

On Domain Localization in Ensemble-Based Kalman Filter Algorithms

TIJANA JANJIĆ AND LARS NERGER

Alfred Wegener Institute, Bremerhaven, Germany

ALBERTA ALBERTELLA

Institute for Astronomical und Physical Geodesy, Munich, Germany

JENS SCHRÖTER

Alfred Wegener Institute, Bremerhaven, Germany

SERGEY SKACHKO

Department of Earth Sciences, University of Quebec in Montreal, Montreal, Canada

(Manuscript received 1 July 2010, in final form 20 February 2011)

ABSTRACT

Ensemble Kalman filter methods are typically used in combination with one of two localization techniques. One technique is covariance localization, or direct forecast error localization, in which the ensemble-derived forecast error covariance matrix is Schur multiplied with a chosen correlation matrix. The second way of localization is by domain decomposition. Here, the assimilation is split into local domains in which the assimilation update is performed independently. Domain localization is frequently used in combination with filter algorithms that use the analysis error covariance matrix for the calculation of the gain like the ensemble transform Kalman filter (ETKF) and the singular evolutive interpolated Kalman filter (SEIK). However, since the local assimilations are performed independently, smoothness of the analysis fields across the subdomain boundaries becomes an issue of concern.

To address the problem of smoothness, an algorithm is introduced that uses domain localization in combination with a Schur product localization of the forecast error covariance matrix for each local subdomain. On a simple example, using the Lorenz-40 system, it is demonstrated that this modification can produce results comparable to those obtained with direct forecast error localization. In addition, these results are compared to the method that uses domain localization in combination with weighting of observations. In the simple example, the method using weighting of observations is less accurate than the new method, particularly if the observation errors are small.

Domain localization with weighting of observations is further examined in the case of assimilation of satellite data into the global finite-element ocean circulation model (FEOM) using the local SEIK filter. In this example, the use of observational weighting improves the accuracy of the analysis. In addition, depending on the correlation function used for weighting, the spectral properties of the solution can be improved.

1. Introduction

The ensemble-based Kalman filter approach has been widely used for data assimilation in both meteorology and oceanography (see, e.g., Houtekamer and Mitchell

1998, 2001; Brankart et al. 2003). In the ensemble Kalman filter algorithms, the forecast error covariance matrix is approximated by a covariance matrix whose rank is 1 less than the number of ensemble members. For computational tractability, the number of ensemble members, and therefore the rank of the covariance matrix is often chosen to be small. This, however, can lead to the divergence of the filter (Houtekamer and Mitchell 1998). To apply the ensemble Kalman filter methods in practice,

Corresponding author address: Tijana Janjić, Alfred Wegener Institute, Bussestrasse 24, D-27570 Bremerhaven, Germany.
E-mail: tijana.janjić.pfander@awi.de

it is therefore often necessary to localize the covariances in order to increase the rank of the analysis error covariance matrix.

In Houtekamer and Mitchell (1998, 2001) the ensemble Kalman filter was used together with a localization applied directly to the forecast error covariance matrix. In the latter paper, the ensemble derived forecast error covariance matrix is Schur multiplied (elementwise multiplied) with a stationary a priori chosen covariance matrix that is compactly supported.

As an alternative to direct forecast error localization, the method of domain localization has been used in several studies (Haugen and Evensen 2002; Penduff et al. 2002; Brusdal et al. 2003; Evensen 2003; Brankart et al. 2003; Ott et al. 2004; Nerger et al. 2006; Hunt et al. 2007; Miyoshi and Yamane 2007). In the domain localization methods, disjoint domains in the physical space are considered as domains on which the analysis is performed. An analysis step is performed independently for each subdomain, using observations not necessarily belonging only to that subdomain. Results of the local analysis steps are pasted together and then the global forecast step is performed.

These methods resemble optimal interpolation (OI), which is also based on a local analysis (Lorenc 1981; Cohn et al. 1998) and much can be learned by applying the substantial body of research that has been performed with OI. For example, it was shown that OI can produce spurious noise in the analysis fields as a result of different sets of observations used on different parts of the model state (Cohn et al. 1998).

In contrast to OI, covariances used in ensemble-based filters are ensemble derived, and therefore nonisotropic and nonstationary. In addition, the covariance used with OI would go to zero with increasing distance. This is not the case in the ensemble Kalman filter (EnKF) because of sampling error. Furthermore, these methods are used with weighting of the observations, or localization of the \mathbf{R} matrix, as described in section 2 (Penduff et al. 2002; Hunt et al. 2007; Nerger and Gregg 2007), which has not been done with OI.

In this work we investigate the impact of this weighting on the analysis and focus on explaining the effects of domain localization in ensemble-based Kalman filter algorithms. Furthermore, we discuss spectral properties of the solution depending on different weighting functions of the observations.

In section 2 we discuss domain localization in the context of ensemble-based Kalman filters and introduce a modification to the algorithm in order to include a Schur product with an isotropic matrix. Then, in section 3 we compare the different localization methods when applied to the Lorenz-40 system and show the beneficial effect of weighting of observations and the technique

introduced in section 2 together with domain localization for ensemble Kalman filter algorithms. In section 4, we apply a domain-localized singular evolutive interpolated Kalman filter (SEIK) (Pham et al. 1998; Pham 2001; Nerger et al. 2006) to a realistic oceanographic problem. Furthermore, we discuss the influence of the different correlation functions used for weighting of the observations within domain localization. We consider the effects on accuracy as well as on spectral properties of the solution. The concluding remarks are presented in the last section.

2. Domain localization

a. Why is domain localization used?

As for OI, one of the major advantages of using domain localization is computational since solving for the analysis update is not performed globally but on much smaller local domains. Accordingly, the updates on the smaller domains can be done independently and therefore in parallel. This gain comes at a price as shown by Cohn et al. (1998) by comparing the OI to global analysis: in case of the estimation of a smooth field, OI can produce boxiness in analysis fields because of its local approximation.

In ensemble-based Kalman filters, the domain localization method is more widely used with algorithms that use the analysis error covariance matrix for the calculation of the gain (Nerger et al. 2006; Hunt et al. 2007; Miyoshi and Yamane 2007) instead of the forecast error covariance matrix. Examples of such methods are the ensemble transform Kalman filter (ETKF; Bishop et al. 2001) and the SEIK (Pham et al. 1998; Pham 2001). These methods used for the analysis update the following formula:

$$\begin{aligned} \mathbf{x}_k^a &= \mathbf{x}_k^f + \mathbf{K}_k(\mathbf{y}_k^o - \mathbf{H}_k\mathbf{x}_k^f), \\ \mathbf{K}_k &= \mathbf{P}_k^a \mathbf{H}_k^T \mathbf{R}_k^{-1}. \end{aligned} \quad (1)$$

Here \mathbf{x}_k^a and \mathbf{x}_k^f represent the analysis and forecast state vector at time t_k , respectively. Observations are denoted with \mathbf{y}_k^o , the observational operator is denoted with \mathbf{H}_k , \mathbf{K}_k is the Kalman gain, \mathbf{P}_k^a is the analysis error covariance matrix, and \mathbf{R}_k denotes the observation error covariance matrix. In these methods an ensemble resampling [in SEIK (Pham 2001) or transformation (Bishop et al. 2001)] is used that ensures that the ensemble statistics represent exactly the analysis state and error covariance matrix. Once the ensemble members are generated, each member is propagated with the full numerical model to time t_{k+1} . The forecast state at time t_{k+1} is the average over the ensemble members. System noise can be added to each ensemble member (see Pham 2001), but it is neglected in the formulation presented in this paper.

In these algorithms, the forecast error covariance matrix is never explicitly calculated. Therefore, direct forecast localization as in Houtekamer and Mitchell (1998, 2001) is not immediately possible. To obtain the effect of direct forecast localization, a simple formula is required to calculate the analysis error covariance matrix corresponding to a localized forecast error covariance matrix in terms of the ensemble members. To this end, one could use the representation of the localized forecast covariance matrix in terms of ensemble members:

$$\mathbf{P}_k^f = \frac{1}{r} \sum_{i=1}^{r+1} [\mathbf{x}^{f,i}(t_k) - \bar{\mathbf{x}}_k^f][\mathbf{x}^{f,i}(t_k) - \bar{\mathbf{x}}_k^f]^T. \quad (2)$$

Here, $\mathbf{x}^{f,i}(t_k)$ are ensemble members $i = 1, \dots, r + 1$ of size n at time t_k and $\bar{\mathbf{x}}_k^f$ is the average over ensemble members: $\bar{\mathbf{x}}_k^f = 1/(r + 1) \sum_{i=1}^{r+1} \mathbf{x}^{f,i}(t_k)$. Now, let \mathbf{C} be a matrix of rank M that is used for the Schur product. Let \mathbf{v}_j represent eigenvectors of matrix \mathbf{C} multiplied with the square root of the corresponding eigenvalue. Then the localized error covariance matrix can be represented as

$$\sum_{i,j=1}^{r+1,M} \mathbf{u}_{i,j} \mathbf{u}_{i,j}^T \quad \text{with} \quad \mathbf{u}_{i,j} = \frac{1}{\sqrt{r}} [\mathbf{x}^{f,i}(t_k) - \bar{\mathbf{x}}_k^f] \circ \mathbf{v}_j, \quad (3)$$

where \circ denotes the element-wise product (Schur product) and the property of the Schur product $(\mathbf{a} \circ \mathbf{c})(\mathbf{b} \circ \mathbf{d})^T = (\mathbf{a}\mathbf{b}^T) \circ (\mathbf{c}\mathbf{d}^T)$ is used. This representation implies that one can use the ensemble $\mathbf{u}_{i,j}$ instead of the ensemble members $\mathbf{x}^{f,i}$ for the calculation of the analysis error covariance matrix with the same formulas as in the original algorithms (see the appendix for the detailed application with the SEIK algorithm). This procedure increases the cost of the algorithm for calculating the analysis error covariance matrix. The increase in cost depends on the rank of the chosen localization matrix \mathbf{C} . However, since the evolution of ensemble members is the most expensive part of these algorithms, this increase in cost may not be significant for low-rank covariance \mathbf{C} .

With this procedure the direct forecast localization is applied without any approximation in the analysis calculations of algorithms involving the analysis error covariance matrix. However, if the rank M of \mathbf{C} is larger than the ratio of the number of observations to the number of ensemble members, it is hard to computationally justify the use of this approach (see the appendix for details).

In the SEIK algorithm (similarly in the ETKF) a resampling of the ensemble is performed by a random transformation matrix under the conditions that the new ensemble mean equals the analysis state and the ensemble covariance matrix equals the analysis error covariance matrix. For the

SEIK filter, the method has been described as second-order exact sampling with linear constraints (Pham 2001). However, this resampling is only possible if the rank of the analysis covariance matrix obtained using the localized forecast error covariance matrix is smaller than the size $r + 1$ of the ensemble (Pham 2001). Since the rank of the analysis error covariance matrix has been most likely increased by the localization, one will not be able to recover this matrix exactly with $r + 1$ ensemble members. However, one can preserve the exact mean and some other prechosen properties of the covariance matrix. For example, one can obtain a new ensemble whose covariance matrix, as calculated from (2), is equal to the analysis error covariance matrix, which corresponds to the forecast error covariance matrix without localization.

b. Basic aspects of domain localization

Let us now consider what happens if one uses domain localization. In an algorithm with domain localization the forecast error covariance matrix is calculated using

$$\mathbf{P}_k^{f,\text{loc}} = \sum_{i,j=1}^{r+1,L} \mathbf{u}_{i,j} \mathbf{u}_{i,j}^T, \quad (4)$$

where $\mathbf{u}_{i,j} = (1/\sqrt{r})[\mathbf{x}^{f,i}(t_k) - \bar{\mathbf{x}}_k^f] \circ \mathbf{1}_{D_j}$ with $j = 1, \dots, L$ and L is the number of subdomains. Here $\mathbf{1}_{D_j}$ is a vector whose elements are 1 if the corresponding point belongs to the domain D_j otherwise they are 0. This forecast error covariance matrix has a block structure and is not isotropic. The localized forecast error covariance matrix in (4) is positive semidefinite because it is a sum of rank-1 positive semidefinite matrices. Equation (4) shows that the domain localization can be represented as a direct localization of a forecast error covariance matrix with a matrix \mathbf{C} that has block structure and is the sum of rank-1 matrices $\mathbf{1}_{D_j} \mathbf{1}_{D_j}^T$. The rank of matrix \mathbf{C} corresponds to the number of subdomains. In other words, in practice the subdomains may have to be made quite small, to ensure that $\text{rank}(\mathbf{C})$ is large enough. This is contrary to the direct forecast error localization where the localization matrix usually has full rank. The fact that \mathbf{C} is not positive definite but semidefinite for domain localization methods can theoretically lead to the following problem. In case that $\text{rank}(\mathbf{C})\text{rank}(\mathbf{P}_k^f) < n$, the matrix $\mathbf{C} \circ \mathbf{P}_k^f$ is singular (Horn and Johnson 1985). This is problematic, if the analysis increment is calculated for very accurate observations.

In domain localization methods, the rank is not increased locally on each subdomain. Accordingly, it is possible to resample exactly on that subdomain in contrast to direct forecast error localization. Because the assimilations are performed independently in each local

region, the smoothness of the analysis fields is of more concern in domain localization methods than with direct forecast error localization. In particular, two neighboring subdomains might produce strongly different analysis estimates when the assimilated observations have gaps, because distinct sets of observations are used for the analyses. To resolve the smoothness problem of the analysis fields, weighting of observations was proposed (Penduff et al. 2002; Hunt et al. 2007; Nerger 2007). This method weights observations depending on their distance from the point at which the assimilation is performed as will be discussed next.

c. Influence of the observation operator on domain localization

Ensemble-based Kalman filters apply the observation operator directly on each ensemble member before localization is applied. Therefore, the localization is usually performed on the matrices $\mathbf{H}_k \mathbf{P}_k^f$ and $\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T$ (Houtekamer and Mitchell 1998, 2001):

$$\mathbf{H}_k \mathbf{P}_k^f = \frac{1}{r} \sum_{i=1}^{r+1} \{ \mathbf{H}_k [\mathbf{x}^{f,i}(t_k)] - \mathbf{H}_k(\mathbf{x}_k^f) \} [\mathbf{x}^{f,i}(t_k) - \mathbf{x}_k^f]^T$$

$$\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T = \frac{1}{r} \sum_{i=1}^{r+1} \{ \mathbf{H}_k [\mathbf{x}^{f,i}(t_k)] - \mathbf{H}_k(\mathbf{x}_k^f) \} \times \{ \mathbf{H}_k [\mathbf{x}^{f,i}(t_k)] - \mathbf{H}_k(\mathbf{x}_k^f) \}^T.$$

Once these matrices are calculated, they are Schur multiplied with the matrices $\mathbf{H}_k \mathbf{C}$ and $\mathbf{H}_k \mathbf{C} \mathbf{H}_k^T$, respectively.

For the domain localization methods, different analysis results can be obtained depending on the treatment of the observations. Two different cases are of interest: Observations can be restricted to the local analysis subdomains only (Ott et al. 2004; Miyoshi and Yamane 2007), or the observational domains can be allowed to be larger and partially overlap (Penduff et al. 2002; Nerger et al. 2006; Hunt et al. 2007). If all the observations in the full domain are used for the analysis in each disjoint subdomain, the algorithm without localization is recovered. This follows from

$$\frac{1}{r} \sum_{i=1}^{r+1} \sum_{j=1}^L \{ \mathbf{H}_k [\mathbf{x}^{f,i}(t_k)] - \mathbf{H}_k(\mathbf{x}_k^f) \} [\mathbf{x}^{f,i}(t_k) \circ \mathbf{1}_{D_j} - \mathbf{x}_k^f \circ \mathbf{1}_{D_j}]^T$$

$$= \frac{1}{r} \sum_{i=1}^{r+1} \{ \mathbf{H}_k [\mathbf{x}^{f,i}(t_k)] - \mathbf{H}_k(\mathbf{x}_k^f) \} [\mathbf{x}^{f,i}(t_k) - \mathbf{x}_k^f]^T = \mathbf{H}_k \mathbf{P}_k^f. \tag{5}$$

If, on the other hand, we restrict observations to the local analysis subdomains the covariance matrix is given by (4).

Let us now consider observations that are limited to a domain D_{m_j} larger than the subdomain D_j on which the analysis is performed. Let $\mathbf{1}_{D_{m_j}}$ be a vector that has a value of 1 if the observation belongs to the domain D_{m_j} and is 0 otherwise. Furthermore, let $D_j \subseteq D_{m_j}$. Again, using the property of the Schur product ($\mathbf{a} \circ \mathbf{c}$) ($\mathbf{b} \circ \mathbf{d}$)^T = ($\mathbf{a} \mathbf{b}^T$) \circ ($\mathbf{c} \mathbf{d}^T$), it is

$$\frac{1}{r} \sum_{i=1}^{r+1} \sum_{j=1}^L \{ \mathbf{H}_k [\mathbf{x}^{f,i}(t_k)] \circ \mathbf{1}_{D_{m_j}} - \mathbf{H}_k(\mathbf{x}_k^f) \circ \mathbf{1}_{D_{m_j}} \} [\mathbf{x}^{f,i}(t_k) \circ \mathbf{1}_{D_j} - \mathbf{x}_k^f \circ \mathbf{1}_{D_j}]^T$$

$$= \frac{1}{r} \sum_{i=1}^{r+1} \sum_{j=1}^L (\mathbf{1}_{D_{m_j}} \mathbf{1}_{D_j}^T) \circ \{ \mathbf{H}_k [\mathbf{x}^{f,i}(t_k)] - \mathbf{H}_k(\mathbf{x}_k^f) \} [\mathbf{x}^{f,i}(t_k) - \mathbf{x}_k^f]^T = \sum_{j=1}^L (\mathbf{1}_{D_{m_j}} \mathbf{1}_{D_j}^T) \circ \mathbf{H}_k \mathbf{P}_k^f.$$

Here, the matrix $\sum_{j=1}^L \mathbf{1}_{D_{m_j}} \mathbf{1}_{D_j}^T$ has entries of 0s and 1s, because the domains D_j are disjoint. The number of diagonals with all values equal to 1, depends on the number of observations in the overlapping subdomains D_{m_j} . In case of domain localization, the analysis in a local subdomain D_j is computed using the corresponding submatrices of $(\mathbf{1}_{D_{m_j}} \mathbf{1}_{D_j}^T) \circ \mathbf{H}_k \mathbf{P}_k^f$ and $\mathbf{1}_{D_{m_j}} \mathbf{1}_{D_j}^T \circ \mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T$.

An obvious modification to this algorithm is to use the matrices $(\mathbf{1}_{D_{m_j}} \mathbf{1}_{D_j}^T) \circ \mathbf{H}_k \mathbf{P}_k^f \circ \mathbf{H}_k \mathbf{C}$ and $\mathbf{1}_{D_{m_j}} \mathbf{1}_{D_j}^T \circ \mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T \circ \mathbf{H}_k \mathbf{C} \mathbf{H}_k^T$. Here, the matrix \mathbf{C} is generated with a support

distance chosen to be not larger than the distance corresponding to the support distance of $\mathbf{1}_{D_{m_j}} \mathbf{1}_{D_j}^T$. Note that we cannot recover the direct forecast error localization by this method, because the solving for analysis in the direct forecast localization is performed globally. Performing the analysis step domain by domain, gives the same result only for block structure matrices, such as those obtained when observations belong to disjoint subdomains. The proposed modification keeps the computational benefits of domain localization, but also

introduces the full rank, positive definite covariance matrix \mathbf{C} to the domain localization method. In the resampling step, however, one generates on each subdomain a local ensemble to represent the analysis covariance matrix corresponding to the forecast covariance matrix without localization with \mathbf{C} . This method will be denoted (SD+Loc) in the next section.

An alternative to this approach is weighting of the observation (Penduff et al. 2002; Hunt et al. 2007; Nerger and Gregg 2007). Its implementation requires for each observation a weight that depends on the distance of the observation from the analysis location. The forecast error covariance matrix does not need to be computed explicitly. The weighting of observations modifies only the observational error covariance matrix \mathbf{R} . With observation localization, the analysis ensemble is generated to represent the analysis covariance calculated with the modified matrix \mathbf{R} . This method will be called method (SD+ObLoc) below. The direct forecast error localization and the weighting of observations are not equivalent, as pointed out by Miyoshi and Yamane (2007) on a simple example. They considered the weighting when only a single observation is assimilated. In this case, the observation error σ_{obs}^2 is modified to $\sigma_{\text{obs}}^2/w_d$, where w_d is a weight that depends on the distance of the observation from the analysis point and can be calculated using any correlation function. Accordingly, in case of weighting of observations, the gain is $w_d(HP^f)^T/(w_dHP^fH^T + \sigma_{\text{obs}}^2)$. In contrast, for direct forecast error localization the gain is $w_d(HP^f)^T/(HP^fH^T + \sigma_{\text{obs}}^2)$. In both methods the same correlation function can be used for w_d . Note, that it is not possible to find two different correlation functions depending only on distance that yield the same gain for both localization methods. The example also indicates that the differences between the method of weighting of observations and direct forecast error localization method will be strongest if the observational error is small compared to HP^fH^T .

3. Experiments with the 40-variable Lorenz model

To examine the localization methods discussed in the previous sections with a small dynamical model, data assimilation experiments with the Lorenz-40 dynamical system of Lorenz and Emanuel (1998) were performed. This nonlinear model has been used to assess ensemble-based assimilation schemes in a number of studies (e.g., Whitaker and Hamill 2002; Ott et al. 2004; Sakov and Oke 2008). The model is governed by 40 coupled ordinary differential equations in a domain with cyclic boundary conditions. The state vector dimension is 40 and the fourth-order Runge–Kutta time integration scheme is used with a time step of 0.05 nondimensional units.

The full state is observed at every time step. The observations are generated from a model trajectory by contaminating the state by uncorrelated normally distributed random noise. For the first set of experiments the observation error has a standard deviation of $\sigma_{\text{obs}} = 1$, while $\sigma_{\text{obs}} = 0.1$ is used for the second set of experiments.

The assimilation is performed over 50 000 time steps after a spinup period of 1000 time steps. A 10-member ensemble is used. To initialize the filter, a sample is generated from the true model trajectory over 60 000 time steps. The initial covariance matrix is estimated via empirical orthogonal function (EOF) analysis such that it corresponds to the nine largest singular values of the computed sample covariance matrix. The initial ensemble is generated by second-order exact sampling. As the performance of the data assimilation algorithms depends on random numbers used in the initial sampling, all experiments are repeated 10 times with different random numbers. The assimilation performance is assessed in terms of root-mean-square (RMS) errors. These are calculated over 50 000 time steps for varying localization radii and forgetting factors (see the appendix for algorithmic details including the forgetting factor) and averaged over 10 experiments with different random numbers.

The following assimilation methods are compared. First, an ensemble square root filter as described in Whitaker and Hamill (2002) is applied. The localization is performed by a Schur product of the ensemble forecast error covariance matrix with a correlation matrix \mathbf{C} given by the piecewise rational fifth-order polynomial function (Gaspari and Cohn 1999). This method will be labeled (GLocEn). The other methods are applied with domain localization (i.e., separate analysis updates for each grid point). In the method labeled (SD+) point-localized assimilation updates are performed using the local SEIK filter and observations within a varying observational radius. This corresponds to the use of the top-hat function for the matrices $\mathbf{H}_k\mathbf{C}$ and $\mathbf{H}_k\mathbf{C}\mathbf{H}_k^T$. The local SEIK filter with observational weighting by the fifth-order polynomial function is used for the method labeled (SD+ObLoc). To apply this method, all 40 available observations are used and the weighting is varied by specifying the support radius for the fifth-order polynomial function. For support radii below 20, this configuration is analogous to setting the influence radius of the observations to the support radius. This is due to the fact that the method would disregard all observations for which the weight is equal to zero. Finally, the method labeled (SD+Loc) uses for each observational subdomain a reformulated local SEIK filter with localization by the Schur product of the local forecast error covariance matrix with the corresponding submatrix of matrix \mathbf{C} . In

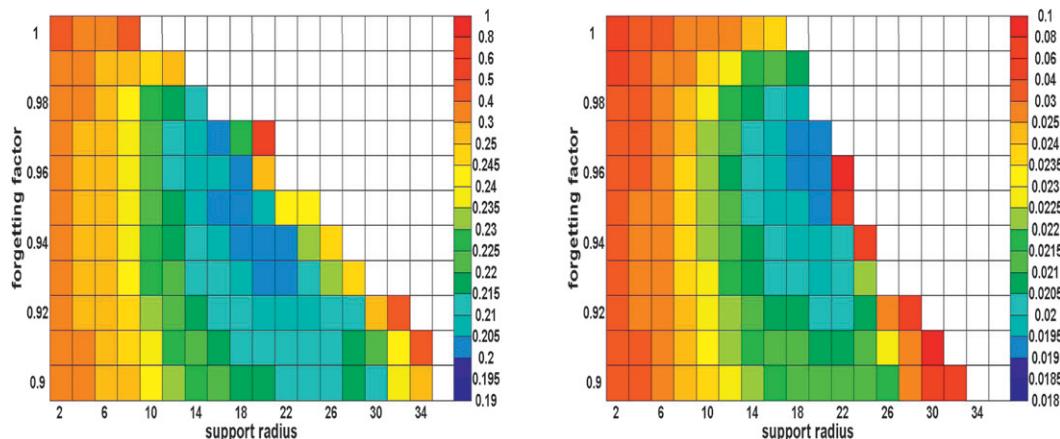


FIG. 1. RMS mean error as a function of the support radius of the localization and the forgetting factor for method GLocEn. Errors obtained with an observational error with standard deviation of (left) $\sigma_{\text{obs}} = 1$ and (right) $\sigma_{\text{obs}} = 0.1$. Please note the different color scales.

this method, the resampling is performed locally using the analysis error covariance matrix corresponding to the forecast error covariance matrix without localization by **C** on the subdomain.

The RMS error for the method (GLocEn) averaged over 10 experiments with different random numbers in the initialization is shown in Fig. 1 as a function of the support radius of the fifth order polynomial and the forgetting factor. Figure 1 shows RMS errors for (left) $\sigma_{\text{obs}} = 1$ and (right) $\sigma_{\text{obs}} = 0.1$. White regions denote parameter values for which the filter diverges, which is defined for errors above 1 for $\sigma_{\text{obs}} = 1$ and above 0.1 for $\sigma_{\text{obs}} = 0.1$. For $\sigma_{\text{obs}} = 1$, the smallest mean RMS error of 0.202 is obtained for a support radius of 18 and a forgetting factor of 0.95. The minimum value as well as the corresponding support radius and forgetting factor vary between the 10 experiments with different random numbers. The optimal support radius and forgetting factor are also influenced by the observation error. For $\sigma_{\text{obs}} = 0.1$ the minimum mean RMS error is 0.0194. It is obtained for a support radius of 20 and a forgetting factor of 0.96.

Figure 2 shows mean RMS errors for the methods using domain localization for the two observational error levels. The top row shows the RMS errors for the method (SD+). The middle row shows corresponding results for the method (SD+ObLoc) and the last row for the method (SD+Loc) with a support radius of 20 grid points for the localization matrix **C**. Let us first compare the methods for the case of $\sigma_{\text{obs}} = 1$. The parameter area corresponding to the convergence of method (SD+) is much smaller than for the other methods. The use of all observations in the method (SD+) corresponds to the use of the SEIK filter without localization, which diverges

for all forgetting factors used here. The smallest mean RMS error for the method (SD+) is 0.220. It is obtained when the observational radius is 6 and the forgetting factor is 0.93.

Weighting of observations increases the range of parameters for which the filter converges. The most accurate results are obtained for an observational radius of 20 and a forgetting factor of 0.93. Here, the minimum error is slightly larger compared to the minimum error of (GLocEn) and its value is 0.203.

Using the method (SD+Loc) for a localization matrix **C** with support radius of 20 grid points, a minimum RMS error value of 0.197 is obtained. The optimal parameters are an observational radius of 10 and a forgetting factor of 0.95. Finally, results obtained with the method (SD+Loc) are shown for a localization matrix **C** with support radius of 20 grid points and varying observational radius and forgetting factor. The accuracy of the results depends on the support radius of the localization matrix **C**. Figure 2 shows the optimal result for $\sigma_{\text{obs}} = 1$. Changing the support radius of the localization matrix **C** will change the range of the parameters for which convergence is achieved, as well as the RMS errors.

If the observational error is decreased to $\sigma_{\text{obs}} = 0.1$, the differences between the methods (GLocEn) and (SD+ObLoc) become evident. The minimum mean RMS errors are 0.0194 for (GLocEn), 0.0220 for (SD+), 0.0205 for (SD+ObLoc), and 0.0188 for (SD+Loc). The methods with the smallest mean RMS errors are (GLocEn) and (SD+Loc). The minimum RMS error obtained for the method (SD+ObLoc) is larger than that for (GLocEn). This difference is due to the different effects of both localization methods as was discussed toward the end of section 2c. However, if one excludes the initial transient

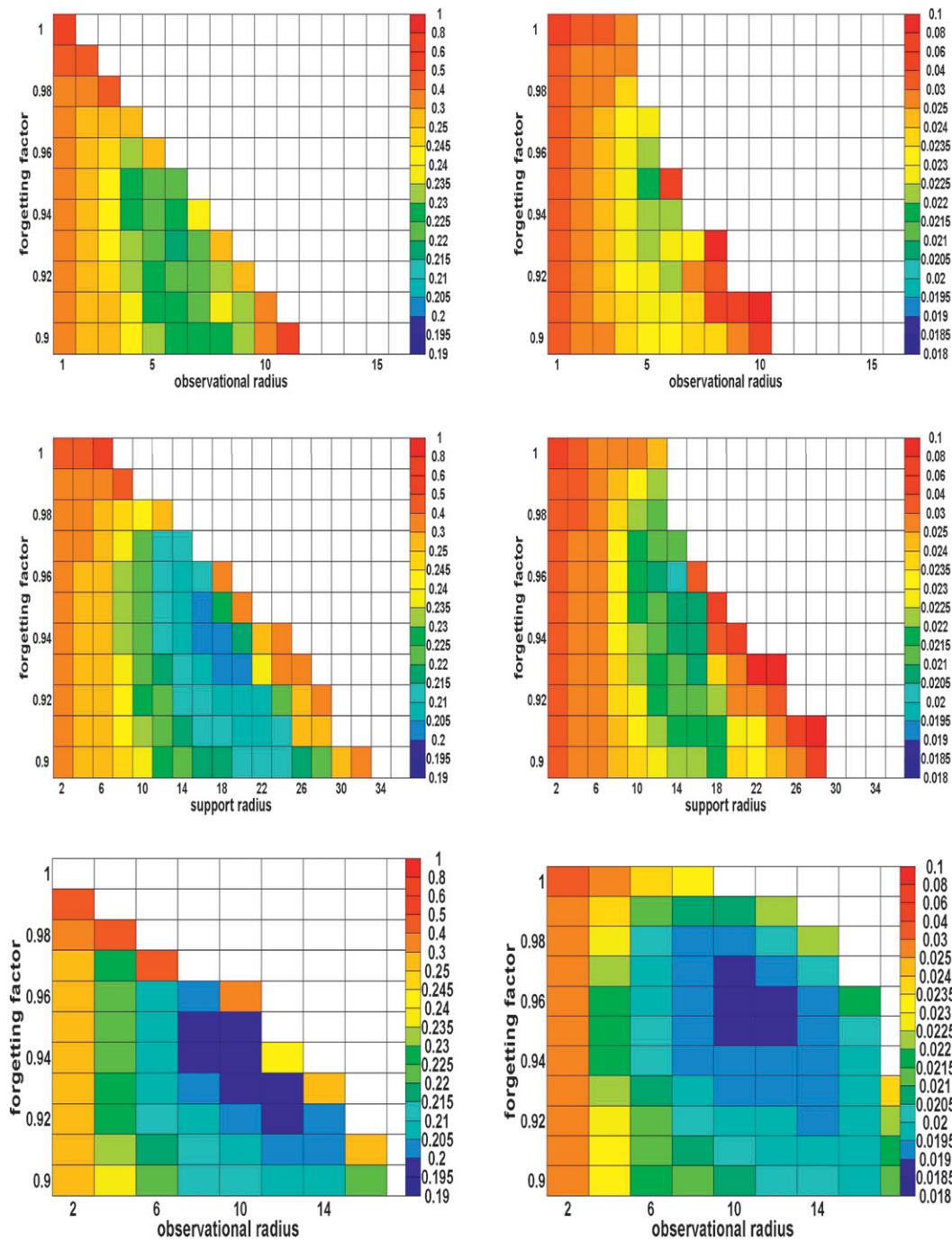


FIG. 2. RMS mean errors as in Fig. 1. Shown are the experiments (top) SD+, (middle) SD+ObLoc, and (bottom) SD+Loc for (left) $\sigma_{\text{obs}} = 1$ and (right) $\sigma_{\text{obs}} = 0.1$.

phase of the assimilation the errors obtained for the two methods are much closer. This is visible in Fig. 3 where the RMS errors are calculated using only the last 48 000 steps of each assimilation experiment. The smaller differences indicate that the asymptotic behavior of both methods is similar. To assess the statistical significance

of the results, t tests were performed. Tested was the hypothesis that the minimum errors obtained over 10 random-number experiments are different. The tests showed that the RMS errors for (GLocEn) and (SD+ObLoc) are significantly different, at 5% significance level, for the mean errors over 48 000 time steps for

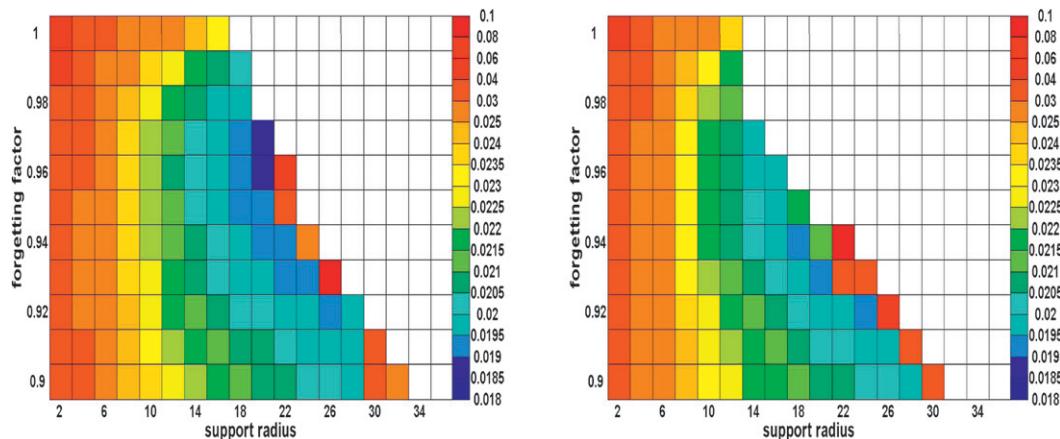


FIG. 3. RMS mean errors for the methods (left) GLocEn and (right) SD+ObLoc for the last 48 000 time steps of experiments with $\sigma_{\text{obs}} = 0.1$.

both $\sigma_{\text{obs}} = 0.1$ and $\sigma_{\text{obs}} = 1$. In addition, the RMS errors are significantly different over 50 000 time steps for $\sigma_{\text{obs}} = 0.1$, but not for $\sigma_{\text{obs}} = 1$.

4. Experiments with a global ocean model

In the previous section we have discussed the performance of the different localization schemes when applied to a small strongly nonlinear dynamical system. In this section, we apply the localization to a realistic oceanographic problem: the assimilation of satellite data into a global ocean circulation model. We focus on the method of weighting of observations with local SEIK filter and study the filtering behavior when different correlation functions for the weighting of observations are applied. The results are assessed in terms of the accuracy of the analysis and forecast fields as well as the spectral properties of the error of the estimated fields.

a. Experimental setup

The experiments conducted here are analogous to Skachko et al. (2008) with a few modifications. Skachko et al. (2008) updated only the sea surface height (SSH) field with a local SEIK filter, while temperature and salinity are updated following the first baroclinic mode in the vertical direction. In contrast, we use the domain localized SEIK filter algorithm (Pham et al. 1998; Pham 2001; Nerger et al. 2006) as implemented within the parallel data assimilation framework (PDAF; Nerger et al. 2005) to update the full model state, consisting of temperature, salinity, SSH, and velocity fields. Furthermore, a model version without adiabatic pressure correction is used, while Skachko et al. (2008) used this correction in order to reduce a systematic drift of the mean surface elevation.

The assimilated observations are dynamical ocean topography (DOT) data. The DOT was obtained by means of a geodetic approach from carefully cross-calibrated multimission-altimetry data [from the *Environmental Satellite (Envisat)*, the *Geosat Follow-On (GFO)*, Jason, and the Ocean Topography Experiment (TOPEX)/Poseidon missions] and the Gravity Recovery and Climate Experiment (GRACE) gravity fields [see Skachko et al. (2008) for more details]. The geoid data are provided as a truncated spherical harmonic series (i.e., in a band-limited global spectral representation on a sphere). The altimetric measurements are given as weighted mean values over the footprint of the radar signal along the ground track of the spacecraft. To ensure that the computed DOT is consistent with the geoid field, spectral consistency between sea surface height fields and the geoid field needs to be achieved (Albertella et al. 2008). The consistency is ensured by applying a Gauss-type filter (Jekeli 1981; Wahr et al. 1998). The filter length is driven by the spatial resolution of the gravity field. For the GRACE-based gravity field model ITG03s (Mayer-Gürr 2007), the filter half-width is typically 240 km. The polar areas, part of the Indonesian region, and the Mediterranean Sea are characterized by low data accuracy due to the presence of ice or complex bottom topography. For this reason, the observational data are substituted in these regions by the values of the (Rio et al. 2005, hereafter RIO05) mean dynamical topography. The data cover the period between January 2004 and January 2005 and are available at 10-day intervals. For combining the different data types, the data are interpolated onto the model grid [see Skachko et al. (2008) and Albertella et al. (2008) for more details]. The interpolation introduces cross correlations between the observations. However, since the same observations are

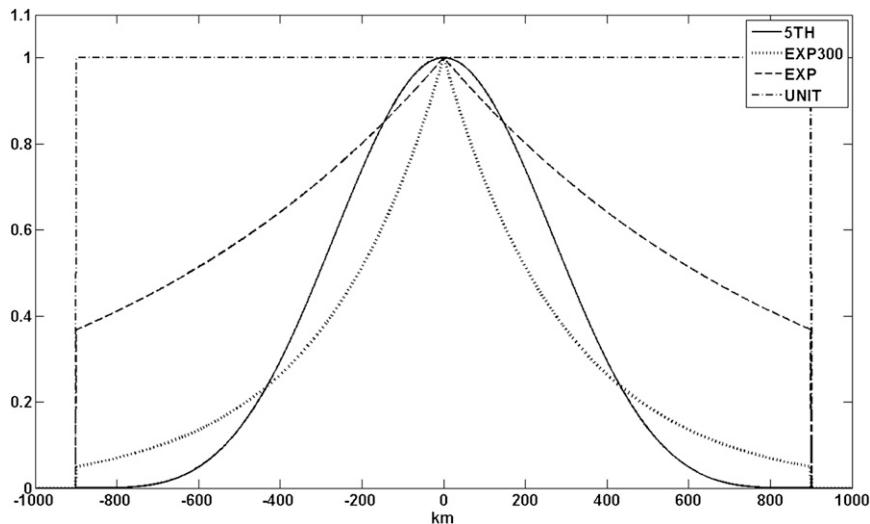


FIG. 4. Correlation functions used for the weighting of observations. The exponential weighting gives more weight to remote observations and less for short distances than the use of the fifth-order polynomial.

used in all experiments presented below and all the observations are contaminated with observational errors in the same way, these errors will not affect the comparison of the weighting functions discussed here.

The study was performed using the finite-element ocean circulation model (FEOM; Wang et al. 2008) configured on a global triangular mesh with the spatial resolution of 1.5° . There are 24 unevenly spaced levels in the vertical direction. FEOM solves the standard set of hydrostatic ocean dynamic primitive equations using continuous linear representations for the horizontal velocity, surface elevation, temperature, and salinity.

To generate the initial error covariance matrix, a model run was performed to produce a set of 10-day forecasts over a year. The initial error covariance matrix is then approximated with a low-rank matrix using the first 12 empirical orthogonal functions of the ensemble. The initial field was the same as in Skachko et al. (2008). The 12 EOFs and the initial state estimate were used to generate the initial ensemble of 13 model states by second-order exact sampling (Pham 2001).

During the assimilation, the analysis of the full ocean field is carried out by applying the local SEIK filter after each 10-day forecast. The analysis for each water column of the model depends only on observations within a specified influence region. In the analysis, a diagonal observation error covariance matrix is used with an error variance of 25 cm^2 analogously to similar studies performed with reduced rank filters (e.g., Penduff et al. 2002). The error variance accounts for the error of altimetry and gravity data as well as mapping errors. Also,

it can partially account for the effects of cross correlations introduced by interpolation (Janjić and Cohn 2006).

b. Influence of the weighting function on estimation errors

Although localization has been widely used in combination with ensemble-based Kalman filters, the reasons for choosing a particular correlation function for localization are not well established. The localization length scales should be related to the scale of the SSH features, which vary in the global ocean depending on latitude and dynamic regime. Currently, ensemble Kalman filter methods use an isotropic correlation function for localization, and rely on ensemble-derived covariances from the model to represent the dynamical properties of the system. For this application, a localization function with length scales varying with the latitude might be beneficial. However, the experiments performed here are concerned with the general differences induced by the weighting functions. For this reason, isotropic correlation functions are used here for localization.

Besides the fifth-order polynomial function an exponential correlation function (Gelb 1974) has been used for weighting observations (Penduff et al. 2002; Nerger and Gregg 2007). In this section, experiments with different correlation functions for the weighting of observations are discussed. Considered are the fifth-order polynomial correlation function from Gaspari and Cohn (1999) with support radius equal to the observational radius (5TH), uniform weighting by one (UNIT), and exponential weighting $e^{-d/L}$ (EXP). Here, d is the distance

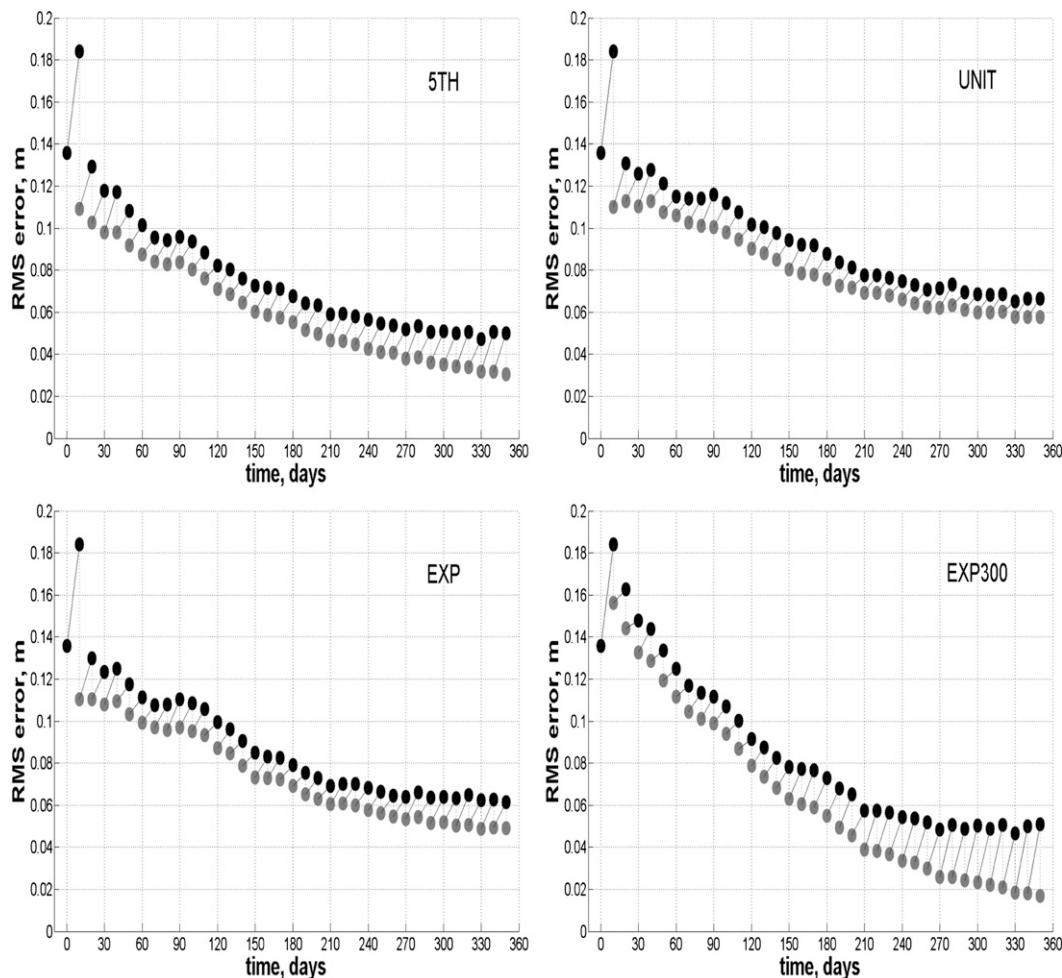


FIG. 5. Evolution of the RMS error of SSH for the global ocean (except zones corresponding to RIO05 data) for the experiments (top left) 5TH, (top right) UNIT, (bottom left) EXP, and (bottom right) EXP300. Black and gray dots represent forecast and analysis errors, respectively. Thin lines connect an analysis error with the subsequent forecast error.

between the analysis point and the observation, while L is the length scale of the weighting. The localization parameters for all three methods have been tuned in an extensive series of tests. In these tests, the observational radius was varied between 2000 and 300 km. For the experiment 5TH, the experiments indicated that an observational radius of 900 km represents an optimal choice to include as many observations as possible and still obtain reasonably accurate results. For the experiments EXP and UNIT we compare the radius of 900 km with smaller radii. The correlation functions used here are shown in Fig. 4. For the detailed discussion, we only consider a selection of relevant cases. In particular, we consider exponential weighting with $L = 900$ km (EXP) and $L = 300$ km (EXP300) as well as unit weighting with radii of 900 km (UNIT), 500 km (UNIT500), and 300 km (UNIT300).

The RMS errors of the SSH field over time obtained with four weighting functions are shown in Fig. 5. After 1 year of assimilation, the experiments 5TH and EXP300 give the most accurate solution with about 0.05-m forecast RMS error, followed by the experiment EXP. Note that the experiment EXP300 shows the largest spread between the analysis and forecast RMS errors. This is the result of the fast decrease of the exponential function for short distances. Thus, the forecast is less stable than in the other experiments due to a higher imbalance (Mitchell et al. 2002; Kepert 2009). Consistent with the experiments performed with the Lorenz-40 model, the application of observational weighting functions improves the accuracy of the forecast field compared to uniform weighting.

The previous experiments showed that for exponential weighting the same forecast accuracy can be obtained as for the experiment 5TH. This, however, required

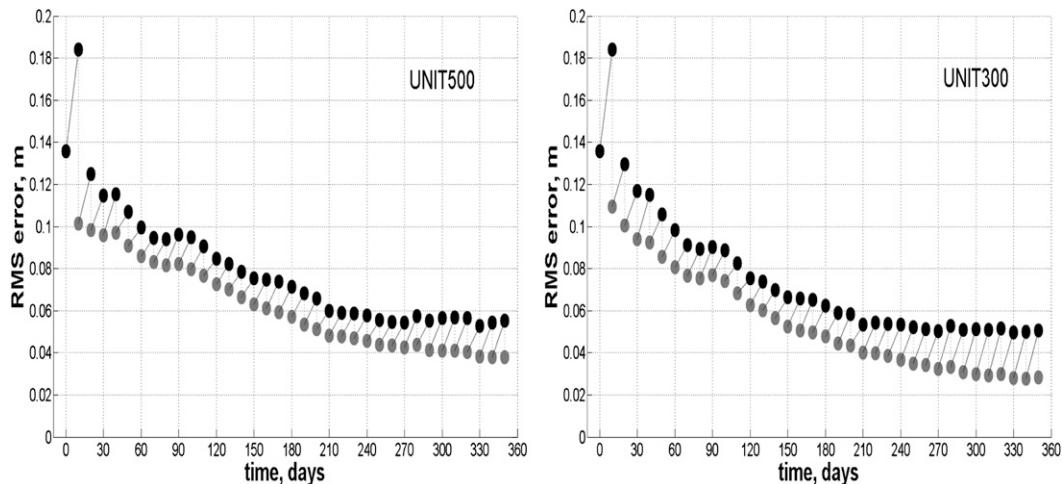


FIG. 6. RMS error of SSH for the global ocean as in Fig. 5 for assimilation with domain localization and unit weighting of observations. The support radius is (left) 500 and (right) 300 km.

a smaller radius L than the observational influence radius. A similar effect can be demonstrated for uniform weighting. As here the weight is constant within the observational radius, the performance is only influenced by the observational radius. Figure 6 shows RMS errors for the SSH fields depending on the observational influence radius for uniform weighting of observations. Comparing Figs. 5 and 6 shows that one can obtain the same accuracy with a 900-km radius in the experiment 5TH, as with an influence radius of 300 km and uniform weighting (UNIT300). However, the spread between the forecast and analysis errors increases as the influence radius is decreased, indicating larger imbalances. This is similar to the experiment EXP300, but for UNIT300 the spread is slightly smaller. As a result of this, the accuracy of the forecasted field is the same for 5TH and UNIT300, while for the analysis the RMS error is smaller for UNIT300. Furthermore, only by disregarding a relatively large number of observations one can obtain estimates that are comparable to the experiment 5TH. This is not a desirable property of the data assimilation scheme.

c. Spectral properties of the estimated fields

The previous section has shown that the modification of the observation error covariance matrix by different weighting functions affects the accuracy of the estimates. We now consider whether the improved accuracy results from additional filtering introduced by the modification of the observation error covariance matrix. For this purpose, we examine spectral properties of the estimation errors depending on the type of the observational weighting function.

Using the global approach (Albertella et al. 2008), the geodetic DOT can be computed on the entire surface of the earth on an equiangular grid of $1^\circ \times 1^\circ$. With this approach, the land areas are filled with an arbitrary function and an iterative procedure is used to smooth the field over land and the land–ocean transition. Using the same procedure, the forecast and analysis fields are extended over the entire earth’s surface. Spherical harmonic analysis up to spherical harmonic degree 180 can then be applied to obtain the harmonic spectrum of each field.

In this study, we only consider the error in the mean DOT obtained by averaging over 10-day outputs. Figure 7 shows the spectral distribution of the mean data field up to spherical harmonic degree $\ell = 45$. In addition, the spectral distribution in terms of the degree ℓ is shown. It is given by $\sum_m (T_{\ell m}^o)^2$ where $T_{\ell m}^o$ are spherical harmonic coefficients of the data. As can be seen from Fig. 7, the mean field has nonzero spectral coefficients between order -10 and 10 and up to spherical harmonic degree 35. The spectral coefficients are negligible beyond these degree and orders.

The spectral error distributions obtained with the four weighting functions displayed in Fig. 4 are shown as a function of the degree in Fig. 8. Namely, the difference

$$\epsilon_\ell^{oi} = \sum_m (T_{\ell m}^o - T_{\ell m}^i)^2$$

is shown as a function of the harmonic degree. Here $T_{\ell m}^i$ are spectral coefficients of the model field and i can be $i = a$ for the analysis result (left panel of Fig. 8) or $i = f$ for the forecast result (right panel). The distribution of ϵ_ℓ^{oi} for the forecast field shows a similar structure as for the analysis. However, the amplitudes are larger for the

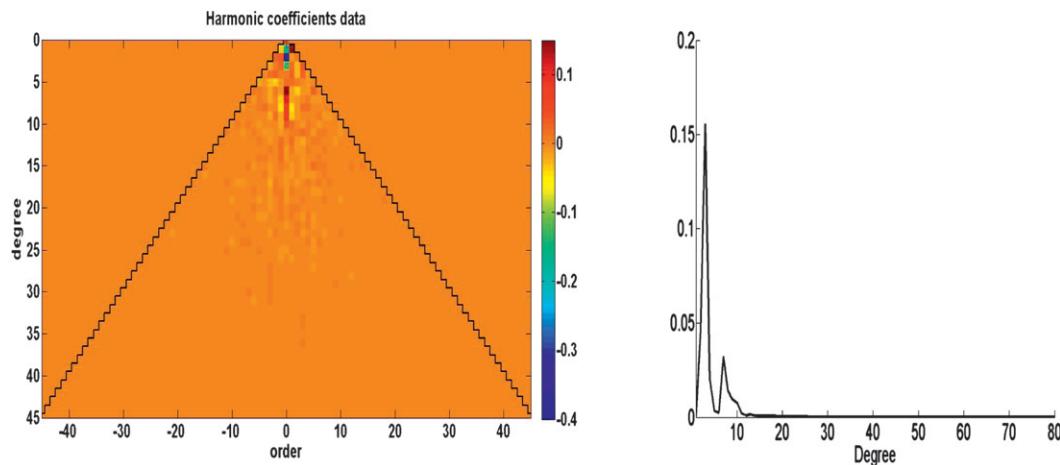


FIG. 7. (left) Logarithmic (base 10) spectral distribution of the mean data field plotted up to spherical harmonic degree 45. (right) Spectral distribution in terms of spherical harmonic degree.

forecast for all methods. Although the RMS errors in the cases EXP and UNIT are similar, the spectral structures show obvious differences. In the case EXP, the error is reduced for larger degrees and increased for smaller degrees (up to spherical harmonic degree 12) compared to the case UNIT. This difference is even more pronounced for the case EXP300. The spectral properties of the error for the methods UNIT, EXP, and EXP300 show that the cutoff distance rather than the structure of the correlation function has a dominating effect on the spectral error distributions above spherical harmonic degree 12. For smaller degrees the structure of the correlation function seems to be dominating. The case 5TH shows relatively evenly distributed error structures for all scales. Note, that the data alone have spectral coefficients only up to spherical harmonic degree 35. Thus, errors above this degree were introduced by the analysis scheme

and were further amplified by the forecast. For the experiments EXP, EXP300, and 5TH, these errors are smaller than for the experiment UNIT. Therefore, the improvement in the accuracy is the result of additional filtering introduced by the modification of the correlation function for the method of weighting of observations.

5. Conclusions

Ensemble Kalman filter methods are typically used with one of two localization techniques: domain localization or direct forecast error localization. For domain localization, the assimilation is split into local domains in which the assimilation updates are performed independently using observations within a chosen distance. Weighting of the observations can be used together with domain localization. Here, domain localization has been

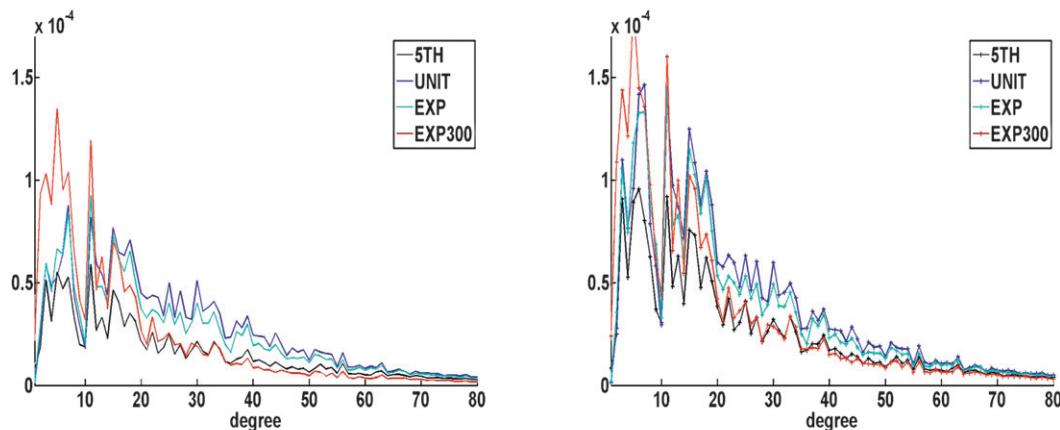


FIG. 8. Logarithm of the spectral difference between (left) analysis and the data and (right) forecast and the data depending on spherical harmonic degree.

investigated and compared to direct forecast error localization. The direct forecast error localization method applies a Schur product of the ensemble forecast covariance matrix with a localization matrix given by a chosen correlation matrix. It was shown that the domain localization is equivalent to the direct forecast error localization using a localization matrix that is positive semidefinite, has block structure and is not isotropic. The rank of this localization matrix depends on the number of subdomains that are used in the assimilation implying that subdomains may have to be made quite small, to ensure that the rank is large enough. In contrast, the matrix usually used in direct forecast error localization has full rank and is positive definite and isotropic.

A new algorithm was presented that introduces a full-rank positive definite matrix in domain localization methods. For each local analysis subdomain, it uses observations from a region that is larger than this subdomain. This is combined with a Schur product with an isotropic matrix for each local forecast covariance matrix (SD+Loc). Numerical experiments using the Lorenz-40 dynamical model showed that the errors obtained with this method are comparable to the direct forecast localization technique. Furthermore, these results were compared to a method of weighting of observations (SD+ObsLoc). In case of the Lorenz-40 model, the weighting of observations was less accurate, than the new method (SD+Loc). However, both ways of introducing the full-rank positive definite matrix in the domain localization algorithms, proved to be beneficial in comparison to the use of domain localization alone as in (SD+). The comparison of the localization methods has been performed here with the highly nonlinear Lorenz-40 model. With a smoother model, the differences in the relative performances might be smaller.

The method of weighting of observations was further examined in the case of assimilation of satellite data into the global finite element ocean model FEOM. Also in this example, the use of the full-rank matrix with observational weighting improves the accuracy of the analysis. Weighting of observations by different correlation functions showed that fifth-order polynomial weighting produced the most accurate forecast results compared to uniform weighting and exponential weighting. In addition, depending on the correlation function chosen for weighting, the spectral properties of the solution can be improved. Weighting of the observations by the fifth-order polynomial produced spectral results that are closest to the data. Note, that the fifth-order piecewise rational function that vanishes at the distance of 900 km is compactly supported and approximates a Gaussian function with a length scale of 246.5 km

(Gaspari and Cohn 1999). Thus, it corresponds well to the accuracy of the spherical harmonic degree 60 of the geoid model. Although the accuracy with exponential weighting and uniform weighting of observations are similar for the forecasted field, the spectral structures show that using the exponential weighting reduces the error for larger degrees and increases the error for smaller degrees compared to uniform weighting. The weighting by fifth-order polynomial shows relatively evenly distributed error structures for all scales.

Acknowledgments. The authors are grateful to Dr. H. Mitchell as well as anonymous reviewers for useful remarks and suggestions. This work has been partially funded under DFG Priority Research Programme SPP 1257 ‘‘Mass Transport and Mass Distribution in the Earth System.’’

APPENDIX

Use of Direct Forecast Localization in the SEIK Algorithm

Here we demonstrate the application of direct forecast localization in the analysis step of the SEIK filter algorithm. In the SEIK algorithm (Pham 2001), the vector \mathbf{x}_k^f of size n is calculated as the average over the ensemble members $\mathbf{x}_k^{f,i}$ with $i = 1, \dots, r + 1$. The $\mathbf{P}_k^{f,SEIK}$ is the ensemble covariance matrix in (2). Since the matrix $\mathbf{P}_k^{f,SEIK}$ has rank r , it can be represented in a factorized form (Pham 2001) as

$$\mathbf{P}_k^{f,SEIK} = \mathbf{L}_k [(r + 1)\mathbf{T}^T\mathbf{T}]^{-1} \mathbf{L}_k^T, \quad (\text{A1})$$

where \mathbf{L}_k is a matrix of size $n \times r$ containing in each column the difference between ensemble members $i = 1, \dots, r$ and the mean \mathbf{x}_k^f . The matrix \mathbf{T} is of size $(r + 1) \times r$, with values $1 - 1/(r + 1)$ on the diagonal and $-1/(r + 1)$ in the off-diagonal elements and the lowermost row. The analysis covariance matrix $\mathbf{P}_k^{a,SEIK}$ is calculated from the factorized form of the forecast error covariance matrix by

$$\mathbf{P}_k^{a,SEIK} = \mathbf{L}_k \mathbf{U}_k \mathbf{L}_k^T, \quad (\text{A2})$$

where \mathbf{U}_k is of size $r \times r$ and given by

$$\mathbf{U}_k = \{\rho[r\mathbf{T}^T\mathbf{T}]^{-1} + (\mathbf{H}_k \mathbf{L}_k)^T \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{L}_k\}^{-1}. \quad (\text{A3})$$

Here, ρ with $0 < \rho \leq 1$ denotes the forgetting factor. It serves to inflate the estimated forecast covariance matrix. It is the inverse of the ‘‘covariance inflation’’ parameter, that is used, for example, in the ETKF. The $\mathbf{P}_k^{a,SEIK}$ is used in its factorized form in (A2) to obtain the analysis according to

$$\begin{aligned} \mathbf{x}_k^a &= \mathbf{x}_k^f + \mathbf{K}_k(\mathbf{y}_k^o - \mathbf{H}_k \mathbf{x}_k^f), \\ \mathbf{K}_k &= \mathbf{P}_k^{a,SEIK} \mathbf{H}_k^T \mathbf{R}_k^{-1} = \mathbf{L}_k \mathbf{U}_k (\mathbf{H}_k \mathbf{L}_k)^T \mathbf{R}_k^{-1}. \end{aligned} \quad (A4)$$

To modify the SEIK algorithm to take into account direct forecast localization, we need to use modified matrices \mathbf{L}_k^S and \mathbf{T}^S in (A3) and (A4). Here \mathbf{L}_k^S contains the difference between the ensemble members \mathbf{u}_{ij} , defined in (3), and their mean. The \mathbf{T}^S denotes a matrix of size $rM + 1 \times rM$, with $\sqrt{rM}(1 - 1/[rM + 1])$ on the diagonal and $-\sqrt{rM}/(rM + 1)$ off diagonal and in the lowermost row. In addition, $r + 1$ needs to be replaced with $rM + 1$ in (A1) and (A3). The equations with modified matrices provide an analysis covariance matrix $\mathbf{P}_k^{a,loc}$ analogous to (A2).

This modification increases the computational cost to compute the analysis. In particular, from (A3) we note that the trivial inversion of a matrix of size $r \times r$ in SEIK algorithm has increased to an inversion of an $rM \times rM$ matrix. One of the reasons for using the analysis error covariance matrix in the calculation of the gain is to avoid the implicit inversion of the matrix $\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k$ that is required in algorithms using the forecast covariance matrix. The matrix to be inverted is of size $p \times p$, where p is the number of observations. Therefore, if the rank M is larger than the ratio of the number of observations to the number of ensemble members, it would be hard to justify the use of this algorithm.

Furthermore, in the SEIK algorithm second-order-exact sampling with linear constraints is used in order to ensure that the new ensemble mean equals the analysis state and the ensemble covariance equals the analysis error covariance at the analysis time. However, second-order-exact resampling is only possible if the rank of the matrix $\mathbf{P}_k^{a,loc}$ is smaller than the number of ensemble members $r + 1$ (Pham 2001). Since we are increasing the rank of the matrix by the localization, we will not be able to exactly represent the matrix $\mathbf{P}_k^{a,loc}$ with $r + 1$ ensemble members. However, we are able to preserve the exact mean and some other prechosen properties of the covariance. For example we can resample such that the analysis ensemble represents the covariance matrix $\mathbf{P}_k^{a,SEIK}$ instead of $\mathbf{P}_k^{a,loc}$.

REFERENCES

Albertella, A., R. Savcenko, W. Bosch, and R. Rummel, 2008: Dynamic ocean topography—The geodetic approach. Tech. Rep. 27, Institute for Astronomical and Physical Geodesy, 53 pp.
 Bishop, C. H., B. J. Etherton, and S. Majumdar, 2001: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.*, **129**, 420–436.
 Brankart, J.-M., C.-E. Testut, P. Brasseur, and J. Verron, 2003: Implementation of a multivariate data assimilation scheme for isopycnic coordinate ocean models: Application to a 1993–1996

hindcast of the North Atlantic Ocean circulation. *J. Geophys. Res.*, **108**, 3074, doi:10.1029/2001JC001198.
 Brusdal, K., J.-M. Brankart, G. Halberstadt, G. Evensen, P. Brasseur, P. van Leeuwen, E. Dombrowsky, and J. Verron, 2003: An evaluation of ensemble based assimilation methods with a layered OGCM. *J. Mar. Syst.*, **40–41**, 253–289.
 Cohn, S. E., A. da Silva, J. Guo, M. Sienkiewicz, and D. Lamich, 1998: Assessing the effects of data selection with the DAO physical-space statistical analysis system. *Mon. Wea. Rev.*, **126**, 2913–2926.
 Evensen, G., 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.*, **53**, 343–367.
 Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757.
 Gelb, A., 1974: *Applied Optimal Estimation*. The M.I.T. Press, 382 pp.
 Haugen, V. E., and G. Evensen, 2002: Assimilation of SLA and SST data into an OGCM for the Indian ocean. *Ocean Dyn.*, **52**, 133–151.
 Horn, R. A., and C. R. Johnson, 1985: *Matrix Analysis*. Cambridge University Press, 561 pp.
 Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796–811.
 —, and —, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129**, 123–137.
 Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230**, 112–126.
 Janjić, T., and S. E. Cohn, 2006: Treatment of observation error due to unresolved scales in atmospheric data assimilation. *Mon. Wea. Rev.*, **134**, 2900–2915.
 Jekeli, C., 1981: Alternative methods to smooth the earth’s gravity field. Department of Geodetic Science and Surveying, Ohio State University, Columbus, OH, Rep. 327, 48 pp.
 Kepert, J., 2009: Covariance localisation and balance in an ensemble Kalman filter. *Quart. J. Roy. Meteor. Soc.*, **135**, 1157–1176.
 Lorenc, A. C., 1981: A global three-dimensional multivariate statistical interpolation scheme. *Mon. Wea. Rev.*, **109**, 701–721.
 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.
 Mayer-Gürr, T., 2007: ITG-GRACE03s: The latest GRACE gravity field solution computed in Bonn. *Joint Grace Science Team and DFG SPP Meeting*, Potsdam, Germany.
 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.
 Miyoshi, T., and S. Yamane, 2007: Local ensemble transform Kalman filter with an AGCM at a T159/L48 resolution. *Mon. Wea. Rev.*, **135**, 3841–3861.
 Nerger, L., and W. W. Gregg, 2007: Assimilation of SeaWiFS data into a global ocean-biogeochemical model using a local SEIK filter. *J. Mar. Syst.*, **68**, 237–254.
 —, W. Hiller, and J. Schröter, 2005: PDAF—The parallel data assimilation framework: Experiences with Kalman filtering. *Use of High Performance Computing in Meteorology: Proceedings of the 11th ECMWF Workshop*, W. Zwiefelhofer and G. Mozdzyński, Eds., World Scientific, 63–83.
 —, S. Danilov, W. Hiller, and J. Schröter, 2006: Using sea level data to constrain a finite-element primitive-equation ocean model with a local SEIK filter. *Ocean Dyn.*, **56**, 634–649.

- Ott, E., and Coauthors, 2004: A local ensemble Kalman filter for atmospheric data assimilation. *Tellus*, **56A**, 415–428.
- Penduff, T., P. Brasseur, C.-E. Testut, B. Barnier, and J. Verron, 2002: A four-year eddy-permitting assimilation of sea-surface temperature and altimetric data in the South Atlantic Ocean. *J. Mar. Res.*, **60**, 805–833.
- Pham, D. T., 2001: Stochastic methods for sequential data assimilation in strongly nonlinear systems. *Mon. Wea. Rev.*, **129**, 1194–1207.
- , J. Verron, and M. C. Roubaud, 1998: A singular evolutive extended Kalman filter for data assimilation in oceanography. *J. Mar. Syst.*, **16**, 323–340.
- Rio, M.-H., P. Schaeffer, J.-M. Lemoine, and F. Hernandez, 2005: Estimation of the ocean mean dynamic topography through the combination of altimetric data, in-situ measurements and GRACE geoid: From global to regional studies. *Proc. GOCINA Int. Workshop*, Luxembourg.
- Sakov, P., and P. R. Oke, 2008: Implications of the form of the ensemble transformation in the ensemble square root filters. *Mon. Wea. Rev.*, **136**, 1042–1053.
- Skachko, S., S. Danilov, T. Janjić, J. Schroeter, D. Sidorenko, R. Savchenko, and W. Bosch, 2008: Sequential assimilation of multi-mission dynamical topography into a global finite-element ocean model. *Ocean Sci.*, **4**, 307–318.
- Wahr, J., F. Bryan, and M. Molenaar, 1998: Time variability of the earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE. *J. Geophys. Res.*, **103** (B12), 30 205–30 229.
- Wang, Q., S. Danilov, and J. Schröter, 2008: Finite element ocean circulation model based on triangular prismatic elements with application in studying the effect of topography representation. *J. Geophys. Res.*, **113**, C05015, doi:10.1029/2007JC004482.
- Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924.