

Intercomparison and Coupling of Ensemble and Four-Dimensional Variational Data Assimilation Methods for the Analysis and Forecasting of Hurricane Karl (2010)

JONATHAN POTERJOY AND FUQING ZHANG

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

(Manuscript received 10 December 2013, in final form 18 April 2014)

ABSTRACT

This study examines the performance of ensemble and variational data assimilation systems for the Weather Research and Forecasting (WRF) Model. These methods include an ensemble Kalman filter (EnKF), an incremental four-dimensional variational data assimilation (4DVar) system, and a hybrid system that uses a two-way coupling between the two approaches (E4DVar). The three methods are applied to assimilate routinely collected data and field observations over a 10-day period that spans the life cycle of Hurricane Karl (2010), including the pregenesis disturbance that preceded its development into a tropical cyclone. In general, forecasts from the E4DVar analyses are found to produce smaller 48–72-h forecast errors than the benchmark EnKF and 4DVar methods for all variables and verification methods tested in this study. The improved representation of low- and midlevel moisture and vorticity in the E4DVar analyses leads to more accurate track and intensity predictions by this system. In particular, E4DVar analyses provide persistently more skillful genesis and rapid intensification forecasts than the EnKF and 4DVar methods during cycling. The data assimilation experiments also expose additional benefits of the hybrid system in terms of physical balance, computational cost, and the treatment of asynoptic observations near the beginning of the assimilation window. These factors make it a practical data assimilation method for mesoscale analysis and forecasting, and for tropical cyclone prediction.

1. Introduction

Data assimilation presents a means of combining observations with knowledge of atmospheric dynamics to study the spatiotemporal evolution of weather systems. Under a strict set of assumptions regarding the error distributions for observed and modeled data, the ensemble Kalman filter (EnKF) and incremental fourdimensional variational data assimilation (4DVar) methods have proven to be the most effective to this date; see Lorenc (2003b), Kalnay et al. (2007), F. Zhang et al. (2011), and Buehner et al. (2010a,b) for theoretical and practical comparisons of these two approaches.

EnKF and 4DVar data assimilation methods are both optimal for linear systems with Gaussian errors, but the process by which each method produces an analysis is quite different. The EnKF uses an ensemble forecast to

DOI: 10.1175/MWR-D-13-00394.1

advance a flow-dependent background error covariance matrix between data assimilation cycles, thus acting as an approximation to the extended Kalman filter (Evensen 1994). 4DVar is a deterministic method that minimizes a cost function with respect to increments from a background state vector that are adjusted to provide the best fit to observations in a time window (Courtier et al. 1994). The sensitivity of the state vector to future observations is estimated using the adjoint of the forecast model and observation operators. While this feature of 4DVar introduces flow-dependent information into the state approximation, the scheme remains limited by the use of a static error covariance matrix at the beginning of the time window. This matrix is often based on climatological information, and formulated to provide an efficient minimization of the cost function (Lorenc 2003a).

A hybrid data assimilation system that uses a two-way coupling between EnKF and 4DVar methods (denoted E4DVar) was introduced by Zhang et al. (2009), and tested over a month-long regional modeling experiment with the Weather Research and Forecasting (WRF) Model (Zhang and Zhang 2012; Zhang et al. 2013). The

Corresponding author address: Fuqing Zhang, Department of Meteorology, The Pennsylvania State University, 503 Walker Building, University Park, PA 16802. E-mail: fzhang@psu.edu

WRF-E4DVar system was found to outperform the stand-alone EnKF and 4DVar methods on a relatively coarse domain with a 90-km horizontal grid spacing. It also proved to be less sensitive to sampling errors than EnKF, given that a mix of static and ensemble co-variance is used to quantify background errors. The Met Office has used a similar hybrid ensemble-4DVar data assimilation system operationally since July 2011, and achieved positive improvements over 4DVar on the global scale (Clayton et al. 2013). Kuhl et al. (2013) also show gains in forecast skill when ensemble information is introduced into the 4DVar data assimilation system used by the U.S. Naval Research Laboratory's global model.

Several advancements in efficiency for the WRF-E4DVar system have allowed for its application at the mesoscale since the studies of Zhang and Zhang (2012) and Zhang et al. (2013). In particular, a multi-incremental hybrid 4DVar method for WRF now minimizes the cost function using a reduced resolution for the tangent linear and adjoint models and ensemble perturbations; see Zhang et al. (2014, manuscript submitted to J. Atmos. Oceanic Technol.) for details regarding additional efficiency improvements in WRF-4DVar. The theoretical benefits of E4DVar for regional-scale data assimilation and modeling have motivated many of the recent improvements. These benefits include the ensemble evolution of background errors between data assimilation cycles, and the implicit evolution of background errors through an observation time window via the tangent linear and adjoint versions of the model (Zhang et al. 2009). Like 4DVar, the analyses rely on a full rank climatological covariance matrix, except the ensemble perturbations provide additional nonstationary, multiscale information regarding the background errors. These two aspects of E4DVar make it a valuable tool for applications in which large numbers of asynoptic observations are available for constraining a dynamical system that evolves at both long and short time scales. Bearing in mind these factors, we introduce the tropical cyclogenesis predictability problem as a suitable test for the relatively new data assimilation system.

The formation of a self-sustained tropical cyclone from a preexisting wave or synoptic-scale disturbance is a multiscale process that poses major challenges for our current state-of-the-art data assimilation systems (Dunkerton et al. 2009; Fang and Zhang 2011). The initialization of a model with accurate vertical vorticity and divergence in the vicinity of the pregenesis cyclone, along with an adequate amount of column moisture, are crucial factors for simulating tropical cyclogenesis. These elements of tropical disturbances have been well documented in numerous studies that investigate developing tropical cyclones (e.g., Hendricks et al. 2004; Reasor et al. 2005; Montgomery et al. 2006; Halverson et al. 2007; Sippel and Zhang 2008, 2010; Wang 2012; Torn and Cook 2013).

More recently, Poterjoy and Zhang (2014, hereafter referred to as PZ14) applied a cycling EnKF data assimilation system to study the genesis and predictability of Hurricane Karl (2010), and examine the effectiveness of field observations in improving analyses of the pregenesis disturbance. Dropsonde measurements that were collected during the Pre-Depression Investigation of Cloud-systems in the Tropics (PREDICT) field campaign were found to be crucial for predicting Karl's genesis in the 30 h leading up to the event. The predictability of Karl in earlier simulations was limited by initial-condition errors in the synoptic and mesoscale vorticity and moisture fields, thus highlighting the need for additional observations and more advanced data assimilation methods.

The current study compares EnKF, 4DVar, and E4DVar data assimilation methods over the life cycle of Karl, using the same observations and model setup as PZ14. We use analyses and forecasts from cycling data assimilation experiments to examine the genesis and rapid intensification of Karl in more detail, and evaluate the performance of the hybrid technique with respect to the benchmark EnKF and 4DVar systems. Unlike Zhang and Zhang (2012) and Zhang et al. (2013), the current study focuses on a single weather event and tests the data assimilation methods for a mesoscale weather application.

The organization of this manuscript is as follows. Section 2 describes the model and data assimilation systems, as well as the design of the experiments. We present the analysis and forecast results from each experiment in sections 3 and 4, and compare analysis increments and initial-condition balance in sections 5 and 6. The last section provides a summary and conclusions.

2. Experiment setup

a. WRF Model and case study

This study uses the Advanced Research WRF, version 3.4.1 (Skamarock et al. 2008), with the same domain configuration and physical parameterizations as PZ14 to simulate the development of Hurricane Karl. The model uses a 251×226 parent domain (black box in Fig. 1) with a spacing of 13.5 km, and a 253×253 storm-following two-way nested domain with a spacing of 4.5 km. Both domains have 35 vertical levels and a 5-hPa model top. Cumulus convection is represented explicitly in both domains, which was found to provide more accurate



FIG. 1. The location of the tropical weather system is indicated in black prior to genesis, blue during tropical storm intensity, and red during hurricane intensity. Dropsonde locations from PREDICT and GRIP flight missions are indicated by the colored circles.

temperature and moisture fields than simulations with parameterized convection in sensitivity experiments. The model simulations span a 10-day period in which Karl developed into a tropical cyclone and rapidly intensified before decaying over land (Fig. 1). The first model forecasts in the data assimilation cycles are initialized at 0600 UTC 8 September, when the pre-Karl disturbance formed from an area of disorganized convection on the northern coast of South America (Stewart 2010). The disturbance propagated westward for 6 days before eventually forming a tropical depression at 1200 UTC 14 September.

Observations of Karl include routinely collected surface and upper-air data from the National Oceanic and Atmospheric Administration (NOAA) Meteorological Assimilation Data Ingest System (MADIS), as well as dropsonde measurements that were taken during two field campaigns. The National Science Foundation (NSF) PREDICT experiment collected data over a 5-day period leading up to genesis (Montgomery et al. 2012). Likewise, the National Aeronautics and Space Administration (NASA) Genesis and Rapid Intensification Processes (GRIP) experiment carried out three pregenesis flight missions, followed by two missions during rapid intensification (Braun et al. 2013). These dropsondes were launched at altitudes between 150 and 200 hPa, and span a region that covers the inner 7.5° (~800 km) of the storm center, as indicated by the dropsonde locations in Fig. 1. As in PZ14, only MADIS and PREDICT observations are assimilated in this study. The GRIP observations are used as an independent source of verification for the analyses, but sounding measurements from all three datasets are used for verifying forecasts.

b. EnKF

The EnKF data assimilation system used in this study was developed by Meng and Zhang (2008a, b), and has been applied since 2008 for the real-time analysis and forecasts of Atlantic tropical cyclones (M. Zhang et al. 2011). It uses an ensemble of model simulations to advance an error distribution for the model state between cycles (Evensen 1994; Houtekamer and Mitchell 1998). Ensemble members are denoted by the vectors $\mathbf{x}_{n,t}^{f}$ for n = 1, 2, 3, ..., N, which hold all prognostic variables at time *t*. Assuming Gaussian errors, the EnKF uses the ensemble-estimated mean $\mathbf{\overline{x}}_{t}^{f}$ and error covariance \mathbf{P}_{t}^{f} to approximate the mean analysis state vector:

$$\overline{\mathbf{x}}_{t}^{a} = \overline{\mathbf{x}}_{t}^{f} + \mathbf{P}_{t}^{f} \mathbf{H}_{t}^{\mathrm{T}} (\mathbf{H}_{t} \mathbf{P}_{t}^{f} \mathbf{H}_{t}^{\mathrm{T}} + \mathbf{R}_{t})^{-1} (\mathbf{y}_{t} - H_{t} \overline{\mathbf{x}}_{t}^{f}).$$
(1)

In Eq. (1), H_t is the observation operator that transforms model variables into observation space, \mathbf{H}_t is a tangent linear version of H_t , and \mathbf{y}_t is an observation state vector with error covariance \mathbf{R}_t . Data can be assimilated serially when observation errors are uncorrelated, in which case the nonlinear H_t is used in place of \mathbf{H}_t in Eq. (1). We use the square root filter described in Whitaker and Hamill (2002) to update the ensemble perturbations around $\overline{\mathbf{x}}_t^a$, and form an ensemble of analysis state vectors $\mathbf{x}_{n,t}^a$ for n = 1, 2, 3, ..., N.

Sampling errors in the ensemble-estimated background error covariance are treated by applying a Gaspari and Cohn (1999) fifth-order correlation function to remove spurious correlations between distant model grid points. For this study, we use a 60-member ensemble with horizontal and vertical localization radii of 900 km and 15 model levels, respectively. We also inflate the analysis covariance after each data assimilation cycle by relaxing 80% of the posterior perturbations back to the prior perturbations, following the "covariance relaxation to the prior" method of Zhang et al. (2004; i.e., only 20% of the posterior perturbations comes from the EnKF update). The inflation compensates for unknown errors that are introduced during the prediction and data assimilation steps of the cycling, which would lead to an underestimation of the true background errors over time.

c. 4DVar

Version 3.4.1 of the WRF data assimilation (WRFDA) package (Barker et al. 2012) is used to perform the 4DVar and E4DVar experiments in this study. The 4DVar method seeks a solution that minimizes the misfit of a control variable to the background state \mathbf{x}_0^b at t = 0(typically taken as a deterministic forecast from the previous cycle) and observations \mathbf{y}_t at times $t = 0, 1, 2, ..., \tau$. For incremental 4DVar, the minimization is carried out with respect to increments $\delta \mathbf{x}_0$ from \mathbf{x}_0^b (Courtier et al. 1994). The function to minimize can be expressed as the sum of background (J_b) and observation (J_o) cost functions:

$$J(\delta \mathbf{x}_{0}) = J_{b}(\delta \mathbf{x}_{0}) + J_{o}(\delta \mathbf{x}_{0}) = \frac{1}{2} \delta \mathbf{x}_{0}^{\mathrm{T}} \mathbf{B}^{-1} \delta \mathbf{x}_{0}$$
$$+ \frac{1}{2} \sum_{t=0}^{\tau} (\mathbf{H}_{t} \mathbf{M}_{t} \delta \mathbf{x}_{0} - \mathbf{d}_{t})^{\mathrm{T}} \mathbf{R}_{t}^{-1} (\mathbf{H}_{t} \mathbf{M}_{t} \delta \mathbf{x}_{0} - \mathbf{d}_{t}),$$
(2)

where **B** is the error covariance matrix for \mathbf{x}_0^b , and \mathbf{M}_t represents an integration of the tangent linear model from time 0 to *t*. Here \mathbf{M}_t is linearized about a nonlinear forecast from the beginning of the assimilation window to time *t*, denoted by $M_t(\mathbf{x}_0^b)$, and uses simplified physical parameterization schemes. The \mathbf{d}_t in Eq. (2) is the innovation at each time along the nonlinear forecast and is given by the difference between observations and $M_t(\mathbf{x}_0^b)$ projected into observation space:

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

Though not shown, the 4DVar system used in this study contains a third term in Eq. (2) that measures the fit to a balanced state using a digital filter; see Eq. (5) of Huang et al. (2009).

We use a 4DVar method that performs the cost function minimization (inner loop) in a lower resolution than the forecast model, while calculating \mathbf{d}_t (outer loop) from model states that are advanced by the full-resolution model (Zhang et al. 2014, manuscript submitted to J. Atmos. Oceanic Technol.). The inner loops in our cycling experiments are performed using a 40.5-km grid spacing, which reduces the computational cost of the minimization by an order of magnitude in test cases. The faster minimization comes from both the lower cost of running the tangent linear and adjoint models and the smaller number of iterations required to reach a userspecified stopping criteria; in this study, we stop the minimization at 1% of the initial gradient norm. The outer loop is run using the same 13.5-km grid spacing that is used by the parent model domain. While multiple outer loops can be beneficial when H_t and M_t are highly nonlinear (Courtier et al. 1994), we apply only one outer loop during this study, owing to the computational cost of each set of inner-loop iterations in 4DVar.

Our 4DVar experiments use the CV5 control variable option in WRFDA, which transforms the prognostic WRF Model variables into streamfunction, unbalanced velocity potential, unbalanced temperature, unbalanced surface pressure, and pseudo-relative humidity. We estimate a domain-specific background error covariance for these control variables using a set of 24- and 12-h forecast differences that were generated over the previous month. Sensitivity experiments were performed using 3DVar to determine whether tuning the length scales and amplitudes for the covariance matrix could make systematic improvements (not shown). The performance benefits of modifying these parameters were found to be too small to warrant any deviations from the default settings.

d. E4DVar

An E4DVar data assimilation system is applied using the same EnKF and 4DVar configurations described above. As illustrated in Fig. 2, the method requires that EnKF and 4DVar run separately with two coupling steps: 1) 4DVar uses $\bar{\mathbf{x}}_0^f$ as the background state (i.e., in place of \mathbf{x}_0^b) and introduces ensemble perturbations as an additional constraint in the cost function, and 2) the hybrid 4DVar solution replaces the posterior mean EnKF analysis (Zhang et al. 2009).

To perform the hybrid variational analysis, $\delta \mathbf{x}_0$ is separated into two terms:

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

where the first and second terms represent the climatological and ensemble components of the analysis



FIG. 2. Schematic of two-way coupled EnKF-4DVar data assimilation system.

increment, respectively [see Eq. (1) of Wang et al. 2008]. In the second term of Eq. (4), each $\mathbf{a}_n \circ \mathbf{x}_{n,0}^{tf}$ is a Schur product between a weighting vector \mathbf{a}_n and the *n*th ensemble perturbation (here $\mathbf{x}_{n,0}^{tf} \equiv \mathbf{x}_{n,0}^{f} - \overline{\mathbf{x}}_{0}^{f}$) in model space. The hybrid variational cost function includes both the climatological and ensemble background information, which are weighted by the scalar coefficients β_c and β_e :

$$J(\delta \mathbf{x}_{0}^{c}, \alpha) = \beta_{c} J_{b}(\delta \mathbf{x}_{0}^{c}) + \beta_{e} J_{e}(\alpha) + J_{o}(\delta \mathbf{x}_{0}^{c}, \alpha)$$

$$= \beta_{c} \frac{1}{2} \delta \mathbf{x}_{0}^{cT} \mathbf{B}^{-1} \delta \mathbf{x}_{0}^{c} + \beta_{e} \frac{1}{2} \alpha^{T} \mathbf{A}^{-1} \alpha$$

$$+ \frac{1}{2} \sum_{t=0}^{\tau} (\mathbf{H}_{t} \mathbf{M}_{t} \delta \mathbf{x}_{0} - \mathbf{d}_{t})^{T} \mathbf{R}_{t}^{-1} (\mathbf{H}_{t} \mathbf{M}_{t} \delta \mathbf{x}_{0} - \mathbf{d}_{t})$$
(5)

In Eq. (5), \mathbf{d}_t is calculated with respect to a model trajectory from $\overline{\mathbf{x}}_0^f$:

$$\mathbf{d}_t = \mathbf{y}_t - H_t[M_t(\mathbf{\bar{x}}_0^f)].$$
(6)

The coefficients β_c and β_e are chosen so that 80% of the increments between each iteration of the minimization come from the ensemble information.¹ The *N* weighting vectors \mathbf{a}_n are concatenated to form control variables α for the ensemble perturbations in Eq. (5) (i.e., $\alpha^T = [\mathbf{a}_1^T, \mathbf{a}_2^T, \dots, \mathbf{a}_N^T]$). These control variables are constrained by \mathbf{A} , which is a block-diagonal matrix that contains *N* identical correlation matrices along the diagonal. As described in Lorenc (2003b), the correlation matrices precondition α for covariance localization, thus increasing the

degrees of freedom during the minimization. See Wang et al. (2007) for additional details regarding the formulation of the hybrid cost function and its theoretical equivalence to other proposed methods.

e. Experiment design

The EnKF, 4DVar, and E4DVar data assimilation systems are cycled every 6 h from 0600 UTC 8 September to 0000 UTC 18 September. The first cycle starts from 12-h forecasts that are initialized from the 1° latitude– longitude Global Data Assimilation System (GDAS) analysis (NCEP/NWS/NOAA/DOC 2000, updated daily) on 1800 UTC 7 September. GDAS data also provide lateral boundary conditions and sea surface temperatures for the parent domain during each cycle. For the ensemble methods, perturbations are sampled from the climatological covariance matrix and added to a GDAS analysis to generate the initial set of model states; lateral boundaries are perturbed in a similar manner during each cycle.

All nonradiance observations from MADIS are assimilated at each cycle, along with PREDICT dropsondes between 1200 UTC 10 September and 1800 UTC 14 September. The MADIS observations include surface data, routine soundings, profiler data, and cloud-tracked winds from Geostationary Operational Environmental Satellites (GOES). In the EnKF experiment, observations that are collected ± 3 h from the analysis time are assumed to be valid at the middle of the observation window. 4DVar and E4DVar, on the other hand, use 1-h bins of observations for fitting a trajectory of data in the window. As in PZ14, the beginning of the cycling period corresponds to the first identification of the pre-Karl disturbance by the National Hurricane Center (NHC), and the last cycle occurs as Karl decays over the Mexican coast. Data assimilation is performed on the 13.5-km

¹These coefficients are not tuned in the current study—the choice of weights comes from sensitivity tests performed in Zhang and Zhang (2012) and Zhang et al. (2013).

domain (Fig. 1) in all experiments. Deterministic forecasts, however, use a nested 4.5-km domain that is interpolated from the parent domain at initialization.

3. Analysis results

a. Storm evolution

This section compares the evolution of the tropical weather system in each of the three cycling data assimilation experiments between 1200 UTC 10 September and 0000 UTC 18 September. The selected period spans all PREDICT and GRIP flight missions, as well as Karl's rapid intensification and decay. We use the region within 3° of the storm center to examine changes in the kinematic and thermodynamic structure of the tropical system. As in PZ14 the pregenesis storm center is estimated by finding the locations that maximize the azimuthal mean winds within 3° of the 950- and 700-hPa centers, and averaging the two estimates. For cases where maximum 10-m winds exceed tropical storm strength (18 m s^{-1}) , a Barnes analysis (Barnes 1964) is performed using the 10-m, 850-hPa, and 700-hPa vorticity fields to find a wind-based storm center.

Figure 3 summarizes the analysis results during this period using mean 950- and 500-hPa vertical vorticity (ζ), 950–500-hPa vortex displacement (tilt) and shear, and 950–500-hPa column relative humidity (CRH; or the ratio of vertically integrated water vapor to vertically integrated saturation water vapor). No systematic differences are found in the low- and midlevel mean ζ (Figs. 3a,b) prior to Karl's genesis, suggesting that the three methods produce qualitatively similar systemscale circulation strengths at these times. Nevertheless, the E4DVar analyses contain the smallest tilt magnitudes and highest values of CRH in the two days before genesis. Both of these factors are suggested to be important precursors to Karl's development (Davis and Ahijevych 2012; PZ14).

To compare the analyses between 0600 UTC 13 September and 12 UTC 14 September in more detail, the filtered (L > 150 km) 950- and 500-hPa ζ and 950– 500-hPa CRH in the E4DVar analyses are subtracted from the filtered EnKF and 4DVar analyses (Fig. 4). The ζ differences in Fig. 4 show that the EnKF and E4DVar experiments produce similar vortex positions, whereas the 4DVar experiment yields a more northerly track of the disturbance; this conclusion follows from the large dipole in the 4DVar-E4DVar ζ differences. Despite the similar estimate of storm position between the two ensemble methods, large differences in the magnitude and horizontal structure of ζ emerge by the 0600 and 1200 UTC 14 September cycles. In particular, the E4DVar case



FIG. 3. Relative vorticity at (a) 950- and (b) 500-hPa, and the 950–500-hPa (c) tilt, (d) vertical shear, and (e) relative humidity plotted every 6 h for EnKF (blue), 4DVar (green), and E4DVar (red) analyses between 1200 UTC 10 Sep and 0000 UTC 18 Sep. The vertical gray dashed line indicates the 0600 UTC 13 Sep time and the black one indicates the true genesis time on 1800 UTC 14 Sep.

produces a more compact vortex in the hours preceding genesis, despite using lower-resolution analysis increments than EnKF.

As indicated in Fig. 3e, the E4DVar data assimilation leads to the highest values of saturation in the vicinity of the developing tropical system. The shaded 950–500-hPa CRH differences in Fig. 4 reveal the spatial extent to which this result is true. At 0600 UTC 13 September, both the EnKF and 4DVar analyses have a dryer lower-to-middle



FIG. 4. Differences between (left) EnKF and E4DVar and (right) 4DVar and E4DVar analyses. The plotted variables are the 950–500-hPa relative humidity (shading), 950-hPa relative vorticity (black contours), and 500-hPa relative vorticity (red contours). The relative vorticity values are filtered to remove wavelengths < 150 km; positive and negative differences are indicated by solid and dashed lines, respectively: (from top to bottom) from 0600 UTC 13 Sep to 1200 UTC 14 Sep.

troposphere than E4DVar for a region that spans the inner 2° of the vortex. As low- and midlevel ζ increases in the proceeding cycles, the EnKF experiment converges toward similar saturation values near the disturbance center. Nevertheless, the EnKF maintains a less saturated environment near the disturbance throughout all cycles. The comparison of CRH near the storm center in the E4DVar and 4DVar cases is complicated by the displacement of the system-scale vorticity centers in the analyses, which produces a mix of positive and negative differences in the center of the verification region. Despite this ambiguity, CRH values in the outer region of the disturbance are clearly larger in the E4DVar analyses.

b. Verification with independent observations

Dropsondes that were collected during the GRIP field campaign are used in this section to verify the EnKF, 4DVar, and E4DVar analyses. These observations were taken directly before genesis (1800 UTC 12 September-0000 UTC 14 September) and during a period of rapid intensification (1800 UTC 16 September-1800 UTC 17 September, Fig. 1), but are not assimilated during the experiments. All measurements that were collected within 3h of the analysis times are considered in this verification; in total, 105 dropsondes are used to perform the calculations. Similar to Davis and Ahijevych (2012), we use storm motion vectors to correct for the horizontal displacement of dropsondes to compare values at synoptic times. In doing so, we use NHC best track positions before and after the verification time, to estimate the translation speeds of dropsondes.

Figure 5 shows vertical profiles of analysis root-meansquare deviations (RMSD), biases, and 95% confidence intervals for the two statistics. While all three experiments perform similarly for the meridional wind (v)verification, systematic difference are found in the zonal wind (u) analyses at these times (Figs. 5a,b). The E4DVar method produces the smallest u-wind RMSDs and biases over most of the troposphere, followed in order by the 4DVar and EnKF analyses. The *u*-wind biases reflect differences in the mean zonal winds, which likely influence the accuracy of track predictions (to be discussed in section 4). Temperature (T) RMSDs in the lower troposphere are comparable to the 1-K errors that are assigned to sounding temperature observations during the data assimilation (Fig. 5c); therefore, we cannot identify any distinctions between the three analyses for this variable.

From the previous section, Figs. 3e and 4 show that E4DVar analyses contain more saturated air near the tropical weather system in the days before and after genesis. Likewise, the q verification in Fig. 5d shows that E4DVar analyses have the smallest RMSDs and biases



FIG. 5. Vertical profiles of analysis bias (dashed) and analysis RMSDs (solid) to GRIP observations for EnKF (blue), 4DVar (green), and E4DVar (red): (a) u, (b) v, (c) T, and (d) q. The 95% confidence interval for each metric is indicated by the region covered by thin dotted lines.

at every altitude. This result confirms that the larger CRH values found in the E4DVar analyses are more representative of Karl's true moisture field at these times.

4. Forecast results

a. Verification with observations and GDAS analyses

This section investigates the short-range (0–72 h) deterministic forecast performance of the three methods with respect to two data sources. We use sounding observations from the PREDICT and GRIP field campaigns as well as conventional radiosonde data within 800 km of the NHC best track center to estimate forecast errors near the tropical weather system. Results from this verification method are plotted in the top row of Fig. 6 and left column of Fig. 7. Likewise, analyses from the GDAS system are used in a separate set of calculations to verify forecast results in the near-storm environment. This verification considers forecast data within 2500 km of the best track data, and is shown in the bottom row of Fig. 6 and right column of Fig. 7. GDAS analyses contain more error sources than the dropsonde measurements, which can be problematic when using these calculations to evaluate forecast errors near the initialization time. These errors, however, are a much smaller factor when verifying forecasts with lead times greater than 24 h.

We summarize the forecast RMSDs to observations and GDAS analyses using deterministic forecasts that are initialized between 0600 UTC 8 September and 0000 UTC 15 September during the cycling period. The RMSDs are averaged vertically from the surface to 200 hPa then temporally over each cycle to estimate a time series for u, v, T, and q forecast errors (Fig. 6). Figure 7 shows vertical profiles of the RMSDs at 24- and 72-h forecast lead times to identify the altitudes at which the largest forecast errors occur.

The time series verification with observations shows similar short-range forecast performance between the EnKF and E4DVar experiments; however, the E4DVar case tends to predict T and q more accurately at lead times greater than 24 h (first row of Fig. 6). For thermodynamic variables, Figs. 7e,g indicate that the E4DVar system provides the largest gains in 72-h forecast performance in the upper troposphere for T and in the lower and midtroposphere for q. The 4DVar experiment performs slightly worse than the two ensemble systems for this verification, with the largest errors occurring below 500 hPa for u, v, and q (first column of Fig. 7). Storm position errors in the 4DVar analyses contribute to this result, which will be discussed in more detail in section 4b.

We use the verification with GDAS analyses to estimate synoptic-scale forecast errors near the tropical storm. As in Figs. 6c,d, the ensemble methods provide smaller RMSDs for thermodynamic variables than the 4DVar experiment at all lead times (Figs. 6g,h). Errors in the 4DVar forecasts, relative to the EnKF and E4DVar cases, are most substantial in the midlevel qfield (Fig. 7g). While the three experiments produce comparable 24-h forecast RMSDs for wind variables, the ensemble methods yield the lowest errors by 72 h for uand v (Figs. 6e, f and Figs. 7b, d). The GDAS verification shows that forecasts from the EnKF experiment produce the closest fit to analysis data in the first 6-24 h of integration. The differences between these two cases, however, may be within the margin of error for the GDAS analyses. Results from the E4DVar experiment are comparable to EnKF during this forecast range, but contain smaller RMSDs by 48h. By 72h, the E4DVar experiment produces the lowest mean forecast RMSDs with respect to the verifying data, with the exception of midlevel T (second column of Fig. 7).



FIG. 6. Deterministic forecast errors compared for EnKF (blue), 4DVar (green), and E4DVar (red). Mean RMSDs are calculated with respect to (top) conventional soundings and field observations within 800 km of the NHC best track center and (bottom) GDAS data within 2500 km of the NHC best track center. Values are averaged between the surface and 200 hPa.

One problematic aspect of the 4DVar simulations is the persistently large 0–12-h forecast errors that occur in both sets of verifications (Fig. 6). The decrease in T and q RMSDs over the first 6 h suggest that imbalances introduced during data assimilation may contribute to the relatively large short-term forecast errors that are found in this experiment. This result is verified in section 6, and is assumed to have a nonnegligible contribution to forecast errors at greater lead times.

b. Track and intensity

This section compares deterministic storm track and intensity forecasts that are generated from the three cycling data assimilation experiments. These forecasts start from the respective analysis times and end shortly after the second landfall of Karl on 0000 UTC 18 September. Figure 8 shows verifications of track and intensity forecasts initialized between 0000 UTC 12 September and 00 UTC 15 September to compare the performance of each method in predicting the genesis event during the three days leading up to Karl's formation. The 10-m wind and position observations of the tropical cyclone are provided from the NHC best track dataset, while the pregenesis storm positions come from the wave-tracking product described in Wang et al. (2012). For the forecast cycles shown in this figure, all three methods capture the correct westward propagation of the tropical system, as well as the intensification (genesis) of Karl before its first landfall over the Yucatan Peninsula. Nevertheless, the data assimilation cycle and lead time at which an



FIG. 7. Vertical profiles of deterministic forecast errors compared for EnKF (blue), 4DVar (green), and E4DVar (red) at 6-h (dashed) and 72-h (solid) lead times. Mean RMSDs are calculated with respect to (left) sounding observations within 800 km of the NHC best track center and (right) GDAS data within 2500 km of the NHC best track center: (a),(b) u; (c),(d) v; (e),(f) T; and (g),(h) q.

accurate genesis prediction is made varies greatly between the three experiments.

Forecasts from the EnKF analyses begin to capture the prelandfall genesis accurately on the 1200 UTC 13 September cycle, though the time at which genesis occurs in these simulations is typically 18-24 h behind the true genesis time of Karl (Figs. 8a,d). Forecasts from the 4DVar analyses accurately produce the prelandfall genesis event as early as the 1200 UTC 12 September cycle, but have a similar time lag in the genesis event (Figs. 8b,e). The error in genesis time for the 4DVar forecasts becomes larger as the cycles progress, indicating a slow mesoscale evolution of the simulated vortex. Forecasts from the E4DVar cycles also yield correct genesis predictions starting from 1200 UTC 12 September; these forecasts, however, simulate the timing for Karl's intensification much more accurately than the EnKF and 4DVar experiments. A comparison of intensity forecasts between 14 and 17 September in Figs. 8d-f demonstrates the extent to which the hybrid method outperforms the benchmark assimilation methods for Karl's genesis and rapid intensification. In general, the superior intensity forecasts from the E4DVar analyses come from the improved timing of genesis in these simulations, which leads to a more accurate prediction for the rapid intensification event that follows. The E4DVar analyses also provide the most accurate estimates of storm position, as verified with track data (Figs. 8a-c). This factor, combined with a better representation of the surrounding environment (Figs. 6e-h), leads to the more accurate track forecasts in this experiment.

One advantage of the ensemble data assimilation methods over 4DVar is their ability to produce probabilistic forecasts. For example, consider the change in deterministic forecast accuracy between 0600 UTC 13 September and 1200 UTC 13 September cycles in the EnKF experiment. In PZ14, ensemble forecasts from these two cycles revealed a large increase in the forecast probability for genesis, which was not obtained in a control experiment that omitted field observations during data assimilation (cf. Figs. 10 and 11 of PZ14). The PREDICT observations at this cycle were found to increase the low- to midlevel circulation in the ensemble analyses, which facilitated Karl's development in the simulations. Ensemble forecasts of storm track and intensity are plotted in Fig. 9 to compare probabilistic forecasts between the EnKF and E4DVar experiments at 0600 UTC 13 September. As in PZ14, we define developing members (red lines in Fig. 9) as forecasts that produce maximum 10-m winds in excess of $18 \,\mathrm{m \, s^{-1}}$ (tropical-storm strength) for three or more consecutive 3-h time stamps between 1800 UTC 14 September and 0000 UTC 16 September. The comparison shows a large



FIG. 8. Deterministic (top) track and (bottom) intensity forecasts from the (left) EnKF, (middle) 4DVar, and (right) E4DVar experiments. Forecasts are colored according to initialization time and NHC best track data are plotted in black.

genesis probability in the E4DVar forecast (i.e., 53/60) at a cycle in which only 22 of the 60 EnKF members develop a tropical storm before landfall. This result occurs despite the fact that the ensemble mean analyses in the two experiments contain comparable average 950-hPa ζ in the cycles leading up to this time (Fig. 3a). Though not shown, the E4DVar and EnKF ensemble analyses also share similar magnitudes of low- to midlevel mean ζ . While the EnKF ensemble analyses require an additional assimilation cycle to intensify the low-level vortex in the model, a large majority of the E4DVar members capture the intensification well before the 1200 UTC 13 September observations. A combination of factors lead to a faster development of the low-level vortex in the E4DVar forecasts, including a less significant displacement between the 950- and 500-hPa vortex centers (Fig. 3c), and a more saturated lower and midtroposphere near the disturbance (Fig. 3e).

c. Data assimilation at the time of genesis

The slow development of Karl in the 4DVar and EnKF simulations is a persistent result in our experiments. A PREDICT flight mission on 14 September collected detailed observations of the storm near the time of genesis, thus providing a fitting case study for examining this issue. Using forecasts from the 1800 UTC 14 September cycle, Fig. 10 compares the 18-h evolution of 950-hPa positive ζ and 950–500-hPa CRH in a 450 × 450 km² box around the storm center. Maximum 10-m wind values (V_{max}) for each forecast time are indicated in the bottom-left of each panel, and filtered (L > 150 km) storm-relative streamlines are plotted for reference. The storm in each analysis has an intensity greater than or equal to a tropical

depression, but only the E4DVar case intensifies the lowlevel vortex immediately after initialization. Despite having the lowest wind speeds, the E4DVar analysis contains a vortex that is more representative of a tropical cyclone than the EnKF and 4DVar cases at this time (e.g., the vortex maintains an annular ring of positive ζ around a relatively dry center). The E4DVar vortex is also embedded in a region of near saturated air, which likely aids in the development of new convective cells near the developing tropical cyclone. By 1200 UTC 15 September, the E4DVar simulation yields a V_{max} of 30.2 ms^{-1} , which is 9 ms^{-1} stronger than the other two cases and more representative of the true hurricane intensity at this time (the observed $V_{\text{max}} = 28.1 \text{ ms}^{-1}$ from NHC best track data).

5. Analysis increments

This section presents a set of experiments that demonstrate how each data assimilation system produces increments near the pregenesis disturbance. We generate 4DVar and E4DVar analyses using the same observations and background field as the cycling EnKF experiment at 1200 UTC 13 September. Observations at this cycle include dropsondes from a PREDICT flight mission that occurred 30 h before Karl's intensification into a tropical storm. The additional field data at this cycle were found in PZ14 to increase the forecast accuracy significantly in the EnKF experiment, which motivates our choice of cycle time. We use the resulting analyses to examine how each method produces increments, given identical information regarding the atmosphere and model state at this cycle (Fig. 11). In the four-dimensional



FIG. 9. Ensemble (top) track and (bottom) intensity forecasts for the (left) EnKF and (right) E4DVar experiments. Ensemble members are initialized at 0600 UTC 13 Sep and run to 0000 UTC 18 Sep. Developing members are red, nondeveloping members are blue, and the best track data are black.

data assimilation cases, the increments are taken to be the difference between the analysis and background state at the middle of the time window.

Clear differences exist between the EnKF and 4DVar increments near the pregenesis disturbance, despite using the same observations and background mean state. The EnKF produces large vortex-scale updates to the wind field at all levels to move the circulation center westward (Figs. 11a,d). The 4DVar increments also act to move the center of the circulation, but the resulting corrections to ζ cover a larger area and have a smaller amplitude (Fig. 11e). To some extent, the E4DVar increments for kinematic variables resemble a combination of the EnKF and 4DVar increments. For example, the 950- and 500-hPa filtered ζ and 850-hPa ζ increments in Figs. 11i,l have characteristics of both the EnKF and 4DVar increments. The ensemble-estimated covariance matrix contains higher variances and shorter correlation length scales near the center of the tropical disturbance, which differ substantially from the month-long climatological errors used by the 4DVar system. The correlation length scales in the climatological covariance are most representative of synoptic-scale errors. As a result, the mesoscale circulation in the 4DVar experiment contains a northward position bias that goes uncorrected during the earlier cycles shown in Fig. 8c.

Increments to thermodynamic variables also reflect major differences in the background error covariance and control variables used by the three data assimilation methods. While the EnKF increments act to decrease 950-hPa θ_v over most of the subdomain plotted in Fig. 11b, the 4DVar and E4DVar increments change sign from negative to positive in the outer $2^{\circ}-3^{\circ}$ of the vortex center (Figs. 11f,j). Direct measurements of the thermodynamic field are not available in the outer region of the circulation on 13 September, owing to the lack of dropsonde data away from the vortex center (Fig. 1). The increments in this region depend largely on cross correlations between θ_{v} and distant values of wind, temperature, and moisture variables near observations. These correlations will vary between the EnKF and E4DVar analyses, owing to the implicit use of time-dependent correlations in the E4DVar assimilation window. Furthermore, the innovations are not equal between these two experiments: observations are assumed to be valid at the same time in EnKF, but they are binned into 1-h intervals and compared to a model forecast from the background field in E4DVar.

Increments to vertically integrated quantities, such as CRH, reveal equally complex differences between the three data assimilation systems. For the EnKF experiment, Fig. 11c shows an increase in CRH in the northern and southern portions of the storm, and a decrease on the eastern side, which results in a largely asymmetric distribution of near-saturated air in the analysis. The corresponding 4DVar increments are smaller and less organized around the low-level cyclone (Figs. 11g,k). E4DVar increments to CRH are similar to the EnKF



FIG. 10. The 950-hPa relative vorticity (shaded every $4 \times 10^{-4} \text{s}^{-1}$ from $1 \times 10^{-4} \text{s}^{-1}$), 950–500-hPa CRH values > 90% (blue contours), and filtered (L > 150 km) 950-hPa storm-relative streamlines (dashed black contours) plotted every 6 h from (left to right) 1800 UTC 14 Sep to 1200 UTC 15 Sep from forecasts initialized from (top) EnKF, (middle) 4DVar, and (bottom) E4DVar analyses on 1800 UTC 14 Sep.

results near the storm center, but contain less substantial negative increments in the outer regions of the vortex. While the 4DVar scheme uses a linearized model and its adjoint to extract flow-dependent information from the observations and model dynamics during the assimilation window, the resulting increments at t = 0 are constrained by the climatological background error covariance. Correlations between water vapor and the model state variables are not included in the control variable option that is used in this study, so θ_v and CRH increments must come from updates to temperature and pressure, or adjustments to specific humidity using the linearized model operators.

Recall that the 4DVar and E4DVar cases use a 40.5-km grid spacing inner loop for performing the analyses. In a second E4DVar experiment, we examine the effects of this approximation by repeating the E4DVar analysis using 13.5-km grid spacing for all stages of the data

assimilation (fourth row of Fig. 11). The E4DVar increments that were generated using 13.5- and 40.5-km inner loops contain subtle differences near the storm; in particular, the CRH increments have opposite sign near the vortex center. Nevertheless, these increments appear to capture qualitatively similar synoptic and meso- α -scale features in the domain.

Similar to Thepáut et al. (1996), we examine the structure functions that produce the four-dimensional data assimilation increments by performing three sets of single-observation experiments for 4DVar and E4DVar and comparing the resulting analysis increments. These experiments use an 850-hPa wind observation at the beginning (t = 0), middle (t = 3), and end (t = 6) of the assimilation window, along with the same background state that is used to produce Fig. 11. Each observation is forced to yield a 1 m s⁻¹ innovation in wind speed, but at different times along the background trajectory.

Figure 12 shows 850-hPa θ_v and ζ analysis increments, along with storm-relative streamlines from the background. Unlike Fig. 11, the increments are taken as the difference between the analysis and background state vectors at the beginning, rather than the middle, of the time window. Increments from an observation at t = 0 provide detailed information regarding the structure of the background error covariance used by the two methods, whereas the cases that assimilate observations at t = 3 and 6 h reveal the effects of the tangent linear and adjoint models.

The impact of the linear error propagation is more apparent in the 4DVar case, owing to the use of a covariance matrix at t = 0 that is independent of the current forecast. For example, consider the 4DVar analysis increments to θ_v in the first column of Fig. 12. The increments from an observation at t = 0 (Fig. 12a) depend entirely on the initial static covariance matrix. These values fail to exceed 0.006 K, which is smaller than the minimum color scale in the figure. As the observation window length increases, the θ_v increments become larger and increasingly more asymmetric (first column of Fig. 12). On the other hand, the E4DVar structure functions change more gradually as the observation is moved to later times in the window. The hybrid control variables use knowledge of past information to describe the background errors at t = 0, and therefore do not require a long time window to develop flow-dependent features from the model dynamics. Differences in how 4DVar and E4DVar use the linear model operators to perform this function are better illustrated by the ζ increments in the second two columns of Fig. 12. Here, the shape of the 4DVar increments evolves from the largely isotropic form that is expected from a 3DVar data assimilation system to the southwest-northeast-oriented dipole that is found in the E4DVar case. Likewise, the E4DVar structure functions contain the same dipole orientation for all three observations.

6. Balance

The second derivative of surface pressure $\partial^2 p_s / \partial t^2$ is calculated at each time step during the first 12 h of model integration to estimate gravity wave activity following model initialization. The 12-h forecast lead time that we use for this comparison corresponds to the time integration that is required to run a forecast from the beginning of the observation time window of the current cycle to the end of the window in the proceeding cycle. Similar to Houtekamer and Mitchell (2005), we use the domain root-mean-square (RMS) $\partial^2 p_s / \partial t^2$ to quantify imbalance in the deterministic forecasts generated during the data assimilation cycles. These calculations are averaged over all cycles for the three methods and compared in Fig. 13a.

The imbalance metric shows a large mass adjustment over the first hour of integration, followed by a decrease in gravity wave activity as the model approaches a more balanced solution. RMS $\partial^2 p_s / \partial t^2$ is largest for the EnKF and E4DVar cases over the first several time steps. This result likely comes from the use of an ensemble mean background field in these two cases. Being a deterministic method, the 4DVar system uses a background field that is a known solution to the model equations, whereas the same does not have to be true for an ensemble mean. Nevertheless, the RMS $\partial^2 p_s / \partial t^2$ in EnKF and E4DVar forecasts converge quickly to the 4DVar case by 1 h, as most of the adjustments occur at small scales. Between 1 and 4h of integration, the pressure adjustment in the EnKF simulations exhibits a large amount of variability between cycles (Fig. 13b), owing to the presence of higher RMS $\partial^2 p_s / \partial t^2$ values in cycles that contain a fully developed tropical cyclone (not shown). The imbalances at these times are largely alleviated in the E4DVar scheme by the four-dimensional assimilation of observations, which place a stronger dynamical constraint on analyses.

Unlike the two ensemble experiments, forecasts from 4DVar analyses tend to undergo a long adjustment period after the initial rapid decrease in RMS $\partial^2 p_s / \partial t^2$. Several properties of the 4DVar scheme tested in this study may introduce imbalances in the analyses. For instance, applying a low-resolution domain for the cost function minimization requires a down sampling of the high-resolution model grid, which may cause highfrequency signals to project onto larger scales when estimating the low-resolution background state. The innovation vectors, as estimated from the high-resolution trajectory, will also contain scales that are not resolved by the low-resolution tangent linear and adjoint model operators. Furthermore, analysis increments are interpolated from the low- to high-resolution grid and added to the original background state after the inner loop, which may introduce physically inconsistent meteorological fields near large gradients in the domain. A more significant factor is the use of a month-long climatological background error estimation in 4DVar, which is expected to be suboptimal for the assimilation of mesoscale observations.

The process of generating a hybrid analysis with multiincremental 4DVar requires the same downsampling and interpolation steps as the standard 4DVar. Despite these sources of imbalance, RMS $\partial^2 p_s / \partial t^2$ decreases the fastest in E4DVar forecasts (Fig. 13a). The use of an ensemble mean background field in the coupled method may reduce the downsampling error described above; however,



FIG. 11. Analysis increments from (a)–(d) EnKF, (e)–(h) 4DVar, (i)–(l) E4DVar with a 40.5-km inner loop, and (m)–(p) E4DVar with a 13.5-km inner loop; all three analyses use the same observations and background field at 1200 UTC 13 Sep. (left to right) Increments of filtered (L > 150 km) relative vorticity, 950-hPa θ_v , CRH, and 850-hPa (unfiltered) relative vorticity. System-relative streamlines from the 950- and 850-hPa altitudes are also overlaid in the second and last two columns, respectively.

we suspect that the largest benefits come from the use of ensemble perturbations in the cost function.

7. Conclusions

This study compares WRF-based EnKF and 4DVar data assimilation methods with a two-way coupled E4DVar scheme for a mesoscale weather application. The three methods assimilate routinely collected observations and field measurements before, during, and after the genesis of Hurricane Karl, an Atlantic storm that was targeted by the 2010 NSF-PREDICT and NASA-GRIP field campaigns. Analyses and short-term (0–72 h) deterministic forecasts are verified using observations in the vicinity of the tropical weather system and GDAS analysis data in the near-storm environment. On average, deterministic forecasts from the E4DVar mean analyses are found to have smaller errors than the benchmark



FIG. 12. Single-observation analysis increments from (first-third columns) 4DVar and (second-fourth columns) E4DVar. (first-second columns) The 850-hPa θ_v increments and (third-fourth columns) 850-hPa relative vorticity increments for observations that are located at (a)-(d) t = 0, (e)-(h) t = 3, and (i)-(l) t = 6 h in the time window. System-relative streamlines at 850 hPa are also overlaid for reference.

EnKF and 4DVar methods at lead times greater than 48 h for all verification methods and variables. Likewise, the EnKF analyses produce smaller analysis and forecast errors than the 4DVar system in the vicinity of the tropical disturbance. While 4DVar has several theoretical and practical advantages over EnKF, its reliance on a static background error covariance is found to be a limitation for producing analyses at the mesoscale.

The more accurate representation of wind and moisture in the E4DVar analyses leads to large improvements in Karl's track and intensity forecasts over EnKF and 4DVar, as verified with NHC best track data. In the data assimilation cycles preceding genesis, forecasts that are initialized from the EnKF and 4DVar analyses typically yield a 12–24-h delay in Karl's initial intensification. E4DVar initial conditions at the same cycles produce a pregenesis disturbance that contains a higher concentration of positive ζ near the large-scale vortex center, as well as more saturated air between 950 and 500 hPa. These factors lead to a shorter spinup of the tropical storm in the model forecasts and a large improvement in the timing of the genesis event.

In addition to reducing forecast errors over the benchmark cases, the E4DVar method produces fewer initial-condition imbalances than the stand-alone EnKF and 4DVar systems. The improved balance comes from the combined effects of using flow-dependent background error covariance and the four-dimensional assimilation of asynchronous observations. The inclusion of ensemble perturbations in the cost function via extended control variables also improves the conditioning of the 4DVar minimization problem. Though not shown, E4DVar required \sim 30% fewer iterations to reduce the gradient to 1% of the original value in our experiments; similar improvements in conditioning are also noted in Wang et al. (2013) for an ensemble-3DVar hybrid system.

Analysis increments from the four-dimensional data assimilation systems reveal complex relationships between the ensemble perturbations and the sensitivity



FIG. 13. Domain RMS $\partial^2 p_s / \partial t^2$ for 12-h deterministic forecasts initialized from EnKF (blue), 4DVar (green), and E4DVar (red) analyses. The (a) mean and (b) standard deviation of RMS $\partial^2 p_s / \partial t^2$ are plotted for values calculated over the data assimilation cycles.

gradients provided from the adjoint model. In some instances, the E4DVar analyses may have little in common with analyses produced by the component 4DVar or EnKF methods (e.g., for variables that depend on moisture). Single-observation experiments also show important details regarding how the 4DVar and E4DVar systems treat observations at different times in the assimilation window. Observations that are located at the beginning of the 4DVar window will be assimilated using less flow-dependent information than observations near the end of the window. This issue is alleviated in E4DVar, owing to the use of ensemble information at the beginning of the assimilation window. Given that E4DVar does not require a long time window to evolve flowdependent analysis increments, this method may also have large benefits for highly nonlinear applications that limit the assimilation window length. Evidence is provided by the slower evolution of structure functions over the 6-h time window in the E4DVar experiment than in the 4DVar experiment (Fig. 12). Nevertheless, additional research is required to understand the full potential of using ensemble information in 4DVar, and the resulting effects on minimization and optimal window length.

Acknowledgments. This work was supported in part by the NOAA Hurricane Forecast Improvement Project (HFIP), Office of Naval Research Grant N000140910526 and the National Science Foundation Grant ATM-0840651. The computing was performed at the Texas Advanced Computing Center. We are thankful for the comments provided by three anonymous reviewers for improving the manuscript.

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.
- 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., 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.
- 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.
- Davis, C. A., and D. A. Ahijevych, 2012: Mesoscale structural evolution of three tropical weather systems observed during PREDICT. J. Atmos. Sci., 69, 1284–1305, doi:10.1175/ JAS-D-11-0225.1.
- Dunkerton, T. J., M. T. Montgomery, and Z. Wang, 2009: Tropical cyclogenesis in a tropical wave critical layer: Easterly waves. *Atmos. Chem. Phys.*, 9, 5587–5646, doi:10.5194/acp-9-5587-2009.
- Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J. Geophys. Res., 99, 10143–10162, doi:10.1029/94JC00572.
- Fang, J., and F. Zhang, 2011: Evolution of multiscale vortices in the development of Hurricane Dolly (2008). J. Atmos. Sci., 68, 103–122, doi:10.1175/2010JAS3522.1.
- 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.
- Halverson, J., and Coauthors, 2007: NASA's Tropical Cloud Systems and Processes experiment: Investigating tropical cyclogenesis and hurricane intensity change. *Bull. Amer. Meteor. Soc.*, 88, 867–882, doi:10.1175/BAMS-88-6-867.
- 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.

- 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.
- —, and —, 2005: Ensemble Kalman filtering. Quart. J. Roy. Meteor. Soc., 131, 3269–3289, doi:10.1256/qj.05.135.
- Huang, X.-Y., 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.
- Kalnay, E., H. Li, T. Miyoshi, S.-C. Yang, and J. Ballabrera-Poy, 2007: 4D-Var or ensemble Kalman filter? *Tellus*, **59A**, 758– 773, doi:10.1111/j.1600-0870.2007.00261.x.
- 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.
- 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.
- Meng, Z., and F. Zhang, 2008a: Test of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part III: Comparison with 3DVAR in a real-data case study. *Mon. Wea. Rev.*, **136**, 522–540, doi:10.1175/2007MWR2106.1.
- —, and —, 2008b: Test of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part VI: Performance over a warm season month of June 2003. *Mon. Wea. Rev.*, **136**, 3671–3682, doi:10.1175/2008MWR2270.1.
- Montgomery, M. T., M. E. Nicholls, T. A. Cram, and M. E. Saunders, 2006: A vortical hot tower pathway to tropical cyclogenesis. *J. Atmos. Sci.*, **63**, 355–386, doi:10.1175/JAS3604.1.
- —, 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.
- NCEP/NWS/NOAA/DOC, 2000: NCEP FNL Operational Model Global Tropospheric Analyses, continuing from July 1999. Research Data Archive, National Center for Atmospheric Research, Computational and Information Systems Laboratory, Boulder, CO. [Available online at http://dx.doi.org/ 10.5065/D6M043C6].
- Poterjoy, J., and F. Zhang, 2014: Predictability and genesis of Hurricane Karl (2010) examined through the EnKF assimilation of field observations collected during PREDICT. J. Atmos. Sci., 71, 1260–1275, doi:10.1175/JAS-D-13-0291.1.
- Reasor, P. D., M. T. Montgomery, and L. F. Bosart, 2005: Mesoscale observations of the genesis of Hurricane Dolly (1996). *J. Atmos. Sci.*, 62, 3151–3171, doi:10.1175/JAS3540.1.
- 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. [Available online at http://www.mmm. ucar.edu/wrf/users/docs/arw_v3_bw.pdf.]

- Stewart, S. R., 2010: Tropical cyclone report Hurricane Karl 14–18 September 2010. Tech. Rep. AL1312010, National Hurricane Center, 17 pp. [Available online at http://www.nhc.noaa.gov/ pdf/TCR-AL132010_Karl.pdf.]
- Thepáut, J.-N., P. Courtier, G. Belaud, and G. Lamaître, 1996: Dynamical structure functions in a four-dimensional variational assimilation: A case study. *Quart. J. Roy. Meteor. Soc.*, **122**, 535– 561, doi:10.1002/qj.49712253012.
- 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., C. Snyder, and T. M. Hamill, 2007: On the theoretical equivalence of differently proposed ensemble-3DVAR hybrid analysis schemes. *Mon. Wea. Rev.*, 135, 222–227, doi:10.1175/ MWR3282.1.
- —, D. M. Barker, C. Snyder, and T. M. Hamill, 2008: 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. F. Parrish, and J. S. W. D. Kleist, 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., 2012: Thermodynamic aspects of tropical cyclone formation. J. Atmos. Sci., 69, 2433–2451, doi:10.1175/JAS-D-11-0298.1.
- —, 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.
- Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, 130, 1913–1924, doi:10.1175/1520-0493(2002)130<1913: EDAWPO>2.0.CO;2.
- Zhang, F., 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.
- —, Y. Weng, J. F. Gamache, and F. D. Marks, 2011: Performance of convection-permitting hurricane initialization and prediction during 2008-2010 with ensemble data assimilation of inner-core airborne Doppler radar observations. *Geophys. Res. Lett.*, **38**, L15810, doi:10.1029/2011GL048469.
- —, M. Zhang, and J. Poterjoy, 2013: E3DVar: Coupling an ensemble Kalman filter with three-dimensional variational data assimilation in a limited-area weather prediction model and comparison to E4DVar. *Mon. Wea. Rev.*, **141**, 900–917, doi:10.1175/MWR-D-12-00075.1.
- Zhang, M., and F. Zhang, 2012: E4DVar: Coupling an ensemble Kalman filter with four-dimensional data assimilation in a limited-area weather prediction model. *Mon. Wea. Rev.*, 140, 587–600, doi:10.1175/MWR-D-11-00023.1.
 - —, —, X.-Y. Huang, and X. Zhang, 2011: Intercomparison of an ensemble Kalman filter with three- and four-dimensional variational data assimilation methods in a limited-area model over the month of June 2003. *Mon. Wea. Rev.*, **139**, 566–572, doi:10.1175/2010MWR3610.1.