Relationships among Four-Dimensional Hybrid Ensemble–Variational Data Assimilation Algorithms with Full and Approximate Ensemble Covariance Localization

CHENGSI LIU

Center for Analysis and Prediction of Storms, University of Oklahoma, Norman, Oklahoma

MING XUE

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

(Manuscript received 26 May 2015, in final form 6 November 2015)

ABSTRACT

Ensemble–variational data assimilation algorithms that can incorporate the time dimension (four-dimensional or 4D) and combine static and ensemble-derived background error covariances (hybrid) are formulated in general forms based on the extended control variable and the observation-space-perturbation approaches. The properties and relationships of these algorithms and their approximated formulations are discussed. The main algorithms discussed include the following: 1) the standard ensemble 4DVar (En4DVar) algorithm incorporating ensemble-derived background error covariance through the extended control variable approach, 2) the 4DEnVar neglecting the time propagation of the extended control variable (4DEnVar-NPC), 3) the 4D ensemble–variational algorithm based on observation space perturbation (4DEnVar), and 4) the 4DEnVar with no propagation of covariance localization (4DEnVar-NPL). Without the static background error covariance term, none of the algorithms requires the adjoint model except for En4DVar. Costly applications of the tangent linear model to localized ensemble perturbations can be avoided by making the NPC and NPL approximations. It is proven that En4DVar and 4DEnVar are mathematically equivalent, while 4DEnVar-NPC and 4DEnVar-NPL are mathematically equivalent. Such equivalences are also demonstrated by single-observation assimilation experiments with a 1D linear advection model. The effects of the non-flow-following or stationary localization approximations are also examined through the experiments.

All of the above algorithms can include the static background error covariance term to establish a hybrid formulation. When the static term is included, all algorithms will require a tangent linear model and an adjoint model. The first guess at appropriate time (FGAT) approximation is proposed to avoid the tangent linear and adjoint models. Computational costs of the algorithms are also discussed.

1. Introduction

Variational data assimilation (DA) (Le Dimet and Talagrand 1986; Talagrand and Courtier 1987) and the ensemble Kalman filter (EnKF; Evensen 1994) are two major advanced approaches for atmospheric DA. The variational DA approach has been successfully used at many operational numerical weather prediction (NWP) centers, first as the three-dimensional variational data assimilation (3DVar) then the four-dimensional variational data assimilation (4DVar) method (e.g., Parrish and Derber 1992; Lorenc et al. 2000; Rabier et al. 2000; Rawlins et al. 2007; Tanguay et al. 2012). Usually, 3DVar uses a static, flow-independent, climatological background error covariance (BEC) that is often spatially homogeneous and anisotropic (e.g., Parrish and Derber 1992; Purser et al. 2003a). Even though some efforts have been made to introduce spatially inhomogeneous, anisotropic BEC into the 3DVar framework (e.g., Wu et al. 2002; Purser et al. 2003b, Fisher 2003), the determination of flow-dependent BEC remains a major challenge.

Compared to 3DVar, the 4DVar method allows the fitting of the model forecast trajectory to observations distributed over a period of time so as to provide more accurate model state estimations that are also more consistent with the prediction model. The use of the model as a strong constraint within the DA system also allows for the retrieval of unobserved state variables from limited observations (e.g., Sun and Crook 1997). 4DVar also implicitly evolves the BEC within the DA window so that it can be flow dependent, but the BEC itself defined at the beginning of the DA window is usually static and is not propagated from one DA window to another (Lorenc 2003).

EnKF estimates flow-dependent BEC from a set of ensemble forecasts and updates the ensemble states based on an optimal linear estimation algorithm (Evensen 1994). Because EnKF estimates and evolves flow-dependent BEC within and through the data assimilation cycles, and does not require tangent linear or adjoint model, EnKF has become increasingly popular within both research and operational communities. EnKF and its variants, including the ensemble transform Kalman filter and local ensemble transform Kalman filter, ensemble adjustment Kalman filter, and ensemble square root filter (Bishop et al. 2001; Anderson 2001; Hunt et al. 2007; Whitaker and Hamill 2002), have now been implemented at a number of operational NWP centers (e.g., Houtekamer and Mitchell, 2005; Whitaker et al. 2008; Hamill et al. 2011a).

The EnKF method can also be extended into the time dimension to assimilate observations distributed over a period of time and such algorithms usually rely on BECs across times to update the state at the analysis time. Four-dimensional EnKF algorithms include those of Hunt et al. (2004), Sakov et al. (2010), and S. Wang et al. (2013), based on variants of EnKF. Computationally, a pure ensemble-based DA system tends to be more scalable than traditional 4DVar because the ensemble forecasting portion of the system that tends to be most expensive can be easily parallelized while the time integrations of the tangent linear and adjoint models involved in the traditional 4DVar have to be performed sequentially across the minimization iterations. Further discussions on the advantages and disadvantages of 4DVar and EnKF can be found in Lorenc (2003), Kalnay et al. (2007a,b), and Gustafsson (2007).

The BEC matrix derived from an ensemble of forecasts that is typically much smaller in size compared to the number of the degrees of freedom of typical NWP models is severely rank deficient. Techniques, such as covariance localization, have been proposed to partially alleviate this problem (e.g., Burgers et al. 1998; Houtekamer and Mitchell 1998; Hamill et al. 2001;

Anderson 2001; Whitaker and Hamill 2002; Evensen 2003). In comparison, the static climatological BEC typically used by 3DVar is full rank. The static BEC, being derived from typically much larger samples often contains useful balance information of less transient flows. From such considerations, the use of "hybrid"[1] BEC that linearly combines static and ensemble-derived flow-dependent BECs had been proposed. The hybrid BEC can overcome limitations of ensemble BEC that spans only the space occupied by the ensemble itself.

Hamill and Snyder (2000) were the first to test this idea within a 3DVar framework, with an algorithm that they called hybrid EnKF–3DVar scheme. They demonstrated this method with a low-resolution quasigeostrophic model and simulated data in a perfect model setting and found that the static BEC is helpful to the analysis when flow-dependent BEC is derived from a small ensemble. For their implementation, the ensemble-derived BEC was explicitly calculated and stored, and combined with the static BEC in their 3DVar framework; while this implementation is attractive for low-dimension problems because it is simple and straightforward, it would become prohibitively expensive for a real NWP model when the ensemble BEC matrix is huge.

Lorenc (2003) proposed a more elegant alternative hybrid algorithm that utilizes a set of extended control variables preconditioned by the ensemble perturbations in the variational cost function. Also, a correlation function is used to localize the ensemble covariance. Wang et al. (2007) demonstrated that this extended control variable formulation is mathematically equivalent to that of Hamill and Snyder (2000) in the 3D framework.

There are some advantages when utilizing the ensemble-derived covariance within a variational framework, as in the aforementioned hybrid algorithms, when compared to pure EnKF. The variational formulation allows for the application of BEC localization in the state space rather than observation space; the latter has problems with observations whose forward operators are nonlocal (Campbell et al. 2010). Additional benefits include easier implementation of equation constraints (Gauthier and Thépaut 2001) and bias correction in the variational framework (Dee and Uppala 2009). In general, variational algorithms that incorporate ensemble-derived BECs are called ensemble–variational or EnVar algorithms, and the hybrid algorithms of Hamill and Snyder

---

[1] In this study, we use the word "hybrid" to mainly refer to the use of a combination of the static and ensemble-derived flow-dependent covariances (i.e., the hybrid covariance). We do not intend to use "hybrid" to refer to any algorithm, although in the literature it had sometimes been used to refer to ensemble–variational algorithms.

(2000) and Lorenc (2003) are such examples. Ensemble-derived BECs have been shown to improve DA in 3DVar frameworks for operational global (Buehner 2005; Hamill et al. 2011b; X. Wang et al. 2013; Kleist and Ide 2015a) and regional (Pan et al. 2014) models.

More recently, EnVar algorithms have been extended into the fourth time dimension to form 4D EnVar algorithms. There are two basic types of such algorithms. One is a direct extension of the 3D extended control variable algorithm of Lorenc (2003) into four dimensions, which can also be considered as introducing the ensemble BEC into a standard 4DVar framework using the extended control variable approach. Being based on the standard 4DVar that involves the integration of tangent linear and adjoint models, we call this algorithm En4DVar. Buehner et al. (2010a,b) applied the extended control variable approach to a 4DVar framework (denoted 4D-Var-Benkf in their papers) and tested with real observations in the Canadian operational global model, and found positive impact of the ensemble-derived BEC. Very recently, a hybrid ensemble–4DVar system is implemented for the Met Office operational global model (Clayton et al. 2013). Zhang and Zhang (2012) also applied the extended control variable approach to a 4DVar framework of a regional research model. Kuhl et al. (2013) implemented En4DVar within the Naval Research Laboratory Atmospheric Variational Data Assimilation System-Accelerated Representer (NAVDAS-AR) data assimilation framework. It was found that the forecast error was significantly reduced by their En4DVar system. These systems were all built on existing 4DVar capabilities that already have an adjoint model, and correspond to the algorithm that we call En4DVar in this paper.

The second type of algorithm formulates a 4D variational cost function that projects the ensemble perturbations to the observation space so that tangent linear model and adjoint model can be avoided in the absence of the static BEC term (Liu et al. 2008). In Liu et al. (2009), a localized matrix is introduced into the algorithm for ensemble covariance localization and Liu and Xiao (2013) further applied it to real data problems. This algorithm was originally called En4DVar but is renamed 4DEnVar in Liu and Xiao (2013) to better distinguish it from algorithms that are more closely linked to the traditional 4DVar and include the integration of an adjoint model (Lorenc 2013). Following Liu and Xiao (2013), 4DEnVar is also used to refer to this type of algorithms in this paper while En4DVar is used to refer to algorithms involving the integration of an adjoint model. In Buehner et al. (2010a,b), a version of the 4DEnVar algorithm of Liu et al. (2008, 2009) was implemented within a global spectral model (called En-4D-Var in their papers) and compared with traditional 4DVar, EnKF,

and 4D-Var-Benkf methods. These data assimilation schemes had a similar performance in the northern extratropics and tropics. In the southern extratropics, En-4D-Var was slightly better than EnKF but slightly worse than 4D-Var-Benkf. Here we point out that the En-4D-Var implementation of Buehner et al. (2010a,b) corresponds to the algorithm that we will call 4DEnVar-NPC in this paper. Buehner et al. (2013) further compared En-4D-Var (corresponding to our 4DEnVar-NPC), 3DVar, and 4DVar for global weather prediction. They found that En-4D-Var is always better than 3DVar and is either similar or better than 4DVar in the tropical troposphere and the winter extratropical regions. Kleist and Ide (2015b) evaluated hybrid 4DEnVar with various initialization techniques within the National Centers for Environmental Prediction Global Data Assimilation System (GDAS) using simulated data. They found that the hybrid 4DEnVar can reduce analysis error for most variables at most levels, especially in the extratropics, compared to hybrid 3DEnVar.

Despite the successful applications of the various 4D ensemble–variational approaches, their relationships, as well as the approximations involved in their implementations, are still unclear. Desroziers et al. (2014) used a generalized variational formulation in terms of a 4D state vector to discuss different possible implementations of 4DEnVar. They proposed two new preconditioned algorithms and compared 4DEnVar and 4DVar for a Burgers equation model. However, approximations related to covariance localization in a 4D space was not discussed in detail in their paper. Fairbairn et al. (2014) pointed out that no explicit localization of the correlations in time was included in their experiments, or in most other implementations of 4DEnVar.

Most recently, Lorenc et al. (2015) compared hybrid-4DVar (which is our hybrid En4DVar) and hybrid-4DEnVar (which is actually our hybrid 4DEnVar-NPC). Hybrid-4DVar was found to perform better than hybrid-4DEnVar in the Met Office global operational system. They suggested the fact that the hybrid-4DVar evolves the static background error covariance within the assimilation window while the hybrid-4DEnVar does not was the main cause of the differences. It is the purpose of this paper to clarify the relationships and understand the approximations involved in the derivations and implementations of various 4D ensemble–variational algorithms. We present various derivative algorithms based on the two approaches in a common framework that can incorporate both static and ensemble-derived BECs. An understanding is also sought on the effects of various approximations made in the algorithms. Furthermore, in order to introduce static BEC to 4DEnVar while still avoiding an adjoint model, we propose the use of the first guess at the appropriate time (FGAT) approximation within the 4D hybrid schemes.

The organization of the paper is as follows. Four 4D ensemble–variational algorithms are first derived and discussed in a common framework in section 2. FGAT formulations for three hybrid algorithms that avoid adjoint model are also introduced. Section 3 provides mathematical proofs of the relationships among En4DVar and 4DEnVar algorithms. Section 4 presents results from single-observation experiments for a one-dimensional linear advection model to demonstrate the behaviors and relationships of the algorithms and illustrate the effects of covariance localization. A summary and some further discussion are provided in the conclusions.

2. Four-dimensional ensemble–variational algorithms

a. Hybrid En4DVar with extended control variable

To efficiently introduce flow-dependent ensemble covariance into a variational DA framework and to allow for the combined use of static and ensemble BECs, Lorenc (2003) proposed the so-called extended control variable method, where the typical transformed control variable (vector), $\mathbf{v}$, found in traditional incremental variational cost function (Courtier et al. 1994) is extended by an additional control variable vector $\boldsymbol{\alpha}$, which in the most general case contains N vectors, $\boldsymbol{\alpha}_i$, of dimension n:

$$\boldsymbol{\alpha} = \begin{bmatrix} \boldsymbol{\alpha}_1 \\ \vdots \\ \boldsymbol{\alpha}_N \end{bmatrix}, \tag{1}$$

where N is the ensemble size and n is the dimension of state vector $\mathbf{v}$ or $\mathbf{x}$.

The final analysis increment for state vector $\mathbf{x}$ at the beginning of the 4D assimilation window is then given by

$$\delta\mathbf{x}_0 = \mathbf{x}_0 - \overline{\mathbf{x}}_b = \beta_1\delta\mathbf{x}_1 + \beta_2\delta\mathbf{x}_2 = \beta_1\delta\mathbf{x}_1 + \beta_2\mathbf{X}'_b \circ \mathbf{A} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}, \tag{2}$$

where $\mathbf{X}'_b$ is a matrix whose columns hold the ensemble perturbation vectors $\mathbf{x}'_i = \mathbf{x}_i - \overline{\mathbf{x}}$ normalized by $\sqrt{N-1}$. In the equation, the overbar denotes ensemble mean; subscripts 0 and b denote the initial time of the 4D assimilation window or the analysis time, and the background, respectively; $\delta\mathbf{x}_1$ and $\delta\mathbf{x}_2$ are the analysis increments related to the static BEC $\mathbf{B}$ and the ensemble-derived BEC $\mathbf{P}_e^b$, respectively; and β_1 and β_2 are the weights given to $\delta\mathbf{x}_1$ and $\delta\mathbf{x}_2$ respectively. Both β_1 and β_2 should have values between 0 and 1, and satisfy equation $\beta_1^2 + \beta_2^2 = 1$. When β_1 is 0, the algorithm uses 100% ensemble BEC, and is

called a pure (not hybrid) ensemble–variational method. If β_1 is 1, the algorithm will degenerate to the traditional variational method using static BEC. Here $\mathbf{A}$ is an N-column matrix whose columns are $\boldsymbol{\alpha}_i$ and $\circ$ denotes the Schur product. The incremental form cost function for a hybrid ensemble 4DVar, or En4DVar, can be written as

$$J(\mathbf{x}_1, \boldsymbol{\alpha}) = \frac{1}{2}\delta\mathbf{x}_1^{\mathrm{T}}\mathbf{B}^{-1}\delta\mathbf{x}_1 + \frac{1}{2}\boldsymbol{\alpha}^{\mathrm{T}}\begin{bmatrix} \mathbf{C} & & 0 \\ & \ddots & \\ 0 & & \mathbf{C} \end{bmatrix}^{-1}\boldsymbol{\alpha}$$
$$+ \frac{1}{2}\sum_{t=1}^{I}[\mathbf{H}_t\mathbf{L}_t\delta\mathbf{x}_0 + \mathbf{d}_t]^{\mathrm{T}}\mathbf{R}^{-1}[\mathbf{H}_t\mathbf{L}_t\delta\mathbf{x}_0 + \mathbf{d}_t]. \tag{3}$$

Here I is the total number of time levels at which observations are available, $\mathbf{H}_t$ is the tangent linear observation operator at observation time t, $\mathbf{L}_t$ is the tangent linear model for propagating initial perturbation $\delta\mathbf{x}_0$ to time t, $\mathbf{R}$ is the observation error covariance matrix, $\mathbf{d}_t$ is the observation innovation vector at time t, and $\mathbf{C}$ is a correlation matrix used to localize the ensemble covariance.

To use the precondition technique in the cost function, we define matrix $\mathbf{U}$ that is related to $\mathbf{B}$ by

$$\mathbf{B} = \mathbf{U}\mathbf{U}^{\mathrm{T}}. \tag{4}$$

Then, the analysis increment is

$$\delta\mathbf{x}_0 = \mathbf{x}_0 - \overline{\mathbf{x}}_b = \beta_1\mathbf{U}\mathbf{v} + \beta_2\mathbf{X}'_b \circ \mathbf{A} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}, \tag{5}$$

and the preconditioned incremental form cost function for hybrid En4DVar can be written as

$$J(\mathbf{v}, \boldsymbol{\alpha}) = \frac{1}{2}\mathbf{v}^{\mathrm{T}}\mathbf{v} + \frac{1}{2}\boldsymbol{\alpha}^{\mathrm{T}}\begin{bmatrix} \mathbf{C} & & 0 \\ & \ddots & \\ 0 & & \mathbf{C} \end{bmatrix}^{-1}\boldsymbol{\alpha}$$
$$+ \frac{1}{2}\sum_{t=1}^{I}[\mathbf{H}_t\mathbf{L}_t\delta\mathbf{x}_0 + \mathbf{d}_t]^{\mathrm{T}}\mathbf{R}^{-1}[\mathbf{H}_t\mathbf{L}_t\delta\mathbf{x}_0 + \mathbf{d}_t], \tag{6}$$

where $\delta\mathbf{x}_0$ is a function of $\mathbf{v}$ and $\boldsymbol{\alpha}$. Equation (6) is the same as Eq. (17) of Lorenc (2003), except for the extension into 4D where the observation term contains multiple times when observations are taken within the 4D DA window. Wang et al. (2007) showed, in a 3D framework, that the extended control variable hybrid formulation of the above form is mathematically equivalent to the more straightforward (but computationally more expensive to use) hybrid algorithm of Hamill and Snyder (2000) that explicitly uses the weighted sum of the static and ensemble BECs in the background term of a variational cost function.

Like the variable transformation used to precondition the static background term in a regular variational cost function [i.e., Eq. (1) above], we define, as first proposed by Buehner (2005), a new control variable, $\tilde{\boldsymbol{\alpha}}$, related to $\boldsymbol{\alpha}$ in cost function of Eq. (6) by

$$\boldsymbol{\alpha} = \begin{bmatrix} \mathbf{C}' & & 0 \\ & \ddots & \\ 0 & & \mathbf{C}' \end{bmatrix} \tilde{\boldsymbol{\alpha}}, \qquad (7)$$

or in an expanded form:

$$\begin{bmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \\ \vdots \\ \boldsymbol{\alpha}_N \end{bmatrix} = \begin{bmatrix} \mathbf{C}' & & 0 \\ & \ddots & \\ 0 & & \mathbf{C}' \end{bmatrix} \begin{bmatrix} \tilde{\boldsymbol{\alpha}}_1 \\ \tilde{\boldsymbol{\alpha}}_2 \\ \vdots \\ \tilde{\boldsymbol{\alpha}}_N \end{bmatrix} = \begin{bmatrix} \mathbf{C}'\tilde{\boldsymbol{\alpha}}_1 \\ \mathbf{C}'\tilde{\boldsymbol{\alpha}}_2 \\ \vdots \\ \mathbf{C}'\tilde{\boldsymbol{\alpha}}_N \end{bmatrix}.$$

Here $\mathbf{C}'$ is the decomposed correlation matrix $\mathbf{C}$ and satisfies $\mathbf{C} = \mathbf{C}'\mathbf{C}'^{\mathrm{T}}$. It is, therefore, analogous to the $\mathbf{U}$ in Eq. (1). According to the definition of $\mathbf{A}$,

$$\mathbf{A} = \begin{bmatrix} \mathbf{C}'\tilde{\boldsymbol{\alpha}}_1 & \mathbf{C}'\tilde{\boldsymbol{\alpha}}_2 & \cdots & \mathbf{C}'\tilde{\boldsymbol{\alpha}}_N \end{bmatrix},$$

therefore,

$$\delta\mathbf{x}_0 = \beta_1\mathbf{U}\mathbf{v} + \beta_2\mathbf{X}'_b \circ \mathbf{A} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \beta_1\mathbf{U}\mathbf{v} + \beta_2 \sum_{i=1}^{N} [\mathbf{x}'_{bi} \circ (\mathbf{C}'\tilde{\boldsymbol{\alpha}}_i)]. \qquad (8)$$

With Eqs. (7) and (8), the En4DVar cost function in Eq. (6) can rewritten in terms of the new extended control variable $\tilde{\boldsymbol{\alpha}}$ as

$$J(\mathbf{v}, \tilde{\boldsymbol{\alpha}}) = \frac{1}{2}\mathbf{v}^{\mathrm{T}}\mathbf{v} + \frac{1}{2}\tilde{\boldsymbol{\alpha}}^{\mathrm{T}}\tilde{\boldsymbol{\alpha}} \\ + \frac{1}{2} \sum_{t=1}^{I} \left[\mathbf{H}_t\mathbf{L}_t \left\{ \beta_1\mathbf{U}\mathbf{v} + \beta_2 \sum_{i=1}^{N} [\mathbf{x}'_{bi} \circ (\mathbf{C}'\tilde{\boldsymbol{\alpha}}_i)] \right\} + \mathbf{d}_t \right]^{\mathrm{T}} \mathbf{R}^{-1} \left[\mathbf{H}_t\mathbf{L}_t \left\{ \beta_1\mathbf{U}\mathbf{v} + \beta_2 \sum_{i=1}^{N} [\mathbf{x}'_{bi} \circ (\mathbf{C}'\tilde{\boldsymbol{\alpha}}_i)] \right\} + \mathbf{d}_t \right]. \qquad (9)$$

In Eq. (9), term $\mathbf{C}'\tilde{\boldsymbol{\alpha}}_i$ represents the application of filter $\mathbf{C}'$ on $\tilde{\boldsymbol{\alpha}}_i$, and can be calculated in a similar way as term $\mathbf{U}\mathbf{v}$ using a recursive filter, spectral filter, or wavelet. Note that the correlation matrix $\mathbf{C}$, or its decomposed form $\mathbf{C}'$, which is used to localize the ensemble covariance is general. For univariate analysis without cross-variable covariance, $\mathbf{C}$ is usually a multidiagonal matrix; matrix $\mathbf{C}$ becomes block multidiagonal in the presence of cross-variable covariances. Usually, the

spatial localization function for individual variables is also used between different variables. In other words, the localization function is only a function of spatial distance, and is the same whether it is between different variables or between the same variable at different locations. This means that all blocks in multidiagonal matrix $\mathbf{C}$ are the same.

The gradients of the En4DVar cost function with respect to $\mathbf{v}$ and $\tilde{\boldsymbol{\alpha}}_i$ are given by

$$\nabla_{\mathbf{v}}J(\mathbf{v}) = \mathbf{v} + \beta_1 \sum_{t=1}^{I} \mathbf{U}^{\mathrm{T}}\mathbf{L}_t^{\mathrm{T}}\mathbf{H}_t^{\mathrm{T}}\mathbf{R}^{-1} \left[\mathbf{H}_t\mathbf{L}_t \left\{ \beta_1\mathbf{U}\mathbf{v} + \beta_2 \sum_{i=1}^{N} [\mathbf{x}'_{bi} \circ (\mathbf{C}'\tilde{\boldsymbol{\alpha}}_i)] \right\} + \mathbf{d}_t \right], \qquad (10)$$

$$\nabla_{\tilde{\boldsymbol{\alpha}}_i}J(\tilde{\boldsymbol{\alpha}}_i) = \tilde{\boldsymbol{\alpha}}_i + \beta_2 \sum_{t=1}^{I} \mathbf{C}'^{\mathrm{T}}\mathbf{x}'_{bi} \circ \mathbf{L}_t^{\mathrm{T}}\mathbf{H}_t^{\mathrm{T}}\mathbf{R}^{-1} \left[\mathbf{H}_t\mathbf{L}_t \left\{ \beta_1\mathbf{U}\mathbf{v} + \beta_2 \sum_{i=1}^{N} [\mathbf{x}'_{bi} \circ (\mathbf{C}'\tilde{\boldsymbol{\alpha}}_i)] \right\} + \mathbf{d}_t \right], \qquad (11)$$

where tangent linear model $\mathbf{L}_t$ and adjoint model $\mathbf{L}_t^{\mathrm{T}}$ are needed. Equations (10) and (11) make En4DVar more expensive than the traditional 4DVar because term $\mathbf{L}_t^{\mathrm{T}}\mathbf{H}_t^{\mathrm{T}}\mathbf{R}^{-1}[\mathbf{H}_t\mathbf{L}_t\{\beta_1\mathbf{U}\mathbf{v} + \beta_2\sum_{i=1}^{N}[\mathbf{x}'_{bi} \circ (\mathbf{C}'\tilde{\boldsymbol{\alpha}}_i)]\} + \mathbf{d}_t]$ includes additional calculations related to the ensemble perturbations and a system is needed to provide the ensemble perturbations. As pointed out earlier, this algorithm is close to the traditional 4DVar in terms of the solution procedure but utilizes ensemble-derived BEC, therefore,

it is called En4DVar. With the combined use of static and ensemble-derived BECs, it is called hybrid En4DVar. When only ensemble BEC is used, the algorithm is called pure En4DVar, or just En4DVar.

b. 4DEnVar-NPC

To avoid needing the tangent linear and adjoint models in the ensemble covariance part of the En4DVar algorithm presented above, two approximations can be introduced:

$$\mathbf{L}_t \sum_{i=1}^{N} [\mathbf{x}'_{bi} \circ (\mathbf{C}' \tilde{\boldsymbol{\alpha}}_i)] = \sum_{i=1}^{N} \mathbf{L}_t [\mathbf{x}'_{bi} \circ (\mathbf{C}' \tilde{\boldsymbol{\alpha}}_i)] \approx \sum_{i=1}^{N} (\mathbf{L}_t \mathbf{x}'_{bi}) \circ (\mathbf{C}' \tilde{\boldsymbol{\alpha}}_i),$$
(12)

$$\mathbf{L}_t \mathbf{x}'_{bi} \approx M_t(\mathbf{x}_{bi}) - \overline{M_t(\mathbf{x}_b)} \equiv [M_t(\mathbf{x}_{bi})]', \quad (13)$$

where M_t is the full nonlinear prediction model to take the background state $\mathbf{x}_{bi}$ to time t. With the approximation in Eq. (12), the tangent linear model evolves perturbation states $\mathbf{x}'_{bi}$ instead of $\mathbf{x}'_{bi} \circ (\mathbf{C}' \tilde{\boldsymbol{\alpha}}_i)$ to time t. Equation (7) shows $\mathbf{C}' \tilde{\boldsymbol{\alpha}}_i$ is the control variable $\boldsymbol{\alpha}_i$ in En4DVar. Therefore, in Eq. (12) the localized control variable $\mathbf{C}' \tilde{\boldsymbol{\alpha}}_i$ is not temporally evolved or time propagated so we give this algorithm label NPC, meaning no

time propagation of the control variable. Further, with Eq. (13), the ensemble forecast perturbations $[M_t(\mathbf{x}_{bi})]'$ given by the nonlinear model integrations are used to replace the time propagation of perturbations by the tangent linear model, $\mathbf{L}_t \mathbf{x}'_{bi}$, avoiding the need for developing and maintaining a tangent linear model and the cost of corresponding integrations given that the nonlinear ensemble needs to be run any way. The absence of "allowance for the flow" in the localization in 4DEnVar is also recognized in a very recent paper of Lorenc et al. (2015).

Substituting Eq. (12) into Eq. (9), and assuming $\beta_1 = 0$ (i.e., assuming no contribution from the static BEC) the cost function and its gradient can be written as

$$J(\tilde{\boldsymbol{\alpha}}) = \frac{1}{2} \tilde{\boldsymbol{\alpha}}^{\mathrm{T}} \tilde{\boldsymbol{\alpha}} + \frac{1}{2} \sum_{t=1}^{I} \left[\mathbf{H}_t \sum_{i=1}^{N} \{ [M_t(\mathbf{x}_{bi})]' \circ (\mathbf{C}' \tilde{\boldsymbol{\alpha}}_i) \} + \mathbf{d}_t \right]^{\mathrm{T}} \mathbf{R}^{-1} \left[\mathbf{H}_t \sum_{i=1}^{N} \{ [M_t(\mathbf{x}_{bi})]' \circ (\mathbf{C}' \tilde{\boldsymbol{\alpha}}_i) \} + \mathbf{d}_t \right], \quad (14)$$

$$\nabla_{\tilde{\boldsymbol{\alpha}}_i} J(\tilde{\boldsymbol{\alpha}}_i) = \tilde{\boldsymbol{\alpha}}_i + \sum_{t=1}^{m} \mathbf{C}'^{\mathrm{T}} [M_t(\mathbf{x}_{bi})]'^{\mathrm{T}} \circ \mathbf{H}_t^{\mathrm{T}} \mathbf{R}^{-1} \left[\mathbf{H}_t \sum_{i=1}^{N} \{ [M_t(\mathbf{x}_{bi})]' \circ (\mathbf{C}' \tilde{\boldsymbol{\alpha}}_i) \} + \mathbf{d}_t \right]. \quad (15)$$

In Eqs. (14) and (15), adjoint model and tangent linear model are avoided.[2] With the NPC approximation mentioned earlier, we name the algorithm more specifically as 4DEnVar-NPC, which is consistent with 4DEnVar in Lorenc et al. (2015). We label this algorithm 4DEnVar instead of En4DVar because it does not involve the integration of the tangent linear model or adjoint, and is closer to EnVar than to 4DVar. When all observations are taken at the analysis time, no time integration of the prediction model is involved and the algorithm becomes 3D and is called 3DEnVar.[3] The issue of propagating the control variable also disappears.

If the observation operator is linear and when spatial localization is configured to be the same, the pure 3DEnVar is the same as the ensemble mean analysis from an EnKF using the same set of ensemble perturbations. Note that even though Eq. (13) is considered an approximation to $\mathbf{L}_t \mathbf{x}'_{bi}$, tangent linear model $\mathbf{L}_t$ actually arises from the linearization of the full nonlinear model M when applying the cost function in Eq. (6) so the use of Eq. (13) to evolve the

ensemble perturbations may actually be preferred. In fact, the same practice is done with EnKF and extended Kalman filter (Evensen 1992) algorithms when the full nonlinear model is used to evolve the ensemble perturbations.

c. 4DEnVar

Liu et al. (2008) proposed a 4D ensemble-based variational algorithm [which they originally called En4D-Var but renamed it 4DEnVar in Liu and Xiao (2013)], whose cost function without localization is given by

$$J(\mathbf{w}) = \frac{1}{2} \mathbf{w}^{\mathrm{T}} \mathbf{w} + \frac{1}{2} \sum_{t=1}^{I} [\mathbf{H}_t \mathbf{L}_t \mathbf{X}'_b \mathbf{w} + \mathbf{d}_t]^{\mathrm{T}} \mathbf{R}^{-1} [\mathbf{H}_t \mathbf{L}_t \mathbf{X}'_b \mathbf{w} + \mathbf{d}_t],$$
(16)

where $\mathbf{w}$ is their control variable. The above was obtained by approximating the BEC matrix by $\mathbf{P}_e^b = \mathbf{X}'_b \mathbf{X}'^{\mathrm{T}}_b$ and performing variable transform $\delta \mathbf{x} = \mathbf{X}'_b \mathbf{w}$ starting from the standard 4DVar cost function. The dimension of control vector $\mathbf{w}$ is N, the ensemble size, and the minimization of the cost function is in an N dimensional space. To avoid using the adjoint model in calculating the gradient of the cost function in Eq. (16), perturbation matrix $\mathbf{X}'_b$ is projected to the observation space by tangent linear model and observation operator:

$$\nabla_{\mathbf{w}} J(\mathbf{w}) = \mathbf{w} + \sum_{t=1}^{I} (\mathbf{H}_t \mathbf{L}_t \mathbf{X}'_b)^{\mathrm{T}} \mathbf{R}^{-1} [(\mathbf{H}_t \mathbf{L}_t \mathbf{X}'_b) \mathbf{w} + \mathbf{d}_t]. \quad (17)$$

The $\mathbf{H}_t \mathbf{L}_t \mathbf{X}'_b$ can be approximated by the nonlinear model:

---

[2] Note that when the static BEC term is included, as in Eqs. (9) and (10), the tangent linear model and adjoint model will again be needed even with the approximations.

[3] Unlike the 4D case, for a 3D formulation, the inclusion of the static BEC term in the cost function does not necessitate the use of a tangent linear model or adjoint model, a 3DEnVar can easily include the static BEC term to use hybrid covariance. For this reason, the algorithm can also be called En3DVar.

$$\mathbf{H}_t\mathbf{L}_t\mathbf{X}'_b \approx \frac{1}{\sqrt{N-1}}\Big\{ H_t[M_t(\mathbf{x}_{b1})] - H_t[\overline{M_t(\mathbf{x}_b)}] \quad H_t[M_t(\mathbf{x}_{b2})] - H_t[\overline{M_t(\mathbf{x}_b)}] \quad \ldots \quad H_t[M_t(\mathbf{x}_{bN})] - H_t[\overline{M_t(\mathbf{x}_b)}] \Big\}$$

or component wise as

$$\mathbf{H}_t\mathbf{L}_t\mathbf{x}'_{bi} \approx H_t[M_t(\mathbf{x}_{bi})] - H_t[\overline{M_t(\mathbf{x}_{bi})}], \qquad (18)$$

so that the tangent linear can be avoided.

The above formulations do not include covariance localization, however, which is needed to reduce the sampling error from small ensembles. Liu et al. (2009) introduced a localized perturbation matrix:

$$\mathbf{Z}'_b = [\, \mathbf{S}'_{b1}\circ\mathbf{C}' \quad \mathbf{S}'_{b2}\circ\mathbf{C}' \quad \cdots \quad \mathbf{S}'_{bN}\circ\mathbf{C}' \,], \qquad (19)$$

where $\mathbf{S}'_{bi}$ is a matrix with n columns (n is the state vector length) and every column is $\mathbf{x}'_{bi}$. Variable $\mathbf{C}'$ here is actually the same as the $\mathbf{C}'$ used earlier. The cost function in Eq. (16) can then be written as

$$J(\mathbf{w}) = \frac{1}{2}\mathbf{w}^{\mathrm{T}}\mathbf{w} + \frac{1}{2}\sum_{t=1}^{I}[(\mathbf{H}_t\mathbf{L}_t\mathbf{Z}'_b)\mathbf{w}+\mathbf{d}_t]^{\mathrm{T}}\mathbf{R}^{-1}[(\mathbf{H}_t\mathbf{L}_t\mathbf{Z}'_b)\mathbf{w}+\mathbf{d}_t],$$
$$(20)$$

and its gradient with respect to control variable $\mathbf{w}$ is

$$\nabla_\mathbf{w}J(\mathbf{w}) = \mathbf{w} + \sum_{t=1}^{I}\{(\mathbf{H}_t\mathbf{L}_t\mathbf{Z}'_b)^{\mathrm{T}}\mathbf{R}^{-1}[(\mathbf{H}_t\mathbf{L}_t\mathbf{Z}'_b)\mathbf{w}+\mathbf{d}_t]\}. \qquad (21)$$

Because of the use of $\mathbf{H}_t\mathbf{L}_t\mathbf{Z}'_b$ in place of matrix $\mathbf{H}_t\mathbf{L}_t\mathbf{X}'_b$, $\mathbf{w}$ now has a length of $n \times N$ instead of N so the computational cost would be much increased. After introducing localized perturbation matrix $\mathbf{Z}'_b$, the number of columns of perturbations matrix $\mathbf{X}'_b$ is extended to $n \times N$, which is equivalent to using a huge ensemble by $\mathbf{Z}'_b$. Still, the adjoint model is avoided. It is a 4D variational algorithm using ensemble derived covariance without involving an adjoint model, so it belongs to the class of 4DEnVar and is called 4DEnVar in Liu and Xiao (2013) [but initially called En4DVar in Liu et a. (2008; 2009)]. With this formulation, because $\mathbf{Z}'_b$ includes not only perturbations but also the localization matrix, the localization matrix is propagated by tangent linear model in $\mathbf{L}_t\mathbf{Z}'_b$ and the approximation of Eq. (18) is not applied to Eq. (20).

Similar to the hybrid En4DVar employing the extended control variable, if the static BEC is also included in this algorithm, its cost function and gradient will be

$$J(\mathbf{v},\mathbf{w}) = \frac{1}{2}\mathbf{v}^{\mathrm{T}}\mathbf{v} + \frac{1}{2}\mathbf{w}^{\mathrm{T}}\mathbf{w} + \frac{1}{2}\sum_{t=1}^{I}[\beta_1\mathbf{H}_t\mathbf{L}_t\mathbf{U}\mathbf{v} + \beta_2(\mathbf{H}_t\mathbf{L}_t\mathbf{Z}'_b)\mathbf{w} + \mathbf{d}_t]^{\mathrm{T}}\mathbf{R}^{-1}[\beta_1\mathbf{H}_t\mathbf{L}_t\mathbf{U}\mathbf{v} + \beta_2(\mathbf{H}_t\mathbf{L}_t\mathbf{Z}'_b)\mathbf{w} + \mathbf{d}_t], \qquad (22)$$

$$\nabla_\mathbf{v}J(\mathbf{v},\mathbf{w}) = \mathbf{v} + \sum_{t=1}^{I}\beta_1\mathbf{U}^{\mathrm{T}}\mathbf{L}_t^{\mathrm{T}}\mathbf{H}_t^{\mathrm{T}}\mathbf{R}^{-1}[\beta_1\mathbf{H}_t\mathbf{L}_t\mathbf{U}\mathbf{v} + \beta_2(\mathbf{H}_t\mathbf{L}_t\mathbf{Z}'_b)\mathbf{w} + \mathbf{d}_t], \quad\text{and} \qquad (23)$$

$$\nabla_\mathbf{w}J(\mathbf{v},\mathbf{w}) = \mathbf{w} + \sum_{t=1}^{I}\beta_2(\mathbf{H}_t\mathbf{L}_t\mathbf{Z}'_b)^{\mathrm{T}}\mathbf{R}^{-1}[\beta_1\mathbf{H}_t\mathbf{L}_t\mathbf{U}\mathbf{v} + \beta_2(\mathbf{H}_t\mathbf{L}_t\mathbf{Z}'_b)\mathbf{w} + \mathbf{d}_t]. \qquad (24)$$

This is a hybrid algorithm whose minimization will require adjoint model $\mathbf{L}_t^{\mathrm{T}}$, which appears in Eq. (23). The analysis increment is

$$\delta\mathbf{x}_0 = \beta_1\mathbf{U}\mathbf{v} + \beta_2\mathbf{Z}'_b\mathbf{w}.$$

The relationship of this solution to the extended control variable solution will be further discussed in the next section.

d. 4DEnVar-NPL

Although with Eqs. (20) and (21), the adjoint model is avoided when calculating the cost function and its

gradient, the computational cost of calculating $\mathbf{L}_t\mathbf{Z}'_b$ is still very high. To reduce the cost, we introduce approximation to $\mathbf{L}_t\mathbf{Z}'_b$:

$$\begin{aligned}\mathbf{L}_t\mathbf{Z}'_b &= [\mathbf{L}_t(\mathbf{S}'_{b1}\circ\mathbf{C}') \quad \mathbf{L}_t(\mathbf{S}'_{b2}\circ\mathbf{C}') \quad \cdots \quad \mathbf{L}_t(\mathbf{S}'_{bN}\circ\mathbf{C}')] \\ &\approx [(\mathbf{L}_t\mathbf{S}'_{b1})\circ\mathbf{C}' \quad (\mathbf{L}_t\mathbf{S}'_{b2})\circ\mathbf{C}' \quad \cdots \quad (\mathbf{L}_t\mathbf{S}'_{bN})\circ\mathbf{C}'] \\ &\approx [(M_t\mathbf{S}_{b1})'\circ\mathbf{C}' \quad (M_t\mathbf{S}_{b2})'\circ\mathbf{C}' \quad \cdots \quad (M_t\mathbf{S}_{bN})\circ\mathbf{C}'] \equiv \mathbf{T}'_t, \end{aligned}$$
$$(25)$$

where $\mathbf{S}_{bi}$ is a matrix with n columns (n is the state vector length) and every column is $\mathbf{x}_{bi}$. Because $\mathbf{S}'_{bi}$ is made up

of $\mathbf{x}'_{bi}$, $\mathbf{L}_t\mathbf{S}'_{bi}$ can be evaluated using the ensemble perturbations according to Eq. (13) without separate integrations of the tangent linear model. With Eq. (25), Eq. (20) can then be rewritten as

$$J(\mathbf{w}) = \frac{1}{2}\mathbf{w}^{\mathrm{T}}\mathbf{w} + \frac{1}{2}\sum_{t=1}^{I}[\mathbf{H}_t\mathbf{T}'_t\mathbf{w} + \mathbf{d}_t]^{\mathrm{T}}\mathbf{R}^{-1}[\mathbf{H}_t\mathbf{T}'_t\mathbf{w} + \mathbf{d}_t].$$
(26)

In the original 4DEnVar cost function in Eq. (20), tangent linear model $\mathbf{L}_t$ propagates the localized perturbations, but with approximations made in Eq. (25), the perturbations are first propagated by the tangent linear model before being acted upon by localization matrix $\mathbf{C}'$. This means that $\mathbf{C}'$ or the localization effect is not propagated by the tangent linear model so we call this algorithm 4DEnVar with no propagation of localization matrix or 4DEnVar-NPL. In practice, further computational saving can be achieved by doing EOF decomposition for the correlation matrix and retaining a limited number of modes n' (Liu et al. 2009). In real NWP data assimilation, n' is about several hundred (Liu et al. 2009; Liu and Xiao 2013). Therefore, the control variables $\mathbf{w}$ is reduced to $n' \times N$ from $n \times N$. Also, Eq. (18) can be used to propagate the perturbations so that the tangent linear model can be avoided.

For completeness, the gradient of the cost function in Eq. (26) with respect to $\mathbf{w}$ is provided by

$$\nabla_{\mathbf{w}}J(\mathbf{w}) = \mathbf{w} + \sum_{t=1}^{I}(\mathbf{H}_t\mathbf{T}'_t)^{\mathrm{T}}\mathbf{R}^{-1}[\mathbf{H}_t\mathbf{T}'_t\mathbf{w} + \mathbf{d}_t]. \quad (27)$$

The relationship of this algorithm with other algorithms will be discussed later.

We note there that even though the NPL approximation was used in the implementation of Liu et al. (2009), and Liu and Xiao (2013), who proposed the original 4DEnVar algorithm, the NPL approximation and its implications were not explicitly described there.

e. 4D hybrid schemes with FGAT

Although the 4DEnVar and 4DEnVar-NPC algorithms with ensemble covariance only do not need an adjoint model, when the static BEC is included, Eq. (10) has to be used, which does involve adjoint model $\mathbf{L}_t^{\mathrm{T}}$. One way to avoid the adjoint model is to borrow the idea of "first guess at appropriate time" or FGAT (e.g., Massart et al. 2010) that has been used within the 3DVar framework for better utilizing observations within a time window.

The FGAT formulation effectively neglects tangent linear model $\mathbf{L}_t$ in the standard 4DVar cost function [cf. Eq. (9)], so that the static BEC portion of the cost function,

$$J(\mathbf{v}) = \frac{1}{2}\mathbf{v}^{\mathrm{T}}\mathbf{v} + \frac{1}{2}\sum_{t=1}^{I}[\mathbf{H}_t\mathbf{L}_t\mathbf{U}\mathbf{v} + \mathbf{d}_t]^{\mathrm{T}}\mathbf{R}^{-1}[\mathbf{H}_t\mathbf{L}_t\mathbf{U}\mathbf{v} + \mathbf{d}_t]$$
(28)

becomes

$$J(\mathbf{v}) = \frac{1}{2}\mathbf{v}^{\mathrm{T}}\mathbf{v} + \frac{1}{2}\sum_{t=1}^{I}[\mathbf{H}_t\mathbf{U}\mathbf{v} + \mathbf{d}_t]^{\mathrm{T}}\mathbf{R}^{-1}[\mathbf{H}_t\mathbf{U}\mathbf{v} + \mathbf{d}_t]. \quad (29)$$

It is also equivalent to assuming $\mathbf{L}_t$ is an identity matrix; therefore, the analysis increment associated with control variable $\mathbf{v}$ is not propagated or changed in time within the DA window. In other words, while the observation innovation vectors $\mathbf{d}_t$ are calculated at their appropriate times against the background trajectory (or first guess), the update to the background trajectory is approximated by applying the same analysis increment obtained at the analysis time, rather than by those propagated to the right times using the tangent linear model. This is the essence of FGAT (e.g., Massart et al. 2010). When the only static BEC terms are involved in Eq. (29), the algorithm is more close to 3DVar than to 4DVar so the algorithm can also be called 3DVar-FGAT, as the FGAT implemented in a 3DVar framework is typically called. The neglect of $\mathbf{L}_t$ in the above also removes the implicit time evolution and therefore flow dependency of static background error covariance within the time window.

Using the FGAT treatment as in Eq. (29), static BEC can be included in the 4DEnVar algorithms without requiring adjoint model. The cost functions of the FGAT version of 4DEnVar-NPC, 4DEnVar, and 4DEnVar-NPL are, respectively,

$$J(\mathbf{v}, \tilde{\boldsymbol{\alpha}}) = \frac{1}{2}\mathbf{v}^{\mathrm{T}}\mathbf{v} + \frac{1}{2}\tilde{\boldsymbol{\alpha}}^{\mathrm{T}}\tilde{\boldsymbol{\alpha}}$$
$$+ \frac{1}{2}\sum_{t=1}^{I}\left[\beta_1\mathbf{H}_t\mathbf{U}\mathbf{v} + \beta_2\mathbf{H}_t\sum_{i=1}^{N}\{[M_t(\mathbf{x}_{bi})]' \circ \mathbf{C}'\mathbf{w}_i\} + \mathbf{d}_t\right]^{\mathrm{T}}\mathbf{R}^{-1}\left[\beta_1\mathbf{H}_t\mathbf{U}\mathbf{v} + \beta_2\mathbf{H}_t\sum_{i=1}^{N}\{[M_t(\mathbf{x}_{bi})]' \circ \mathbf{C}'\mathbf{w}_i\} + \mathbf{d}_t\right], \quad (30)$$

$$J(\mathbf{v}, \mathbf{w}) = \frac{1}{2}\mathbf{v}^T\mathbf{v} + \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}\sum_{t=1}^{I}[\beta_1\mathbf{H}_t\mathbf{U}\mathbf{v} + \beta_2\mathbf{H}_t\mathbf{L}_t\mathbf{Z}'_b\mathbf{w} + \mathbf{d}_t]^T\mathbf{R}^{-1}[\beta_1\mathbf{H}_t\mathbf{U}\mathbf{v} + \beta_2\mathbf{H}_t\mathbf{L}_t\mathbf{Z}'_b\mathbf{w} + \mathbf{d}_t], \tag{31}$$

$$J(\mathbf{v}, \mathbf{w}) = \frac{1}{2}\mathbf{v}^T\mathbf{v} + \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}\sum_{t=1}^{I}[\beta_1\mathbf{H}_t\mathbf{U}\mathbf{v} + \beta_2\mathbf{H}_t\mathbf{T}'_t\mathbf{w} + \mathbf{d}_t]^T\mathbf{R}^{-1}[\beta_1\mathbf{H}_t\mathbf{U}\mathbf{v} + \beta_2\mathbf{H}_t\mathbf{T}'_t\mathbf{w} + \mathbf{d}_t]. \tag{32}$$

We call the corresponding algorithms 4DEnVar-NPC-FGAT, 4DEnVar-FGAT, and 4DEnVar-NPL-FGAT. The first and third also avoid the time integration of tangent linear model, and the minimization of the cost functions can be performed by providing an ensemble of nonlinear model background forecasts (in the form of 4D states spanning the time window) and a set of observations within the time window. No further time integration of any form is involved.[4] 4DEnVar-NPC-FGAT and 4DEnVar-NPL-FGAT are, therefore, practical hybrid 4D ensemble–variational algorithms that do not require an adjoint model.

We note there that the use of the FGAT strategy to avoid adjoint model in a hybrid 4DEnVar system is also used in a very recent paper of Lorenc et al. (2015). Such a choice is perhaps not surprising given that a key benefit of 4DEnVar compared to 4DVar or En4DVar is the elimination of the need to develop and maintain an adjoint of a full NWP model. Such an idea was also presented in Liu and Xue (2013). Related to this issue, in the generalized 4D hybrid framework of Desroziers et al. (2014), it was mentioned in passing that the climatological background error could be static in time (within the assimilation window). This is equivalent to making the FGAT approximation but their paper did not directly refer to the FGAT terminology. In Kleist and Ide (2015b), it was also pointed that FGAT-like approximation can be made in an En4DVar system.

f. Temporal localization in 4D ensemble-based algorithms

One aspect that has not been discussed so far is the need for ensemble BEC localization in time, which arises because of the existence of noise in the covariances calculated between ensemble perturbation states of different times due to the limited ensemble size. Temporal localization should in general be applied in all 4D ensemble-based algorithms, and has been used in, for example, the 4D ensemble square root filter algorithm

of S. Wang et al. (2013). Placing the analysis time at the middle of assimilation window (when the algorithm does not involve adjoint model integration) can help somewhat, as was done in Liu et al. (2009).

Temporal localization in 4DEnVar and En4DVar algorithms can be achieved by multiplying the localization matrix $\mathbf{C}$ by temporal localization coefficient ρ_t, as is effectively done in S. Wang et al. (2013), where ρ_t is a function of the time separation between the analysis and observation times. In practice, $\mathbf{C}'$ in Eq. (9) would be replaced by $\sqrt{\rho_t}\mathbf{C}'$. When the separation is larger than the temporal localization radius, the coefficient goes to zero. We note here that Bishop and Hodyss (2007, 2009) proposed adaptive covariance localization algorithms that attempt to address localization issues related to both spatial and temporal separations. These methods are so far still too expensive to implement practically, but do try to address some of the flow-following localization issues raised in our paper.

3. Equivalence among the ensemble–variational algorithms

We show in this section that some of the ensemble–variational algorithms presented above are actually mathematically equivalent.

a. Equivalence of En4DVar and 4DEnVar

From the hybrid 4DEnVar cost function in Eq. (20), $\mathbf{Z}'_b\mathbf{w}$ can be written as

$$\mathbf{Z}'_b\mathbf{w} = (\mathbf{S}'_{b1} \circ \mathbf{C}', \mathbf{S}'_{b2} \circ \mathbf{C}', \cdots, \mathbf{S}'_{bN} \circ \mathbf{C}')\begin{bmatrix}\mathbf{w}_1 \\ \mathbf{w}_2 \\ \vdots \\ \mathbf{w}_N\end{bmatrix}$$

$$= \sum_{i=1}^{N}(\mathbf{S}'_{bi} \circ \mathbf{C}')\mathbf{w}_i$$

$$= \sum_{i=1}^{N}\mathbf{x}'_{bi} \circ (\mathbf{C}'\mathbf{w}_i). \tag{33}$$

The length of $\mathbf{w}_i$ is the same as the dimension of each column of $\mathbf{C}'$. Substituting Eq. (33) into the hybrid 4DEnVar cost function in Eq. (22) gives

---

[4] In principle, one can implement iterative outer loops as with the regular variational algorithms; in that case, additional nonlinear ensemble integrations will be needed for each outer loop iteration.

$$J(\mathbf{v}, \mathbf{w}) = \frac{1}{2}\mathbf{v}^{\mathrm{T}}\mathbf{v} + \frac{1}{2}\mathbf{w}^{\mathrm{T}}\mathbf{w}$$
$$+ \frac{1}{2}\sum_{t=1}^{I}\left[\beta_1 \mathbf{H}_t \mathbf{L}_t \mathbf{U}\mathbf{v} + \beta_2 \mathbf{H}_t \mathbf{L}_t \sum_{i=1}^{N}(\mathbf{x}'_{b1}\circ\mathbf{C}'\mathbf{w}_i) + \mathbf{d}_t\right]^{\mathrm{T}}\mathbf{R}^{-1}\left[\beta_1 \mathbf{H}_t \mathbf{L}_t \mathbf{U}\mathbf{v} + \beta_2 \mathbf{H}_t \mathbf{L}_t \sum_{i=1}^{N}(\mathbf{x}'_{b1}\circ\mathbf{C}'\mathbf{w}_i) + \mathbf{d}_t\right]. \quad (34)$$

Therefore, the hybrid 4DEnVar cost function in Eq. (22) based on the 4DEnVAR originally proposed by Liu et al. (2009) (but with the addition of the static BEC term) is mathematically identical to the hybrid En4DVar cost function in Eq. (9) formulated based on the extended control variable method proposed by Lorenc (2003) (but extended to 4D), with **w** in Eq. (22) being equivalent to $\tilde{\boldsymbol{\alpha}}$ in Eq. (9).

Both schemes apply the full covariance localization without approximation, and the tangent linear model is required for their full implementation. If the static background error covariance is not involved, an adjoint model is not required by 4DEnVar but is still required by En4DVar. The need for an adjoint model due to the static BEC terms can be removed by applying the FGAT approximation, as discussed earlier.

b. Equivalence of 4DEnVar-NPC and 4DEnVar-NPL

In this section, we examine the relationship between 4DEnVar-NPC and 4DEnVar-NPL, which are approximate formulations to En4DVar and 4DEnVar, respectively. From the 4DEnVar-NPC cost function in Eq. (26), $\mathbf{T}'_t\mathbf{w}$ can be written as

$$\mathbf{T}'_t\mathbf{w} = \left[(M_t\mathbf{S}_{b1})'\circ\mathbf{C}' \quad (M_t\mathbf{S}_{b2})'\circ\mathbf{C}' \quad \cdots \quad (M_t\mathbf{S}_{bN})'\circ\mathbf{C}'\right]\begin{bmatrix}\mathbf{w}_1\\\mathbf{w}_2\\\vdots\\\mathbf{w}_N\end{bmatrix} = \sum_{i=1}^{N}[M_t(\mathbf{x}_{bi})]'\circ(\mathbf{C}'\mathbf{w}_i). \quad (35)$$

With Eq. (13), substituting Eq. (35) into the 4DEnVar-NPL cost function in Eq. (22), we obtain

$$J(\mathbf{w}) = \frac{1}{2}\mathbf{w}^{\mathrm{T}}\mathbf{w} + \frac{1}{2}\sum_{t=1}^{I}\left[\mathbf{H}_t\sum_{i=1}^{N}\{[M_t(\mathbf{x}_{bi})]'\circ\mathbf{C}'\mathbf{w}_i\} + \mathbf{d}_t\right]^{\mathrm{T}}\mathbf{R}^{-1}\left[\mathbf{H}_t\sum_{i=1}^{N}\{[M_t(\mathbf{x}_{bi})]'\circ\mathbf{C}'\mathbf{w}_i\} + \mathbf{d}_t\right]. \quad (36)$$

Therefore, the 4DEnVar-NPL cost function in Eq. (26) and the 4DEnVar-NPC cost function in Eq. (14) are identical, with **w** in Eq. (26) being equivalent to $\tilde{\boldsymbol{\alpha}}$ in Eq. (14).

We have therefore proven that 4DEnVar-NPC and 4DEnVar-NPL are mathematically equivalent, and the approximations made in these algorithms to avoid tangent linear and adjoint models are also effectively the same; they both sacrifice the time-evolving or flow-following aspect of the covariance localization.

All algorithms and equivalent proof can be rewritten in a nonpreconditioning variational format, just like the formula for the Gridpoint Statistical Interpolation analysis system (GSI; Wu et al. 2002; Kleist et al. 2009). The equivalent proofs are still the same, which we will not address repeatedly here.

4. Single-observation tests with a one-dimensional linear advection model

The previous section demonstrated through mathematical derivation the equivalence of two groups of ensemble–variational algorithms with and without covariance localization approximations. In this section, we further demonstrate through simple numerical experiments other such equivalences. For all the experiments, the static BEC is excluded, and, therefore, we focus on the treatment of the ensemble covariance and the associated covariance localization effects. For this reason, 4DEnVar-FGAT is not considered here.

The dynamic system is governed by a one-dimensional (1D) linear advection equation

$$\frac{\partial u}{\partial t} + U\frac{\partial u}{\partial x} = 0,$$

with periodic boundary conditions. The initial condition is defined by

$$u_0 = \sin(x).$$

The model domain width is 2π and is discretized uniformly at $2\pi/100$ grid spacing. The default advection speed U is $2\pi/3$ and the time step size is 0.001. We
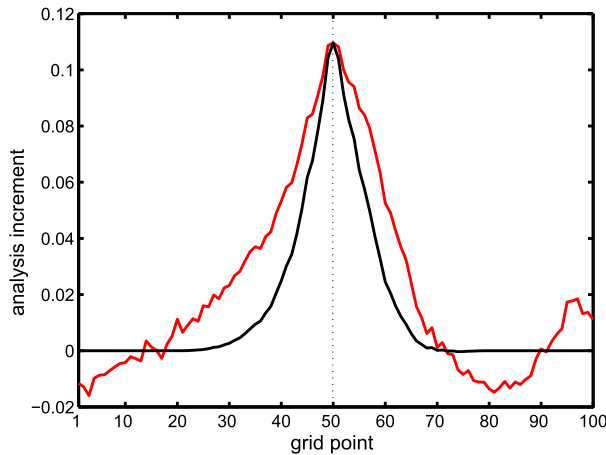
FIG. 1. Analysis increment at the initial time when assimilating a single observation located at gridpoint 50 at the initial time. The thin dotted line denotes the observation location, the red line indicates the analysis increment without localization, and the black line is for the analysis increment with localization.

FIG. 2. Analysis increments at the initial time assimilating a single observation located at the 50th grid point at the ending time, when no localization is used (red line), when using flow-following localization (black line), and when using non-flow-following localization (blue line).

integrate from the initial condition for 160 steps (the ending time) to create a truth simulation. This period is also our 4D DA window. A single observation is taken and assimilated in our DA experiments, and is located at the 50th grid, either at the initial or ending time. The observation error variance is assumed to be 0.01. The initial background state is created by adding one sample of spatially correlated, normally distributed perturbations with an error variance of 0.1 to the truth at the initial time. The observation innovations at the initial and ending times are 0.1. The initial 50-member ensemble was created by adding to this background state 50 realizations of the normally distributed perturbations with the spatial correlation structures and an error variance of 0.1. While the ensemble size of 50 is typical of ensemble data assimilation with NWP models, in this application, 50 is considered large compared to the degree of freedom of the current problem, which is about 50.

The following function is used to construct the spatially correlated random perturbations; it is a compactly supported second-order autoregressive function (Liu and Rabier 2003):

$$\rho(s) = \begin{cases} \left(1 + \dfrac{s}{s_0}\right) e^{-s/s_0} \left(1 - \dfrac{s}{s_1}\right) & s \leq s_1, \\ 0 & s > s_1 \end{cases}$$

where s stands for the distance between two data points, s_0 is the decorrelation scale, and s_1 is the localization radius beyond which the correlation becomes zero. Here, the localization radius is set to be 1.8 and the

decorrelation scale is 0.6. This same function is also used for spatial covariance localization.

The observation at the initial or ending time is assimilated by each of the 4D ensemble–variational methods discussed above. The tangent linear and adjoint models of the advection equation needed by the En4DVar are developed based on the discrete model. For this linear system, the tangent linear model is the same as the finite-difference prediction model. Because all algorithms without localization have the same cost function as in Eq. (16), a reference analysis without localization (Fig. 1) is obtained by minimizing Eq. (16). In all cases, analysis at the initial time (i.e., at the beginning of the DA window) is sought, just like in a standard 4DVar. When the single observation is located at the initial time, there is no need for time integration. Therefore, all algorithms degenerate to 3D algorithms. In this case, it is easy to find that all cost functions are the same assuming that the same spatial localization is applied at the initial time. Figure 1 shows that there is significant noise at places far from the observation when no localization is performed because of the covariance sampling noise. When using a spatial localization, a clean Guassian-shaped analysis increment symmetric around the observation point is obtained (black line in Fig. 1), and the increments obtained by En4DVar and 4DEnVar are the same; they are indistinguishable in Fig. 1.

When the single observation is located at the end of the DA window, the largest analysis increment at the initial time should be found at about (160 × 0.001 × 2π/3)/(2π/100) = 5.33 grid points upstream of the observation location based on the advection speed. Such a solution is rather accurately reproduced by the
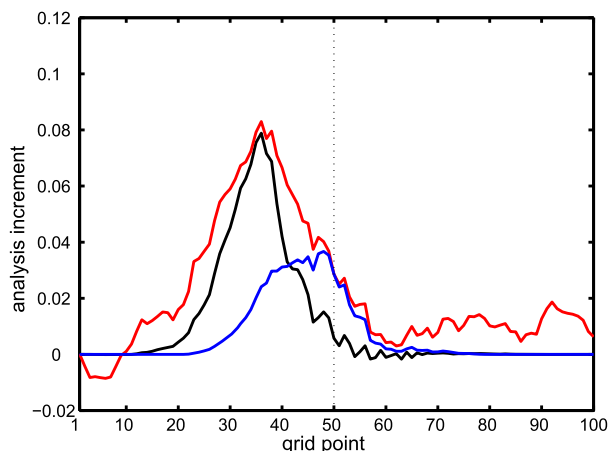
FIG. 3. As in Fig. 2, but for doubling the advection speed to $4\pi/3$.

FIG. 4. As in Fig. 2, but for half of the localized radius and the error spatial decorrelation scale of the initial ensemble perturbations.

algorithms employing full flow-following localization (i.e., En4DVar and 4DEnVar) (black line in Fig. 2); the analysis increment is very similar to that of initial time observation except for the upstream displacement. When no localization is applied, the increment given by Eq. (16) is again broader and contains significant noise away from the observation. When a non-flow-following (or stationary) spatial localization is applied in 4DEnVar-NPC and 4DEnVar-NPL, whose localization function is centered at the 50th grid point and has a localization radius of 1.8, the remote noisy increment is suppressed, but increment maximum is also reduced because of the amplitude-reduction effect of localization away from the center of localization function. Still, because the algorithm is 4D, the peak of the increment is correctly shifted upstream of the observation location, but by a smaller distance than the true solution, again due to the localization effect.

As can be seen, consistent with the mathematical proofs, the analyses of En4DVar and 4DEnVar are the same while the analyses of 4DEnVar-NPC and 4DEnVar-NPL are the same. When the advection speed is lower or the required spatial localization scale is larger, the effects of localization approximations would have smaller effects and vice versa. When the advection speed is doubled in the next set of experiments, the upstream shift of the correct increment peak is also doubled, as given by the flow-following algorithms (black line in Fig. 3). Limited by the non-flow-following localization centered at the 50th grid point, the analysis increment upstream of grid point 40 is severely damped. When the spatial localization scale (as well as the spatial decorrelation scale used by the initial perturbations in our tests) is halved, the analysis increment becomes smaller and the negative effects of the non-flow-following localization become larger (blue line in Fig. 4). In practice, when no better
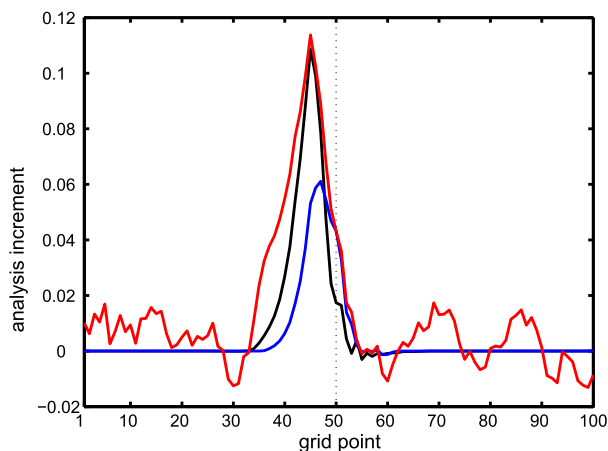
solution is available (i.e., when a flow-following localization algorithm is not available or too expensive to implement), the optimal radius of non-flow-following localization used in the approximated 4D algorithms is expected to be larger than that of the corresponding 3D algorithm in order to better accommodate the propagation effects. Flow-following localization in ensemble data assimilation remains an unsolved problem that requires more research.

5. Summary and discussion

Flow-dependent BECs have been shown to be valuable for improving the quality of state estimation for atmospheric and oceanic as well as other geophysical flows in the past decade. Introducing ensemble-derived flow-dependent BEC into 3D and 4D variational DA frameworks has advantages. Instead of directly calculating the ensemble covariance matrix and using it inside a variational framework, which is computationally impractical for full atmospheric models, Lorenc (2003) proposed an extended control variable approach that introduces the ensemble BEC through an additional extended control variable "background" term in a 3DVar cost function (named En3DVar or 3DEnVar here). The formulation can be easily extended into a 4DVar framework (named En4DVar). The ensemble covariance localization can be achieved by introducing a correlation matrix that "preconditions" the extended control variable and its effect can be achieved by applying a recursive filter. The need to apply tangent linear and adjoint models in En4DVar also carries high computational costs.

An alternative approach for utilizing ensemble BEC within a 4D variational framework was proposed by Liu

TABLE 1. A comparison of four-dimensional ensemble–variational algorithms without a static background error term.

	En4DVar	4DEnVar-NPC	4DEnVar	4DEnVar-NPL
Derivation of cost function formulation	Using α control variable to introduce ensemble covariance into the variational cost function		Projecting ensemble perturbations into observation space to avoid applying an adjoint model to ensemble perturbations of model state. The control variable **w** is effectively the same as the α control variable.	
Covariance localization	Flow-following, can use recursive filter	Non-flow-following, can use recursive filter	Flow-following, use EOF-decomposed correlation matrix	Non-flow-following, use EOF-decomposed correlation matrix
Adjoint model	Needed, integrated every iteration	Not needed	Not needed	Not needed
Tangent linear model	Needed, integrated every iteration	Not needed, using ensemble forecast perturbations	Needed	Not needed, using ensemble forecast perturbations
Computational cost	Comparable to the cost of 4DVar, without counting the cost of ensemble forecasting	Larger than the cost of 3DVar-FGAT because the control variable size is extent to $n \times N$, but far less than 4DVar without counting the cost of ensemble forecasting	Expensive, with EOF approximation, control variables size is $n' \times N$ and need $n' \times N$ ensemble forecasts without approximation	Similar 3DVar-FGAT with the control variable size $n' \times N$, but far less than 4DVar without counting the cost of ensemble forecasting
Relationships	Mathematically equivalent to 4DEnVar	Mathematically equivalent to 4DEnVar-NPL	Mathematically equivalent to En4DVar	Mathematically equivalent to 4DEnVar-NPC

et al. (2008; 2009), which projects the ensemble perturbations to the observation space so that the tangent linear and adjoint models can be avoided. Their original formulations (called 4DEnVar here) did not include the static BEC part in the 4D cost function. Liu et al. (2009) introduced a large localization matrix to modify the ensemble perturbations before they are used so as to achieve ensemble covariance localization, a procedure that is also computationally very expensive.

In this paper, the Liu et al. (2009) formulation is extended to include the static BEC to form a hybrid system. This observation-space-perturbation 4DEnVar formulation is compared with the extended control variable En4DVar formulation. It is shown that before any approximation is made with the localization treatment, the two formulations are mathematically equivalent. The control variable **w** introduced by Liu et al. (2009) is the same as the transformed extended control variable $\tilde{\boldsymbol{\alpha}}$ in the extended control variable formulation.

Approximations are then introduced into the En4DVar algorithm based on the extended control variable so that tangent linear model integrations on the extended control variable are avoided, and the time evolution of the ensemble perturbations are provided by the ensemble forecasts. This approximate formulation is called 4DEnVar-NPC, because there is no propagation of the extended control variable and the algorithm does not inherently require an adjoint model (except when static BEC is included in the hybrid formulation). Approximations which avoid separate integrations on localized

ensemble perturbations are introduced into the original 4DEnVar algorithm, resulting in the approximate 4DEnVar-NPL formulation, where label NPL indicates no propagation of the localization matrix in time. This paper proves that 4DEnVar-NPC and 4DEnVar-NPL are also mathematically equivalent.

All algorithms can include static BEC to form hybrid algorithms but this inclusion would make adjoint model integration necessary. To address this issue, the FGAT approximation is introduced to the static BEC portion of the hybrid En4DVar and 4DEnVar formulations, including the 4DEnVar-NPC, 4DEnVar, and 4DEnVar-NPL algorithms. This approximation avoids the adjoint model while still allowing for the use of observations distributed over a time window in the static portion of the cost function (they can already be used in the ensemble BEC portion), although the formulation is no longer truly 4DVar in the traditional sense. With the FGAT approximation, the static BEC, unlike in the traditional 4DVAR, is no longer implicitly evolved in time or flow following within the assimilation window. A comparison of the pure 4D ensemble–variational algorithms discussed in this paper is given in Table 1.

Single-observation tests for a one-dimensional linear advection system are performed to confirm the mathematical equivalence of the algorithms, and to examine the effects of the localization approximations. When the flow speed is low or the desired BEC localization scale is large, the effects of the non-flow-following localization approximations are smaller and vice versa.

Attempts have been made in the literature (e.g., Bishop and Hodyss 2007, 2009, 2011) to realize flow-dependent covariance localization in an ensemble framework, but all schemes proposed so far are computationally very expensive. Ota et al. (2013) estimated observation impact by a flow-following localization within EnKF framework. However, how to realize effective and efficient flow-following covariance localization in a 4D variational DA framework is still a major and important area for future research, and neglecting the following aspect of covariance localization will remain an important source of approximations before effective solutions are found.

REFERENCES

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

Bishop, C. H., and D. Hodyss, 2007: Flow-adaptive moderation of spurious ensemble correlations and its use in ensemble-based data assimilation. *Quart. J. Roy. Meteor. Soc.*, **133**, 2029–2044, doi:10.1002/qj.169.

——, and ——, 2009: Ensemble covariances adaptively localized with ECO-RAP. Part 1: Tests on simple error models. *Tellus*, **61A**, 84–96, doi:10.1111/j.1600-0870.2008.00371.x.

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

——, 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.

Buehner, M., 2005: Ensemble-derived stationary and flow-dependent background error covariances: Evaluation in a quasi-operation NWP setting. *Quart. J. Roy. Meteor. Soc.*, **131**, 1013–1043, doi:10.1256/qj.04.15.

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

——, J. Morneau, and C. Charette, 2013: Four-dimensional ensemble-variational data assimilation for global deterministic weather prediction. *Nonlinear Processes Geophys.*, **20**, 669–682, doi:10.5194/npg-20-669-2013.

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

Campbell, W. F., C. H. Bishop, and D. Hodyss, 2010: Vertical covariance localization for satellite radiances in ensemble Kalman filters. *Mon. Wea. Rev.*, **138**, 282–290, doi:10.1175/2009MWR3017.1.

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. Thepaut, 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.

Dee, D. P., and S. Uppala, 2009: Variational bias correction of satellite radiance data in the ERA-Interim reanalysis. *Quart. J. Roy. Meteor. Soc.*, **135**, 1830–1841, doi:10.1002/qj.493.

Desroziers, G., J.-T. Camino, and L. Berre, 2014: 4DEnVar: Link with 4D state formulation of variational assimilation and different possible implementations. *Quart. J. Roy. Meteor. Soc.*, **140**, 2097–2110, doi:10.1002/qj.2325.

Evensen, G., 1992: Using the extended Kalman filter with a multilayer quasi-geostrophic ocean model. *J. Geophys. Res.*, **97**, 17 905–17 924, doi:10.1029/92JC01972.

——, 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Montre Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10 143–10 162, doi:10.1029/94JC00572.

——, 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.*, **53**, 343–367, doi:10.1007/s10236-003-0036-9.

Fairbairn, D., S. R. Pring, A. C. Lorenc, and I. Roulstone, 2014: A comparison of 4D-Var with ensemble data assimilation methods. *Quart. J. Roy. Meteor. Soc.*, **140**, 281–294, doi:10.1002/qj.2135.

Fisher, M., 2003: Background error covariance modelling. *Proc. ECMWF Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean*, Reading, United Kingdom, ECMWF, 45–53.

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

Gustafsson, N., 2007: Discussion on '4D-Var or EnKF?' *Tellus*, **59A**, 774–777, doi:10.1111/j.1600-0870.2007.00262.x.

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.

——, J. S. Whitaker, and C. Snyder, 2001: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.*, **129**, 2776–2790, doi:10.1175/1520-0493(2001)129<2776:DDFOBE>2.0.CO;2.

——, ——, M. Fiorino, and S. G. Benjamin, 2011a: Global ensemble predictions of 2009's tropical cyclones initialized with an ensemble Kalman filter. *Mon. Wea. Rev.*, **139**, 668–688, doi:10.1175/2010MWR3456.1.

——, ——, D. T. Kleist, M. Fiorino, and S. G. Benjamin, 2011b: Predictions of 2010's tropical cyclones using the GFS and ensemble-based data assimilation methods. *Mon. Wea. Rev.*, **139**, 3243–3247, doi:10.1175/MWR-D-11-00079.1.

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.

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

——, E. Kostelich, and I. Syzunogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230**, 112–126, doi:10.1016/j.physd.2006.11.008.

Kalnay, E., H. Li, M. Takemsa, S.-C. Yang, and J. Ballabrera-Poy, 2007a: 4DVAR or ensemble Kalman filter. *Tellus*, **59A**, 758–773, doi:10.1111/j.1600-0870.2007.00261.x.

——, ——, ——, ——, and ——, 2007b: Response to the discussion on "4-D-Var or EnKF?" by Nils Gustafsson. *Tellus*, **59A**, 778–780, doi:10.1111/j.1600-0870.2007.00263.x.

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

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

——, D. F. Parrish, J. C. Derber, R. Treadon, W.-S. Wu, and S. Lord, 2009: Introduction of the GSI into the NCEP Global Data Assimilation System. *Wea. Forecasting*, **24**, 1691–1705, doi:10.1175/2009WAF2222201.1.

Kuhl, D., T. E. Rosmond, C. H. Bishop, J. McClay, 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.

Le Dimet, F. X., and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus*, **38A**, 97–110, doi:10.1111/j.1600-0870.1986.tb00459.x.

Liu, C., and Q. Xiao, 2013: An ensemble-based four-dimensional variational data assimilation scheme. Part III: Antarctic applications with advanced research WRF using real data. *Mon. Wea. Rev.*, **141**, 2721–2739, doi:10.1175/MWR-D-12-00130.1.

——, and M. Xue, 2013: A unified framework for four-dimensional ensemble-variational hybrid data assimilation: Relationships among ensemble variational algorithms with full and approximate ensemble covariance localization. *Proc. Sixth WMO Symp. on Data Assimilation*, College Park, MD, WMO.

——, Q. Xiao, and B. Wang, 2008: An ensemble-based four-dimensional variational data assimilation scheme. Part I: Technical formulation and preliminary test. *Mon. Wea. Rev.*, **136**, 3363–3373, doi:10.1175/2008MWR2312.1.

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

Liu, Z., and F. Rabier, 2003: The potential of high-density observations for numerical weather prediction: A study with simulated observations. *Quart. J. Roy. Meteor. Soc.*, **129**, 3013–3035, doi:10.1256/qj.02.170.

Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP—A comparison with 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **129**, 3183–3204, doi:10.1256/qj.02.132.

——, 2013: Recommended nomenclature for EnVar data assimilation methods. Research Activities in Atmospheric and Oceanic Modeling, WGNE, 2 pp. [Available online at http://www.wcrp-climate.org/WGNE/BlueBook/2013/individual-articles/01_Lorenc_Andrew_EnVar_nomenclature.pdf.]

——, and Coauthors, 2000: The Met Office global 3-dimensional variational data assimilation scheme. *Quart. J. Roy. Meteor. Soc.*, **126**, 2991–3012, doi:10.1002/qj.49712657002.

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

Massart, S., B. Pajot, A. Piacentini, and O. Pannekoucke, 2010: On the merits of using a 3D-FGAT assimilation scheme with an outer loop for atmospheric situations governed by transport. *Mon. Wea. Rev.*, **138**, 4509–4522, doi:10.1175/2010MWR3237.1.

Ota, Y., J. C. Derber, E. Kalnay, and T. Miyoshi, 2013: Ensemble-based observation impact estimates using the NCEP GFS. *Tellus*, **65A**, 20038, doi:10.3402/tellusa.v65i0.20038.

Pan, Y., K. Zhu, M. Xue, X. Wang, M. Hu, S. G. Benjamin, S. S. Weygandt, and J. S. Whitaker, 2014: A GSI-based coupled EnSRF–En3DVar hybrid data assimilation system for the operational rapid refresh model: Tests at a reduced resolution. *Mon. Wea. Rev.*, **142**, 3756–3780, doi:10.1175/MWR-D-13-00242.1.

Parrish, D., and J. Derber, 1992: The National Meteorological Center's spectral statistical interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763, doi:10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2.

Purser, R. J., W.-S. Wu, D. F. Parrish, and N. M. Roberts, 2003a: Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: Spatially homogeneous and isotropic Gaussian covariances. *Mon. Wea. Rev.*, **131**, 1524–1535, doi:10.1175//1520-0493(2003)131<1524:NAOTAO>2.0.CO;2.

——, ——, ——, and ——, 2003b: Numerical aspects of the application of recursive filters to variational statistical analysis. Part II: Spatially inhomogeneous and anisotropic general covariances. *Mon. Wea. Rev.*, **131**, 1536–1548, doi:10.1175//2543.1.

Rabier, F., H. Järvinen, E. Klinker, J.-F. Mahfouf, and A. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Quart. J. Roy. Meteor. Soc.*, **126**, 1143–1170, doi:10.1002/qj.49712656415.

Rawlins, F., S. P. Ballard, K. J. Bovis, A. M. Clayton, D. Li, G. W. Inverarity, A. C. Lorenc, and T. J. Payne, 2007: The Met Office global four-dimensional variational data assimilation scheme. *Quart. J. Roy. Meteor. Soc.*, **133**, 347–362, doi:10.1002/qj.32.

Sakov, P., G. Evensen, and L. Bertino, 2010: Asynchronous data assimilation with the EnKF. *Tellus*, **62A**, 24–29, doi:10.1111/j.1600-0870.2009.00417.x.

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

Talagrand, O., and P. Courtier, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. *Quart. J. Roy. Meteor. Soc.*, **113**, 1311–1328, doi:10.1002/qj.49711347812.

Tanguay, M., L. Fillion, E. Lapalme, and M. Lajoie, 2012: Four-dimensional variational data assimilation for the Canadian Regional Deterministic Prediction System. *Mon. Wea. Rev.*, **140**, 1517–1538, doi:10.1175/MWR-D-11-00160.1.

Wang, S., M. Xue, and J. Min, 2013: A four-dimensional asynchronous ensemble square-root filter (4DEnSRF) algorithm and tests with simulated radar data. *Quart. J. Roy. Meteor. Soc.*, **139**, 805–819, doi:10.1002/qj.1987.

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

——, D. Parrish, D. Kleist, and J. Whitaker, 2013: GSI 3DVar-based ensemble–variational hybrid data assimilation for NCEP Global Forecast System: Single-resolution experiments. *Mon. Wea. Rev.*, **141**, 4098–4117, doi:10.1175/MWR-D-12-00141.1.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924, doi:10.1175/1520-0493(2002)130<1913: EDAWPO>2.0.CO;2.

——, ——, X. Wei, Y. Song, and Z. Toth, 2008: Ensemble data assimilation with the NCEP Global Forecast System. *Mon. Wea. Rev.*, **136**, 463–482, doi:10.1175/2007MWR2018.1.

Wu, W.-S., R. J. Purser, and D. F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.*, **130**, 2905–2916, doi:10.1175/1520-0493(2002)130<2905:TDVAWS>2.0.CO;2.

Zhang, M., and F. Zhang, 2012: E4DVar: Coupling an ensemble Kalman filter with four-dimensional variational data assimilation in a limited-area weather prediction model. *Mon. Wea. Rev.*, **140**, 587–600, doi:10.1175/MWR-D-11-00023.1.