SPECIAL
4D-Var ↔ EnKF
COLLECTION

Adaptive Ensemble Covariance Localization in Ensemble 4D-VAR State Estimation

CRAIG H. BISHOP AND DANIEL HODYSS

Naval Research Laboratory, Monterey, California

(Manuscript received 1 March 2010, in final form 3 December 2010)

ABSTRACT

An adaptive ensemble covariance localization technique, previously used in ''local'' forms of the ensemble Kalman filter, is extended to a global ensemble four-dimensional variational data assimilation (4D-VAR) scheme. The purely adaptive part of the localization matrix considered is given by the element-wise square of the correlation matrix of a smoothed ensemble of streamfunction perturbations. It is found that these purely adaptive localization functions have spurious far-field correlations as large as 0.1 with a 128-member ensemble. To attenuate the spurious features of the purely adaptive localization functions, the authors multiply the adaptive localization functions with very broadscale nonadaptive localization functions. Using the Navy's operational ensemble forecasting system, it is shown that the covariance localization functions obtained by this approach adapt to spatially anisotropic aspects of the flow, move with the flow, and are free of far-field spurious correlations. The scheme is made computationally feasible by (i) a method for inexpensively generating the square root of an adaptively localized global four-dimensional error covariance model in terms of products or modulations of smoothed ensemble perturbations with themselves and with raw ensemble perturbations, and (ii) utilizing algorithms that speed ensemble covariance localization when localization functions are separable, variable-type independent, and/or large scale. In spite of the apparently useful characteristics of adaptive localization, single analysis/forecast experiments assimilating 583 200 observations over both 6- and 12-h data assimilation windows failed to identify any significant difference in the quality of the analyses and forecasts obtained using nonadaptive localization from that obtained using adaptive localization.

1. Introduction

General background on the need for localization in ensemble based data assimilation can be found in Houtekamer and Mitchell (2001) and Hamill et al. (2001). [See Evensen (2003) for a review of ensemble Kalman filter research.] When Nonadaptive Ensemble Covariance Localization (NECL) is used in ensemble data assimilation (DA), raw ensemble covariances are attenuated by a function that only depends on the physical distance between the covarying error variables. To try and improve on this type of approach, Bishop and Hodyss (2007, 2009a,b) have introduced a variety of Adaptive Ensemble Covariance Localization (AECL) techniques. These approaches allow the localization or attenuation functions to widen (narrow) as the true error correlation function widens (narrows). They also provide localization functions that move with the flow so that observations of variables at the end of a (long) data assimilation time window can be used to correct upstream variables at the beginning of the data assimilation window.

In the idealized DA models examined in Bishop and Hodyss (2007, 2009a,b), AECL only showed a significant benefit over NECL in systems in which either (i) the width of the true error correlation function was a strongly varying function of space or time and/or (ii) the true error correlation function moved a distance greater than or equal to the width of the optimally tuned nonadaptive localization function. As such, an outstanding question is whether the error correlation functions associated with modern numerical weather prediction (NWP) models exhibit enough variability for AECL to be significantly superior to NECL. This study begins to address this question by comparing AECL and NECL performance in experiments using the Navy Operational Global Atmospheric Prediction System (NOGAPS; Hogan and Rosmond, 1991).

Corresponding author address: Craig H. Bishop, Naval Research Laboratory, Marine Meteorology Division, 7 Grace Hopper Ave., Stop 2, Building 702, Room 212, Monterey, CA 93943-5502.
E-mail: bishop@nrlmry.navy.mil

DOI: 10.1175/2010MWR3403.1

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **DEC 2010**	2. REPORT TYPE	3. DATES COVERED **00-00-2010 to 00-00-2010**
4. TITLE AND SUBTITLE **Adaptive Ensemble Covariance Localization In Ensemble 4D-VAR State Estimation**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Naval Research Laboratory,Monterey,CA,93943-**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Monthly Weather Review, Apr 2011. Vol. 139, Iss. 4; pg. 1241, 15 pgs

14. ABSTRACT
An adaptive ensemble covariance localization technique, previously used in ??local?? forms of the ensemble Kalman filter, is extended to a global ensemble four-dimensional variational data assimilation (4D-VAR)scheme. The purely adaptive part of the localization matrix considered is given by the element-wise square of the correlation matrix of a smoothed ensemble of streamfunction perturbations. It is found that these purely adaptive localization functions have spurious far-field correlations as large as 0.1 with a 128-member ensemble. To attenuate the spurious features of the purely adaptive localization functions, the authors multiply the adaptive localization functions with very broadscale nonadaptive localization functions. Using the Navy?s operational ensemble forecasting system, it is shown that the covariance localization functions obtained by this approach adapt to spatially anisotropic aspects of the flow, move with the flow, and are free of far-field spurious correlations. The scheme is made computationally feasible by (i) a method for inexpensively generating the square root of an adaptively localized global four-dimensional error covariance model in terms of products or modulations of smoothed ensemble perturbations with themselves and with raw ensemble perturbations and (ii) utilizing algorithms that speed ensemble covariance localization when localization functions are separable, variable-type independent, and/or large scale. In spite of the apparently useful characteristics of adaptive localization, single analysis/forecast experiments assimilating 583 200 observations over both 6- and 12-h data assimilation windows failed to identify any significant difference in the quality of the analyses and forecasts obtained using nonadaptive localization from that obtained using adaptive localization.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**16**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

To illustrate AECL methods within the context of a global numerical weather prediction model, Bishop and Hodyss (2007, 2009b) incorporated them into a large observation volume local ensemble Kalman filter (Ott et al. 2004; Hunt et al. 2007). These local observation volumes enabled sophisticated AECL schemes to be implemented in a cost-efficient way. This paper, together with the companion paper Bishop et al. (2011), extends this work by showing how AECL can also be cost efficiently included in global ensemble four-dimensional variational data assimilation (4D-VAR). The motivation for this extension to a nonlocal global framework is as follows.

Local observation volumes are inappropriate for long-time window data assimilation because errors are liable to propagate out of observation volumes. Local approaches limit the effectiveness of some variational techniques for bias correction and the estimation of forecast and observation error variances. There is a growing interest in the use of localized ensemble covariances in three-dimensional variational data assimilation (3D-VAR) and 4D-VAR schemes evident in papers by Lorenc (2003), Buehner (2005), Buehner et al. (2010a,b), Wang et al. (2007), Liu et al. (2009), and others. Buehner et al. (2010a,b) found that an ensemble 4D-VAR scheme that used NECL but did not use a tangent linear model or adjoint, outperformed a version of the operational 4D-VAR scheme both in the tropics and Southern Hemisphere, but not in the northern extratropics.

The promise of improving 4D-VAR with ensemble covariances is a particularly strong motivation for this study because at the authors' research laboratory, a major effort has been underway for the last 9 yr to create the world's first operational weak constraint 4D global variational data assimilation system called the Naval Research Laboratory (NRL) Atmospheric Variational Data Assimilation System–Accelerated Representer (NAVDAS-AR; Xu et al. 2005). NAVDAS-AR became the operational data assimilation scheme for global model atmospheric forecasting in September of 2009. NAVDAS-AR is the only observation space-based 4D-VAR system currently used for operational weather forecasting. El Akkraoui et al. (2008) refer to this observation space form as the dual form of 4D-VAR.

While the NAVDAS-AR code was under construction many aspects of it were in a rapid state of flux. In addition, the authors were unsure about how best to incorporate AECL and/or NECL within NAVDAS-AR. For these reasons, a decision was made to create a prototype observation space ensemble 4D-VAR using both NECL and AECL outside of the developing NAVDAS-AR code so that experimentation on the prototype could inform the final implementation of NECL and AECL within NAVDAS-AR. This paper together with a companion paper (Bishop et al. 2011) reports on the results of these efforts. This paper's aims are to give the first demonstration of the incorporation of flow-adaptive ensemble covariance localization in a global 4D-VAR algorithm and to describe a new ensemble covariance localization that blends nonadaptive localization and adaptive localization. The aim of the companion paper is to show how assumptions about localization functions such as separability, variable-type independence, and large length scales can be used to greatly reduce the cost of ensemble covariance localization within 4D-VAR.

Section 2 gives the (new) partially adaptive localization approach and describes how to incorporate it in a 4D-VAR scheme. Section 3 describes experiments that serve to test and tune the palette of localization functions provided by the partially adaptive localization approach, while section 4 provides a comparison of the similarities and differences of optimally tuned adaptive and nonadaptive localization functions. Conclusions follow in section 5.

2. Partially Adaptive Ensemble Covariance Localization

a. *Definition of forecast error covariance matrix using PAECL*

The Partially Adaptive Ensemble Covariance Localization (PAECL) forecast error covariance model considered in this paper uses both an adaptive localization matrix $\underline{\mathbf{C}}_A$ and a nonadaptive localization matrix $\underline{\mathbf{C}}_N$ such that

$$\underline{\mathbf{P}}_P^f = \underline{\mathbf{P}}^f \odot \underline{\mathbf{C}}_A \odot \underline{\mathbf{C}}_N = \underline{\mathbf{Z}}\underline{\mathbf{Z}}^{\mathrm{T}} \odot (\underline{\mathbf{Z}}_s\underline{\mathbf{Z}}_s^{\mathrm{T}} \odot \underline{\mathbf{Z}}_s\underline{\mathbf{Z}}_s^{\mathrm{T}}) \odot (\underline{\mathbf{W}}\underline{\mathbf{W}}^{\mathrm{T}}). \tag{1}$$

where $\odot$ indicates the element-wise matrix product. In (1), $\underline{\mathbf{P}}^f = \underline{\mathbf{Z}}\underline{\mathbf{Z}}^{\mathrm{T}} = \sum_{k=1}^{K} \underline{\mathbf{z}}_k\underline{\mathbf{z}}_k^{\mathrm{T}}$ is the raw sample covariance matrix of a K member ensemble, $\underline{\mathbf{z}}_k$ is the kth column of the square root of this matrix (typically, $\underline{\mathbf{z}}_k = \underline{\mathbf{x}}_k/\sqrt{K-1}$ where $\underline{\mathbf{x}}_k$ is the kth ensemble perturbation about the ensemble mean). The underlines in the above terms indicate that they pertain to a set of states separated in time. For example, if a 12-h data assimilation window was being considered with a 1-h discretization in time then $\underline{\mathbf{x}}_k^{\mathrm{T}} = [\mathbf{x}_k^{\mathrm{T}}(0), \mathbf{x}_k^{\mathrm{T}}(1), \mathbf{x}_k^{\mathrm{T}}(2), \ldots, \mathbf{x}_k^{\mathrm{T}}(12)]$ where $\mathbf{x}_k^{\mathrm{T}}(t_i)$ gives the n-vector describing the state of the kth ensemble perturbation at the ith discrete time t_i defining the data assimilation window. The matrix $\underline{\mathbf{C}}_A = (\underline{\mathbf{Z}}_s\underline{\mathbf{Z}}_s^{\mathrm{T}} \odot \underline{\mathbf{Z}}_s\underline{\mathbf{Z}}_s^{\mathrm{T}})$ is the AECL matrix; where $\underline{\mathbf{Z}}_s\underline{\mathbf{Z}}_s^{\mathrm{T}} = \sum_{i=1}^{L} \underline{\mathbf{z}}_i^s\underline{\mathbf{z}}_i^{s\mathrm{T}}$ is the sample correlation matrix of a smoothed normalized ensemble. Details of a generalized method for obtaining $\underline{\mathbf{Z}}_s$ are given

in the appendix. While $\underline{\mathbf{C}}_A$ allows for variable-dependent localization, in this paper we suppress the capability because, using the techniques discussed in Bishop et al. (2011), doing so reduces the cost of AECL by more than a factor of 4. Here, the attenuation of the raw covariance between two variables depends only on the position of the two variables in space–time and not on their variable type. This type of variable-type independent localization was achieved by letting the smoothed $\underline{\mathbf{z}}_i^s$ fields be based solely on horizontally and vertically smoothed streamfunction perturbations. We also experimented with letting the smoothed $\underline{\mathbf{z}}_i^s$ fields be based on smoothed equivalent potential temperature perturbations. It was found that AECL DA performance was insensitive to this choice. Note that the number of smoothed members L may be different to the number K of raw members.

Since $\underline{\mathbf{Z}}_s\underline{\mathbf{Z}}_s^{\mathrm{T}}$ is constructed from smoothed raw ensemble perturbations, its correlation functions will tend to have larger (shorter) length scales in regions where the raw ensemble correlation length scales are large (short). In addition, the correlation functions of $\underline{\mathbf{Z}}_s\underline{\mathbf{Z}}_s^{\mathrm{T}}$ will tend to propagate in a similar way to the raw ensemble correlation functions. Finally, note that the correlation functions associated with the adaptive localization $(\underline{\mathbf{Z}}_s\underline{\mathbf{Z}}_s^{\mathrm{T}}) \odot (\underline{\mathbf{Z}}_s\underline{\mathbf{Z}}_s^{\mathrm{T}})$ can be widened or narrowed simply by respectively increasing or decreasing the amount of smoothing used to create $\underline{\mathbf{Z}}_s$.

The matrix $\underline{\mathbf{C}}_N = \underline{\mathbf{W}}\underline{\mathbf{W}}^{\mathrm{T}} = \sum_{m=1}^{M}\underline{\mathbf{w}}_m\underline{\mathbf{w}}_m^{\mathrm{T}}$ is the NECL matrix, $\underline{\mathbf{W}}$ is its square root and $\underline{\mathbf{w}}_m$ denotes the mth column of $\underline{\mathbf{W}}$. We assume that each column $\underline{\mathbf{w}}_m$ is represented in terms of a separable product of a horizontal structure function and a vertical structure function. The horizontal structure functions are given by spherical harmonics while the vertical structure functions are given by cosine functions. The 4D state $\underline{\mathbf{w}}_m$ is obtained by replicating the mth 3D-structure function for each of the variables and times that define the dimension and ordering of the 4D state under consideration. This replication procedure means that no localization is performed in time or with respect to intervariable covariances. In general, the broader the localization functions defined by $\underline{\mathbf{C}}_N$, the fewer the number of columns M in $\underline{\mathbf{W}}$.

Note that PAECL in (1) reduces to pure AECL when every element of $\underline{\mathbf{C}}_N$ is equal to 1; in other words, AECL comes from PAECL when $\underline{\mathbf{W}} = \underline{\mathbf{w}}_1 = \mathbf{1}$ $(M = 1)$ so that $\underline{\mathbf{C}}_N = \mathbf{1}\mathbf{1}^{\mathrm{T}}$, where $\mathbf{1}$ is a vector of ones as long as the state vector. Alternatively, (1) delivers pure NECL when $\underline{\mathbf{Z}}_s = \underline{\mathbf{z}}_1^s = \mathbf{1}$ $(L = 1)$ so that $\underline{\mathbf{C}}_A = \mathbf{1}\mathbf{1}^{\mathrm{T}}$.

From the square root theorem given in Bishop and Hodyss (2009b) and its more general form given in Bishop et al. (2011), (1) may be written in the following form:

$$\underline{\mathbf{P}}_P^f = \sum_{k=1}^{K}\sum_{j=1}^{L}\sum_{i=1}^{L}\sum_{m=1}^{M}(\underline{\mathbf{z}}_k \odot \underline{\mathbf{z}}_j^s \odot \underline{\mathbf{z}}_i^s \odot \underline{\mathbf{w}}_m)(\underline{\mathbf{z}}_k \odot \underline{\mathbf{z}}_j^s \odot \underline{\mathbf{z}}_i^s \odot \underline{\mathbf{w}}_m)^{\mathrm{T}}. \tag{2}$$

In other words, the forecast error covariance model $\underline{\mathbf{P}}_P^f = \underline{\mathbf{Z}}_P\underline{\mathbf{Z}}_P^{\mathrm{T}}$ to be used in this paper is equal to the covariance of a KL^2M member ensemble of perturbations of the form $\underline{\mathbf{z}}_k \odot \underline{\mathbf{z}}_j^s \odot \underline{\mathbf{z}}_i^s \odot \underline{\mathbf{w}}_m$ all stored as columns of the matrix $\underline{\mathbf{Z}}_P$. Each of these members represents an element-wise product of a scaled raw ensemble member $\underline{\mathbf{z}}_k$ with the product of two smoothed and normalized ensemble members $\underline{\mathbf{z}}_j^s \odot \mathbf{z}_i^s$ and a nonflow-dependent structure function $\underline{\mathbf{w}}_m$. These function products amount to modulations of the raw ensemble and it is for this reason that we refer to the KL^2M ensemble implied by (2) as a *modulation* ensemble.

Whether the primal or dual form of 4D-VAR is implemented, variational schemes require the repeated evaluation of matrix vector products like $\underline{\mathbf{Z}}_P\mathbf{a}$ and $\underline{\mathbf{Z}}_P^{\mathrm{T}}\mathbf{b}$. It may be shown that

$$\underline{\mathbf{Z}}_P\mathbf{a} = \sum_{k=1}^{K}\sum_{j=1}^{L}\sum_{i=1}^{L}(\underline{\mathbf{z}}_k \odot \underline{\mathbf{z}}_j^s \odot \underline{\mathbf{z}}_i^s) \odot (\underline{\mathbf{W}}\hat{\mathbf{a}}_{ijk}) \quad \text{and}$$
$$(\underline{\mathbf{Z}}_P^{\mathrm{T}}\mathbf{b})_{ijk} = \underline{\mathbf{W}}^{\mathrm{T}}[(\underline{\mathbf{z}}_k \odot \underline{\mathbf{z}}_j^s \odot \underline{\mathbf{z}}_i^s) \odot \mathbf{b}], \tag{3}$$

where $\hat{\mathbf{a}}_{ijk}$ and $(\underline{\mathbf{Z}}_P^{\mathrm{T}}\mathbf{b})_{ijk}$ are both M vectors listing all of the elements of $\mathbf{a}$ and $\underline{\mathbf{Z}}_P^{\mathrm{T}}\mathbf{b}$, respectively, which are associated with the same values of i, j, and k such that $\mathbf{a}^{\mathrm{T}} = [\hat{\mathbf{a}}_{111}^{\mathrm{T}}, \hat{\mathbf{a}}_{211}^{\mathrm{T}}, \ldots, \hat{\mathbf{a}}_{ijk}^{\mathrm{T}}, \ldots, \hat{\mathbf{a}}_{LLK}^{\mathrm{T}}]$ and $(\underline{\mathbf{Z}}_P^{\mathrm{T}}\mathbf{b})^{\mathrm{T}} = [(\underline{\mathbf{Z}}_P^{\mathrm{T}}\mathbf{b})_{111}^{\mathrm{T}}, (\underline{\mathbf{Z}}_P^{\mathrm{T}}\mathbf{b})_{211}^{\mathrm{T}}.., (\underline{\mathbf{Z}}_P^{\mathrm{T}}\mathbf{b})_{ijk}^{\mathrm{T}}.., (\underline{\mathbf{Z}}_P^{\mathrm{T}}\mathbf{b})_{LLK}^{\mathrm{T}}]$. The term $\underline{\mathbf{W}}\hat{\mathbf{a}}_{ijk}$ may be associated with a smoothly varying set of weights for the structure given by $(\underline{\mathbf{z}}_k \odot \underline{\mathbf{z}}_j^s \odot \underline{\mathbf{z}}_i^s)$. The companion paper (Bishop et al. 2011) shows how the separable and "replicated" nature of the columns of $\underline{\mathbf{W}}$ can be used to rapidly evaluate the right-hand sides of (3).

Since $\underline{\mathbf{z}}_j^s \odot \underline{\mathbf{z}}_i^s = \underline{\mathbf{z}}_i^s \odot \underline{\mathbf{z}}_j^s$, (2) reduces to the more computationally efficient form:

$$\begin{aligned}\underline{\mathbf{P}}_P^f &= \sum_{k=1}^{K}\sum_{j=1}^{L}\sum_{m=1}^{M}(\underline{\mathbf{z}}_k \odot \underline{\mathbf{z}}_j^s \odot \underline{\mathbf{z}}_j^s \odot \underline{\mathbf{w}}_m)(\underline{\mathbf{z}}_k \odot \underline{\mathbf{z}}_j^s \odot \underline{\mathbf{z}}_j^s \odot \underline{\mathbf{w}}_m)^{\mathrm{T}} + 2\sum_{k=1}^{K}\sum_{j=1}^{L-1}\sum_{i=j+1}^{L}\sum_{m=1}^{M}(\underline{\mathbf{z}}_k \odot \underline{\mathbf{z}}_j^s \odot \underline{\mathbf{z}}_i^s \odot \underline{\mathbf{w}}_m)(\underline{\mathbf{z}}_k \odot \underline{\mathbf{z}}_j^s \odot \underline{\mathbf{z}}_i^s \odot \underline{\mathbf{w}}_m)^{\mathrm{T}} \\ &= \tilde{\underline{\mathbf{Z}}}_P\tilde{\underline{\mathbf{Z}}}_P^{\mathrm{T}}.\end{aligned} \tag{4}$$

Thus, if all K members of the raw ensemble were linearly independent and all L members of the smoothed ensemble were linearly independent and all M columns of $\underline{\mathbf{W}}$ were independent, then (4) indicates that the modulation ensemble contains up to $MK[L + L(L - 1)/2] = MK[L(L + 1)/2]$ linearly independent members. The matrix $\underline{\tilde{\mathbf{Z}}}_P$ lists these members. We never have to hold all of these members in memory as they can easily be produced as needed via element-wise products.

On the computers currently available at the larger numerical weather prediction centers, it is feasible to generate 128 member ensembles of 6–18-h forecasts in a timely manner. Similarly, it is feasible to smooth each of these forecasts in a timely manner to obtain $L = K = 128$ smooth members. Hence, ensemble sizes for which $[K^2(K + 1)/2]M = [128^2 \times 129/2]M = 1\,056\,768M$ are readily obtained. Global numerical weather prediction models will soon be running at 10-km horizontal resolution with 100 vertical levels and about 10 variables at each grid point. Such models will have approximately 10^{10} variables. Thus, with a $K = 128$ member ensemble at this resolution and a $\underline{\mathbf{W}}$ with $M = 10\,000$ independent structure functions, $\underline{\mathbf{Z}}_P$ could have as many linearly independent columns as there were model variables—even at this very high resolution. In this way, (4) removes Lorenc's (2003) concern that ensemble-based forecast error covariance models might be rank deficient and unable to fit large numbers of very accurate observations.

b. Global variational solution for analysis correction

While we experimented with both the primal and dual forms of 4D-VAR and found that they both gave similar results, the results shown here pertain to the dual or observation space form of 4D-VAR of which NAVDAS-AR is an example. The dual form can be derived by noting that the minimum error variance estimate $\underline{\mathbf{x}}^a$ of the true 4D state $\underline{\mathbf{x}}$ given the forecast $\underline{\mathbf{x}}^f$ and observations $\underline{\mathbf{y}}$ is

$$\begin{aligned}\underline{\mathbf{x}}^a &= \underline{\mathbf{x}}^f + \underline{\mathbf{P}}_P^f\mathbf{H}^{\mathrm{T}}(\mathbf{H}\underline{\mathbf{P}}_P^f\mathbf{H}^{\mathrm{T}} + \underline{\mathbf{R}})^{-1}[\underline{\mathbf{y}} - H(\underline{\mathbf{x}}^f)]\\ &= \underline{\mathbf{x}}^f + \underline{\mathbf{P}}_P^f\mathbf{H}^{\mathrm{T}}\underline{\mathbf{R}}^{-1/2}(\underline{\mathbf{R}}^{-1/2}\mathbf{H}\underline{\mathbf{P}}_P^f\mathbf{H}^{\mathrm{T}}\underline{\mathbf{R}}^{-1/2} + \mathbf{I})^{-1}\underline{\mathbf{R}}^{-1/2}[\underline{\mathbf{y}}\\ &\quad - H(\underline{\mathbf{x}}^f)] = \underline{\mathbf{x}}^f + \underline{\mathbf{P}}_P^f\tilde{\mathbf{H}}^{\mathrm{T}}(\tilde{\mathbf{H}}\underline{\mathbf{P}}_P^f\tilde{\mathbf{H}}^{\mathrm{T}} + \mathbf{I})^{-1}\widetilde{[\underline{\mathbf{y}} - H(\underline{\mathbf{x}}^f)]}\end{aligned} \tag{5}$$

assuming that the forecast and observation error covariance matrices are accurate (see e.g., Xu et al. 2005; Daley and Barker 2001). The tildes in the last line of (5) indicate that the matrices/vectors have been premultiplied by $\underline{\mathbf{R}}^{-1/2}$. To avoid the high computational cost of the matrix inverse indicated in (5), Daley and Barker (2001) note that since the gradient of the quadratic form

$$J = \frac{1}{2}\mathbf{b}^{\mathrm{T}}(\tilde{\mathbf{H}}\underline{\mathbf{P}}_P^f\tilde{\mathbf{H}}^{\mathrm{T}} + \mathbf{I})\mathbf{b} - \mathbf{b}^{\mathrm{T}}\widetilde{[\underline{\mathbf{y}} - H(\underline{\mathbf{x}}^f)]} \tag{6}$$

with respect to $\mathbf{b}$ is

$$\begin{aligned}\frac{\partial J}{\partial \mathbf{b}} &= (\tilde{\mathbf{H}}\underline{\mathbf{P}}_P^f\tilde{\mathbf{H}}^{\mathrm{T}} + \mathbf{I})\mathbf{b} - \widetilde{[\mathbf{y} - H(\mathbf{x}^f)]}\\ &= (\tilde{\mathbf{H}}\underline{\mathbf{Z}}_P\underline{\mathbf{Z}}_P^{\mathrm{T}}\tilde{\mathbf{H}}^{\mathrm{T}} + \mathbf{I})\mathbf{b} - \widetilde{[\mathbf{y} - H(\mathbf{x}^f)]},\end{aligned} \tag{7}$$

the vector $\mathbf{b}_{\min}$ that minimizes (6) makes $\partial J/\partial \mathbf{b} = 0$ so that

$$\mathbf{b}_{\min} = (\tilde{\mathbf{H}}\underline{\mathbf{P}}_P^f\tilde{\mathbf{H}}^{\mathrm{T}} + \mathbf{I})^{-1}\widetilde{[\mathbf{y} - H(\mathbf{x}^f)]} \tag{8}$$

and (5) is equivalent to

$$\underline{\mathbf{x}}^a = \underline{\mathbf{x}}^f + \underline{\mathbf{P}}_P^f\tilde{\mathbf{H}}^{\mathrm{T}}\mathbf{b}_{\min}. \tag{9}$$

The vector $\mathbf{b}_{\min}$ was obtained by minimizing (6) using the conjugate gradient technique. The conjugate gradient technique exhibited steady convergence for all experiments reported here. To ensure tight convergence in all experiments, 128 iterations of the conjugate gradient inner loop were performed to find $\mathbf{b}_{\min}$.

3. Experimental setup for tuning of localization functions

a. Generation of truth run and pseudo-observations

We consider an idealization in which we let the true state of the atmosphere $\underline{\mathbf{x}}^t$ be defined by a 30–42-h forecast using NOGAPS run at horizontal spectral resolution T119 with 30 vertical levels (T119L30). The NAVDAS analysis at 0000 UTC 23 June 2005 was used to initialize this truth run. NAVDAS (Daley and Barker 2001) is the 3D-VAR form of NAVDAS-AR and it was used for operational Navy weather forecasting until September 2009 when it was replaced by its 4D-VAR counterpart NAVDAS-AR (Xu et al. 2005). The first guess $\underline{\mathbf{x}}^f$ was taken to be a 6–18-h forecast valid at the same time as the 30–42-h forecast used to generate the truth. The initial condition for this first guess was the NAVDAS analysis 0000 UTC 24 June 2005. Since we wish to compare the effect of adaptive localization in both 6- and 12-h DA windows, this same 6–18-h forecast will be used as the first guess for a 6-h DA window running from 0900 to 1500 UTC and for a 12-h DA window running from 0600 to 1800 UTC.

Figure 1 gives the vertical variation of global horizontal averages of the root-mean-square (rms) error of this first guess in wind and temperature at 1200 UTC. The rms wind errors exceed 4 m s^{-1} at the jet level and

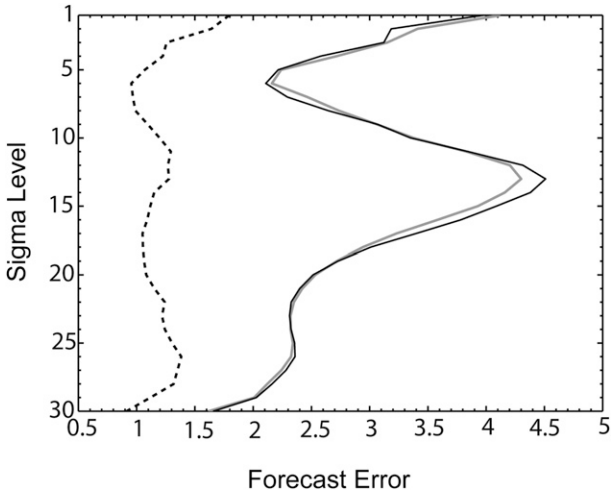

FIG. 1. The vertical profile of the root-mean-square error of the prior first guess at 1200 UTC in zonal wind (gray line), meridional wind (black line), and temperature (dashed line). The vertical axis indicates the models vertical σ-level index (level 30 is the earth's surface) while the horizontal axis indicates speed (m s^{-1}) for the wind components and temperature (K) for the temperature.

in the stratosphere while rms temperature errors exceed 1 K throughout much of the depth of the atmosphere. Note that this "error" is entirely due to the NAVDAS assimilation of observations over the 24-h period preceding the initialization of the first-guess run. Hence, the structure of the difference between the two states is partially determined by the quasi-isotropic error correlation functions of NAVDAS and partially determined by the atmospheric perturbation dynamics of both the atmosphere and NOGAPS.

Pseudo-observations can be created at any desired location by adding random numbers representative of "observation error" to variables from the true state selected for "observation." Two types of observational network are considered. The first pertains to a 6-h DA window and features 194 400 observations of zonal and meridional wind (u, v) and temperature T at 0900, 1200, and 1500 UTC for a total of 583 200 observations. We shall hereafter refer to this observational network as the 6-h network. The second network is almost identical except it pertains to a 12-h DA window and features 194 400 observations of zonal and meridional wind (u, v) and temperature T at 0600, 1200, and 1800 UTC and will be referred to as the 12-h network.

The standard deviation of observation error was 2 m s^{-1} for the wind components and 1 K for temperature. For simplicities sake and to check whether adaptive localization provides any clear systematic benefit when the observing network is quasi-isotropic and quasi-homogeneous, we let the observations subsample the T119L30 grid at every third point in both the zonal and meridional directions. In addition, no observations were included between the South Pole and 80°S and the North Pole and 80°N. In the vertical, observations were at every third sigma level (for a total of 10 levels) beginning at the sigma level immediately above the terrain pressure.

b. Generation of ensemble

A 128-member ensemble was generated using the technique described in McLay et al. (2010) that is a modification of the ensemble transform (ET) technique described in McLay et al. (2008). The modification is that in McLay et al. (2010) transformation coefficients are computed for five distinct latitudinal bands and then interpolated between the centers of the bands whereas in McLay et al. (2008) only one set of global transformation coefficients are computed. As pointed out in McLay et al. (2008), the ET technique (Bishop and Toth 1999) is a variation on Toth and Kalnay's (1997) breeding technique and as such must be cycled for a few days in order to "breed" growing structures. To this end, the 128-member ET ensemble was cycled for 6 days preceding the first DA window.

Because of the way they are generated, the ET ensemble perturbation structure is entirely independent of the error covariance functions used to define the NAVDAS forecast error covariance model. Consequently, the ET perturbations are not particularly well configured to correct the difference between the experiments first guess and pseudotruth because this difference is partially determined by the NAVDAS forecast error covariance functions. Mismatches between the distribution from which forecast errors are drawn and the distribution from which ensemble perturbations are drawn must also occur in operational ensemble DA systems when known and unknown sources of model error are not accurately accounted for. Hence, the mismatch between the distribution from which our forecast error is drawn and that from which our ensemble perturbations is drawn has some very crude similarities to what occurs in operational ensemble DA schemes. When such distribution mismatches are present, ensemble covariance localization has two roles: to suppress spurious ensemble correlations and to create variance in directions of error that are not well sampled by the ensemble generation process.

c. Tuning of localization functions

We compare three types of localization: NECL, where $\underline{\mathbf{P}}_P^f = \underline{\mathbf{Z}}\underline{\mathbf{Z}}^{\mathrm{T}} \odot \underline{\mathbf{W}}\underline{\mathbf{W}}^{\mathrm{T}}$; AECL, where $\underline{\mathbf{P}}_P^f = \underline{\mathbf{Z}}\underline{\mathbf{Z}}^{\mathrm{T}} \odot \underline{\mathbf{Z}}_s\underline{\mathbf{Z}}_s^{\mathrm{T}} \odot \underline{\mathbf{Z}}_s\underline{\mathbf{Z}}_s^{\mathrm{T}}$; and PAECL, where $\underline{\mathbf{P}}_P^f = \underline{\mathbf{P}}^f \odot \underline{\mathbf{C}}_A \odot \underline{\mathbf{C}}_N = \underline{\mathbf{Z}}\underline{\mathbf{Z}}^{\mathrm{T}} \odot (\underline{\mathbf{Z}}_s\underline{\mathbf{Z}}_s^{\mathrm{T}} \odot \underline{\mathbf{Z}}_s\underline{\mathbf{Z}}_s^{\mathrm{T}}) \odot (\underline{\mathbf{W}}\underline{\mathbf{W}}^{\mathrm{T}})$. The width of the correlation functions associated with the NECL matrix

$\underline{\mathbf{C}}_N = \underline{\mathbf{W}}\underline{\mathbf{W}}^{\mathrm{T}}$ are adjusted via two length parameters that control the width of the horizontal and vertical functions that define the columns of $\underline{\mathbf{C}}_N$, while the width of the correlation functions associated with the AECL matrix $\underline{\mathbf{C}}_A = \underline{\mathbf{Z}}_s\underline{\mathbf{Z}}_s^{\mathrm{T}} \odot \underline{\mathbf{Z}}_s\underline{\mathbf{Z}}_s^{\mathrm{T}}$ is controlled by varying two smoothing parameters that control the strength of the horizontal and vertical smoothing applied to turn the raw ensemble perturbations $\underline{\mathbf{Z}}_K$ into the smoothed ensemble perturbations $\underline{\mathbf{Z}}_s$. In principle, these parameters can be tuned to optimize any given cost function. We chose to tune them to optimize the cost function

$$J_{\mathrm{observed}}(t) = \frac{1}{3}\left\{\frac{\overline{[u_{\mathrm{err}}^{\mathrm{post}}(t)]^2}}{\overline{[u_{\mathrm{err}}^{\mathrm{prior}}(12)]^2}} + \frac{\overline{[v_{\mathrm{err}}^{\mathrm{post}}(t)]^2}}{\overline{[v_{\mathrm{err}}^{\mathrm{prior}}(12)]^2}} + \frac{\overline{[T_{\mathrm{err}}^{\mathrm{post}}(t)]^2}}{\overline{[T_{\mathrm{err}}^{\mathrm{prior}}(12)]^2}}\right\} \tag{10}$$

at the analysis time $t = 12$, where $\overline{[u_{\mathrm{err}}^{\mathrm{post}}(t)]^2}$, $\overline{[v_{\mathrm{err}}^{\mathrm{post}}(t)]^2}$ and $\overline{[T_{\mathrm{err}}^{\mathrm{post}}(t)]^2}$ give the global average of the square of the posterior error of the state estimate in the troposphere of zonal wind u, meridional wind v, and temperature T at 1200 UTC after the assimilation of observations from the 6-h network. At the optimization time $t = 1200$ UTC, $\overline{[u_{\mathrm{err}}^{\mathrm{post}}(t)]^2}$, $\overline{[v_{\mathrm{err}}^{\mathrm{post}}(t)]^2}$, and $\overline{[T_{\mathrm{err}}^{\mathrm{post}}(t)]^2}$ pertain to the error of the analysis obtained from (9). At later times, $t > 1200$ UTC, $\overline{[u_{\mathrm{err}}^{\mathrm{post}}(t)]^2}$, $\overline{[v_{\mathrm{err}}^{\mathrm{post}}(t)]^2}$, and $\overline{[T_{\mathrm{err}}^{\mathrm{post}}(t)]^2}$ pertain to the error of the nonlinear forecast initialized with the 1200 UTC analysis. The symbols, $\overline{[u_{\mathrm{err}}^{\mathrm{post}}(t)]^2}$, $\overline{[v_{\mathrm{err}}^{\mathrm{post}}(t)]^2}$, and $\overline{[T_{\mathrm{err}}^{\mathrm{post}}(t)]^2}$ give the global average in the troposphere of the square of the prior error or first-guess error of the 12-h forecast initialized at 0000 UTC 24 June 2005 of zonal wind u, meridional wind v, and temperature T before the assimilation of observations. Hence, $J_{\mathrm{observed}}(t)$ gives the average fractional reduction in the error variance of the observed variables due to the assimilation of observations. Although not used for tuning the localization functions, we also measured the fractional reduction in error variance of the unobserved variables using

$$J_{\mathrm{unobserved}}(t) = \frac{1}{2}\left\{\frac{\overline{[(p_s)_{\mathrm{err}}^{\mathrm{post}}(t)]^2}}{\overline{[(p_s)_{\mathrm{err}}^{\mathrm{prior}}(12)]^2}} + \frac{\overline{[q_{\mathrm{err}}^{\mathrm{post}}(t)]^2}}{\overline{[q_{\mathrm{err}}^{\mathrm{prior}}(12)]^2}}\right\}. \tag{11}$$

where the meaning of the overbars, subscripts, and superscripts are the same as in (10), but p_s and q denote the unobserved variables of surface pressure and specific humidity, respectively.

To ensure that global averages gave equal weight to areas with the same surface area, gridpoint errors were multiplied by the cosine of their latitude before inclusion in the global sum in order to account for the fact that the distance between NOGAPS Gaussian grid points decreases with the cosine of latitude. To make sure that the errors pertained primarily to the troposphere, only the lower 23 model levels corresponding to the atmosphere around 100 hPa and below were included.

As previously mentioned, each of the localization methods was tuned to minimize the state estimation error at the analysis time for the 6-h data assimilation window. To see the effect of lengthening the observation window, the localization tunings that minimized analysis error variance of the observed variables for the 6-h network were then used to assimilate observations from the 12-h network.

Figure 2a depicts the evolution of the state estimation error as measured by $J_{\mathrm{observed}}(t)$ [(9)] from the analysis time out to a 5-day forecast for all of the localization times and for both the 6- and 12-h observational networks. Figure 2b depicts the evolution of the corresponding measure of the error in the unobserved variables $J_{\mathrm{unobserved}}(t)$. Figures 2a,b show that after tuning the differing localization methods NECL, AECL, and PAECL all give very similar 5-day forecast errors for both observed and nonobserved variables and for both the 6- and 12-h observational networks. All of the approaches profoundly reduce state estimation error. Figure 2a (Fig. 2b) shows that it takes about 96 h (84 h) for the forecast error variance in the observed (unobserved) variables to exceed the corresponding forecast error variance of the first guess. For both the 6- and 12-h networks, the analysis error obtained from AECL is somewhat larger than that from NECL and PAECL. However, for the 12-h network in particular, the 5-day forecast error from AECL is smaller than NECL and PAECL. However, none of these differences are large enough to be statistically significant.

An interesting feature of Fig. 2b is that the error of the forecast of the unobserved terrain pressure and humidity variables decreases for the first 12 h of the forecast. We speculate that this improvement is partially a result of the dynamics of the nonlinear model bringing the unobserved variables into an improved dynamical balance with the observed variables. As noted by Kepert (2009), while one could expect analyses made with unlocalized ensemble covariances to satisfy linear balance constraints at large scales, when localization is employed, imbalance is to be expected at all scales.

Using 128 Power 5 + IBM processors at 1.9 GHz, it was found that it took 53 and 73 min to obtain the 1200 UTC analysis using NECL and PAECL, respectively. This time was the same for both the 6- and 12-h DA windows because the total number of observations was the same

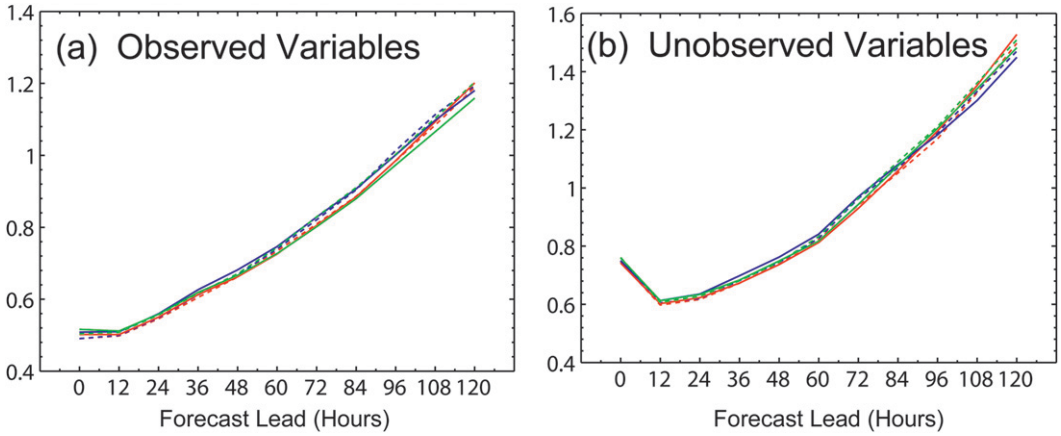

FIG. 2. Normalized root-mean-square posterior error as a function of forecast time. Blue, red, and green curves pertain to nonadaptive, partially adaptive, and purely adaptive localization. Dashed and solid curves pertain to the 6- and 12-h data assimilation window, respectively. (a) The results for the mean of the normalized observed variables u, v, and T [(10)]. (b) The corresponding results for the normalized unobserved variables p_s and q [(11)].

for both of these experiments. The relatively small difference in computational cost is due to the fact that the tight nonadaptive localization (Fig. 6) required by NECL corresponds to a nonadaptive localization matrix $\underline{\mathbf{W}}$ that has many more columns than the number of columns in the $\underline{\mathbf{W}}$ corresponding to the very broad localization (Fig. 5) used for PAECL. As noted in Bishop et al. (2011), the cost of localization can be reduced if one defines the localization matrix on a coarser-resolution grid than the raw ensemble perturbations. This method of reducing computational cost was not employed in our calculations. If we had employed a reduced-resolution grid that had an order of magnitude fewer points than our T119L30 grid, then one would expect an order of magnitude reduction in the amount of time required for each analysis. Were it not for the fact that the optimal tuning for NECL gave the NECL covariance matrix many more columns than that of the PAECL matrix, the margin by which NECL was faster than PAECL would have been even greater.

Since our adaptive localization functions are built from ensemble perturbations, the quality of the adaptive localization is closely linked to the quality of the ensemble perturbations. If assimilating real data, this quality depends on how well the ensemble generation scheme accounts for all sources of error. Accounting for unknown sources of model error is very difficult. In our experiment, we modeled the error-ensemble mismatch using distinctly different methods to create a pseudo-first-guess error and the ensemble perturbations. The mismatch created by our approach may be greater than that between actual forecast error and ensemble perturbations from state-of-the-art ensemble generation schemes. Consequently, the potential benefit of adaptive localization could be greater than that suggested by our results.

The results from Bishop and Hodyss (2007, 2009a) indicate that adaptive localization is only likely to deliver large benefits when true error correlation functions propagate a significant distance relative to the width of the localization functions and/or when the true error correlation length scales varies markedly from one location to another. If the primary reason for the neutrality of our results is short error propagation distances over 6–12-h DA windows and weak spatial inhomogeneity of the true error correlation length scale, then the neutrality of our results is unlikely to change with an improved ensemble model of the forecast error distribution.

4. The PAECL palette of localization functions

As previously mentioned, adaptive localization is likely to be beneficial if covariance functions move a significant distance relative to the localization length scale and/or if the true error correlation length scale is highly inhomogeneous. In this section, we compare and contrast the effect of adaptive and nonadaptive localization functions at points on the globe where such characteristics are likely to be found.

To find a point where the true error correlation function is likely to exhibit significant propagation, we searched for the maximum value of zonal wind at σ–level 15 (around 400 hPa) at 1200 UTC. The location of this maximum value in zonal wind was found to be at 40°S, 90°E. This (winter) region is typically characterized by strong baroclinic zonal jets so it was no surprise to find the strongest zonal wind at this point. Figure 3 gives horizontal and vertical perspectives of the unlocalized

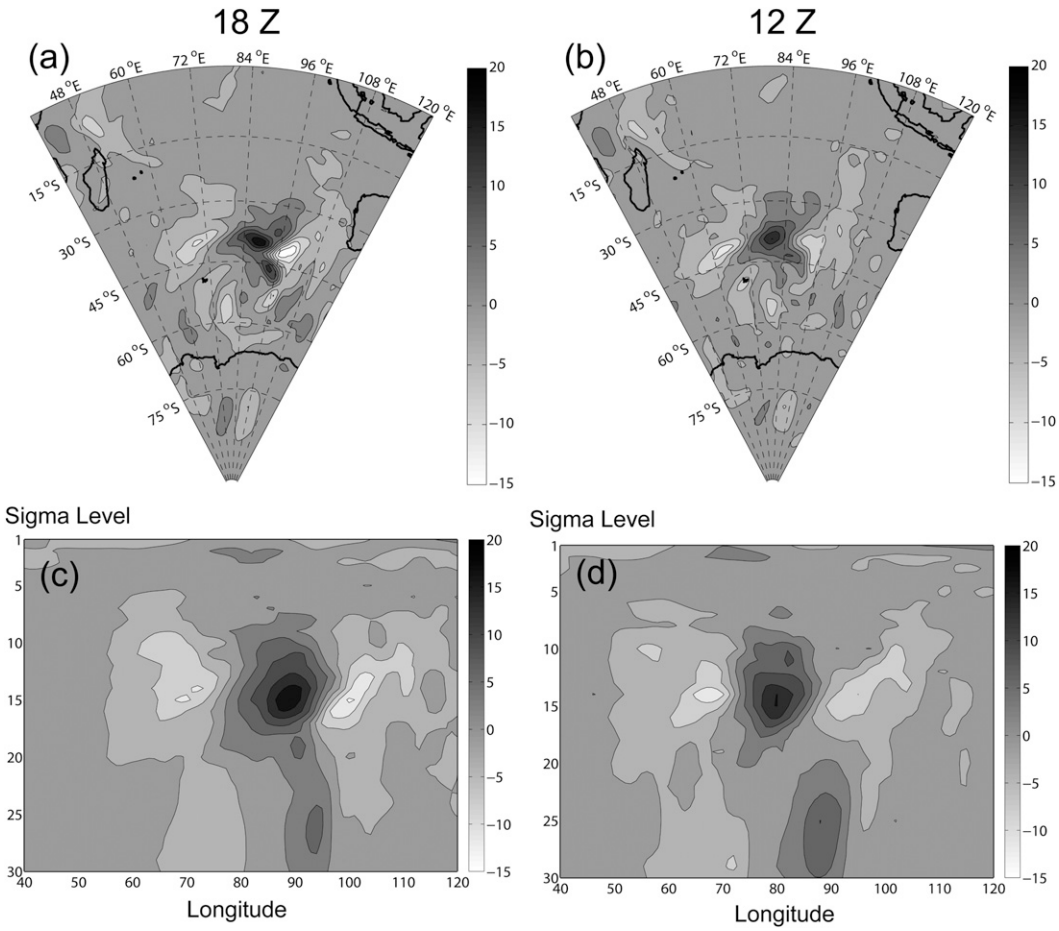

FIG. 3. Unlocalized ensemble covariance function of meridional wind at 1800 and 1200 UTC with 1800 UTC meridional wind variables at 40°S, 90°E and σ-level 15 (about 400 hPa). The ensemble has 128 members. The horizontal cross sections at σ-level 15 of the covariance function at (a) 1800 and (b) 1200 UTC. The zonally oriented vertical cross sections at 40°S of the covariance function at (c) 1800 and (d) 1200 UTC.

ensemble covariance function of meridional wind at 1800 (Figs. 3a,c) and 1200 UTC (Figs. 3b,d) with a single 1800 UTC meridional wind variable located at 40°S, 90°E. The function corresponds to a column of the unlocalized ensemble covariance matrix $\underline{\mathbf{P}}^f = \underline{\mathbf{Z}}\underline{\mathbf{Z}}^{\mathrm{T}}$. This particular column was chosen because its diagonal element corresponds to a local maximum in the westerly zonal wind. The structure of this function is identical to that of the single observation correction that would be obtained from assimilating an observation at 1800 UTC of meridional wind at the model grid point at 40°S, 90°E and σ-level 15 (about 400 hPa). The horizontal cross section of the 1800 UTC covariance function shown in Fig. 3a features a positive region centered around 40°S, 90°E and σ-level 15 with negative regions to the east and west of this central region. Because this pattern is not unlike that of a baroclinic wave packet (Zimin et al. 2003), it is conceivable that the wavelike pattern of covariance stretching from 65° to 105°E is not spurious. However, it would be more difficult to argue that the undulations between 50° and 60°S together with those equatorward of 15°S represent nonspurious covariances. Figure 3b shows that the 1200 UTC covariance function is broadly similar to the 1800 UTC covariance function except that it lies 8°–10°W (upstream) of the 1800 UTC covariance function. This is unsurprising given the strong westerly winds associated with this location.

Figures 3c,d give the corresponding vertical structure at 1800 and 1200 UTC, respectively. Interestingly, from the surface to sigma level 15, the functions tilt westward with increasing height in a manner that is qualitatively similar to a growing baroclinic wave. Also note that upper-level features lie a greater distance to the west of lower-level features at 1200 than at 1800 UTC. This change of vertical tilt with time is reminiscent of nonmodal baroclinic wave growth (Farrell 1989).

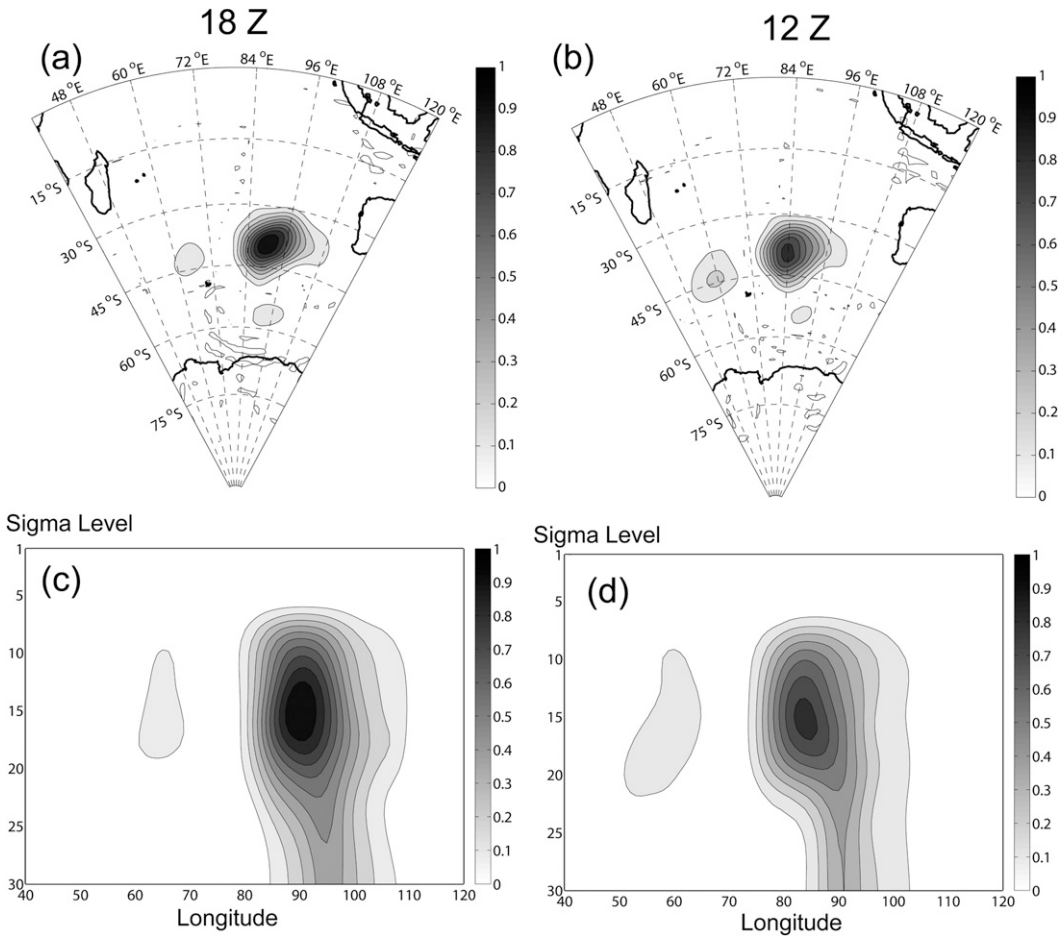

FIG. 4. The AECL function for the raw covariance function shown in Fig. 3 is shown. This localization function is the element-wise square of the correlation function of a 128-member ensemble of smoothed and normalized streamfunction fields.

Figure 4 gives the AECL localization function corresponding to Fig. 3. Figure 4a shows that at the 1800 UTC time, the 0.1 contour of the adaptive localization function stretches from 30° to 50°S and from about 80° to 108°E. Figure 4b shows that the function to moderate the raw ensemble covariances at 1200 UTC is similar to that for 1800 UTC except the pattern is shifted about 6°W of the 1800 UTC localization function. This movement is consistent though a little slower than the apparent movement of the raw covariance function shown in Fig. 3. Figures 4c,d show that the adaptive localization function tilts westward with height and, hence, is well suited to preserving the westward tilt of the raw covariance function shown in Fig. 3.

What Fig. 4 does not show is that the AECL localization function has oscillations in the Northern Hemisphere with isolated maxima as large as 0.1. The nonadaptive part $\underline{\mathbf{C}}_N = \underline{\mathbf{W}}\underline{\mathbf{W}}^{\mathrm{T}}$ of the PAECL scheme was introduced to attenuate such far-field oscillations of the AECL. For this purpose, the localization functions associated with $\underline{\mathbf{C}}_N = \underline{\mathbf{W}}\underline{\mathbf{W}}^{\mathrm{T}}$ were designed to have no vertical variation but a very broad horizontal variation. The structure of the column of $\underline{\mathbf{C}}_N = \underline{\mathbf{W}}\underline{\mathbf{W}}^{\mathrm{T}}$ corresponding to Figs. 3 and 4 that optimized the DA performance of PAECL is shown in Fig. 5. Note that this function is broad enough to preserve the significant parts of the AECL localization function shown in Fig. 4 and most of the raw covariance function (Fig. 3) that is likely to be nonspurious at both 1800 and 1200 UTC.

The full PAECL localization function (column of $\underline{\mathbf{C}}_A \odot \underline{\mathbf{C}}_N$) corresponding to Fig. 3 is not shown here because it is almost identical to that shown in Fig. 4. This is because the nonadaptive part (Fig. 5) is so much broader than the adaptive part (Fig. 4). As such, we use Fig. 4 as a proxy for the full PAECL localization function.

It is of interest to compare this proxy for the PAECL localization function with the purely NECL localization

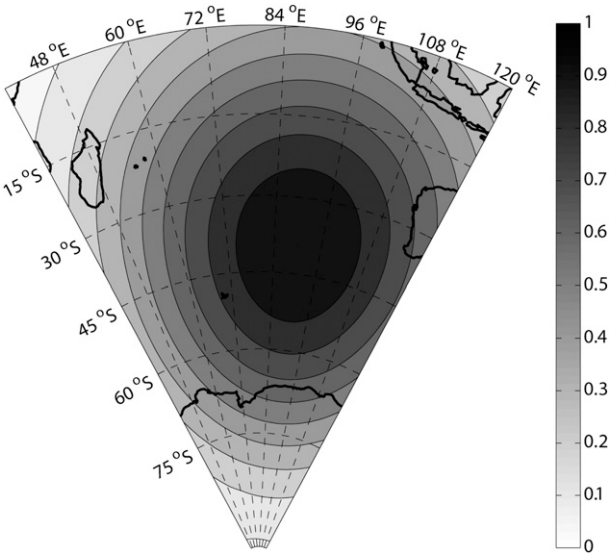

FIG. 5. This is the very broadscale horizontal nonadaptive localization function that is used to attenuate spurious far-field features of the adaptive localization function shown in Fig. 4. Note that this nonadaptive localization function has no vertical variation.

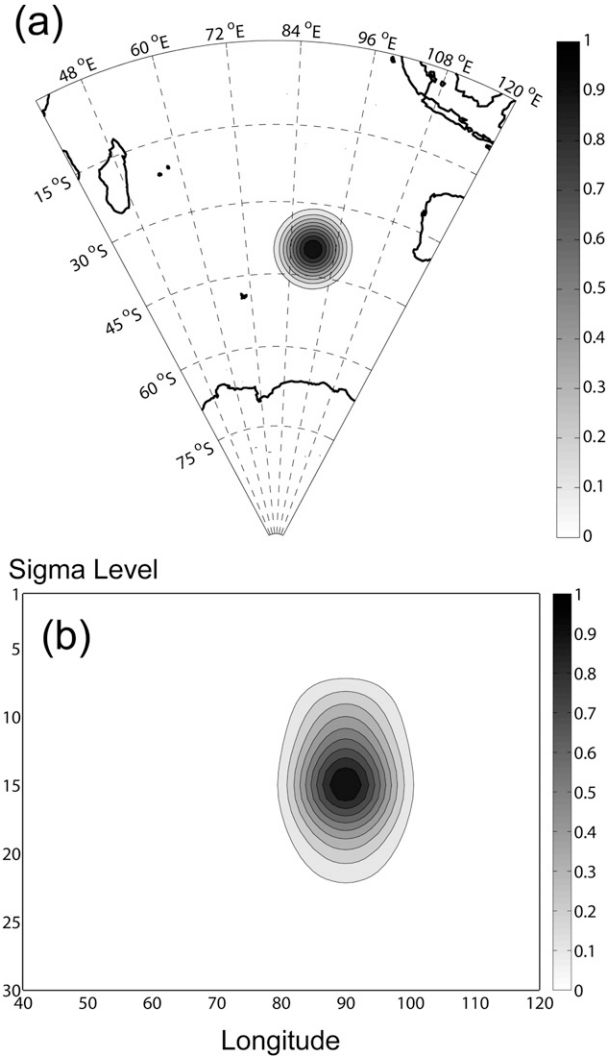

FIG. 6. As in Figs. 4a,c, but the NECL function is shown. (The function is the same at 1800 and 1200 UTC.)

function shown in Fig. 6. This nonadaptive localization function has been tuned to minimize the area-weighted average of analysis error. It has a slightly reduced horizontal scale relative to PAECL having a 0.1 contour stretching 32°–48°S and 80°–100°E. Comparison of Figs. 6b and 4c, however, shows that it has a significantly smaller vertical scale than the PAECL function. Note that the nonadaptive localization function is the same at both 1200 and 1800 UTC.

Figures 7 and 8 give the raw ensemble covariance functions localized with PAECL and NECL, respectively. Figure 7 shows that PAECL preserves much of the positive central lobe and the eastern negative lobe of the raw ensemble covariance function at both 1800 and 1200 UTC. At both 1800 and 1200 UTC, PAECL allows the central positive lobe to have a larger magnitude than the eastern negative lobe—as it is in the raw covariance function. In contrast, because the 1200 UTC nonadaptive localization function does not lie to the east of the 1800 UTC nonadaptive localization function, NECL forces the negative eastern lobe to have a larger magnitude than the positive lobe at 1200 UTC. In addition, PAECL preserves much more of the vertical structure of the raw ensemble covariance than NECL. However, we must recall that the covariance functions displayed in Figs. 7 and 8 pertain to a point where the ensemble mean zonal wind is maximized; hence, most points on the globe would be in less need of adaptive localization than the point shown in Figs. 7 and 8. This may be why we did not find a significant benefit from adaptive localization.

While Figs. 3 to 8 give insight into how PAECL can accommodate moving error fields, Fig. 9 compares and contrasts the wide variety of 3D localization function structures provided by PAECL with those given by NECL localization for 4 points in the Northern Hemisphere. Figure 9 shows that, on average, the horizontal extent of the PAECL localization functions is similar to that of the NECL localization functions while the vertical extent of PAECL localization functions is greater than that of the NECL localization function.

The localization functions that minimize area-averaged root-mean-square analysis error in our experiments are considerably narrower than those used by Buehner et al. (2010a,b). In their EnKF and (for the sake of consistency)

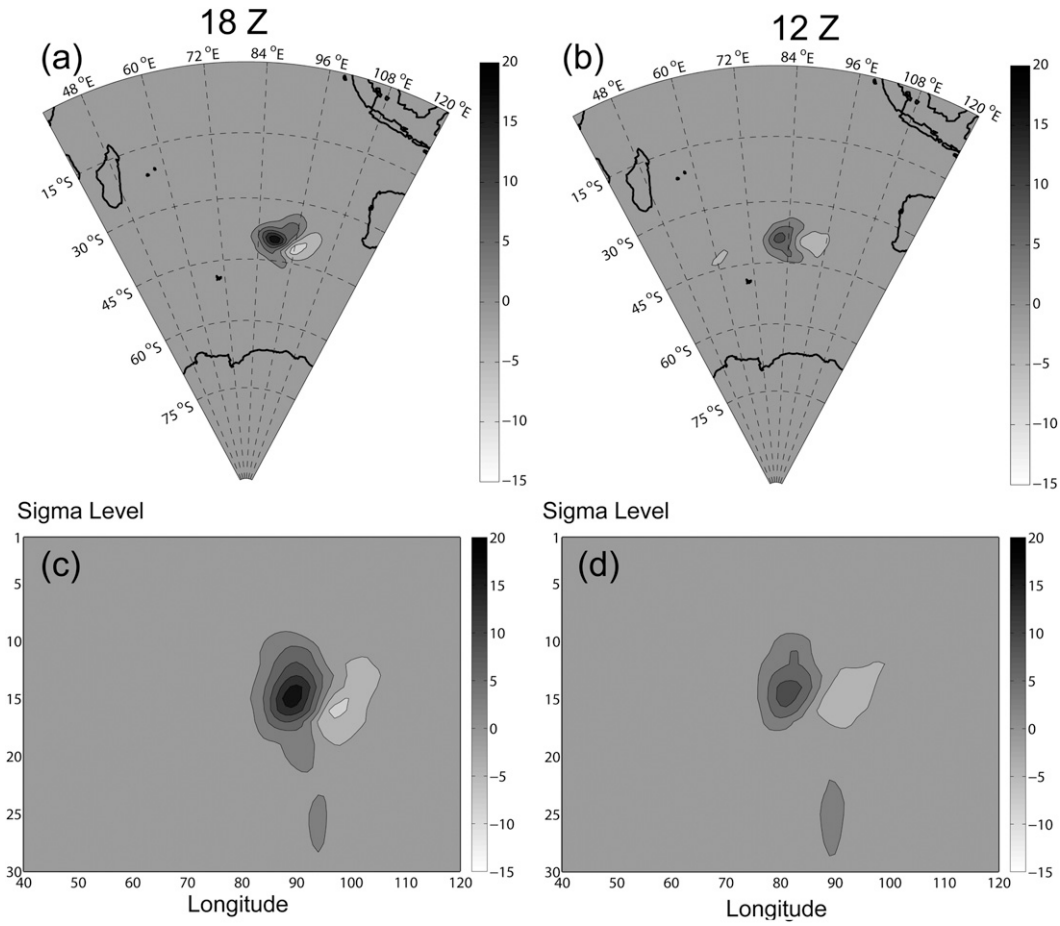

FIG. 7. As in Fig. 3, but the raw ensemble covariance function has been localized with the PAECL.

the various flavors of ensemble 4D-VAR considered in their study, the horizontal localization function goes to 0.2 of its maximum value 1400 km away from the function maximum and to zero 2800 km away from the function maximum. In contrast, Figs. 4 and 6 indicate that both our adaptive and nonadaptive localization functions are below 0.1 of their maximum values about 1000 km away from their respective function maximums. The 2800-km localization limit used in Buehner et al. (2010a,b) was first arrived at as part of the study by Mitchell et al. (2002) on ensemble size and balance issues in the Environment Canada (EC) stochastic EnKF. Researchers at EC have steadily improved their EnKF since that time [see Houtekamer et al. (2009) for a more recent description of the system]. However, no intense effort has been made since Mitchell et al. (2002) to retune their EnKF localization limit as it was felt that the few experiments that have been performed with differing localization limits did not justify an extensive tuning effort for this parameter (H. L. Mitchell 2010, personal communication).

In reflecting on why our optimal localization limit is so different to that obtained by Mitchell et al. (2002), it is worth recalling that their EnKF assimilates observations in sequential batches. The same localization function is applied to each batch regardless of how effectively preceding batches of observations have removed error from large scales. This invariance of the localization function to previous observations assimilated means that the correction obtained from sequential assimilation will be somewhat different to that which would be obtained by assimilating all of the observations in one large batch—as is done in our ensemble 4D-VAR. With this in mind, we speculate that the difference in horizontal correlation length scale may be a result of (i) fundamental differences in the optimal scale of localization functions for ensemble 4D-VAR and the sequential EnKF, (ii) the high density and quasi-isotropic nature of our idealized network favoring shorter localization length scales than the more anisotropic network found in the real atmosphere, (iii) our area-averaged measure of analysis error giving more weight to the

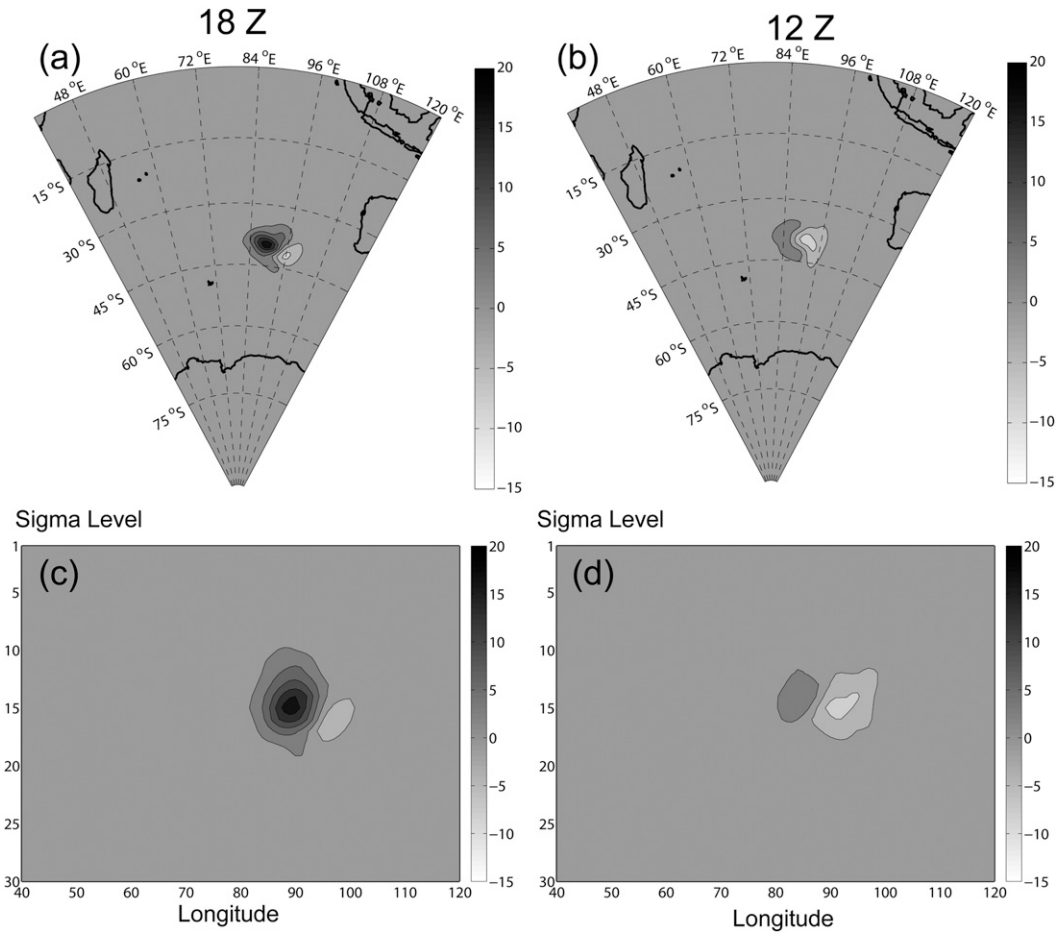

FIG. 8. As in Fig. 7, but the raw ensemble covariance function has been localized with the NECL.

accuracy of analyses at tropical grid points than the various measures used by Mitchell et al. (2002) to tune their localization length scale, and/or (iv) McLay et al.'s (2010) ensemble generation technique causing ensemble perturbations to decorrelate over shorter distances than those of ensemble perturbations generated by the EnKF.

5. Conclusions

A new ensemble covariance localization technique that blends adaptive localization with nonadaptive localization has been presented. The method includes nonadaptive ensemble covariance localization and purely adaptive localization as special cases. It was shown how a square root theorem enables the scheme to be incorporated into global 4D variational algorithms. To demonstrate the technique, both nonadaptive and adaptive ensemble covariance localization approaches were tuned to minimize globally averaged mean square analysis error with both 6- and 12-h data assimilation windows. It was found that while the adaptive localization functions produced showed a high degree of adaptability in both space and time, they were not significantly better than nonadaptive localization functions at reducing forecast or analysis error.

Bishop and Hodyss (2007, 2009a) found clear benefits from adaptive localization when the ensemble perturbations were drawn from precisely the same distribution as the distribution of forecast errors. In our experiment, the error-ensemble mismatch was modeled by using distinctly different methods to create a pseudo-first-guess error and the ensemble perturbations. It is possible that the mismatch created by our approach is greater than that between actual forecast error and ensemble perturbations from state-of-the-art ensemble generation schemes. If so, then the potential benefit of adaptive localization is greater than that suggested by our results.

Other possible reasons for the failure of adaptive localization to show benefit in our experiments include the

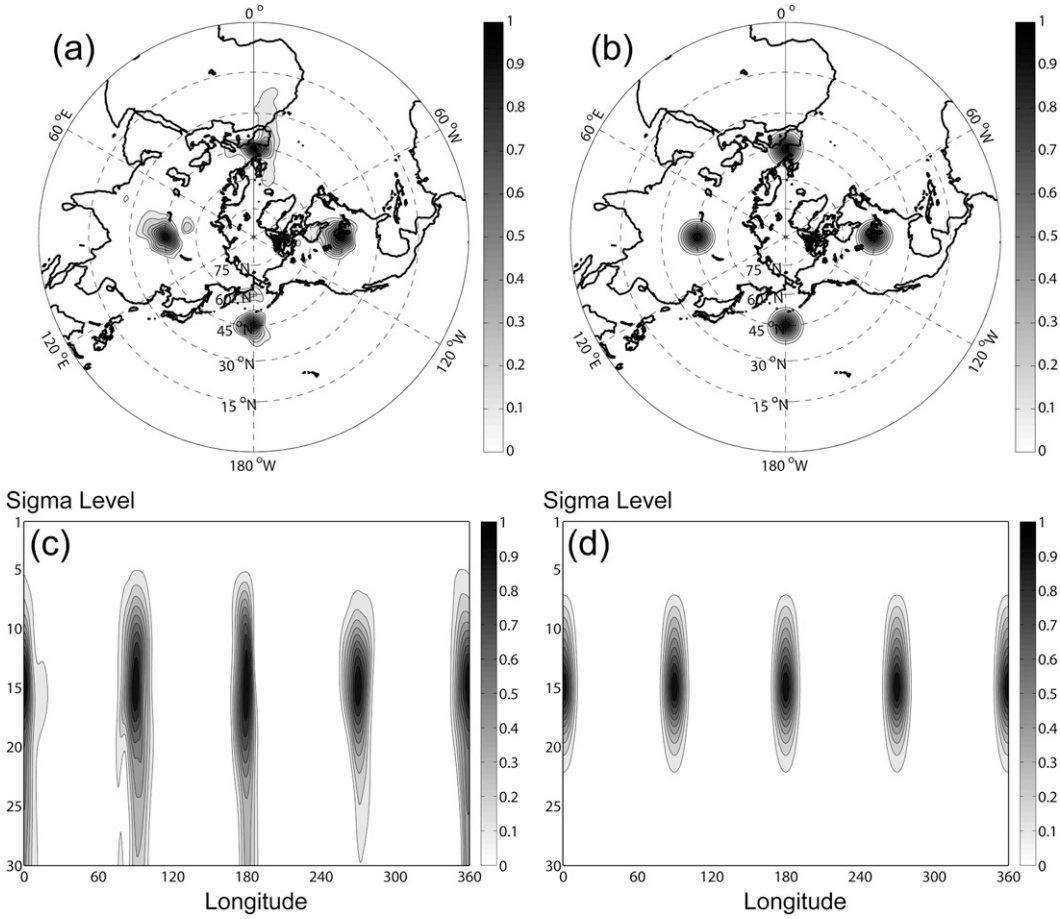

FIG. 9. Comparison of structure of optimally tuned (a) adaptive and (b) nonadaptive localization functions at 1200 UTC for 4 grid points in the Northern Hemisphere all situated at the same latitude (45°N) and σ-level 15 (about 400 hPa), but all separated by 90° of longitude (0°, 90°, 180°, 270°E) and 45°N at model σ-level 15 (about 400 hPa). The corresponding vertical structure of the (c) adaptive and (d) nonadaptive localization functions along the 45°N latitude circle.

possibility that (i) the movement of error correlation functions over 6- and 12-h data assimilation windows is not large enough for adaptive localization to show a clear advantage, and/or (ii) the spatial inhomogeneity of the true error correlation functions is not large enough for adaptive localization to show a benefit. If the primary reason for the neutrality of our results is short error propagation distances and weak spatial inhomogeneity of the true error correlation length scale over 6–12-h DA windows, then the neutrality of our results is unlikely to change with an improved ensemble model of the forecast error distribution.

The single case studied here is not enough to reveal whether or not there is a small, but significant difference between adaptive and nonadaptive localization in our idealized system. However, because of the somewhat contrived nature of our first-guess error and the mismatch of our ensemble covariances with this first-guess error, identifying a small but significant difference from repeated experiments would be of little relevance to the design of candidate operational ensemble DA schemes.

Future research will be aimed at incorporating both adaptive and nonadaptive ensemble covariances into the Navy's operational 4D-VAR scheme (NAVDAS-AR) and then performing cycling data assimilation and forecasting experiments. Hopefully, these experiments will reveal significant (but probably small) differences in the performance of adaptive and nonadaptive localization.

Acknowledgments. CHB and DH gratefully acknowledge financial support from ONR Project Element 0602435N, Project BE-435-003, and ONR Grant N0001407WX30012.

APPENDIX

Creating Smooth Normalized Perturbations

The smoothed ensemble members are obtained from the following steps:

1) Create ensemble perturbations that are normalized by their ensemble standard deviation; that is, if x_{ijtq} denotes the perturbation of the qth variable type at the tth output time at the jth model grid point of the ith ensemble member, normalize this variable using $\tilde{x}_{ijtq} = x_{ijtq}/\sqrt{\frac{1}{K-1}\sum_{i=1}^{K}(x_{ijtq})^2}$.
2) Smooth over variable types by performing a weighted average over variable type using $\hat{x}_{ijtq} = \sum_{q'=1}^{n_{\varpi}} w_q^{q'} \tilde{x}_{ijtq'}^s$, where $\sum_{q'=1}^{n_{\varpi}} w_q^{q'} = 1$. Let the vector of variable smoothed ensemble perturbations resulting from this procedure be denoted by $\hat{\underline{\mathbf{x}}}_i$. Considerable time can be saved by using univariate localization functions. Univariate adaptive localization is obtained by letting $w_{q=i}^{q'} = w_{q=j}^{q'}$ for any variable type i and j.
3) Smooth the variables $\hat{\underline{\mathbf{x}}}_i$ over space to obtain $\hat{\underline{\mathbf{x}}}_i^s$ by first smoothing in the horizontal and then in the vertical or vice versa.

If $\hat{x}_{ijtq}^s$ denotes the element of $\hat{\underline{\mathbf{x}}}_i^s$ corresponding to the qth variable type at the tth output time at the jth model grid point of the ith ensemble member, normalize this variable using $z_{ijtq}^s = [\hat{x}_{ijtq}^s/\sum_{i=1}^{K}(\hat{x}_{ijtq}^s)^2]$. The vector $\underline{\mathbf{z}}_i^s$ listing the variables obtained through this procedure gives the ith column of the smoothed normalized ensemble matrix $\underline{\mathbf{Z}}_s$ which makes $\underline{\mathbf{Z}}_s\underline{\mathbf{Z}}_s^{\mathrm{T}} = \sum_{i=1}^{L}\underline{\mathbf{z}}_i^s\underline{\mathbf{z}}_i^{s\mathrm{T}}$ be a correlation matrix.

In this paper, we used a univariate localization based solely on the streamfunction; in other words, for all values of q, we set $w_q^{q'} = 0$ if q' indicated a nonstreamfunction variable, but $w_q^{q'} = 1$ if q' indicated the streamfunction variable. We also used a slight variation of the above generalized approach in which the part of $\underline{\mathbf{z}}_i^s$ associated with the terrain pressure p_s was given by the smoothed and normalized value of the streamfunction on the model's fifth-lowest vertical level (approximately 900 mb). We did not use the lowest level because streamfunction tends to zero at the lowest model level.

REFERENCES

Bishop, C. H., and Z. Toth, 1999: Ensemble transformation and adaptive observations. *J. Atmos. Sci.,* **56,** 1748–1765.

——, 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.

——, and ——, 2009a: Ensemble covariances adaptively localized with ECO-RAP, Part 1: Tests on simple error models. *Tellus,* **61,** 84–96.

——, and ——, 2009b: Ensemble covariances adaptively localized with ECO-RAP, Part 2: A strategy for the atmosphere. *Tellus,* **61,** 97–111.

——, ——, P. Steinle, H. Simms, A. M. Clayton, A. C. Lorenc, D. M. Barker, and M. Buehner, 2011: Efficient ensemble covariance localization in variational data assimilation. *Mon. Wea. Rev.,* **139,** 573–580.

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.

——, 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.

——, ——, ——, ——, 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.

Daley, R., and E. Barker, 2001: NAVDAS: Formulation and diagnostics. *Mon. Wea. Rev.,* **129,** 869–883.

El Akkraoui, A., P. Gauthier, S. Pellerin, and S. Buis, 2008: Intercomparison of the primal and dual formulations of variational data assimilation. *Quart. J. Roy. Meteor. Soc.,* **134,** 1015–1025.

Evensen, G., 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.,* **53,** 343–367.

Farrell, B. F., 1989: Optimal excitation of baroclinic waves. *J. Atmos. Sci.,* **46,** 1193–1206.

Hamill, T. M., 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.

Hogan, T. F., and T. E. Rosmond, 1991: The description of the Navy Operational Global Atmospheric Prediction System's spectral forecast model. *Mon. Wea. Rev.,* **119,** 1786–1815.

Houtekamer, P. L., and H. L. Mitchell, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.,* **129,** 123–137.

——, ——, and X. Deng, 2009: Model error representation in an operational ensemble Kalman filter. *Mon. Wea. Rev.,* **137,** 2126–2143.

Hunt, B. R., E. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D,* **230,** 112–126.

Kepert, J. D., 2009: Covariance localisation and balance in an Ensemble Kalman Filter. *Quart. J. Roy. Meteor. Soc.,* **135,** 1157–1176.

Liu, C., Q. Xiao, and B. Wang, 2009: An ensemble-based four-dimensional variational data assimilation scheme. Part II: Observing System Simulation Experiments with Advanced Research WRF (ARW). *Mon. Wea. Rev.,* **137,** 1687–1704.

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.

McLay, J. G., C. H. Bishop, and C. A. Reynolds, 2008: Evaluation of the ensemble transform analysis perturbation scheme at NRL. *Mon. Wea. Rev.,* **136,** 1093–1108.

——, ——, and ——, 2010: A local formulation of the ensemble-transform (ET) analysis perturbation scheme. *Wea. Forecasting,* **25,** 985–993.

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.

Ott, E., and Coauthors, 2004: A local ensemble Kalman filter for atmospheric data assimilation. *Tellus,* **56A,** 415–428.

Toth, Z., and E. Kalnay, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.,* **125,** 3297–3319.

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.

Xu, L., T. Rosmond, and R. Daley, 2005: Development of NAVDAS-AR: Formulation and initial tests of the linear problem. *Tellus,* **57A,** 546–559.

Zimin, A. V., I. Szunyogh, D. J. Patil, B. R. Hunt, and E. Ott, 2003: Extracting envelopes of Rossby wave packets. *Mon. Wea. Rev.,* **131,** 1011–1017.