243.1.

Noh, Y., W.-G. Cheon, and S.-Y. Hong, 2003: Improvement of the K-profile model for the planetary boundary layer based on

large eddy simulation data. *Bound.-Layer Meteor.*, **107**, 401–427, doi:10.1023/A:1022146015946.

Nolan, D. S., 2007: What is the trigger for tropical cyclogenesis? *Aust. Meteor. Mag.*, **56**, 241–266.

——, and M. G. McGauley, 2012: Tropical cyclogenesis in wind shear: Climatological relationships and physical processes. *Cyclones: Formation, Triggers and Control*, K. Oouchi and H. Fudeyasu, Eds., Nova Science Publishers, 1–34.

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, 2014a: Inter-comparison and coupling of ensemble and variational data assimilation approaches for the analysis and forecasting of Hurricane Karl (2010). *Mon. Wea. Rev.*, **142**, 3347–3364, doi:10.1175/MWR-D-13-00394.1.

——, and ——, 2014b: Tropical cyclogenesis predictability examined through ensemble analyses and forecasts for Hurricane Karl (2010) during PREDICT. *J. Atmos. Sci.*, **71**, 1260–1275, doi:10.1175/JAS-D-13-0291.1.

——, and ——, 2015: Systematic comparison of four-dimensional data assimilation methods with and without a tangent linear model using hybrid background error covariance: E4DVar versus 4DEnVar. *Mon. Wea. Rev.*, **143**, 1601–1621, doi:10.1175/MWR-D-14-00224.1.

Rabier, F., H. Järvinen, 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.

Rappin, E. D., and D. S. Nolan, 2012: The effect of vertical shear orientation on tropical cyclogenesis. *Quart. J. Roy. Meteor. Soc.*, **138**, 1035–1054, doi:10.1002/qj.977.

Schecter, D. A., M. T. Montgomery, and P. D. Reasor, 2002: A theory for the vertical alignment of a quasigeostrophic vortex. *J. Atmos. Sci.*, **59**, 150–168, doi:10.1175/1520-0469(2002)059<0150:ATFTVA>2.0.CO;2.

Sippel, J. A., and F. Zhang, 2008: A probabilistic analysis of the dynamics and predictability of tropical cyclogenesis. *J. Atmos. Sci.*, **65**, 3440–3459, doi:10.1175/2008JAS2597.1.

——, and ——, 2010: Factors affecting the predictability of Hurricane Humberto (2007). *J. Atmos. Sci.*, **67**, 1759–1778, doi:10.1175/2010JAS3172.1.

Skamarock, W. C., and Coauthors., 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., doi:10.5065/D68S4MVH.

Stewart, S. R., 2010: Tropical cyclone report Hurricane Karl (AL1312010) 14–18 September 2010. Tech. Rep. AL1312010, National Hurricane Center, 17 pp. [Available online at http://www.nhc.noaa.gov/data/tcr/AL132010_Karl.pdf.]

Sun, J., and N. A. Crook, 1997: Dynamical and microphysical retrieval from Doppler radar observations using a cloud model and its adjoint. Part I: Model development and simulated data experiments. *J. Atmos. Sci.*, **54**, 1642–1661, doi:10.1175/1520-0469(1997)054<1642:DAMRFD>2.0.CO;2.

Torn, R. D., and D. Cook, 2013: The role of vortex and environmental errors in genesis forecasts of Hurricanes Danielle and Karl (2010). *Mon. Wea. Rev.*, **141**, 232–251, doi:10.1175/MWR-D-12-00086.1.

Wang, X., and T. Lei, 2014: GSI-based four-dimensional ensemble-variational (4DEnsVar) data assimilation: Formulation and single-resolution experiments with real data for NCEP Global Forecast System. *Mon. Wea. Rev.*, **142**, 3303–3325, doi:10.1175/MWR-D-13-00303.1.

——, D. Barker, C. Snyder, and T. M. Hamill, 2008a: A hybrid ensemble transform Kalman filter (ETKF)-3DVAR data assimilation scheme for WRF. *11th Symp. on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface (IOAS-AOLS)*, San Antonio, TX, Amer. Meteor. Soc., P4.10. [Available online at https://ams.confex.com/ams/87ANNUAL/techprogram/paper_116592.htm.]

——, D. M. Barker, C. Snyder, and T. M. Hamill, 2008b: A hybrid ETKF–3DVar data assimilation scheme for the WRF model. Part I: Observation system simulation experiment. *Mon. Wea. Rev.*, **136**, 5116–5131, doi:10.1175/2008MWR2444.1.

——, 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.

Wang, Z., M. T. Montgomery, and C. Fritz, 2012: A first look at the structure of the wave pouch during the 2009 PREDICT-GRIP dry runs of the Atlantic. *Mon. Wea. Rev.*, **140**, 1144–1163, doi:10.1175/MWR-D-10-05063.1.

Zhang, F., and D. Tao, 2013: Effects of vertical wind shear on the predictability of tropical cyclones. *J. Atmos. Sci.*, **70**, 975–983, doi:10.1175/JAS-D-12-0133.1.

——, C. Snyder, and J. Sun, 2004: Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter. *Mon. Wea. Rev.*, **132**, 1238–1253, doi:10.1175/1520-0493(2004)132<1238:IOIEAO>2.0.CO;2.

——, 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 data assimilation in a limited-area weather prediction model. *Mon. Wea. Rev.*, **140**, 587–600, doi:10.1175/MWR-D-11-00023.1.

Zhang, X., X.-Y. Huang, and N. Pan, 2013: Development of the upgraded tangent linear and adjoint of the Weather Research and Forecasting (WRF) Model. *J. Atmos. Oceanic Technol.*, **30**, 1180–1188, doi:10.1175/JTECH-D-12-00213.1.

——, ——, J. Liu, J. Poterjoy, Y. Weng, F. Zhang, and H. Wang, 2014: Development of an efficient regional four-dimensional variational data assimilation system for WRF. *J. Atmos. Oceanic Technol.*, **31**, 2777–2794, doi:10.1175/JTECH-D-13-00076.1.

Zupanski, M., 2005: Maximum likelihood ensemble filter: Theoretical aspects. *Mon. Wea. Rev.*, **133**, 1710–1726, doi:10.1175/MWR2946.1.