Nudging, Ensemble, and Nudging Ensembles for Data Assimilation in the Presence of Model Error

LILI LEI* AND JOSHUA P. HACKER

National Center for Atmospheric Research,[+] Boulder, Colorado

ABSTRACT

Objective data assimilation methods such as variational and ensemble algorithms are attractive from a theoretical standpoint. Empirical nudging approaches are computationally efficient and can get around some amount of model error by using arbitrarily large nudging coefficients. In an attempt to take advantage of the strengths of both methods for analyses, combined nudging-ensemble approaches have been recently proposed. Here the two-scale Lorenz model is used to elucidate how the forecast error from nudging, ensemble, and nudging-ensemble schemes varies with model error. As expected, an ensemble filter and smoother are closest to optimal when model errors are small or absent. Model error is introduced by varying model forcing, coupling between scales, and spatial filtering. Nudging approaches perform relatively better with increased model error; use of poor ensemble covariance estimates when model error is large harms the nudging-ensemble performance. Consequently, nudging-ensemble methods always produce error levels between the objective ensemble filters and empirical nudging, and can never provide analyses or short-range forecasts with lower errors than both. As long as the nudged state and the ensemble-filter state are close enough, the ensemble statistics are useful for the nudging, and fully coupling the ensemble and nudging by centering the ensemble on the nudged state is not necessary. An ensemble smoother produces the overall smallest errors except for with very large model errors. Results are qualitatively independent of tuning parameters such as covariance inflation and localization.

1. Introduction

Objective data assimilation combines observations with a forecast to produce an analysis that is the best estimate of the state by some norm, which is most often minimum variance or maximum likelihood. Empirical data assimilation uses intuitive spatial and/or temporal functions in combination with tuning parameters to provide an analysis judged satisfactory by a user-dependent metric. Model states resulting from either approach can be used as initial conditions for a numerical prediction.

No single data assimilation methodology always provides the best forecast, and different strengths can be combined. Recent literature (surveyed below) makes it clear that concepts from a variety of ensemble filters, variational minimization techniques, incremental analysis update (IAU), and nudging may benefit future systems. With that in mind, we aim to examine performance variations in nudging and ensemble filters as they relate to varying magnitudes of model inadequacy.

The ensemble Kalman filter (EnKF; Evensen 1994; Burgers et al. 1998), offered as a Monte Carlo alternative to the extended Kalman filter, is a discrete representation of the linear Kalman–Bucy filter (Kalman 1960) that samples forecast distributions with an ensemble of predictions from a (possibly) nonlinear model. Jazwinsky (1970) develops filters for stochastic dynamical systems, providing the relationship between the discrete Kalman filter and nudging. A vector stochastic differential equation (which can be either nonlinear or linear) describes system evolution. The stochastic term is the product of a matrix function and Gaussian white noise (the differential of Brownian motion). Because the

---

white noise is independent of the system's state, it has properties similar to innovation statistics when the observations are random variables. Discretizing the stochastic differential equation relies on Gaussian probability distributions, and the discrete Kalman filter is then optimal for a perfect model. The result is the more familiar form for data assimilation.

Nudging, also known as Newtonian relaxation, is a continuous data assimilation method. It relaxes the model state toward observations by adding new terms, proportional to the difference between observations and model state, to the prognostic equations (Hoke and Anthes 1976). It can be seen in Jazwinsky (1970) that the nudging term is analogous to the stochastic term in a stochastic differential equation for the dynamical system as long as the observations are assumed random. In numerical weather prediction (NWP) models, the nudging coefficients are typically tuned manually. Those coefficients are analogous to the matrix function multiplying the white noise in the stochastic differential equation; in the optimal solution it depends on the instantaneous uncertainty in the dynamical system, which includes model error. Because nudging does not consider that uncertainty, it can be tuned to be resistant to model error. In practice, models can be subjectively tuned to fit observations arbitrarily closely and/or to give skillful forecasts. Small relaxation terms in nudging can also avoid large increments associated with sequential methods.

Although empirical, nudging has been widely used in NWP and air quality studies (e.g., Leidner et al. 2001; Deng et al. 2004; Garvert et al. 2005; Deng and Stauffer 2006; Schroeder et al. 2006; Otte 2008a,b; Dixon et al. 2009; Ballabrera-Poy et al. 2009). Nudging is also a critical component in forming proposal densities under the equivalent weights particle filter proposed for initializing high-dimensional systems (van Leeuwen 2010; Ades and van Leeuwen 2015).

Recent work at operational centers suggests that nudging may prove to be a viable component in future hybrid systems. Current versions of the Met Office's hybrid 4D ensemble–variational data assimilation (4DEnVar) system (Clayton et al. 2013; Lorenc et al. 2015) are using IAU (Bloom et al. 1996) for initialization. Lorenc et al. (2015) applied a four-dimensional IAU (4DIAU) in their hybrid 4DEnVar. The 4DIAU adds to the traditional IAU the propagation of analysis increments in an assimilation window. Nudging and IAU share the trait of slowly introducing increments via additional terms in the model equations. But unlike nudging, IAU typically has the potentially undesirable property of artificially forcing the model without explicit consideration of differences between an observation, or analysis, and the model state (i.e., an innovation) at the precise time the forcing is applied.

The IAU could be adapted so that observation bins correspond directly with the model time steps, and reapplied closer to the observation time, but that would be impractical. Nudging is more amenable to observing networks with heterogeneous reporting times. Although research is needed to understand the impact, it is conceivable that nudging would prove a useful replacement to IAU in the 4DEnVar.

To combine the strengths of the EnKF and nudging, a nudging-EnKF (NE) was proposed by Lei et al. (2012a,b,c). The NE retains the flow-dependent error covariances of the EnKF, but eliminates temporal discontinuities resulting from sequential analysis updates that can lead to spurious oscillations affecting short-range forecasts. The NE constructs flow-dependent and time-varying nudging coefficients from the EnKF, and includes the effects from nonzero off-diagonal elements in the nudging coefficients. The NE also avoids the need for manual tuning of nudging coefficients.

One motivation for the NE approach is to reduce the imbalance caused by the intermittent update from EnKF. Bergemann and Reich (2010) proposed a "mollified" EnKF (MEnKF) to damp the spurious high-frequency response that can result from the sequential EnKF. The MEnKF can be viewed as a continuous formulation of the analysis step in the EnKF, and can also be viewed as a nudging method where the nudging coefficients are determined from the ensemble. In variational methods, the imbalance can be eliminated by including a digital-filter penalty term (Gauthier and Thepaut 2001) as a weak constraint, or applying the tangent-linear normal-mode constraint (TLNMC; Kleist et al. 2009). Other methods to address possibly deleterious imbalances, separate from the assimilation, include normal mode initialization (Machenhauer 1977; Baer and Tribbia 1977) and digital filtering (Lynch and Huang 1992; Huang and Lynch 1993).

Use of ensemble covariances in a nudging-ensemble method is a step toward an optimal filter, while avoiding introduction of spurious modes may be a practical advantage. But it is clear that nudging, as typically implemented for NWP and also in the NE, does not give a solution to Bayes's theorem. Bayes's theorem, which provides the optimal posterior distribution, guarantees a finite shift from prior to posterior mean in the system state at analysis time, when observations are presented. Only by the more general stochastic formulation presented by Jazwinsky (1970) can Bayes's theorem be satisfied by a nudging process, given linear dynamics and a forward operator, Gaussian error statistics, and perfectly described error distributions.

Lei et al. (2012 a,b) applied the NE in the Lorenz three-variable system (Lorenz 1963) and a two-dimensional

shallow-water model, with a focus on producing a continuous analysis. An NE-based analysis with the Weather Research and Forecasting (WRF) Model (Skamarock et al. 2008) was then used to drive an atmospheric transport and dispersion model (Lei et al. 2012c). Results showed that NE consistently gave lower analysis error than one of either EnKF or nudging (but generally not both). When imbalances from updates were large, and flow-dependent covariances were important, the NE showed the lowest short-range forecast errors. That work did not examine the role of model inadequacy, and was also limited to investigations of analysis errors. Optimal methods in data assimilation are judged on the forecast as opposed to the analyses. The key extensions to the past work are to examine forecast errors rather than analysis errors, and how they vary with model errors.

We know that optimal methods such as the ensemble filter become suboptimal when an unknown model inadequacy is present, or the model inadequacy is not properly handled in the assimilation. Often, that is mitigated by inflating covariances so that the model's prior (forecast) distribution has a stronger overlap with the observation likelihood (e.g., Anderson and Anderson 1999; Whitaker and Hamill 2002, 2012). Conversely, nudging is blind to model inadequacy, and deleterious effects can be manually mitigated to achieve an arbitrarily close fit to observations. A systematic inspection of the NE method and its components, each tuned under varying magnitudes of model inadequacy, reveal clear behavior that may not be surprising, but can inform and lend interpretation to results obtained from the myriad of more complex hybrids proposed and under development.

The two-scale model III described by Lorenz (2005, hereafter L05), with a state size of 960 variables, provides the dynamical system of choice. It allows realistic assimilation experiments that assume a perfect model, and can be used to consider model errors by varying a number of parameters or eliminating the fast scale.

The structure of this paper is as follows. Section 2 briefly describes the data assimilation methods used in this study, which include the EnKF, the ensemble Kalman smoother (EnKS; e.g., Evensen and van Leeuwen 2000), nudging, and NE. The L05 model and experimental design are discussed in section 3. Section 4 presents the results and discussion. The conclusions are summarized in section 5.

2. Data assimilation methods

In this section the ensemble and nudging methods are briefly reviewed. References for more details are included and easy to find. Let $\mathbf{x}$ and f be the state vector and tendency function of a dynamical system, respectively. The evolution of the system can be written as

$$\frac{d\mathbf{x}}{dt} = f(\mathbf{x}). \tag{1}$$

Evensen (1994) proposed that the background error covariance in the Kalman filter be estimated from an ensemble of model forecasts by evolving the analysis distribution at a previous time through (1). The update equation for the ensemble mean of EnKF can be written as

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^b + \mathbf{K}(\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^b), \tag{2}$$

where the overbar denotes ensemble mean; superscripts b and a denote background and analysis, respectively; $\mathbf{y}^o$ is the observation vector; and $\mathbf{H}$ is a linearized observation operator that derives observable quantities from the model state. The EnKF gain matrix $\mathbf{K}$ is

$$\mathbf{K} = \mathbf{P}^b\mathbf{H}^T(\mathbf{H}\mathbf{P}^b\mathbf{H}^T + \mathbf{R})^{-1}, \tag{3}$$

where $\mathbf{R}$ is the observation error variance, and $\mathbf{P}^b$ is the background error covariance that is estimated from the ensemble forecasts. Because a model is used to provide a sample element from the forecast error distribution, model inadequacy results in flawed error estimates and a suboptimal filter.

Various implementations of the EnKF are in the literature (e.g., Tippett et al. 2003; Bishop et al. 2001; Anderson 2001). Here we choose the deterministic ensemble adjustment Kalman filter (EAKF; Anderson 2001), and expect that results are independent of the specific update method. With the assumption that observation errors are uncorrelated, the observations can be assimilated serially (Houtekamer and Mitchell 2001). Making this assumption, at an analysis time the EAKF computes an observation increment and then computes the corresponding state variable increment of this observation for each ensemble member, and repeats this procedure for the following observations using the ensemble that has been updated with previous observations.

As an extension of the EnKF, the EnKS (Cohn et al. 1994; Evensen and van Leeuwen 2000) assimilates observations in the future at an analysis time by using the estimated temporal correlation from the ensemble model forecasts at forward lag times. A fixed-lag EnKS uses the EnKF solution as the first guess for analysis, thus the lag-0 implementation is the same as the EnKF. For lag $l > 0$, the increment for state variable at observation time k given the future observation with time lag l is

$$\overline{\mathbf{x}}_{k|k+l}^a = \overline{\mathbf{x}}_{k|k+l-1}^a + \mathbf{K}_{k|k+l}(\mathbf{y}_{k+l}^o - \mathbf{H}_{k+l}\overline{\mathbf{x}}_{k+l|k+l-1}^b), \quad (4)$$

where

$$\mathbf{K}_{k|k+l} = \mathbf{P}_{k|k+l-1,k+l}^{ab}\mathbf{H}_{k+l}^T(\mathbf{H}_{k+l}\mathbf{P}_{k+l|k+l-1}^b\mathbf{H}_{k+l}^T + \mathbf{R}_{k+1})^{-1}. \tag{5}$$

The subscript notation $m|n$ refers to quantity at observation time m, which incorporates information of all observations within the window up to time n. The term $\mathbf{P}_{k|k+l-1,k+l}^{ab}$ denotes the covariance matrix between analysis perturbation (with superscript a) at time k given observations up to time $k + l - 1$ and forecast perturbation (with superscript b) at time $k + l$.

Unlike the sequential EnKF method and the smoother, nudging applies a small correction over several time steps to gradually relax the model state toward observations. Artificial nudging terms, appended to the model dynamical equations, are proportional to the difference between observations and model state. The basic form of a dynamical assimilation system with nudging terms can be written as

$$\frac{d\mathbf{x}}{dt} = f(\mathbf{x}) + Gw_sw_t(\mathbf{y}^o - \mathbf{H}\mathbf{x}), \tag{6}$$

where G is the nudging coefficient that determines the strength of relaxation; and w_s and w_t are spatial and temporal weighting coefficients, respectively, that determine how to spread the innovation in space and time.

The optimal filter for a continuous or discrete system (Jazwinsky 1970) can give guidance for the coefficients in (6), and lead to nudging on an instantaneous time scale. Here observations are not available every time step and nudging without a time window would require large nudging coefficients (G) for the model to follow the observed evolution of the system. The coefficients in real atmospheric applications are instead tuned manually to give the lowest analysis or forecast error according to a set of chosen metrics. Nudging is thus an empirical method for data assimilation. It is also blind to model inadequacy, and a nudging system can be tuned to function well when model error is present. Here the nudging scheme is tuned to produce the lowest 6-h forecast error, which is consistent with the ensemble Kalman filter and smoother that are also optimized on the 6-h forecasts (see a further discussion in section 4).

To use flow-dependent background error covariances available in the EnKF, but maintain model states that evolve smoothly through an analysis time, an NE method was proposed by Lei et al. (2012 a,b). An ensemble filter and a nudging data assimilation system are maintained in parallel, and the Kalman gain $\mathbf{K}$ is folded into the nudging coefficient. The NE provides a continuous nudging state with flow-dependent background error covariance embedded in the nudging coefficients, which can be used in other applications [e.g., to drive an atmospheric transport and dispersion model as in Lei et al. (2012c)]. The NE method is equivalent to the optimal filter in Jazwinsky (1970) when the nudging time window is 0, with linear dynamics, linear forward operator, Gaussian error statistics, and perfectly described error distributions. By using the ensemble covariance, it is closer to an optimal filter than empirical nudging, but it does not give a solution to Bayes's theorem with nonzero nudging time window.

The product of the nudging coefficient and spatial weight in (6) is given by

$$Gw_s = \frac{\mathbf{K}}{\sum\limits_{t=t^o-\tau_N}^{t^o} w_t \Delta t}, \tag{7}$$

where t is model time, t^o denotes observation time, Δt is time step, and τ_N is half of the nudging time window. Matrix $\mathbf{P}^b$ is used to deliver $\mathbf{K}$, and thus the nudging coefficient and spatial weighting are estimated from the ensemble statistics. Consequences of model inadequacy manifesting in the ensemble filter are introduced to the nudging scheme via use of the ensemble statistics.

Two NE methods, differing by the extent of the coupling between the filter and the nudging, are considered here. One can recenter the ensemble mean on the nudging analysis when an ensemble analysis (posterior) is available (NE2). This can be thought of as two-way coupling because the nudged state affects the ensemble. The other skips the recentering; the ensemble provides information to the nudging but receives nothing in return. We can denote this as one-way coupling (NE1). Results later show that errors are relatively insensitive to one-way or two-way coupling, compared to differences from the other methods.

3. Model and experimental design

a. The L05 model

The two-scale model III described in L05 is flexible in a number of parameters, and allows realistic assimilation experiments that assume a perfect model, or can include model errors. The model also degenerates to a one-scale model by eliminating the fast scale. A summary of the models III and II in L05 follows.

The two-scale model III is written as

$$\frac{dZ_n}{dt} = [X,X]_{K,n} + b^2[Y,Y]_{1,n} + c[Y,X]_{1,n} \\ - X_n - bY_n + F. \tag{8}$$

Here Z is the integration variable, and X and Y are slow and fast variables. The construction of X and Y through Z is defined below. Coefficient b determines the relative frequency and amplitude of Y compared to X. Coupling coefficient c gives the strength of coupling between X and Y. Term F is a constant forcing. For perfect-model experiments, the parameters b, c, and F are chosen to be 10, 3, and 15, respectively. The subscript n indexes $N =$ 960 grid points, which gives grid spacing of 0.375°. The constant $K = 32$ is chosen to be much smaller than N, and another constant $J = K/2$ when K is even and $(K - 1)/2$ when K is odd.

To formulate the advection term in (8), a modified summation Σ', which is like the ordinary summation except that the first and last terms are divided by 2, is used. The advection term is given by

$$[X, X]_{K,n} = -W_{n-2K}W_{n-K} + \sum_{j=-J}^{J}{}' W_{n-K+j}X_{n+K+j}/K,$$ (9)

where $W_n = \sum_{i=-J}^{J}{}' X_{n-i}/K$, and Σ' is replaced by Σ when K is odd.

By choosing a smoothing scale I, and a pair of constants α and β, the separation of scales is achieved by

$$X_n = \sum_{i=-I}^{I}{}' (\alpha - \beta|i|)Z_{n+i}, \quad \text{and}$$ (10a)

$$Y_n = Z_n - X_n.$$ (10b)

The constants α and β are chosen such that X_n will equal Z_n whenever Z varies quadratically over the interval $n - I$ through $n + I$. The smoothing scale I is chosen as 12 for the truth model.

When the fast variable Y is eliminated, model III in (8) degrades to single scale model II given as

$$\frac{dX_n}{dt} = [X, X]_{K,n} - X_n + F.$$ (11)

Equations (8) and (11) are integrated with the fourth-order Runge–Kutta scheme with a nondimensional time step of 0.001 that is, by error-growth scaling arguments, roughly equivalent to 432 s.

b. Experimental design

Initial conditions are drawn from a large set of independent states. The nature run (truth) uses one state, and is integrated for 30 days. The known forward operator supplies synthetic observations of the truth for assimilation. Even grid points comprise the observation network (480 observing locations). Synthetic observations are created

every 50 time steps ($\sim$6 h) by adding random perturbations drawn from a normal distribution, with mean 0 and variance 1.0, to the true values at the observed locations.

Each data assimilation method described in section 2 is evaluated in both perfect-model and imperfect-model assimilation experiments. In perfect-model experiments, model III produces the nature run, and is also the assimilating model. For imperfect-model experiments, model error is included by varying parameters in the assimilating model, including the forcing term F, smoothing scale I, and coupling strength c. Finally, in an example of an extremely poor model, the fast variable is retained in model III for the nature run, but eliminated from the assimilating model (model II).

All experiments use an ensemble size of 40. Spurious error correlation between an observation and a state variable can be mitigated by localizing (tapering) the covariances with the Gaspari and Cohn (1999) fifth-order piecewise polynomial function [GC; their Eq. (4.10)]. A single real parameter determines the width of the GC localization function. To ensure ensemble spread and prevent filter divergence, constant multiplicative covariance inflation is applied (Anderson and Anderson 1999). Error covariance inflation is applied in EnKS only to the prior estimate of the EnKF (Khare et al. 2008). The EnKS and NE use the same inflation and localization as the EnKF. Localization and inflation parameters are manually tuned for each experiment by running a range of inflation factors and localization function widths.

Spatial weights (w_s) in nudging are specified as an isotropic quadratic function with a fixed radius of influence (Stauffer and Seaman 1994). The temporal nudging coefficients w_t are a trapezoidal function of time as defined by Stauffer and Seaman (1994), with a half time window specified by τ_N. The trapezoidal function is also used to nudge with the ensemble covariances in the NE methods. Localization (i.e., radius of influence), the half time window, and the nudging coefficient G are also manually tuned for the nudging experiments.

The first 10 days of each experiment are discarded to avoid transients. The last 20 days are used for evaluation. The root-mean-square error (RMSE) of the ensemble mean, compared to the true state, is used for verification. Results are shown for the RMSE of the prior 6-h forecasts.

4. Results

In perfect-model assimilation experiments using model III and imperfect-model assimilation experiments using model II, parameters are manually tuned to produce the smallest 6-h prior RMSE. Manually tuned parameters for EnKF include those for inflation and localization, which are also used in the EnKS, NE1, and

TABLE 1. Values of tunable quantities for the perfect-model experiments, giving the lowest 6-h forecast RMSE for the nudging, EnKF, EnKS, NE2, and NE1. Localization and inflation tuned for EnKF are also used for EnKS and NE schemes. Time quantities τ_N and l are given in time-step units.

Scheme	Localization scale	Inflation	$\tau_N/2$	G	l
Nudging	16	—	6	4	—
NE2	256	1.05	1	—	—
NE1	256	1.05	2	—	—
EnKF	256	1.05	—	—	—
EnKS	256	1.05	—	—	2

NE2 experiments. Lag l for the EnKS is tuned. Nudging requires coefficient G, radius of influence w_s, and time window τ_N (also applied to the NE). Table 1 gives values for perfect-model experiments that result in the lowest RMSE. The nudging time window is short, especially for the NE methods. This is consistent with Lei et al. (2012a), and is possibly due to the highly nonlinear system in which nudging may not work well when the system is at transitions, since the model state may evolve too quickly to be corrected by the gradual nudging forcing.

Lei et al. (2012a,b,c) argued that nudging can avoid dynamical imbalances arising from sequential assimilation and unbalanced analysis increments, and used that to motivate the NE approach. While nudging may in fact mitigate spurious dynamical responses, estimation theory (whether Bayesian or frequentist) applied to a discrete dynamical system requires temporal discontinuities as might be given by an optimal sequential update. Here we simply point out that in the L05 system, nudging does show less indication of a spurious dynamic response compared to the sequential introduction of increments. Results presented later will confirm that finite state shifts arising from sequential updates are not of primary concern.

Average kinetic energy tendency in L05, following assimilation times, measures accelerations (Fig. 1). Tendencies for nudging are smaller than both the filter and smoother for the first 12 h of the forecast. Tendencies of NE methods are greater at first, but decay faster. We speculate that the NE methods produce greater tendencies because the nudging state in the NE evolves according to its dynamics while being nudged back to the time-zero analysis distribution from the ensemble. Nudging by itself does not suffer from this effect because it is tuned to avoid it.

During the last 20 days, a 16-day forecast is executed from each analysis, with RMSE shown in Fig. 2. Forecasts from the nudged analyses are each an individual realization, and saturate at $\sqrt{2}$ times the errors in the ensemble-mean forecasts from the EnKF and EnKS ensemble analyses.
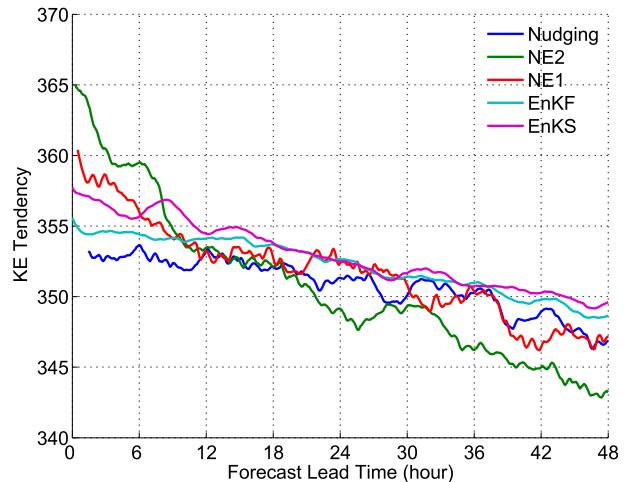

FIG. 1. Kinetic energy tendency immediately following an assimilation time. Curves for nudging and the NEs begin at the end of the nudging window, after the assimilation times from the sequential methods.

The EnKS is the most optimal in perfect-model experiments (Fig. 2a), with smaller RMSE than the EnKF at analysis time and throughout the forecast. Both show smaller errors than any of the nudging-based schemes at all lead times. At short lead times the two NE methods produce RMSEs between nudging and the EnKF. The error differences at initialization time can be seen more clearly in Fig. 3.

Eliminating the fast variable from the assimilating model, parameters in each assimilation method are again manually tuned to produce the smallest 6-h prior RMSE. For the ensemble-based methods, inflation is greater and the localization width smaller. Nudging coefficients are also greater to account for the model error. The lag l and time window τ_N often change from 2 to 1. The empirical nudging scheme shows its resistance to model inadequacy (Fig. 2b). The EnKS still shows smaller RMSE than the EnKF, but both have greater RMSE than the nudging during the first three forecast days. Because the nudging weights are from a flawed ensemble, the NE suffers from model inadequacy much like the filter and smoother. The overall performance of the NE is similar regardless of one-way or two-way coupling, and the short-time skill will become clearer below.

Performance variability across a range of model error can be explored more thoroughly by varying model parameters. Lacking detailed knowledge of the model response to each parameter, we seek qualitatively similar assimilation results from varying different parameters. The 6-h forecast RMSE is the metric, and the L05 paper provides some guidance for interpretation.

First varying the forcing F, the RMSE for different values of F in the assimilating model is shown in Fig. 3.
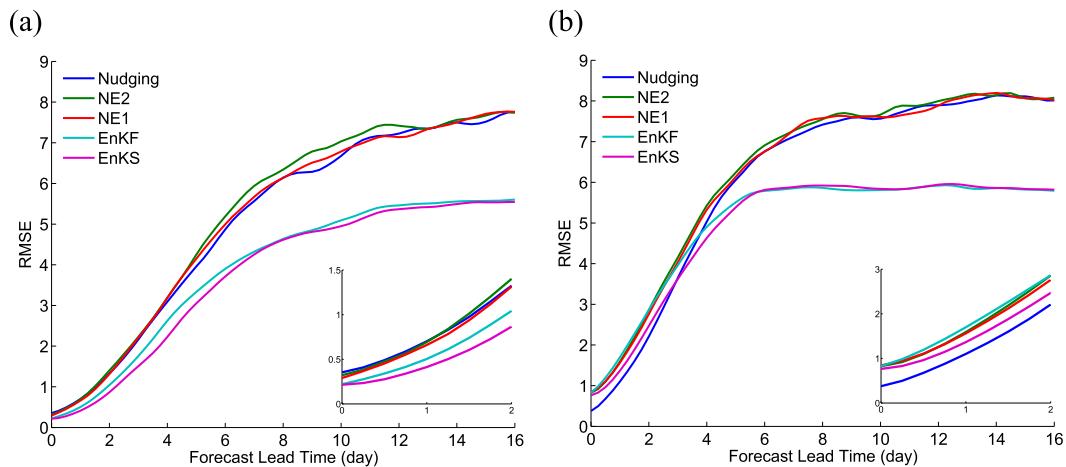
FIG. 2. RMSE by forecast lead times for different data assimilation methods. (a) Perfect-model assimilation experiments using model III and (b) imperfect-model assimilation experiments using model II. Inset panels enlarge the first two days.

To account for model error, inflation and localization parameters are tuned for each value of F to produce the smallest 6-h forecast RMSE for that particular experiment. Perfect-model results ($F = 15$) here augment those shown in Fig. 2a.

In Fig. 3 the ensemble filter has lower error compared to nudging when F is close to 15. As model error increases, the error in the predictions with the ensemble filter increases more steeply than in the predictions from nudging. The NE methods maintain some advantages from the nudging at large model error levels, but suffer from the use of suboptimal covariance estimates. A treatment for the NE methods at large model error levels will be discussed later.

The errors of NE methods generally lie between errors from objective filter and empirical nudging. The differences between the NE and the EnKF results are smaller with positive F errors than with negative F errors; any explanation would at this point be speculation. NE1 has slightly smaller errors than NE2, but the differences are much smaller than between the NE experiments and the others. Thus, the NE appears relatively insensitive to one-way or two-way coupling. The smoother maintains the lowest error levels throughout the range of errors in F. But for a range of incorrect values of F, a skill transition from the ensemble filter to the nudging approach is clear as model error increases.

A statistical significant test using a bootstrap resampling with replacement (Efron and Tibshirani 1993) is applied on the errors from different experiments in Fig. 3 (not shown in the figure). The null hypothesis that the error differences between the NE methods and nudging are zero can be rejected at the 95% significance level. The

same is true for error differences between the NE methods and the EnKF.

The results are qualitatively unchanged when the data assimilation parameters are not retuned to account for varying levels of model error, but the cross-over between the ensemble filter and nudging is nearer to the perfect-model case along the model error axis (not shown). That is, unsurprisingly, the ensemble filter can tolerate more model error when covariance localization and inflation is tuned to account for the model error.

Results from varying the smoothing scale I and coupling strength c are similar to each other, and within reasonable ranges show qualitative similarity to varying
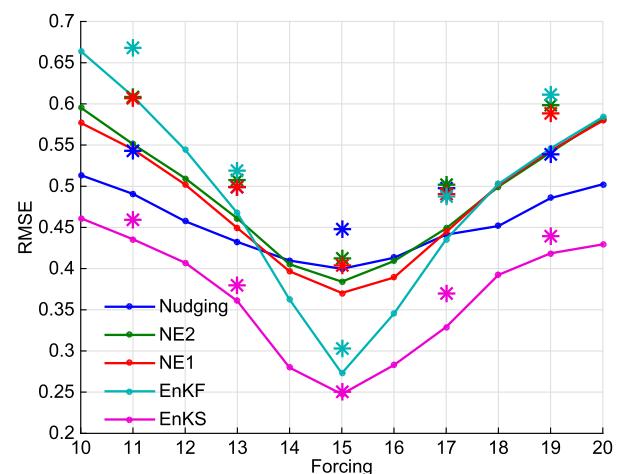

FIG. 3. Prior (6-h forecast) RMSE for different data assimilation methods, varying by the forcing term F in the assimilating model. The correct (true) $F = 15$. The asterisks show the prior RMSE for experiments with observations every fourth grid point rather than every second grid point.
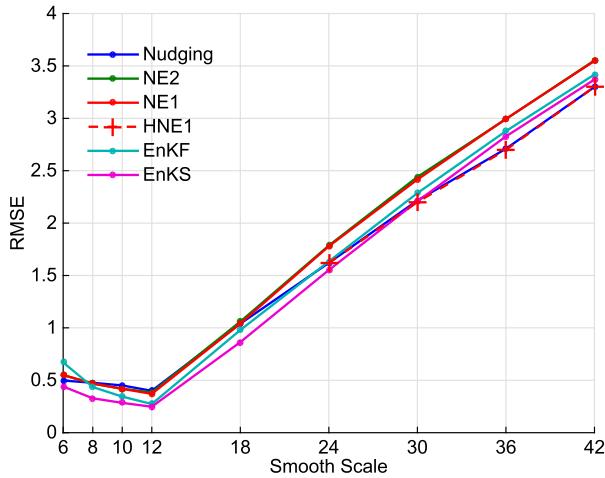
FIG. 4. Prior (6-h forecast) RMSE of different data assimilation methods with varying the smoothing scale I. The correct (true) $I = 12$.
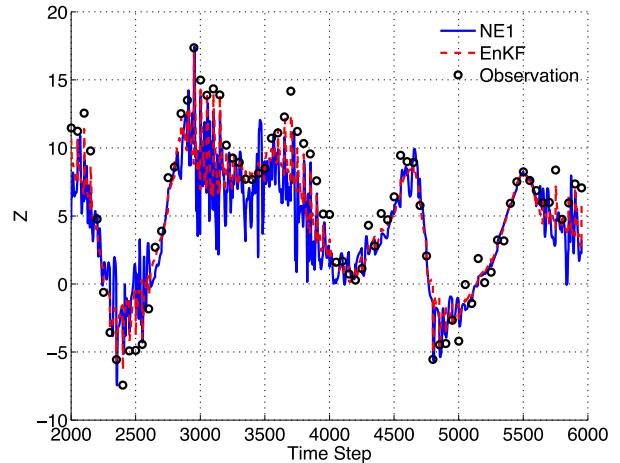

FIG. 5. Nudged state for NE1 (blue solid) and ensemble mean of EnKF (red dashed) at gridpoint 40 for the imperfect model with $I = 42$ during the last 20 days of assimilation. Circles are observations from the nature run with $I = 12$. Ordinate value Z is the L05 state variable.

F. Smoothing scale $I = 12$ is the perfect-model case, and corresponds to earlier results. As in Fig. 3, Fig. 4 shows the RMSE with varying I, and localization and inflation parameters are again tuned for each integer value of I.

The results with decreasing I are similar to those with decreasing F. The ensemble filter goes from best to worst (further from optimal), and nudging overcomes model error to be the most accurate for large model errors (small I). The NE methods give errors between the filter and nudging. The smoother maintains the lowest error levels. The error differences between the NE methods and nudging are significant at the 95% level. The error differences between the NE methods and EnKF are also significant at the 95% level.

As I in the assimilating model increases from the true value, a cross-over where the nudging outperforms the ensemble filter remains evident. The nudging produces a smaller error than the smoother when I is greater than 30. The NE methods produce larger errors than both nudging and the ensemble filter when I is larger than 18. When $I > 18$, the slow waves are eliminated, leaving nothing but the highly chaotic fast scale in the assimilating model (L05). This fundamental dynamical change prevents us from evaluating large positive errors in the smoothing scale. The range for reasonable dynamics is $10 < I < 18$. At $I < 6$ and $I > 42$, the state is pulled so far off its attractor by the observations that the model becomes numerically unstable when integrated.

One new twist is that the errors from the NE methods are greater than the errors from both the nudging and the filter for large I, resulting from the dynamics of the model at large I. The model has greater variance at small, fast scales than exists in the truth. The nudged state of NE1 at grid point 40 with $I = 42$ experiences

severe high-frequency fluctuations between observation times (Fig. 5). The sequence of observations lacks these scales, and appears random in the state prior to assimilation. Poor covariance estimates from the ensemble lead to poor nudging coefficients. Also the deterministic nudging state of the NE lacks the random-error smoothing present in the ensemble. Deficiencies in both schemes combine synergistically, and the NE is inferior to both individual methods. This contrasts to the canonical case of a model that is more diffusive than nature, but it serves to demonstrate an unusual behavior.

One possible way to improve the NE methods is to incorporate a hybrid covariance, rather than a purely ensemble-based covariance estimate, to form the nudging coefficient. This is akin to the hybrid variational–ensemble methods (e.g., Hamill and Snyder 2000; Lorenc 2003; Wang 2010). Hybrid variational–ensemble methods have been shown to help mitigate sampling errors from finite-sized ensembles (e.g., Etherton and Bishop 2004; Wang et al. 2007, 2009). The hybrid covariances for nudging is written as

$$Gw_s = \frac{1}{\displaystyle\sum_{t=t^o-\tau_N}^{t^o} w_t \Delta t} [\alpha \mathbf{K} + (1-\alpha)\mathbf{B}], \qquad (12)$$

where $\mathbf{K}$ is the sample Kalman gain from the ensemble and $\mathbf{B}$ is a static background error covariance matrix (here elements are simply given by the empirical nudging coefficients). The weighting parameter α in (12) determines the contribution from static and flow-dependent error covariance estimates, and requires tuning. It is clear

that (12) and (7) are equivalent for $\alpha = 1$. For $\alpha = 0$, (12) gives empirical nudging coefficients.

The NE method with hybrid covariances (called HNE) is tested for varying values of I. One-way coupling of HNE (HNE1) is examined; we expect the insensitivity of NE to one- or two-way coupling to be present here. With small model error (I close to 12), large values of α (approximately 0.75), heavily weighting the ensemble covariances, leads to similar error to NE1 (not shown for clarity in the figure). For large I, giving a model that produces poor covariance estimates from the ensemble and then poor nudging coefficients, smaller α (0.875 for $I = 24$ and 1.0 for $I > 24$) leads to lower errors, HNE1 errors converge to errors from nudging, as shown by the red dashed line in Fig. 4.

To test the sensitivity to observation networks, a sparser network with observations every fourth grid point is used with varying the forcing F. The RMSE of these experiments are shown with asterisks in Fig. 3. For each experiment, the RMSE with fewer observations is larger than that with more observations. Ranked RMSE among the experiments is consistent with those having observations on even grid points. Thus, the results with observations on even grid points would be expected to generally hold for different observation networks.

5. Conclusions

The impacts of model error on ensemble and empirical data assimilation methods are elucidated with the L05 model. For a range of model errors, an ensemble smoother always produces smaller errors than an ensemble filter. The filter has smaller errors than empirical nudging when the model error is small, and vice versa. Imbalances introduced through either large sequential updates or large nudging coefficients do not necessarily deteriorate forecasts in L05.

Nudging-ensemble (NE) methods produce errors between the objective filter and empirical nudging, except under extremely large errors in one particular model parameter where flawed covariances and lack of smoothing combine to make the NE poorest. The NE appears relatively insensitive to centering the ensemble on the nudged state when the nudging and ensemble independently provide analyses sufficiently close to each other.

In the L05 model, the NE methods cannot be the basis for a forecasting system expected to outperform one based on one of only nudging or ensemble methods. When model inadequacy dominates, the empirical nudging will outperform objective methods. When model error is a small part of forecast error, the objective methods will be superior. To draw more general conclusions, experiments with more realistic models are needed.

The NE method can be improved when model error is large by including a static background error covariance to form a hybrid covariance. Results here show that the use of static covariances, weighted heavily for large model error, lead to results that agree closely with empirical nudging. Further improvements may be realized by using the EnKS instead of the EnKF, replacing the Kalman gain in (7) with that from the EnKS [(5)]. The time-varying Kalman gain would be used in the nudging coefficients, and the innovations still computed at the observation time. Alternately, the NE method could use the EnKF with perturbed observations. Each ensemble member is nudged to perturbed observations. In these cases the need for separate ensemble and nudged states disappears.

Although none of the data assimilation systems can be optimal, the perfect-model results are consistent with theory. Namely, the filter and smoother are optimal for Gaussian error statistics, linear dynamics and forward operator, and no sampling error; the case of small model error is a small departure from the perfect-model case, and the objective methods remain clearly superior there.

REFERENCES

Ades, M., and P. J. van Leeuwen, 2015: The equivalent-weights particle filter in a high-dimensional system. *Quart. J. Roy. Meteor. Soc.,* **141,** 484–503, doi:10.1002/qj.2370.

Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.,* **129,** 2884–2903, doi:10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2.

——, and S. L. Anderson, 1999: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.,* **127,** 2741–2758, doi:10.1175/1520-0493(1999)127<2741:AMCIOT>2.0.CO;2.

Baer, F., and J. Tribbia, 1977: On complete filtering of gravity modes through nonlinear initialization. *Mon. Wea. Rev.,* **105,** 1536–1539, doi:10.1175/1520-0493(1977)105<1536:OCFOGM>2.0.CO;2.

Ballabrera-Poy, J., E. Kalnay, and S.-C. Yang, 2009: Data assimilation in a system with two scales—Combing two initialization techniques. *Tellus,* **61A,** 539–549, doi:10.1111/j.1600-0870.2009.00400.x.

Bergemann, K., and S. Reich, 2010: A mollified ensemble Kalman filter. *Quart. J. Roy. Meteor. Soc.,* **136,** 1636–1643, doi:10.1002/qj.672.

Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.,* **129,** 420–436, doi:10.1175/1520-0493(2001)129<0420:ASWTET>2.0.CO;2.

Bloom, S. C., L. L. Takacs, A. M. da Silva, and D. Ledvina, 1996: Data assimilation using incremental analysis updates. *Mon.*

Wea. Rev., **124,** 1256–1271, doi:10.1175/1520-0493(1996)124<1256: DAUIAU>2.0.CO;2.

Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998: Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.,* **126,** 1719–1724, doi:10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2.

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.

Cohn, S., N. S. Sivakumaran, and R. Todling, 1994: A fixed-lag Kalman smoother for retrospective data assimilation. *Mon. Wea. Rev.,* **122,** 2838–2867, doi:10.1175/1520-0493(1994)122<2838: AFLKSF>2.0.CO;2.

Deng, A., and D. R. Stauffer, 2006: On improving 4-km mesoscale model simulations. *J. Appl. Meteor. Climatol.,* **45,** 361–381, doi:10.1175/JAM2341.1.

——, N. L. Seaman, G. K. Hunter, and D. R. Stauffer, 2004: Evaluation of interregional transport using the MM5-SCIPUFF system. *J. Appl. Meteor.,* **43,** 1864–1886, doi:10.1175/JAM2178.1.

Dixon, M., Z. Li, H. Lean, N. Roberts, and S. Ballard, 2009: Impact of data assimilation on forecasting convection over the United Kingdom using a high-resolution version of the Met Office Unified Model. *Mon. Wea. Rev.,* **137,** 1562–1584, doi:10.1175/ 2008MWR2561.1.

Efron, B., and R. J. Tibshirani, 1993: *An Introduction to the Bootstrap.* Chapman and Hall, 436 pp.

Etherton, B. J., and C. H. Bishop, 2004: Resilience of hybrid ensemble/3DVar analysis schemes to model error and ensemble covariance error. *Mon. Wea. Rev.,* **132,** 1065–1080, doi:10.1175/ 1520-0493(2004)132<1065:ROHDAS>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.

Garvert, M. F., B. A. Colle, and C. F. Mass, 2005: The 13–14 December 2001 IMPROVE-2 event. Part I: Synoptic and mesoscale evolution and comparison with a mesoscale model simulation. *J. Atmos. Sci.,* **62,** 3474–3492, doi:10.1175/JAS3549.1.

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.

Gauthier, P., and J.-N. Thepaut, 2001: Impact of the digital filter as a weak constraint in the preoperational 4DVAR assimilation system of Meteo-France. *Mon. Wea. Rev.,* **129,** 2089–2102, doi:10.1175/1520-0493(2001)129<2089:IOTDFA>2.0.CO;2.

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

Hoke, J. E., and R. A. Anthes, 1976: The initialization of numerical models by a dynamical initialization technique. *Mon. Wea. Rev.,* **104,** 1551–1556, doi:10.1175/1520-0493(1976)104<1551: TIONMB>2.0.CO;2.

Houtekamer, P. L., and H. L. Mitchell, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.,* **129,** 123–137, doi:10.1175/1520-0493(2001)129<0123: ASEKFF>2.0.CO;2.

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

Jazwinsky, A. H., 1970: *Stochastic Processes and Filtering Theory.* Academic Press, 376 pp.

Kalman, R. E., 1960: A new approach to linear filtering and prediction problems. *J. Basic Eng.,* **82,** 35–45, doi:10.1115/ 1.3662552.

Khare, S. P., J. L. Anderson, T. J. Hoar, and D. Nychka, 2008: An investigation into the application of an ensemble Kalman smoother to high-dimensional geophysical systems. *Tellus,* **60A,** 97–112, doi:10.1111/j.1600-0870.2007.00281.x.

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, R. M. Errico, and R. Yang, 2009: Improving incremental balance in the GSI 3DVar analysis system. *Mon. Wea. Rev.,* **137,** 1046–1060, doi:10.1175/2008MWR2623.1.

Lei, L., D. R. Stauffer, S. E. Haupt, and G. S. Young, 2012a: A hybrid nudging- ensemble Kalman filter approach to data assimilation. Part I: Application in the Lorenz system. *Tellus,* **64A,** 18484, doi:10.3402/tellusa.v64i0.18484.

——, ——, and A. Deng, 2012b: A hybrid nudging-ensemble Kalman filter approach to data assimilation. Part II: Application in a shallow-water model. *Tellus,* **64A,** 18485, doi:10.3402/ tellusa.v64i0.18485.

——, ——, and ——, 2012c: A hybrid nudging-ensemble Kalman filter approach to data assimilation in WRF/DART. *Quart. J. Roy. Meteor. Soc.,* **138,** 2066–2078, doi:10.1002/qj.1939.

Leidner, S. M., D. R. Stauffer, and N. L. Seaman, 2001: Improving short-term numerical weather prediction in the California coastal zone by dynamic initialization of the marine boundary layer. *Mon. Wea. Rev.,* **129,** 275–294, doi:10.1175/ 1520-0493(2001)129<0275:ISTNWP>2.0.CO;2.

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

——, N. Bowler, A. Clayton, S. 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., 1963: Deterministic non-periodic flow. *J. Atmos. Sci.,* **20,** 130–141, doi:10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.

——, 2005: Designing chaotic models. *J. Atmos. Sci.,* **62,** 1574–1587, doi:10.1175/JAS3430.1.

Lynch, P., and X.-Y. Huang, 1992: Initialization of the HIRLAM model using a digital filter. *Mon. Wea. Rev.,* **120,** 1019–1034, doi:10.1175/1520-0493(1992)120<1019:IOTHMU>2.0.CO;2.

Machenhauer, B., 1977: On the dynamics of gravity oscillations in a shallow water model with applications to normal mode initialization. *Contrib. Atmos. Phys.,* **50,** 253–271.

Otte, T. L., 2008a: The impact of nudging in the meteorological model for retrospective air quality simulations. Part I: Evaluation against national observation networks. *J. Appl. Meteor. Climatol.,* **47,** 1853–1867, doi:10.1175/2007JAMC1790.1.

——, 2008b: The impact of nudging in the meteorological model for retrospective air quality simulations. Part II: Evaluating collocated meteorological and air quality observations. *J. Appl. Meteor. Climatol.,* **47,** 1868–1887, doi:10.1175/2007JAMC1791.1.

Schroeder, A. J., D. R. Stauffer, N. L. Seaman, A. Deng, A. M. Gibbs, G. K. Hunter, and G. S. Young, 2006: Evaluation of a high-resolution, rapidly relocatable meteorological nowcasting and prediction system. *Mon. Wea. Rev.,* **134,** 1237– 1265, doi:10.1175/MWR3118.1.

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.]

Stauffer, D. R., and N. L. Seaman, 1994: Multiscale four-dimensional data assimilation. *J. Appl. Meteor.,* **33,** 416–434, doi:10.1175/1520-0450(1994)033<0416:MFDDA>2.0.CO;2.

Tippett, M., J. Anderson, C. Bishop, T. Hamill, and J. Whitaker, 2003: Ensemble square root filters. *Mon. Wea. Rev.,* **131,** 1485–1490, doi:10.1175/1520-0493(2003)131<1485:ESRF>2.0.CO;2.

van Leeuwen, P. J., 2010: Nonlinear data assimilation in geosciences: An extremely efficient particle filter. *Quart. J. Roy. Meteor. Soc.,* **136,** 1991–1999, doi:10.1002/qj699.

Wang, X., 2010: Incorporating ensemble covariance in the gridpoint statistical interpolation variational minimization: A mathematical framework. *Mon. Wea. Rev.,* **138,** 2990–2995, doi:10.1175/2010MWR3245.1.

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

——, ——, ——, and ——, 2009: A comparison of the hybrid and EnSRF analysis schemes in the presence of model errors due to unresolved scales. *Mon. Wea. Rev.,* **137,** 3219–3232, doi:10.1175/2009MWR2923.1.

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.

——, and ——, 2012: Evaluating methods to account for system errors in ensemble data assimilation. *Mon. Wea. Rev.,* **140,** 3078–3089, doi:10.1175/MWR-D-11-00276.1.