

MIT Open Access Articles

An Adjoint-Based Adaptive Ensemble Kalman Filter

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Song, Hajoon, Ibrahim Hoteit, Bruce D. Cornuelle, Xiaodong Luo, and Aneesh C. Subramanian. "An Adjoint-Based Adaptive Ensemble Kalman Filter." *Monthly Weather Review* 141, no. 10 (October 2013): 3343–3359. © 2013 American Meteorological Society

As Published: <http://dx.doi.org/10.1175/MWR-D-12-00244.1>

Publisher: American Meteorological Society

Persistent URL: <http://hdl.handle.net/1721.1/87992>

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of Use: Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



An Adjoint-Based Adaptive Ensemble Kalman Filter

HAJOON SONG*

Ocean Sciences Department, University of California, Santa Cruz, Santa Cruz, California

IBRAHIM HOTEIT

King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

BRUCE D. CORNUELLE

Scripps Institution of Oceanography, University of California, San Diego, San Diego, California

XIAODONG LUO

King Abdullah University of Science and Technology, Thuwal, Saudi Arabia, and International Research Institute of Stavanger, Bergen, Norway

ANEESH C. SUBRAMANIAN

Scripps Institution of Oceanography, University of California, San Diego, San Diego, California

(Manuscript received 21 August 2012, in final form 14 March 2013)

ABSTRACT

A new hybrid ensemble Kalman filter/four-dimensional variational data assimilation (EnKF/4D-VAR) approach is introduced to mitigate background covariance limitations in the EnKF. The work is based on the adaptive EnKF (AEnKF) method, which bears a strong resemblance to the hybrid EnKF/three-dimensional variational data assimilation (3D-VAR) method. In the AEnKF, the representativeness of the EnKF ensemble is regularly enhanced with new members generated after back projection of the EnKF analysis residuals to state space using a 3D-VAR [or optimal interpolation (OI)] scheme with a preselected background covariance matrix. The idea here is to reformulate the transformation of the residuals as a 4D-VAR problem, constraining the new member with model dynamics and the previous observations. This should provide more information for the estimation of the new member and reduce dependence of the AEnKF on the assumed stationary background covariance matrix. This is done by integrating the analysis residuals backward in time with the adjoint model. Numerical experiments are performed with the Lorenz-96 model under different scenarios to test the new approach and to evaluate its performance with respect to the EnKF and the hybrid EnKF/3D-VAR. The new method leads to the least root-mean-square estimation errors as long as the linear assumption guaranteeing the stability of the adjoint model holds. It is also found to be less sensitive to choices of the assimilation system inputs and parameters.

*Current affiliation: Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Corresponding author address: Hajoong Song, Ocean Sciences Department, University of California, Santa Cruz, 1156 High St., Santa Cruz, CA 95064.
E-mail: hsong6@ucsc.edu

1. Introduction

Data assimilation (DA) aims at determining the best possible estimate of the state of a system by combining information from observations and a model forecast according to their respective uncertainties (Ghil and Malanotte-Rizzoli 1991). Techniques based on the four-dimensional variational data assimilation (4D-VAR) approach (Lewis and Derber 1985; Dimet and Talagrand

1986) and the ensemble Kalman filter (EnKF) approach (Evensen 1994; Tippett et al. 2003) are now recognized as the most promising assimilation methods. The 4D-VAR produces the model trajectory that best fits the data over a given period of time by adjusting a set of control parameters. The EnKF optimally blends model outputs and observations according to their respective uncertainties. Continuous progress in computing resources recently enabled the implementation of these methods with state-of-the-art atmospheric and oceanic applications. We refer to Klinker et al. (2000), Houtekamer et al. (2005), Köhl et al. (2007), Carton and Giese (2008), and Hoteit et al. (2010, 2013) to cite a few.

Several studies discussed and compared the strengths and weaknesses of these two approaches (Lorenc 2003; Caya et al. 2005). 4D-VAR methods are mainly known for generating dynamically consistent state estimates within the period of validity of the tangent linear model (Hoteit et al. 2010). Their performance, however, strongly depends on the specification of the background covariance matrix that represents the prior uncertainties about the controls (Weaver et al. 2003). Constructing the background covariance is still the subject of intensive research and various methods have been proposed to model and parameterize this matrix (Parrish and Derber 1992; Daley 1991; Weaver et al. 2003). These assumptions are, however, not always appropriate, and more importantly the resulting background matrix is not flow dependent in the sense that there is still no available efficient variational method to update the background uncertainty in time.

EnKF methods operate sequentially every time a new observation is available. The update of the background covariance matrix in time is carried out through the integration of an ensemble of states representing the uncertainties about the prior (or forecast) with the nonlinear model. Accounting for model deficiencies and use of large ensembles are important factors in obtaining accurate estimates of the background covariance with an EnKF. However, prior knowledge about the nature and the statistics of model uncertainties are generally lacking to properly take into account model errors in the EnKF (Hamill and Whitaker 2005; Hoteit et al. 2007), and computational resources are still lacking for implementing the filters with large ensembles. Using a small ensemble means that the EnKF starts from a prior space that is almost certainly too small and distorted. As the filter proceeds, this space shrinks and can drift even further from truth. Small ensembles also mean rank deficiency and spurious correlations that could prevent the filter's correction from efficiently fitting the observations (Houtekamer and Mitchell 1998; Hamill and Snyder 2000). This problem is often mitigated by covariance

localization that "artificially" increases the effective rank of the background matrix. Strong localization may, however, distort the dynamical balance of the analysis and may lead to a bad forecast (Mitchell et al. 2002). This latter problem has been recently investigated by Kepert (2009).

Recently, the assimilation community has become strongly interested in developing hybrid methods that combine the variational and filtering approaches. The idea is to develop new assimilation schemes that could potentially incorporate the advantages from both approaches. Existing hybrid methods can be basically classified into two main categories; either following the hybrid EnKF/three-dimensional variational data assimilation (3D-VAR) [or ENKF/optimal interpolation (OI)], which augments the EnKF covariance by the stationary background covariance \mathbf{B} of a variational system as a way to reduce the impact of ensemble sampling errors on the EnKF (Hamill and Snyder 2000; Lorenc 2003; Buehner 2005; Wang et al. 2007), or using the flow-dependent covariance matrix of the EnKF as the background matrix for the 4D-VAR problem (Evensen and van Leeuwen 2000; Hunt et al. 2004; Liu et al. 2008; Zhang et al. 2009). For convenience we will refer to these approaches as the 3D and 4D hybrid methods. Augmenting the ensemble by \mathbf{B} in an ad hoc manner in the 3D hybrid approach is not optimal, but might at least improve the EnKF correction by including some of the omitted features in the rank-deficient EnKF covariance. The main problem with 4D hybrid methods is that they require running both the EnKF and 4D-VAR and this can be algorithmically and computationally quite demanding.

In this work we propose a different approach to combine the good features of the EnKF and the 4D-VAR. It is based on a new hybrid scheme that has been recently introduced by Song et al. (2010), called the adaptive ensemble Kalman filter (AEnKF). The idea behind the AEnKF is to adaptively improve the representativeness of the EnKF ensemble by "enriching" it with new members. The new members are generated after every analysis cycle by back projecting the analysis residuals, onto the state space using a 3D-VAR (or OI) assimilation system. The use of information contained in the residuals to enrich the ensemble was already investigated by Ballabrera-Poy et al. (2001) and Lermusiaux (2007) in the context of reduced Kalman filters. Cumulative errors in the EnKF background covariance can be seen by the increase in residuals, and in particular structures in the residuals. These contain information about the missing part of the background covariance that prevented the EnKF from fitting the data, typically model errors and the null space of the ensemble (Song et al. 2010). In contrast

to the 3D hybrid approach, the AEnKF uses the analysis step residuals to target specific directions of the pre-selected background matrix to enrich the EnKF ensemble. This should reduce the addition of unnecessary structures to the EnKF background. Moreover, the EnKF and 3D-VAR analysis steps are applied separately, which offers more numerical and implementation flexibility.

As with the 3D hybrid approach, the AEnKF behavior depends on the stationary background matrix, which is often not well known. The AEnKF was also found to be sensitive to the amount of available observations for efficient reconstruction of the residuals in the state space (Song et al. 2010). Here we further develop the idea of the AEnKF and propose to generate the new members from a 4D-VAR assimilation system. We refer to this approach as the 4D-AEnKF.

4D-VAR analysis lets the model–data misfit choose the descent directions used to fit the data. These vectors are created using the assumed background error covariance. In an analogy to filtering, each member along a descent direction from an iterative fit can be considered to be an adaptive ensemble element. The idea is then to use the 4D-VAR method to transform excessively large residuals left by the incomplete ensemble into descent directions, or new ensemble elements that improve the fit to observations and diversify the ensemble. Reformulating the selection process of the new members as a 4D-VAR problem allows inclusion of more information from the model dynamics and the previous observations, and reduces dependence on the specified stationary background. This would further provide a dynamically consistent new member that is more suitable for forecasting.

The paper is organized as follows. After briefly recalling the characteristics of the AEnKF, we describe the 4D-AEnKF approach in section 2. Results of numerical experiments with the Lorenz-96 model (Lorenz and Emanuel 1998) are then presented and discussed in section 3, followed by a general discussion to conclude in section 4.

2. The 4D adaptive ensemble Kalman filter

a. Review of the adaptive ensemble Kalman filter

The AEnKF is introduced by Song et al. (2010) as an adaptive approach to mitigate the background covariance limitations in the EnKF.

The hypothesis motivating the AEnKF is that null space of the ensemble may grow in time and will manifest itself as increasing residuals. The idea is then to use the residuals to estimate corrections to the model state and use these as new ensemble members. This was

demonstrated to significantly enhance the EnKF performance in the case of small ensembles and the presence of model errors.

The algorithm of the AEnKF is based on that of the EnKF and has the same succession of a forecast step to integrate the analysis ensemble forward in time and an analysis step to correct the ensemble every time a new observation is available. After every analysis step, new members are generated by solving a 3D assimilation problem and then added to the analysis ensemble before a new forecast step takes place. At any time, the state is estimated as the mean of the current filter ensemble.

We present the adaptive approach in the framework of the standard (stochastic) EnKF as described by Evensen (2003). However, the method can be however easily implemented in the framework of any deterministic, or square root EnKF (Tippett et al. 2003; Hoteit et al. 2002). Assuming an ensemble of state estimates is available at an analysis time, the forecast step consists of integrating the ensemble members forward with the model to the time of the next available observation. The forecast ensemble is then corrected using the observation with the Kalman analysis step:

$$\mathbf{x}_i^a = \mathbf{x}_i^f + \mathbf{K}(\mathbf{d}_i - \mathcal{H}\mathbf{x}_i^f), \quad (1)$$

where \mathbf{x}_i^f and \mathbf{x}_i^a are the i th forecast and analysis ensemble members, respectively; \mathbf{d}_i is the observation vector, perturbed with a realization of independent random noise generated from the probability distribution of the observational errors (Burgers et al. 1998); and \mathcal{H} is the observational operator relating the state to the observation. Here \mathbf{K} is the Kalman gain matrix given by

$$\mathbf{K} = \mathbf{P}^{\text{cr}}(\mathbf{P}^{\text{pr}} + \mathbf{R})^{-1}, \quad (2)$$

where \mathbf{P}^{cr} is the sample cross covariance between the background ensemble and its projection on the observation space, \mathbf{P}^{pr} is the sample covariance matrix of the background ensemble projected on the observation space (Evensen 2003), and \mathbf{R} is the observational error covariance matrix.

Instead of adding a stationary covariance matrix \mathbf{B} to the EnKF sample covariance matrix as in the hybrid EnKF/3D-VAR, the AEnKF generates new ensemble members first through the minimization of the cost function of a 3D assimilation problem (OI or 3D-VAR)

$$J(\delta\mathbf{x}) = \frac{1}{2}\delta\mathbf{x}^T\mathbf{B}^{-1}\delta\mathbf{x} + \frac{1}{2}(\mathbf{r} - \mathcal{H}\delta\mathbf{x})^T\mathbf{R}^{-1}(\mathbf{r} - \mathcal{H}\delta\mathbf{x}) \quad (3)$$

with respect to $\delta\mathbf{x}$. The residual vector \mathbf{r} is the difference between the filter analysis \mathbf{x}^a and the observations:

$$\mathbf{r} = \mathbf{d} - \mathcal{H}(\mathbf{x}^a). \tag{4}$$

The analytical solution for (3) is

$$\delta\mathbf{x}^e = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}\mathbf{r}. \tag{5}$$

A new ensemble member $\mathbf{x}^{a,e}$ is then taken as

$$\mathbf{x}^{a,e} = \mathbf{x}^a + \beta\delta\mathbf{x}^e, \tag{6}$$

with β a factor controlling the distance of the new member from the ensemble mean (Song et al. 2010).

Replacing an ensemble member will of course change the ensemble distribution (e.g., mean and covariance), rather than preserving it. This is also true for the hybrid approach where, roughly speaking, a full matrix is added to the ensemble distribution. In the proposed approach, the significance of these changes, which can be easily quantified, depends on the residuals and the specified static background covariance matrix. For instance, the sample mean after adding the new member becomes

$$\tilde{\mathbf{x}} = \frac{1}{N}(\mathbf{x}^a + \beta\delta\mathbf{x}^e) + \frac{1}{N} \sum_{i=1, i \neq k}^N \mathbf{x}_i, \tag{7}$$

$$= \frac{1}{N}(\mathbf{x}^a + \beta\delta\mathbf{x}^e) - \frac{1}{N}(\mathbf{x}^a + \mathbf{x}'_k) + \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i, \quad \text{and} \tag{8}$$

$$= \left(\mathbf{x}^a - \frac{1}{N}\mathbf{x}'_k\right) + \frac{1}{N}\beta\delta\mathbf{x}^e, \tag{9}$$

where N is the ensemble size and \mathbf{x}'_k is the perturbation of the k th ensemble member that is removed. We argue that these changes may be desirable, even necessary, in certain situations where the EnKF background covariance is not well estimated, which could happen from uncertainty omitted from the original ensemble subspace and/or unanticipated model error. In this case, it would make sense to enrich the filter ensemble with new members sampled not from the ensemble distribution, but from its ‘‘complementary part,’’ which we estimate here from the statistics of the residuals back projected into the state space.

Figure 1 illustrates an optimal choice for the value of β in the ideal situation where there is no observation error, and the truth \mathbf{x}^t is known to us. In practice, the above two conditions are not satisfied, therefore, one may have to rely on some ad hoc criterion. In the present work, we set $\beta = 1$ in all our experiments based on the results of Song et al. (2010), who reported that the AEnKF was not strongly sensitive to the value of β in the small ensemble case. Further tuning of this parameter is expected to

State space

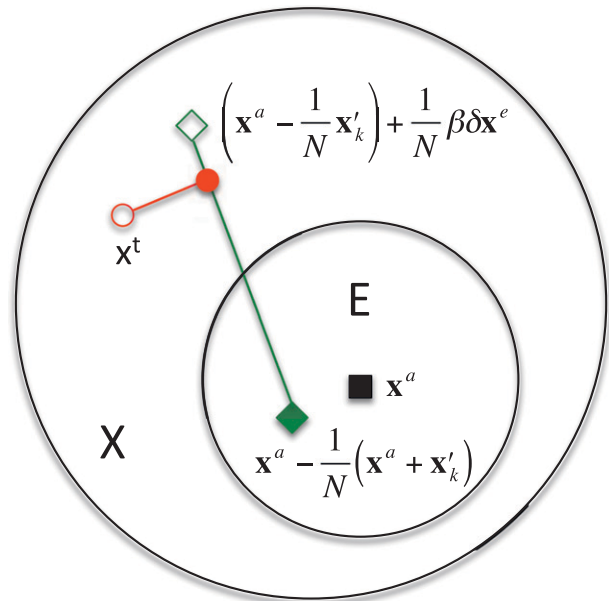


FIG. 1. The position of the ensemble sample mean before and after the new member $\mathbf{x}^{a,e}$ has been added. The original sample mean \mathbf{x}^a (square) and its adjacent point $\mathbf{x}^a - (1/N)(\mathbf{x}^a + \mathbf{x}'_k)$ (closed diamond) after dropping the member $\mathbf{x}_k = \mathbf{x}^a + \mathbf{x}'_k$ are in the subspace E generated by the original ensemble, while the truth \mathbf{x}^t (open circle) lies outside. After adding $\mathbf{x}^{a,e}$, the new sample mean (open diamond) becomes $\tilde{\mathbf{x}} = [\mathbf{x}^a - (1/N)\mathbf{x}'_k] + (1/N)\beta\delta\mathbf{x}^e$. By tuning the coefficient β , one can make the distance $\|\tilde{\mathbf{x}} - \mathbf{x}^t\|_2$ less than $\|\mathbf{x}^a - \mathbf{x}^t\|_2$. The minimum of $\|\tilde{\mathbf{x}} - \mathbf{x}^t\|_2$ is achieved if $\tilde{\mathbf{x}}$ reaches the filled circle, such that $\tilde{\mathbf{x}} - \mathbf{x}^t$ is orthogonal to $\tilde{\mathbf{x}} - \mathbf{x}^a + (1/N)\mathbf{x}'_k$.

improve the behavior of the adaptive scheme. One may certainly use more sophisticated criteria. For instance, similar to Li and Reynolds (2007), one may adopt some iterative search algorithm to choose β such that the norm $\|\mathbf{d} - \mathcal{H}\tilde{\mathbf{x}}\|$ of the new residual, where $\tilde{\mathbf{x}}$ is the mean of the ensemble, is less than that at the previous searching step. This type of iterative searching, however, may incur substantial computational cost, especially in the case of nonlinear observational operators.

More members can be generated from the a posteriori distribution of $\delta\mathbf{x}^e$, or by using the conjugate gradient descent directions in an iteration minimizing J as discussed in Song et al. (2010). To avoid growth of the ensemble from adding new members, some ‘‘old’’ members may be dropped from the ensemble. Here we follow Song et al. (2010) and remove the closest member(s) to the mean after every analysis step. Distances between the ensemble mean and the members were determined by the Euclidean norm normalized by the standard deviations of the model variables computed from a long model run. One can also show, using some simple algebra, that this

choice imposes minimal changes to the mean and the covariance of the original ensemble.

b. The 4D adaptive ensemble Kalman filter

To reduce the dependence of the AEnKF on the specification of \mathbf{B} and to include more information from the model dynamics and previous data in the estimation of the new member, we propose a 4D variational formulation of the AEnKF, that we will call 4D-AEnKF. More specifically, the idea behind the 4D-AEnKF is to replace the 3D cost function in (3) by the following 4D cost function:

$$J_{4D}[\delta\mathbf{x}(t_{i-n})] = \frac{1}{2}\delta\mathbf{x}(t_{i-n})^T\mathbf{B}^{-1}\delta\mathbf{x}(t_{i-n}) + \frac{1}{2}\sum_{j=i-n}^i\alpha_j[\mathbf{r}_j - \mathbf{G}_j\delta\mathbf{x}(t_{i-n})]^T \times \mathbf{R}_j^{-1}[\mathbf{r}_j - \mathbf{G}_j\delta\mathbf{x}(t_{i-n})]. \quad (10)$$

The parameter α_j allows for varying weight of the different time levels, which will be explored in the examples later, where earlier data are not always used in the experiments. Here we followed the incremental formulation of the 4D-VAR approach (Courtier et al. 1994) to define J_{4D} in which the matrix \mathbf{G}_j propagates the perturbation $\delta\mathbf{x}^e(t_{i-n})$ from time t_{i-n} to time t_i in the observation space and is given by

$$\mathbf{G}_j = \mathbf{H}_j\mathbf{M}_{j,i-n}, \quad (11)$$

with $\mathbf{M}_{j,i-n}$ as the tangent linear model of the transition operator $\mathcal{M}_{j,i-n}$ integrating the state between t_{i-n} and t_j . As in (6), the solution $\delta\mathbf{x}^e(t_{i-n})$ of (10) is then added to the analysis $\mathbf{x}^a(t_{i-n})$ as in the AEnKF to form a new member at time t_{i-n} :

$$\mathbf{x}^{a,e}(t_{i-n}) = \mathbf{x}^a(t_{i-n}) + \beta\delta\mathbf{x}^e(t_{i-n}). \quad (12)$$

As discussed in section 2a, the value of β was to set to 1 in all the experiments presented in this study. The member $\mathbf{x}^{a,e}(t_{i-n})$ is next integrated forward in time with the nonlinear model to obtain the new ensemble member $\mathbf{x}^{a,e}(t_i)$ at the current time t_i . As in the AEnKF, the 4D-AEnKF augments the EnKF ensemble with this new member, before starting a new forecast step. The algorithms of the AEnKF and 4D-AEnKF are depicted in Fig. 2.

This 4D formulation of the problem reads as if we are looking for a new member $\delta\mathbf{x}^e(t_{i-n})$ in the past time t_{i-n} (and not at the current time t_i as in the AEnKF) that provides information from the model dynamics and the observations about the part of the correction subspace that was not well captured by the EnKF ensemble in the

n most recent analysis steps. Integrating $\mathbf{x}^{a,e}(t_{i-n})$ forward with the nonlinear model should provide a better and dynamically consistent new member to start the new analysis step.

The analytical solution $\delta\mathbf{x}^e(t_{i-n})$ of (10) is given by (Courtier et al. 1994)

$$\delta\mathbf{x}^e(t_{i-n}) = \left(\mathbf{B}^{-1} + \sum_{j=i-n}^i\alpha_j\mathbf{G}_j^T\mathbf{R}_j^{-1}\mathbf{G}_j\right)^{-1}\sum_{j=i-n}^i\alpha_j\mathbf{G}_j^T\mathbf{R}_j^{-1}\mathbf{r}_j. \quad (13)$$

Computing $\delta\mathbf{x}^e(t_{i-n})$ requires specifying \mathbf{G}_j^T , and therefore the integration of the adjoint model $\mathbf{M}_{i-n,j}^T$. The backward integration of the residuals is therefore only efficient within the period of validity of the tangent linear model (Weaver et al. 2003). Moreover, computing the analytical solution requires the inversion of the matrix $(\mathbf{B}^{-1} + \sum_{j=i-n}^i\alpha_j\mathbf{G}_j^T\mathbf{R}_j^{-1}\mathbf{G}_j)$, which can become numerically problematic when the system dimension is large. In this case, one can either solve this problem by performing the optimization in the observational space (if the number of observations is not too large) as suggested in the representer method (Bennett 2002), or using an iterative optimization algorithm, such as the steepest descent or the conjugate gradient algorithm as it is commonly done in realistic 4D-VAR applications (Fisher 1998; Lorenc 2003). Both of these methods use the adjoint model to compute the gradients of the cost function with respect to the control.

Note that it is possible to generate more than one member in the 4D-AEnKF after every analysis step following similar ideas to those presented in the AEnKF (Song et al. 2010). For instance, several descent directions could be used during the optimization of (10) if desired. The 4DVAR has no error estimates, so the weighting of the descent directions in the ensemble is heuristically chosen. Another way would be to sample several realizations from the distribution of the residuals as described in Song et al. (2010) and then integrate these backward in time with the adjoint before integrating them forward with the model to generate the new members (exactly as it is done for one member). This obviously could become computationally demanding as each new ensemble member would require forward and backward integrations of the dynamical model and its adjoint.

3. Numerical experiments

a. Model description and settings

We use the Lorenz-96 (L96) model (Lorenz and Emanuel 1998) to test and evaluate the behavior of the 4D-AEnKF and to compare its performance to the

Adaptive EnKF (AEnKF)

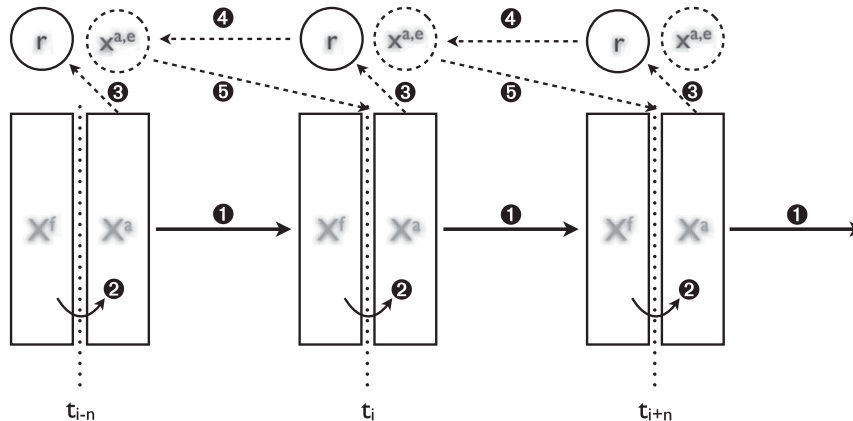
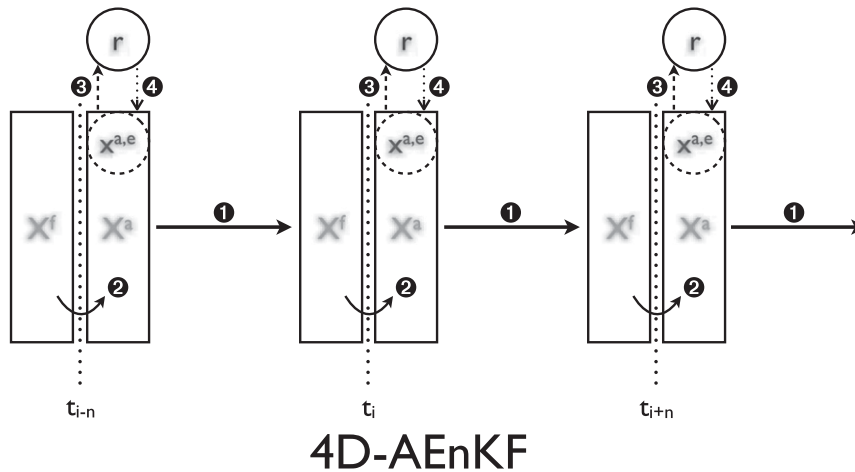


FIG. 2. Diagram describing the algorithms of (top) the AEnKF and (bottom) the 4D-AEnKF; \mathbf{X}^f and \mathbf{X}^a denote the forecast ensemble and analysis ensemble, respectively. In the AEnKF, the ensemble members are first integrated forward with the model (1), then updated with incoming observations (2), exactly as in the EnKF. The residual \mathbf{r} at the analysis time is computed (3) before getting back projected into a new ensemble member in the state space (4). A new filtering cycle then begins. The 4D-AEnKF follows a very similar procedure, except that the residual computed at the analysis time is integrated backward with the adjoint model (4) before getting integrated forward with the model (5) to generate a new ensemble member.

EnKF, AEnKF, and the hybrid EnKF/3D-VAR. This model is widely used in the assimilation community to test EnKF-based schemes and provides a repeatable performance benchmark (Whitaker and Hamill 2002). L96 was designed to mimic the evolution of an atmospheric variable in time and is described by the following highly nonlinear set of differential equations $j = 1, 2, \dots, L$:

$$\frac{dx(j, t)}{dt} = x(j-1, t)[x(j+1, t) - x(j-2, t)] - x(j, t) + F. \quad (14)$$

The model was implemented here in its most common form. We therefore considered $L = 40$ variables, forcing term $F = 8$, and periodic boundary conditions, i.e., $x(-1, t) = x(L-1, t)$, $x(0, t) = x(L, t)$, and $x(L+1, t) = x(1, t)$. For $F = 8$, disturbances propagate from low to high indices (from “west” to “east”) and the model behaves chaotically (Lorenz and Emanuel 1998). L96 and its tangent linear model (and adjoint) were discretized using a Runge–Kutta fourth-order integration (Sandu 2006) with a time step $t = 0.05$, which corresponds to 6 h in real-world time.

We follow Song et al. (2010) to generate the filter initial conditions and the back-projection matrix \mathbf{B} . The model was first integrated forward without assimilation for a time period equivalent to several years in real-world time. The mean and the covariance from the run were used for the initial ensemble mean and the matrix \mathbf{B} , respectively. The starting ensemble members were generated by adding independent Gaussian random perturbations with unit variance to each model variable.

All tested assimilation schemes were implemented with covariance inflation and covariance localization using the Gaspari–Cohn fifth-order correlation function as described by Whitaker and Hamill (2002). For a given length scale, the correlation between two grid points becomes zero if the distance between those points is greater than twice of the length scale (Hamill et al. 2001). It is important to notice here that the proposed adaptive scheme should also increase the spread of the ensemble because the members that are closest to the ensemble mean, which contribute the least to the ensemble spread, are replaced by new members that presumably account for model and undersampling errors. There are other EnKF formulations that do not use multiplicative covariance inflation and/or do not require covariance inflation (Houtekamer and Mitchell 1998; Houtekamer et al. 2009; Bocquet 2011).

In the hybrid scheme, the estimated background error covariance $\tilde{\mathbf{P}}$ is defined as

$$\tilde{\mathbf{P}} = (1 - \gamma)\mathbf{P} + \gamma\mathbf{B}. \quad (15)$$

After a series of tests in which $\gamma \in \{0, 0.1, 0.2, \dots, 0.9, 1\}$, we used $\gamma = 0.1$ as it yielded the smallest root-mean-squared error (RMSE). Assimilation experiments were carried out in the presence of undersampling and model errors by using relatively small ensembles with 10 members, and incorrect forcing $F = 6$ in the forecast model, respectively. Note that the forcing term is constant in the Lorenz-96 model and therefore does not appear in the adjoint model. This means that the adjoint model is the same for the true and perturbed models. This partly explains why the new member carried appropriate information about the model dynamics despite the introduction of an important forcing error in the forecast model. In this study, we considered two scenarios to test the filters performances: experiments including only sampling error, and a more general case including both sampling and model errors. An additional experiment was also performed under only model error, similar to Song et al. (2010). The results from this experiment are consistent with those published in Song et al. (2010) showing the improved filtering performance by the adaptive methods (not shown). The proposed

adaptive schemes were also implemented with the ensemble transform Kalman filter (ETKF) under the same scenarios as the EnKF and similar improvements to those reported here were obtained.

Observations were sampled every four time steps (which is equivalent to 1 day in real-world time) instead of every time step as is typically considered in assimilation/filtering studies with the L96 model to test the filters in more typical and challenging situations when data are not available every model time step for assimilation. The filters were evaluated under three different sampling strategies in which the observations were considered available for all, half, and quarter of the model variables using constant sampling intervals. Assimilation experiments were performed over a period of 1115 days (or 4460 model steps), but only the last 3 years were considered in the analysis of the results after excluding an early spinup period of about 20 days. For a given inflation factor and covariance localization length scale, each filter run was repeated 10 times, each with randomly drawn initial ensemble and observational errors, and the average RMSEs over these 10 runs were reported to reduce statistical fluctuations. Several longer (more than 100 000 model steps) assimilation runs were also performed to test the impact of the adaptive schemes on the long behavior of the EnKF. The resulting RMSEs from these runs are very close (within 1%) to those obtained with the 3-yr runs.

As discussed in section 2b, the proposed 4D-AEnKF scheme would only be efficient if the residuals were integrated backward within the valid period of the tangent linear model. Below we describe the tangent linear model validation test and study its behavior.

b. Validation of the tangent linear model assumption

For a given state vector \mathbf{x} and a perturbation \mathbf{x}' , we first integrate the two state vectors \mathbf{x} and $\mathbf{x} + \mathbf{x}'$ with the nonlinear model \mathcal{M} forward in time, and compute the difference $\Delta\mathbf{x} = \mathcal{M}(\mathbf{x} + \mathbf{x}') - \mathcal{M}(\mathbf{x})$, which measures the time evolution of perturbation in the nonlinear system. We also integrate the perturbation \mathbf{x}' with the tangent linear model \mathbf{M} . Then the difference between the outputs of these runs, $d\mathbf{x} = \Delta\mathbf{x} - \mathbf{M}\mathbf{x}'$, represents the nonlinear terms associated with the same perturbation. A measure of the growth of the nonlinearities during the integration period can be obtained by taking the ratio ρ between the length of $d\mathbf{x}$ and $\Delta\mathbf{x}$:

$$\rho = \frac{\sqrt{d\mathbf{x}^T d\mathbf{x}}}{\sqrt{\Delta\mathbf{x}^T \Delta\mathbf{x}}}. \quad (16)$$

The ratio ρ is zero if the system is linear. If ρ is greater than 1, this indicates that the nonlinear part strongly affects the perturbation growth.

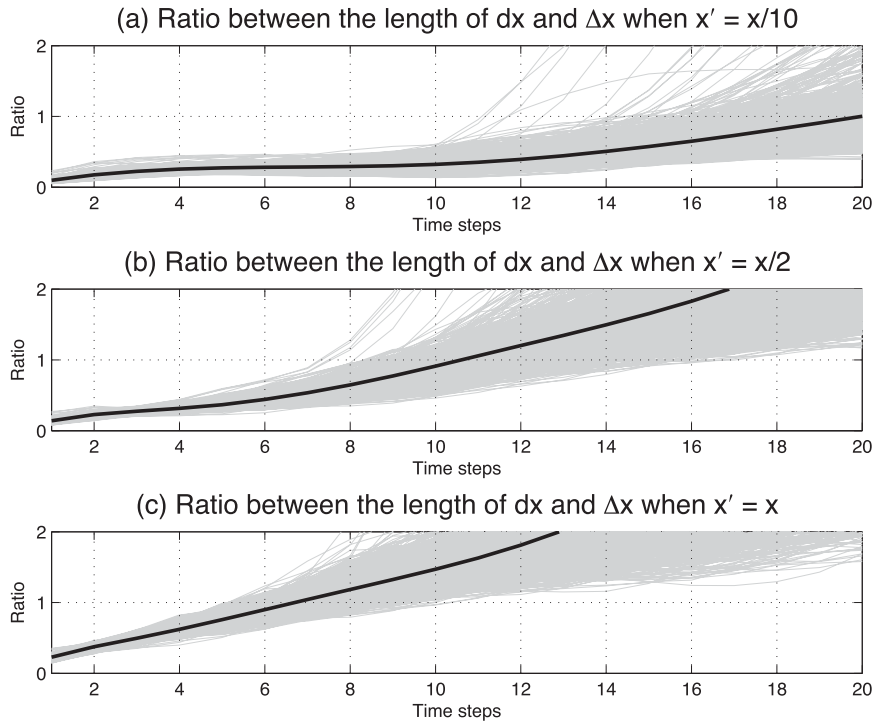


FIG. 3. Ratio between the length of dx and Δx in time, where Δx and dx represent the difference between two nonlinear trajectories started from \mathbf{x} and $\mathbf{x} + \mathbf{x}'$, and the time evolution of the nonlinear part for \mathbf{x}' . Hence, the ratio shows the growth of the nonlinear part with respect to the growth of the perturbation in time. Three different initial perturbations were chosen for testing: (a) $\mathbf{x}' = \mathbf{x}/10$, (b) $\mathbf{x}' = \mathbf{x}/2$, and (c) $\mathbf{x}' = \mathbf{x}$. The results are plotted in black lines with variability computed from 1000 realizations.

Figure 3 plots the ratio ρ as it results from runs with three different perturbation sizes $\mathbf{x}' = \mathbf{x}/10$, $\mathbf{x}' = \mathbf{x}/2$, and $\mathbf{x}' = \mathbf{x}$. A total of 1000 runs were performed with different initialization for each perturbation size to reduce statistical fluctuations and determine variability, and these and their mean are plotted in gray and black, respectively, in Fig. 3. When the size of the perturbation is relatively small ($\mathbf{x}' = \mathbf{x}/10$), the mean value of the ratio ρ remains less than 0.5 after 18 time integration steps (about 4.5 days) as can be seen in Fig. 3a. As expected, the ratio grows faster with larger perturbations, and for $\mathbf{x}' = \mathbf{x}/2$, it becomes close to 1 after only 10 time steps (Fig. 3b). The nonlinear part becomes even more significant if the perturbation is of the same size as the state (Fig. 3c). However, considering the variability, and assuming that the residuals are usually smaller than the states, which is a reasonable assumption for a well-behaved assimilation system, one can assume that the linear assumption remains valid for at least four time steps (1 day), which is the observation frequency in our experiments. Based on these results and unless specified otherwise, we created the new member after integrating the residuals backward with the adjoint model for four time steps.

c. 4D-AEnKF versus EnKF, AEnKF, and 3D-hybrid

1) CASE WITH ONLY SAMPLING ERROR

In the first experiment we study the behavior of the 4D-AEnKF and evaluate its performance against those of the EnKF, AEnKF, and 3D-hybrid in the presence of only sampling errors. The filters use the same forecast model as the true mode, so no model errors were included. The ensemble size was set to 10 and the new member in the 4D-AEnKF was estimated at the previous analysis time step after integrating the adjoint model backward for four time steps. Two different implementations of the 4D-AEnKF were tested, using or not the observations at time t_{i-4} in the 4D cost function of (10) of the 4D-AEnKF (i.e., setting α_{i-4} to either 1 or 0). In the following we refer to the 4D-AEnKF that does not include the previous data at time t_{i-4} as AD-AEnKF.

Hybrid methods generally reduce the sensitivity of the filter to the inflation factor and error covariance localization length scale as can be seen from Fig. 4. This is consistent with our above discussion about the adaptive scheme also introduces some sort of inflation to the background covariance. The minimum RMSE values of

the EnKF and the hybrid methods are generally similar. In the full observation case, the minimum RMSE of the EnKF is smaller than the 3D-hybrid and AD-AEnKF, although their minimum RMSE can be reduced by optimizing their parameters γ and β , respectively.

When fewer observations are assimilated into the system, the AD-AEnKF and 4D-AEnKF provide the lower RMSEs compared to the other filters. They also reduce the sensitivity of the filters to the tuning parameters. Interestingly, the rate of the RMSE reduction by adjoint-based adaptive filters is the greatest when only a quarter of the variables are observed. This suggests that the model dynamics can enrich the ensemble by spreading the information in the residual from the observed variables to the unobserved variables in a dynamically consistent manner.

2) GENERAL CASE WITH SAMPLING AND MODEL ERRORS

In this experiment we study the behavior of the 4D-AEnKF in the presence of both sampling and model errors. As in the previous experiment, the ensemble size was set to 10. The model error is introduced by setting $F = 6$ in the filters forecast model. The minimum RMSE of the EnKF was obtained using a larger range of inflation factors than those tested in the previous experiment. We therefore extended the range of tested inflation values.

To reduce the number of experiments and save computing time, we evaluated the RMSE of the filters varying the inflation factor with different upper values (2, 1.5, and 1.25, respectively) for the all, half, and quarter observations scenarios. These ranges should be sufficient to understand the general behavior of the different filters, including also the minimum RMSE (best performance) of each filter.

Figure 5 plots the RMSE resulting from the different filtering schemes with the three observation scenarios as a function of inflation factor and covariance localization length scale. All hybrid methods generally improve upon the performance of the EnKF. In the full observations case scenario, the EnKF with a well-tuned inflation factor and localization length scale can perform as well as the hybrid schemes. However, the hybrid schemes are more robust for nonoptimal values of inflation factor and localization length scale. In all cases, the lowest RMSE is obtained with the 4D-AEnKF. More tuning of the weighting factor β is also expected to further improve the performance of the 4D-AEnKF.

Both AD-AEnKF and 4D-AEnKF have generally lower RMSE than the AEnKF. The results also suggest that the proposed adaptive schemes clearly enhance the robustness of the EnKF, especially in the dense

observation scenarios. In the quarter observation case, the robustness of the EnKF is not much improved, though the minimum RMSEs are still obtained with the 4D-AEnKF. More observation means more information for both faster reduction of ensemble spread and better estimation of a new member. In all observation scenarios, the adaptive filters also reached their minimum RMSE at lower inflation values than the EnKF and the hybrid EnKF/3D-VAR, and their RMSEs are generally better than those obtained with the EnKF RMSE with larger inflation factors. This supports the claim that the new members also improve the ensemble spread. Overall and as expected, including the data at t_{i-4} in the generation of the new member are beneficial for the adaptive schemes, so that the best performances were obtained with the 4D-AEnKF.

To analyze the impact of the newly generated member on the distribution of the ensemble of the 4D-AEnKF, the time evolution of the first model state variable $\mathbf{x}(1, t)$ is shown for the 4D-AEnKF ensemble members over a 21-day period (between days 70 and 90) in Fig. 6 before (forecast) and after (analysis) applying the correction step. Plots are shown for the case where observations of all model variables were assimilated, localization length scale 10.95, and inflation factor 1.01. Black dots in Fig. 6a represent the position of the new members after they have been integrated from the previous analysis time to the current time. White dots in Fig. 6b indicate the positions of the new members. Following the algorithm of the 4D-AEnKF, white dots in Fig. 6b at day t_{i-4} are integrated with the L96 model and become the black dots in Fig. 6a at day t_i . The plots show good examples of how the behavior of the EnKF can be improved by the new member created using the adjoint model. For instance, at day 76, all ensemble members but the new one are located around the value 7, while the true state is close to 10 (Fig. 6a). The new member, which has been integrated from days 75 to 76, has a value that is close to the true state. This new member was created one day earlier (white dot on day 75 in Fig. 6b) such that the ensemble forecast better represents the distribution of the forecast state. As a result, the new member increases the uncertainty and brings the ensemble and the analysis ensemble mean at day 76 closer to the true state (Fig. 6b). Another time where improvement is clear is day 79. The integrated new member is closer to the true state than the other ensemble members, resulting in a better analysis. Although the newly created member at day 78 is in fact farther away from the true state at the previous time, it was generated so as to improve the forecast at the next filtering step.

The reliability of the filter ensemble can be assessed using the rank histogram (Hamill and Snyder 2000).

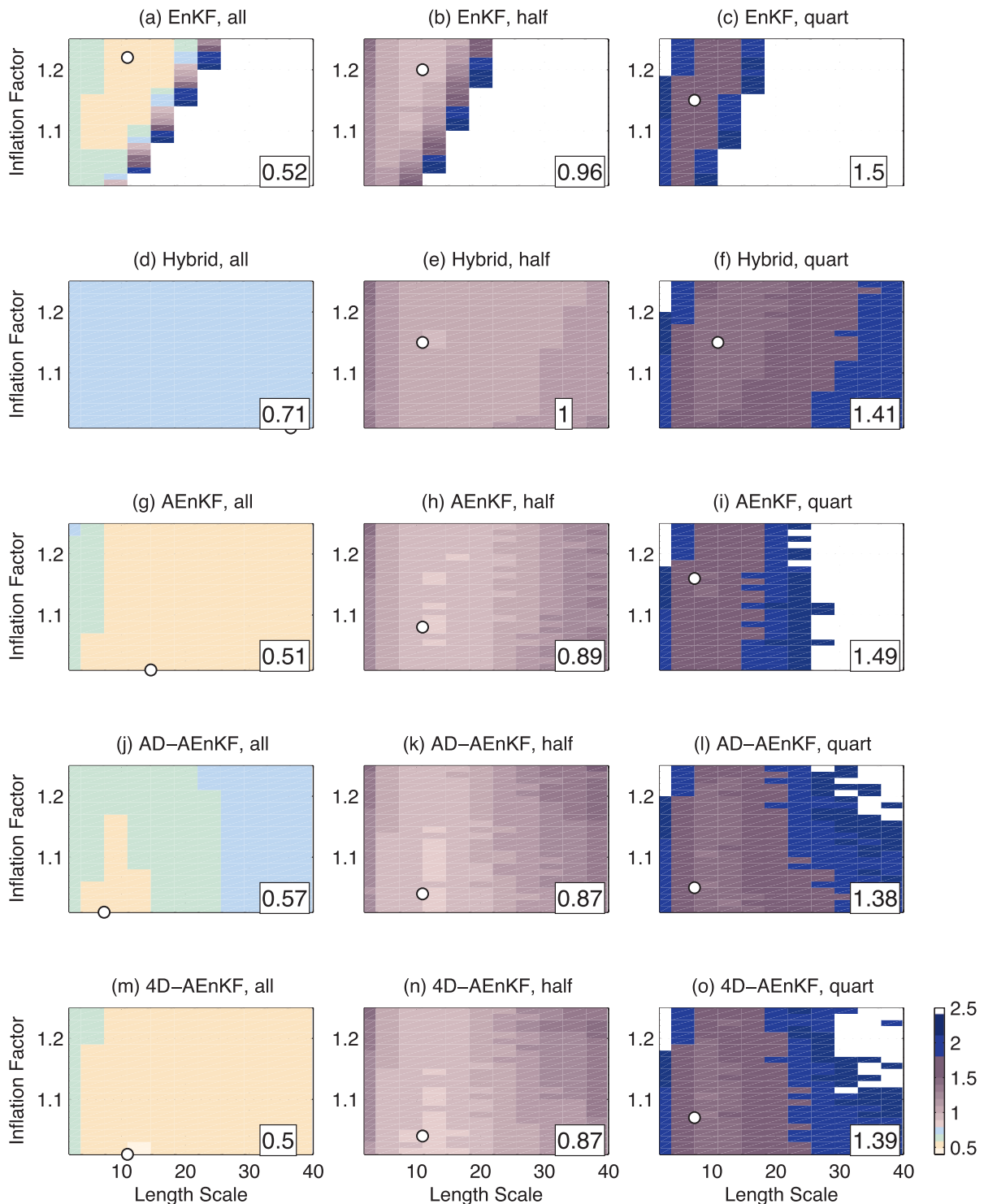


FIG. 4. RMSE averaged over time and all variables as a function of inflation factor and covariance localization length scale from (a)–(c) EnKF, (d)–(f) hybrid 3DVAR/EnKF, (g)–(i) AEnKF, (j)–(l) AD-AEnKF, and (m)–(o) 4D-AEnKF. Only sampling error was introduced. The AD-AEnKF and 4D-AEnKF both run four backward time steps to create a new member, but the AD-AEnKF uses the residual at time t_i while the 4D-AEnKF uses the residuals at time t_i and t_{i-4} . All filters assimilated the observations from the three different strategies. The RMSE from the filters assimilating (left) all, (middle) half, and (right) quarter of variables. White dots in each panel indicate the location of minimum RMSE with the value of the minimum shown in the bottom-right corner.

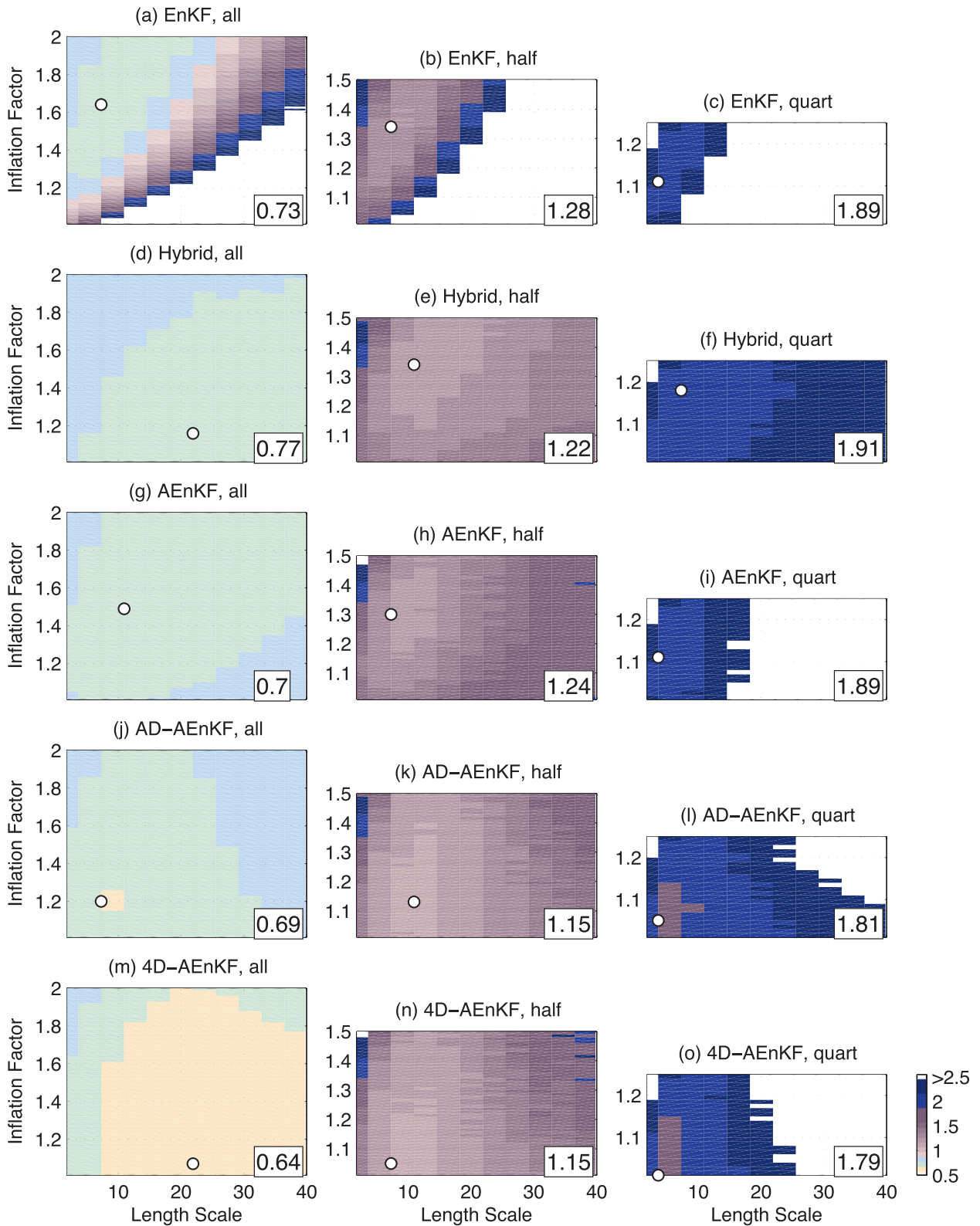


FIG. 5. As in Fig. 4, but both sampling and model errors were introduced. The maximum values of inflation factors for all, half, and quarter observations scenarios are 2.0, 1.50, and 1.25, respectively.

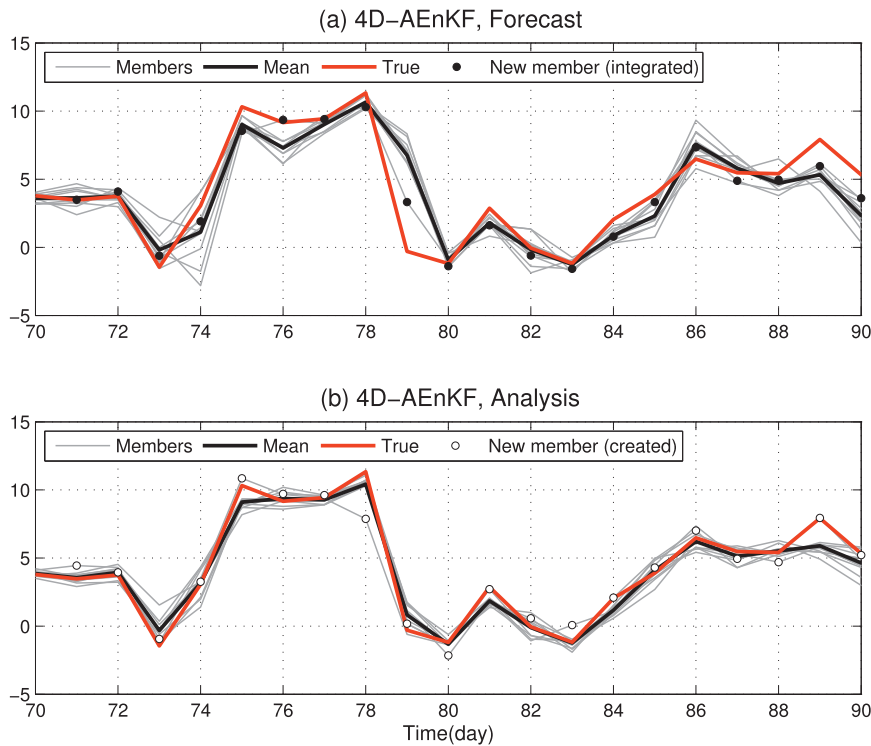


FIG. 6. Time evolution of the first model variable $x(1, t)$ between days 70 and 90 in the true state (red line), the mean of ensemble members or filter's estimates (black line), and 10 ensemble members (gray lines), as it results from (a) 4D-AEnKF forecast and (b) 4D-AEnKF analysis with 1.01 inflation factor and 10.95 error covariance localization length scale. Closed circles in (a) indicate the positions of the new members that were created one day earlier [marked as open circles in (b)] before integrating them with the model to the analysis time.

Ideally, the value of a truth has an equal chance to occur in any of the possible ranks relative to the sorted ensemble (from low to high). Over many samples, this is reflected by a flat histogram. Nonuniformity in rank histograms usually suggests potential problems in the ensemble. For example, an ensemble with insufficient spread or biased state will have a rank histogram with higher values at one or both edges (U shaped) while an ensemble with excessive spread will have a rank histogram with low values at the edges (Anderson 1996; Hamill and Snyder 2000).

The rank histograms from the EnKF, hybrid EnKF/3D-VAR, AEnKF, AD-AEnKF, and 4D-AEnKF are shown in Fig. 7 for the three different observation scenarios and the combination of localization length scale and inflation factor that yield the best-state estimates for each filtering scheme. One can first notice that the rank histograms are generally tilted to the right, probably associated with the bias in the forecast model used in assimilation resulting from incorrect (biased) forcing. For all filters, the slopes of the rank histograms show a tendency to increase as fewer observations are

assimilated. The rank histogram of the hybrid EnKF/3D-VAR is not very different from that of the EnKF. The rank histograms of the EnKF with the adaptive schemes are flatter. This suggests that the adaptive schemes reduce the impact of the forcing bias on the filter ensemble. The more the residuals are integrated backward and data are included, the less the impact of the forcing bias is visible. The 4D-AEnKF has small ensemble variance likely because it constrains the new ensemble member with more data from the previous analysis time, and because of the power of the 4D-VAR scheme in fitting the data (here the residuals). However, the 4D-AEnKF still provides the best-state estimate in the experiments in terms of lowest RMSE. The large bars at the edges of the rank histograms might suggest occasional existence of "outliers" likely caused by the large weight carried by the one, and only, member that we used to represent the null space of the filter ensemble. These should become more consistent with the rest of the ensemble members after they are integrated with the model dynamics during the forecast steps, and as more data become available in time. Sampling of more

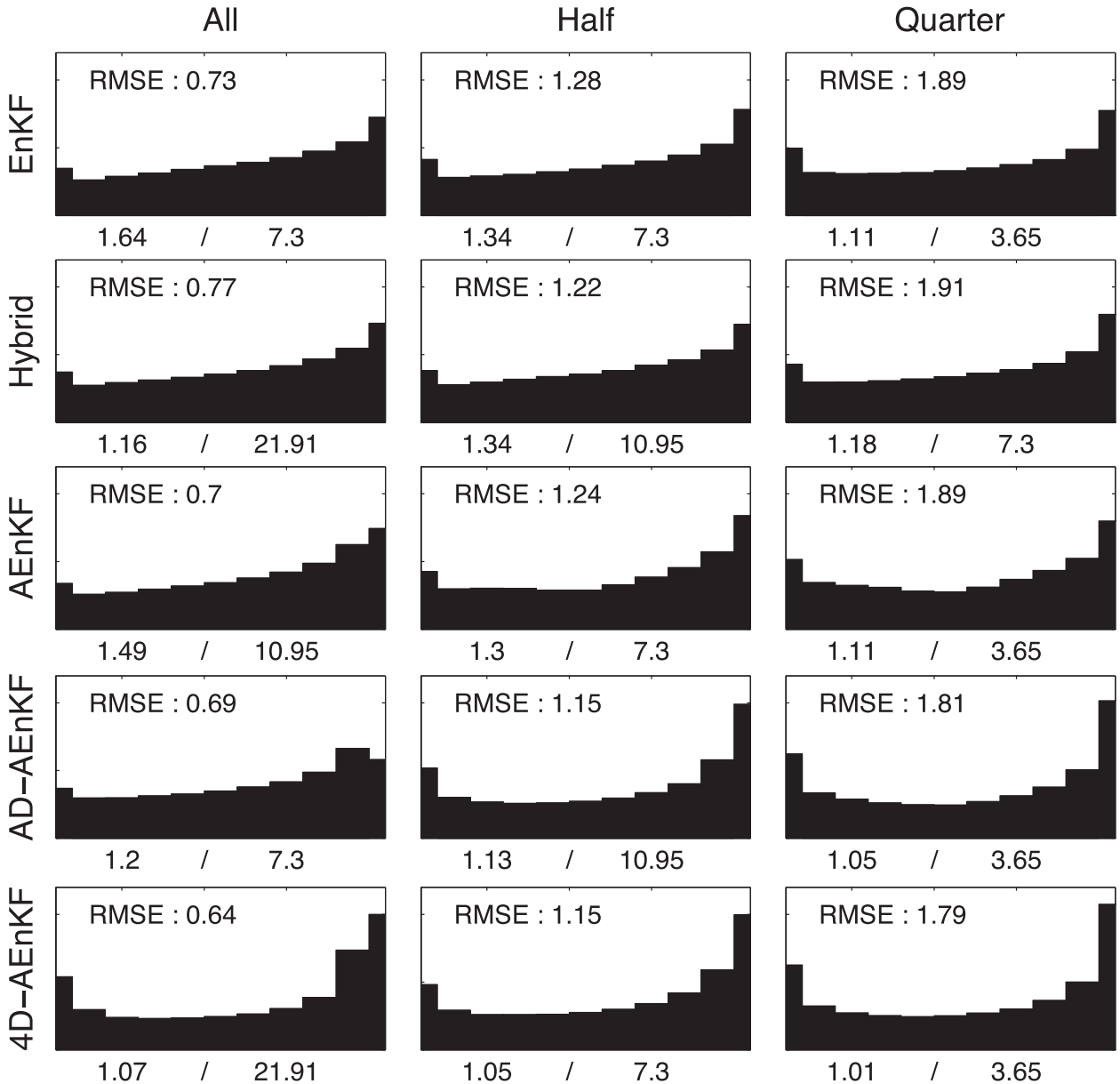


FIG. 7. Rank histograms as calculated from the ensembles of the EnKF, hybrid 3DVAR/EnKF, AEnKF, AD-AEnKF, and 4D-AEnKF for the three different observation scenarios. Results are shown for the combination of inflation factor and localization length scale (left and right values printed below each plot) that yield the best-state estimates for each filter. The corresponding RMSEs are also shown in each panel.

new members at every analysis cycle should also help, but this would generally increase computational cost.

d. Sensitivity of 4D-AEnKF to the number of adjoint steps and \mathbf{B}

A set of experiments were performed to study the impact of the 4D-AEnKF setup on performance. Several assimilation experiments were run, varying the number of adjoint backward integration steps of the back-projected residuals. We also tested different choices of the

stationary background covariance matrices \mathbf{B} (used in the variational 3D-VAR or 4D-VAR components of the system). To implement this, we computed different approximations $\hat{\mathbf{B}}_r$ of the sample covariance matrix \mathbf{B} used in the previous experiment by selecting only the r eigenvectors of \mathbf{B} associated with the first few largest eigenvalues. These low-rank (r) matrices of the following form:

$$\hat{\mathbf{B}}_r = \mathbf{L}_r \mathbf{\Sigma}_r \mathbf{L}_r^T, \tag{17}$$

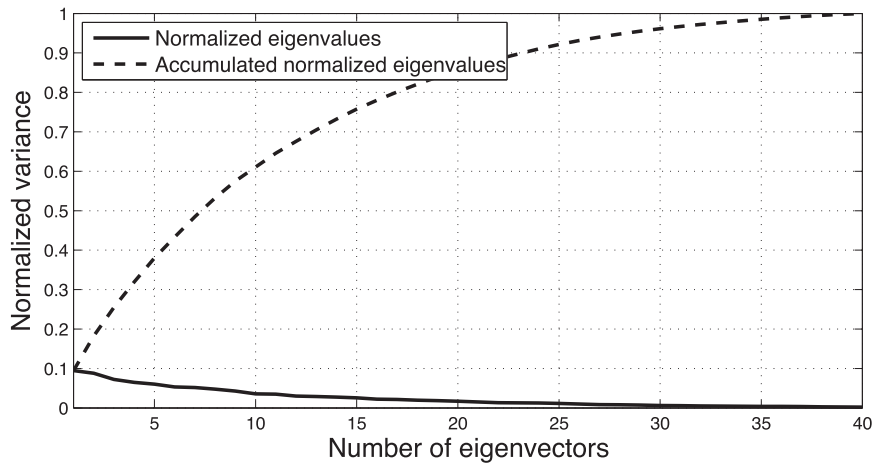


FIG. 8. Normalized eigenvalues of the error covariance matrix \mathbf{B} from the climatology. Solid line and dashed line show the normalized eigenvalues and accumulated normalized eigenvalues, respectively.

were then used as stationary background covariances in the filters. In this equation, \mathbf{L}_γ is the matrix whose columns are the r eigenvectors associated with the r largest eigenvalues of \mathbf{B} , and $\mathbf{\Sigma}_\gamma$ is the $r \times r$ diagonal matrix with the first r eigenvalues. Of course, the more eigenvectors selected (or the larger r) the closer $\hat{\mathbf{B}}_r$ is to \mathbf{B} . For instance, one can see from the spectra of the eigenvalues plotted in Fig. 8 that at least seven eigenvalues are needed to explain half of the total variance of \mathbf{B} .

One can already expect from (13) that the more backward integration steps are taken the more the new residual is constrained with dynamics and data, which should reduce dependence on the stationary background error covariance \mathbf{B} . Thus, one goal of this exercise is to study whether longer backward integrations of the new member can reduce dependence on the quality of \mathbf{B} . However, as discussed before, this should be only true for a limited number of backward integration steps within the range of validity of the linear approximation. Here we limited the number of backward steps to four to make sure that the linear assumption holds and the adjoint is stable for all new ensemble members, in line with our discussion in section 3b. More backward steps should in general improve the performance of the 4D-AEnKF, but incurs the risk of making the adjoint unstable (Hoteit et al. 2005) in certain cases. One could monitor the behavior of the adjoint to select the “optimal” number of backward steps for each newly generated member. This was not done in this experiment for ease of presentation and discussion of the results. As will be clear from the next paragraph, the presented results are enough to study the behavior of the 4D-AEnKF with more backward steps.

Figure 9 plots the RMSE as it results from the 4D-AEnKF as a function of the number of backward time

steps (x coordinate) and the number of eigenvectors that were used to approximate $\hat{\mathbf{B}}$ (y coordinate) with the three different observation scenarios. Of course, the RMSE resulting from zero backward time steps (on the leftmost side) is the RMSE of the AEnKF. Since very low ranks give little freedom for the back projection of the residuals, preventing improvement of the ensemble, a full rank matrix $\hat{\mathbf{B}} + \varepsilon \mathbf{I}$ was used instead of $\hat{\mathbf{B}}$ for the generation of the new members in the AEnKF. With $\varepsilon = 0.001$, results are shown from the combination of localization length scale and inflation factor that yield the best overall state estimates with all, half, and quarter observation cases. These were set as (1.14, 21.91), (1.04, 7.30), and (1.05, 3.65), respectively. As one can expect, the RMSE always decreases in all observation scenarios as the “approximated” $\hat{\mathbf{B}}$ gets closer to \mathbf{B} , as the latter is expected to provide a better representation of the background covariance than \mathbf{B} . The RMSE also generally decreases with more backward integration time steps (of course only within the range of validity of the tangent linear assumption). It is also clearly seen in Fig. 9 that more backward integration reduces the sensitivity of the RMSE to the number of eigenvectors used. This suggests that in general the model dynamics and previous observations can help compensate for a poorly modeled background covariance matrix \mathbf{B} .

4. Discussion

Four-dimensional variational data assimilation (4D-VAR) and ensemble Kalman filters (EnKF) are advanced data assimilation schemes that are being extensively used by the ocean and atmospheric community. Each approach, however, has its own strengths and weaknesses.

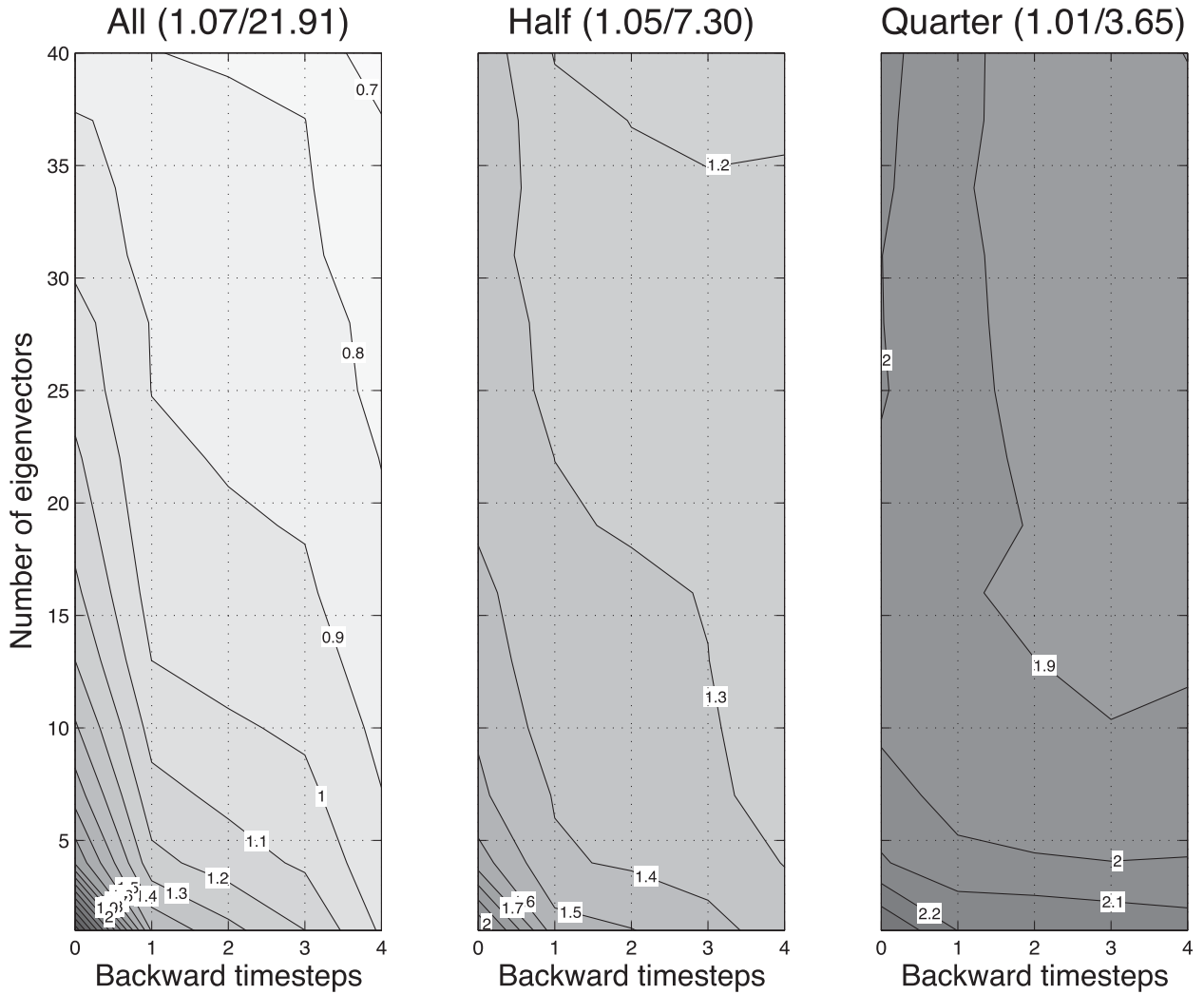


FIG. 9. RMSE averaged over time and all variables as a function of number of backward time steps and number of eigenvectors that were used to approximate \mathbf{B} . The 4D-AEnKF is implemented with 10 members, and observations from three different network scenarios (all variables, every second variable, and every fourth variable) were assimilated at every fourth model time step (or 1 day in real time). A full rank matrix $\mathbf{B} + \epsilon \mathbf{I}$ was used instead of \mathbf{B} for the generation of the new members.

The EnKF provides an efficient algorithm to update the background covariance forward in time by integrating the prior uncertainty as an ensemble of state vectors. Because of limitation in computing resources, however, the EnKF can only integrate small ensembles, meaning that the prior space of uncertainty is almost certainly too small, and distorted in amplitudes. Moreover, as the filter proceeds, this space shrinks and can drift from truth.

There has been strong interest recently in building hybrid schemes that combine the strengths of each approach. New methods have been proposed either by augmenting the EnKF covariance by the stationary background covariance \mathbf{B} of a variational system (hybrid EnKF/3D-VAR), or by using the flow-dependent covariance matrix of the EnKF as the background matrix

for the 4D-VAR problem (hybrid EnKF/4D-VAR). In this paper, we proposed a new direction based on the use of the 4D-VAR method to transform large residuals left behind after an EnKF correction step into new ensemble members along the descent directions, which improve the fit to observations and diversify the ensemble. Indeed, 4D-VAR analysis lets the model-data misfit choose the descent directions used to fit the data. In an analogy to filtering, each descent direction from an iterative fit can then be considered to be a possible adaptive ensemble element.

The new approach is based on the adaptive EnKF (AEnKF), which uses a 3D assimilation (3D-VAR or OI) system and a chosen stationary background covariance to back project the residuals into new EnKF ensemble

members. The idea is then to generalize the 3D assimilation system to a 4D-VAR system; hence, it was referred to as 4D-AEnKF. In contrast with the AEnKF that creates the new member at the current analysis step, the 4D-AEnKF creates the new member in the past so that model dynamics and more data can be included in the estimation process of the new member. The 4D formulation of the AEnKF involves integrating the residuals backward in time with the adjoint model. This should reduce the dependence of the AEnKF on the stationary background covariance matrix and provide more information for better estimation of the new member.

The proposed adaptive schemes should be viewed as a new auxiliary tool that could be used to improve performances when an EnKF is not behaving well because the filter ensemble does not accurately represent the prior distribution. In the ideal case, when an EnKF is behaving well, the residuals are small, and a newly generated ensemble member should be close to the ensemble mean. Therefore, the addition of a new member will not introduce useful information to the ensemble, but even in this case, results from different numerical experiments suggest that the use of the proposed scheme will not degrade the performance of an EnKF operating well under ideal conditions. One could come up with an additional procedure to check if this happens, by for instance comparing the residuals with the observational error. If the residuals are smaller than the observational error, then one may judge that the filter ensemble is good enough and therefore not generate any new members.

The adaptive schemes can be applied to any EnKF scheme simultaneously with other existing auxiliary tools such as inflation and localization. They are simple to implement, but as with the hybrid EnKF-variational schemes, require the specification of a stationary background covariance matrix. The 4D version further requires an adjoint model. For implementation, one should add one OI/3D-VAR step to the 3D version and one adjoint step to the 4D version, and therefore the proposed schemes should not incur significant extra computational burden when implemented with small systems. In cases where iterative methods are necessary to solve the linear problem in (13), the extra computational cost will depend on the number of iterations needed.

The 4D-AEnKF was tested with the L96 model in the presence of both sampling and model errors. L96 provides a benchmark setting to test and evaluate the performance of a new assimilation scheme and the 4D-AEnKF proved to be successful with this model. In the experiments, for all tested cases, the adaptive schemes always enhance the EnKF behavior and, in general, the 4D-AEnKF improves upon the performance of the AEnKF. Furthermore, the backward integration of the residuals enhances the

robustness of the AEnKF and decreases dependence on the stationary background covariance as long as the tangent linear assumption is valid. The benefit of the adaptive scheme is less significant in the case of coarse observation networks, since there is less available information for the back projection of the residuals into the state space. In this case the role of the stationary background covariance in the back-projection scheme and the background integration of the residuals with the adjoint become more important.

The proposed adaptive schemes were found effective in enhancing the EnKF behavior in our test studies with the L96 model. This preliminary application was a necessary step before trying realistic applications it provided encouraging results. We recently tested the scheme with a quasigeostrophic model and the preliminary results were encouraging. More work is required to study and understand the behavior of the proposed schemes with realistic oceanic and atmospheric data assimilation problems. We are currently working on implementing and testing these methods with a realistic ocean model.

Acknowledgments. Cornuelle was supported through NOAA Grant NA17RJ1231 (SIO Joint Institute for Marine Observations). The authors thank Dr. Sakov and two anonymous reviewers for valuable comments and suggestions, which significantly improved the manuscript.

REFERENCES

- Anderson, J. L., 1996: A method for producing and evaluating probabilistic forecasts from ensemble model integrations. *J. Climate*, **9**, 1518–1530.
- Ballabrera-Poy, J., P. Brasseur, and J. Verron, 2001: Dynamical evolution of the error statistics with the SEEK filter to assimilate altimetric data in eddy-resolving ocean models. *Quart. J. Roy. Meteor. Soc.*, **127**, 233–253.
- Bennett, A. F., 2002: *Inverse Modeling of the Ocean and Atmosphere*. Cambridge University Press, 234 pp.
- Bocquet, M., 2011: Ensemble Kalman filtering without the intrinsic need for inflation. *Nonlinear Processes Geophys.*, **18**, 735–750.
- Buehner, M., 2005: Ensemble-derived stationary and flow-dependent background error covariances: Evaluation in a quasi-operational NWP setting. *Quart. J. Roy. Meteor. Soc.*, **131**, 1013–1043.
- Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998: Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.*, **126**, 1719–1724.
- Carton, J. A., and B. S. Giese, 2008: A reanalysis of ocean climate using Simple Ocean Data Assimilation (SODA). *Mon. Wea. Rev.*, **136**, 2999–3017.
- Caya, A., J. Sun, and C. Snyder, 2005: A comparison between the 4DVAR and the ensemble Kalman filter techniques for radar data assimilation. *Mon. Wea. Rev.*, **133**, 3081–3094.
- Courtier, P., J. N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.*, **120**, 1367–1387.

- Daley, R., 1991: *Atmospheric Data Analysis*. Cambridge University Press, 457 pp.
- Dimet, F. X. L., and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus*, **38A**, 97–110.
- Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99** (C5), 10 143–10 162.
- , 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.*, **53**, 343–367.
- , and P. J. van Leeuwen, 2000: An ensemble Kalman smoother for nonlinear dynamics. *Mon. Wea. Rev.*, **128**, 1852–1867.
- Fisher, M., 1998: Minimization algorithms for variational data assimilation. *Proc. Recent Developments in Numerical Methods for Atmospheric Modelling*, Reading, United Kingdom, ECMWF, 364–385.
- Ghil, M., and P. Malanotte-Rizzoli, 1991: Data assimilation in meteorology and oceanography. *Advances in Geophysics*, Vol. 33, Academic Press, 141–266.
- Hamill, T. M., and C. Snyder, 2000: A hybrid ensemble Kalman filter–3D variational analysis scheme. *Mon. Wea. Rev.*, **128**, 2905–2919.
- , and J. S. Whitaker, 2005: Accounting for the error due to unresolved scales in ensemble data assimilation: A comparison of different approaches. *Mon. Wea. Rev.*, **133**, 3132–3147.
- , —, and C. Snyder, 2001: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.*, **129**, 2776–2790.
- Hoteit, I., D. T. Pham, and J. Blum, 2002: A simplified reduced-order Kalman filtering and application to altimetric data assimilation in tropical Pacific. *J. Mar. Syst.*, **36**, 101–127.
- , B. Cornuelle, A. Köhl, and D. Stammer, 2005: Treating strong adjoint sensitivities in tropical eddy-permitting variational data assimilation. *Quart. J. Roy. Meteor. Soc.*, **131**, 3659–3682.
- , G. Triantafyllou, and G. Korres, 2007: Using low-rank ensemble Kalman filters for data assimilation with high dimensional imperfect models. *J. Numer. Anal. Ind. Appl. Math.*, **2**, 67–78.
- , B. Cornuelle, and P. Heimbach, 2010: An eddy-permitting, dynamically consistent adjoint-based assimilation system for the tropical Pacific: Hindcast experiments in 2000. *J. Geophys. Res.*, **115**, C03001, doi:10.1029/2009JC005437.
- , T. Hoar, G. Gopalakrishnan, N. Collins, J. Anderson, B. Cornuelle, A. Köhl, and P. Heimbach, 2013: A MITgcm/DART ensemble analysis and prediction system with application to the Gulf of Mexico. *Dyn. Atmos. Oceans*, **63**, 1–23.
- Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796–811.
- , —, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen, 2005: Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. *Mon. Wea. Rev.*, **133**, 604–620.
- , —, and X. Deng, 2009: Model error representation in an operational ensemble Kalman filter. *Mon. Wea. Rev.*, **137**, 2126–2143.
- Hunt, B. R., and Coauthors, 2004: Four-dimensional ensemble Kalman filtering. *Tellus*, **56A**, 273–277.
- Kepert, J. D., 2009: Covariance localisation and balance in an ensemble Kalman filter. *Quart. J. Roy. Meteor. Soc.*, **135**, 1157–1176.
- Klinker, E., F. Rabier, G. Kelly, and J. F. Mahfouf, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. III: Experimental results and diagnostics with operational configuration. *Quart. J. Roy. Meteor. Soc.*, **126**, 1191–1215.
- Köhl, A., D. Stammer, and B. Cornuelle, 2007: Interannual to decadal changes in the ECCO global synthesis. *J. Phys. Oceanogr.*, **37**, 313–337.
- Lermusiaux, P. F. J., 2007: Adaptive modeling, adaptive data assimilation and adaptive sampling. *Physica D*, **230**, 172–196.
- Lewis, J. M., and J. C. Derber, 1985: The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus*, **37A**, 309–322.
- Li, G., and A. Reynolds, 2007: An iterative ensemble Kalman filters for data assimilation. *SPE Annual Tech. Conf. and Exhibition*, Anaheim, CA, Society of Petroleum Engineers.
- Liu, C., 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.
- Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP—A comparison with 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **129**, 3183–3203.
- Lorenz, E. N., and K. A. Emanuel, 1998: Optimal sites for supplementary weather observations: Simulation with a small model. *J. Atmos. Sci.*, **55**, 399–414.
- Mitchell, H. L., P. L. Houtekamer, and G. Pellerin, 2002: Ensemble size, balance, and model-error representation in an ensemble Kalman filter. *Mon. Wea. Rev.*, **130**, 2791–2808.
- Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763.
- Sandu, A., 2006: On the properties of Runge-Kutta discrete adjoints. *Lecture Notes in Computer Science*, V. N. Alexandrov et al., Eds., Computational Science-ICCS 2006, Vol. 3994, Springer, 550–557.
- Song, H., I. Hoteit, B. D. Cornuelle, and A. C. Subramanian, 2010: An adaptive approach to mitigate background covariance limitations in the ensemble Kalman filter. *Mon. Wea. Rev.*, **138**, 2825–2845.
- Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble square root filters. *Mon. Wea. Rev.*, **131**, 1485–1490.
- Wang, X., C. Snyder, and T. M. Hamill, 2007: On the theoretical equivalence of differently proposed ensemble–3DVAR hybrid analysis schemes. *Mon. Wea. Rev.*, **135**, 222–227.
- Weaver, A. T., J. Vialard, and D. L. T. Anderson, 2003: Three- and four-dimensional variational assimilation with a general circulation model of the tropical Pacific Ocean. Part I: Formulation, internal diagnostics, and consistency checks. *Mon. Wea. Rev.*, **131**, 1360–1378.
- Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924; Corrigendum, **134**, 1722.
- Zhang, F., M. Zhang, and J. A. Hansen, 2009: Coupling ensemble Kalman filter with four-dimensional variational data assimilation. *Adv. Atmos. Sci.*, **26**, 1–8.