On serial observation processing in localized ensemble Kalman filters

Lars Nerger*

Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research, Bremerhaven, Germany

* *Corresponding author address:* Lars Nerger, Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research, Am Handelshafen 12, 27570 Bremerhaven, Germany.

E-mail: lars.nerger@awi.de

ABSTRACT

Ensemble square root filters can either assimilate all observations that are available at a given time at once, or assimilate the observations in batches or one at a time. For large-scale models, the filters are typically applied with a localized analysis step. This study demonstrates that the interaction of serial observation processing and localization can destabilize the analysis process and examines under which conditions the instability becomes significant. The instability results from a repeated inconsistent update of the state error covariance matrix that is caused by the localization. The inconsistency is present in all ensemble Kalman filters, except the classical ensemble Kalman filter with perturbed observations. With serial observation processing, its effect is small in cases when the assimilation changes the ensemble of model states only slightly. However, when the assimilation has a strong effect on the state estimates, the interaction of localization and serial observation processing can significantly deteriorate the filter performance. In realistic large-scale applications, when the assimilation changes the states only slightly and when the distribution of the observations is irregular and changing over time, the instability is likely not significant.

1. Introduction

Ensemble square-root Kalman filters are an efficient deterministic variant of the original Ensemble Kalman Filter (EnKF, Evensen 1994; Burgers et al. 1998). Common members of this class of filters are the Ensemble Transform Kalman filter (ETKF, Bishop et al. 2001), the Ensemble Adjustment Kalman Filter (EAKF, Anderson 2001, 2003), and the Ensemble Square-root Kalman filter with serial processing of observations (EnSRF, Whitaker and Hamill 2002). Recently, also the Singular "Evolutive" Interpolated Kalman (SEIK) filter (Pham et al. 1998a; Pham 2001) and the newly developed Error-subspace Transform Kalman filter (ESTKF, Nerger et al. 2012b) have been classified as ensemble square-root filters (Nerger et al. 2012b). All ensemble square-root Kalman filters express the analysis equation of the Kalman filter in a square-root form combined with an explicit transformation of the state ensemble (see Tippett et al. 2003). Most filter methods are formulated to assimilate all observations synchronously. However, the EAKF and the EnSRF are typically described to assimilate single observations serially, which increases the efficiency of these filter formulations. Further, both algorithms are algorithmically identical in case of serial observation processing. For example, the DART assimilation system (Anderson et al. 2009) provides an EAKF with serial observation processing.

Localization of covariance matrices in ensemble-based Kalman filters is required for data assimilation into large-scale models, because the typical ensemble size is limited to the order of 10 to 100 states, which is much smaller than the degrees of freedom of the models. By damping long-distance covariances, localization stabilizes the analysis update of the filter and increases the rank of the forecast covariance matrix as well as the local number of degrees of freedom for the analysis. The localization is either applied to the forecast covariance matrix, here denoted covariance localization (CL) (Houtekamer and Mitchell 1998, 2001), or to the observation error covariance matrix

(Hunt et al. 2007), here denoted observation localization (OL). The relation of both localization methods was the focus of several recent studies (Sakov et al. 2010; Greybush et al. 2011; Janjić et al. 2011). Further, Nerger et al. (2012a) proposed a method, denoted regulated localization, to make the localizing effect of OL and CL comparable. OL is typically applied in algorithms that do not explicitly compute the forecast error covariance matrix like the LETKF (Hunt et al. 2007), the SEIK filter, and the ESTKF. In contrast, the EAKF and the EnSRF compute elements of the forecast covariance matrix and apply CL. While the filters that apply OL assimilate all available observations at once, the EAKF and EnSRF methods that use CL perform a serial assimilation of single observations.

This study examines the interaction between CL and serial processing of observations in detail and demonstrates that it can destabilize of the analysis update. It is known in the community (e.g. C. Snyder, personal communication) that the serial processing of observations can lead to the situation that the actual analysis result depends on the order in which the observations are assimilated. This dependence is caused by the fact that the update equation for the state error covariance matrix is not fulfilled when localization is applied. This was already noted by Whitaker and Hamill (2002), but there is yet no publication that studies the effect of the inconsistent update of the state error covariance matrix. Whitaker et al. (2008) used the observation ordering to develop a variant of the EnSRF in which the observations are assimilated in an order of decreasing impact to the assimilation. The motivation for this scheme was described to be that it allows for an adaptive observation thinning algorithm by omitting observations that insignificantly reduce the estimated state error variance. Whitaker et al. (2008) also compared the assimilation performance of the EnSRF with the LETKF when applied with a global atmospheric model and found only small differences. Similarly, Holland and Wang (2013) compared the LETKF with the EnSRF without particular observation ordering for the assimilation with a simplified atmospheric model.

They found only small differences in the state estimates with slightly smaller errors in the LETKF estimates.

While the previous studies found small differences between the estimates of LETKF and EnSRF it is unclear which conditions influence the differences and whether there are conditions under which larger differences can occur. To some extent the differences in the state estimates are a result of different localization strengths in the OL and CL schemes for the same localization function (see Miyoshi and Yamane 2007). Here, this difference will be reduced by using for OL the regulated localization function by Nerger et al. (2012a). The instability that can result from the interaction of localization and serial observation processing is demonstrated and examined in numerical experiments with the small Lorenz-96 model (Lorenz 1996; Lorenz and Emanuel 1998). To compare the different effects of serial and synchronous assimilation of the observations, the two widely used filter methods EnSRF and LETKF are applied. For a direct examination of the influence of serial observation processing also a formulation of the EnSRF that assimilates all observations at once is applied. While this formulation is too costly to be applied in large-scale systems, it can be used with the small Lorenz-96 model.

The study is organized as follows: The EnSRF and the LETKF will be reviewed together with their localizations in section 2. The section also discusses the reasons for the inconsistent update of the covariance matrix. The configuration of the twin experiments with the Lorenz-96 model are described in section 3. The filter instability is demonstrated in time-mean results in section 4. The interaction of the localization and serial observation processing is further examined in Section 5, while Section 6 examines the effect of the order in which the observations are assimilated. In Section 7 the relevance of the findings with regard to real atmospheric and oceanographic applications is discussed. Finally, conclusions are drawn in section 8.

2. Filter algorithms

This section reviews the EnSRF with CL (Whitaker and Hamill 2002) as a typical method using serial observation processing and the LETKF using OL (Hunt et al. 2007), which uses synchronous assimilation.

All ensemble-based Kalman filters use an ensemble of m vectors $\mathbf{x}^{a(\alpha)}, \alpha = 1, \ldots, m$, of model state realizations of dimension n,

$$\mathbf{X}_k = \{\mathbf{x}_k^{a(1)}, \ldots, \mathbf{x}_k^{a(m)}\} \; , \tag{1}$$

to represent the state estimate and its uncertainty at some time t_k. The state estimate is given by the ensemble mean

$$\overline{\mathbf{x}_k^a} = \frac{1}{m} \sum_{\alpha=1}^{m} \mathbf{x}_k^{a(\alpha)}, \quad \overline{\mathbf{X}_k^a} := \{\overline{\mathbf{x}_k^a}, \ldots, \overline{\mathbf{x}_k^a}\} \tag{2}$$

where the superscript 'a' denotes the analysis. The uncertainty of the state estimate is described by the ensemble covariance matrix

$$\mathbf{P}_k^a = \frac{1}{m-1} (\mathbf{X}_k'^a)(\mathbf{X}_k'^a)^T \; . \tag{3}$$

where the prime denotes the matrix $\mathbf{X}_k'^a := \mathbf{X}_k^a - \overline{\mathbf{X}_k^a}$ of ensemble perturbations. The data assimilation procedure is initialized with an ensemble $\mathbf{X}_0^a$ that is generated based on some initial estimates of the state and the error covariance matrix. To compute a forecast, all ensemble members are integrated by the fully dynamical model resulting in the forecast ensemble $\mathbf{X}_k^f$. In the following, the time index 'k' is omitted as in the analysis step of the filters all quantities refer to the same time.

a. The EnSRF

Whitaker and Hamill (2002) proposed an ensemble square-root Kalman filter with serial processing of observations (EnSRF). In this filter, the state estimate and the ensemble perturbations are updated iteratively in a loop over all individual observations. This method is motivated by the fact that for a single observation the formulation of Potter (see Maybeck 1979, Sec. 7.3) can be applied to update the state error covariance matrix. This formulation is particularly efficient because matrix inversions, required for multiple observations, reduce to the inverse of a single number.

Let the subscript (i) indicate quantities at the i'th iteration of the loop over single observations. Likewise, the subscript denotes the index of the scalar observation assimilated at the i'th iteration. The state estimate is updated according to

$$\overline{\mathbf{x}^a}_{(i)} = \overline{\mathbf{x}^f}_{(i)} + \mathbf{K}_{(i)} \left(\mathbf{y}^o_{(i)} - \mathbf{H}_{(i)} \overline{\mathbf{x}^f}_{(i)} \right) \tag{4}$$

with the Kalman gain $\mathbf{K}_{(i)}$ of size $n \times 1$ given by

$$\mathbf{K}_{(i)} = \mathbf{P}^f_{(i)} \mathbf{H}^T_{(i)} \left(\mathbf{H}_{(i)} \mathbf{P}^f_{(i)} \mathbf{H}^T_{(i)} + \mathbf{R}_{(i)} \right)^{-1} . \tag{5}$$

Here, $\mathbf{H}_{(i)}$ is the observation operator for observation i. $\mathbf{y}^o_{(i)}$ is i'th element of the observation vector of size p and $\mathbf{R}$ is the observation error covariance matrix. To allow for the serial observation processing, $\mathbf{R}$ has to be diagonal.

For a single observation, the matrices $\mathbf{H}\mathbf{P}^f\mathbf{H}^T$ and $\mathbf{R}$ are scalars and $\mathbf{P}^f\mathbf{H}^T$ is a vector of size n. The matrix of ensemble perturbations is updated according to

$$\mathbf{X}'^a_{(i)} = \mathbf{X}'^f_{(i)} - \tilde{\mathbf{K}}_{(i)} \mathbf{H}_{(i)} \mathbf{X}'^f_{(i)} \tag{6}$$

with

$$\tilde{\mathbf{K}}_{(i)} = \left(1 + \sqrt{\frac{\mathbf{R}_{(i)}}{\mathbf{H}_{(i)} \mathbf{P}^f_{(i)} \mathbf{H}^T_{(i)} + \mathbf{R}_{(i)}}} \right)^{-1} \mathbf{K}_{(i)} . \tag{7}$$

The factor in front of the gain $\mathbf{K}_{(i)}$ reduces the Kalman gain for the update of the ensemble perturbations. This reduction is required for statistical consistency as without it the analysis error variances would be underestimated unless an ensemble of perturbed observations would be used (Burgers et al. 1998). A forgetting factor (Pham et al. 1998b) to inflate the covariances can be applied in this formulation by replacing $\mathbf{X}'^f$ by $\rho^{-1/2}\mathbf{X}'^f$ once before the loop over the single observations. The forgetting factor is the older concept of covariance inflation, which is frequently described in terms of the inflation factor $\alpha = \rho^{-1/2}$. Equations (4) to (7) are then applied in the loop over all observations available at an analysis time. In the first iteration, $\overline{\mathbf{x}^f_{(1)}}$ and $\mathbf{P}^f_{(1)}$ are given by the mean and covariance matrix of the ensemble forecast. In subsequent iterations of the loop, the analysis state and covariance matrix of the previous iteration serve as the forecast quantities.

While the EnSRF is usually applied with serial observation processing, it can also be formulated to assimilate all observations at once. In this case, Eqns. (4) to (6) are applied with the full vector $\mathbf{y}^o$ of observations and the corresponding observation operator. Following Whitaker and Hamill (2002), the reduced Kalman gain for the update of the ensemble perturbations defined by Eq. (7) is replaced by

$$\tilde{\mathbf{K}} = \mathbf{P}^f\mathbf{H}^T\left(\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R}\right)^{-\frac{T}{2}}\left[\left(\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R}\right)^{\frac{1}{2}} + \mathbf{R}^{\frac{1}{2}}\right]^{-1}. \tag{8}$$

For large-scale systems the evaluation of Eq. (8) would be very costly as matrices of size $p \times p$ have to be inverted. In the practical implementation used in numerical experiments, the matrix square-roots are implemented as the unique symmetric square root, which is also used for the LETKF. Below, this variant of the EnSRF will be referred to as EnSRF-bulk.

The localization of the EnSRF is performed as CL by multiplying the forecast state covariance matrix $\mathbf{P}^f$ element-wise with a correlation matrix $\mathbf{D}$ of compact support. As the full $\mathbf{P}^f$ will be very large for high-dimensional models, the localization is often applied in the observation space to the matrices $\mathbf{P}^f\mathbf{H}^T$ and $\mathbf{H}\mathbf{P}^f\mathbf{H}^T$. For a single observation, $\mathbf{H}\mathbf{P}^f\mathbf{H}^T$ reduces to the single value of the estimated observed state variance at the location of the observation. Accordingly, $\mathbf{H}\mathbf{P}^f\mathbf{H}^T$ is not modified for the EnSRF. However, the local analysis uses the modified vector

$$\left(\mathbf{P}^f\mathbf{H}^T\right)^{loc}_{(i)} = \mathbf{D}^{PH}_{(i)} \circ \left(\mathbf{P}^f\mathbf{H}^T\right)_{(i)} \tag{9}$$

where $\circ$ denotes the element-wise product. $\mathbf{D}_{(i)}^{PH}$ is a weight vector, which is a column of the correlation matrix $\mathbf{D}$ projected onto the observation space.

In the experiments performed below, the localization matrix $\mathbf{D}$ will be constructed using a 5th-order polynomial that mimics a Gaussian function but has compact support (Gaspari and Cohn 1999, shortly GC99). The localization is determined by the support radius at which the value of the function reaches zero.

b. The LETKF

The LETKF was introduced by Hunt et al. (2007) as a localized variant of the ETKF (Bishop et al. 2001). The LETKF applies a localized analysis with OL. Here, the LETKF is reviewed following Nerger et al. (2012a), which provides a particularly efficient formulation of the algorithm.

For the global ETKF, the forecast ensemble is projected onto the space of ensemble perturbations of dimension m by

$$\mathbf{X}^{f'} := \mathbf{X}^f \mathbf{T}. \tag{10}$$

The projection matrix $\mathbf{T}$ has size $m \times m$ and its elements are defined by:

$$\mathbf{T}_{i,j} := \begin{cases} 1 - \frac{1}{m} & \text{for } i = j \\ -\frac{1}{m} & \text{for } i \neq j \end{cases} \tag{11}$$

For the analysis update, the transform matrix $\mathbf{A}$ of size $m \times m$ is defined by

$$\mathbf{A}^{-1} := \rho(m-1)\mathbf{I} + (\mathbf{H}\mathbf{X}^{f'})^T \mathbf{R}^{-1} \mathbf{H}\mathbf{X}^{f'} \tag{12}$$

where $\mathbf{I}$ is the identity and ρ with $0 < \rho \leq 1$ is the forgetting factor (Pham et al. 1998b) that is used to implicitly inflate the forecast error covariance estimate. Using $\mathbf{A}$, the analysis covariance matrix is given by

$$\mathbf{P}^a = \mathbf{X}^{f'} \mathbf{A} \left(\mathbf{X}^{f'}\right)^T. \tag{13}$$

The analysis state estimate is computed from the forecast as

$$\overline{\mathbf{x}^a} = \overline{\mathbf{x}^f} + \mathbf{X}^{f'}\overline{\mathbf{w}} \tag{14}$$

where the weight vector $\overline{\mathbf{w}}$ of size m is given by

$$\overline{\mathbf{w}} := \mathbf{A}\left(\mathbf{H}\mathbf{X}^{f'}\right)^T \mathbf{R}^{-1}\left(\mathbf{y} - \mathbf{H}\overline{\mathbf{x}^f}\right) \ . \tag{15}$$

The ensemble is now transformed as

$$\mathbf{X}^a = \overline{\mathbf{X}^a} + \sqrt{m-1}\mathbf{X}^{f'}\mathbf{C}. \tag{16}$$

Here, $\mathbf{C}$ is the symmetric square root of $\mathbf{A}$. It is computed from the singular value decomposition $\mathbf{USV} = \mathbf{A}^{-1}$ such that $\mathbf{C} = \mathbf{US}^{-1/2}\mathbf{U}^T$. Using the definition of $\mathbf{X}^{f'}$ in Eq. (10) one can avoid to explicitly compute $\mathbf{X}^{f'}$, which leads to a very efficient algorithm in the typical situation that both the state dimension and the number of observations are much larger than the ensemble size. Namely, $\mathbf{HX}^{f'}$ in Eq. (14) can be computed as $(\mathbf{HX}^f)\mathbf{T}$. Further, in Eq. (16), the term $\mathbf{X}^{f'}\mathbf{C}$ can be computed as $\mathbf{X}^f(\mathbf{TC})$, which is a much cheaper operation than computing $\mathbf{X}^{f'}$ explicitly.

To obtain the LETKF as a localized form of the ETKF, the analysis and the ensemble transformation are performed in a loop through disjoint local analysis domains. In the simplest case, each single grid point is independently updated. For each local analysis domain, the observations are weighted by their distance from this domain using an element-wise product of the matrix $\mathbf{R}^{-1}$ with a localization matrix $\tilde{\mathbf{D}}$. $\tilde{\mathbf{D}}$ is usually constructed from a correlation function with compact support, like the GC99 function. Thus, observations beyond a certain distance obtain zero weight and can be neglected for the local analysis update. Using the subscript σ to denote the local analysis domain and δ to denote the domain of the corresponding observations of non-zero weight, the LETKF can be written as

$$\overline{\mathbf{x}^a_\sigma} = \overline{\mathbf{x}^f_\sigma} + \mathbf{X}^{f'}_\sigma\overline{\mathbf{w}_\delta} \ , \tag{17}$$

$$\overline{\mathbf{w}_\delta} = \mathbf{A}_\delta (\mathbf{H}_\delta \mathbf{X}^{f'})^T \left(\tilde{\mathbf{D}}_\delta \circ \mathbf{R}_\delta^{-1}\right) \left(\mathbf{y}_\delta - \mathbf{H}_\delta \overline{\mathbf{x}^f}\right), \tag{18}$$

$$\mathbf{A}_\delta^{-1} = \rho_\delta (m-1)\mathbf{I} + (\mathbf{H}_\delta \mathbf{X}^{f'})^T \left(\tilde{\mathbf{D}}_\delta \circ \mathbf{R}_\delta^{-1}\right) \mathbf{H}_\delta \mathbf{X}^{f'} \ , \tag{19}$$

$$\mathbf{X}_\sigma^a = \overline{\mathbf{X}_\sigma^a} + \sqrt{m-1}\mathbf{X}_\sigma^{f'} \mathbf{C}_\delta \ , \tag{20}$$

where the matrix $\mathbf{C}_\delta$ is the symmetric square root of $\mathbf{A}_\delta$.

In the experiments described below, the localization matrix $\tilde{\mathbf{D}}_\delta$ is constructed using the GC99 function as for the EnSRF. Note, that $\tilde{\mathbf{D}}_\delta$ is not a correlation matrix, because the diagonal elements vary with the distance. The effective localization length will be different from the prescribed support radius for OL (Nerger et al. 2012a). To make the effective localization lengths in the EnSRF with CL and the LETKF with OL comparable, the regulated localization function introduced by Nerger et al. (2012a) is used for the LETKF. The function ensures that the localization effect in the analysis step is identical for CL and OL in case of a single observation. For multiple observations, the exact function depends on the number of observations, but the function for a single observation can be used as an approximation. For a given localization function d^{CL} used for CL (e.g. the 5th-order polynomial of GC99), the regulated weight function for assimilating a single observation with OL is

$$d^{OLR} = \frac{d^{CL}\sigma_R^2}{HP^fH^T + \sigma_R^2}\left(1 - \frac{d^{CL}HP^fH^T}{HP^fH^T + \sigma_R^2}\right)^{-1} . \tag{21}$$

Here, HP^fH^T is the single element of the matrix $\mathbf{HP}^f\mathbf{H}$ corresponding to the single observation. σ_R^2 is the observation error variance. In the local analysis of the LETKF, several observations within the support radius around a local analysis domain are assimilated at once. A weight is computed for each observation, with the term HPH^T being computed as the square of the corresponding row of $\mathbf{H}_\delta \mathbf{X}^{f'}$ divided by $m-1$.

c. Inconsistency of the covariance update with localization

Whitaker and Hamill (2002) noted that the update equation for the state covariance matrix in the EnSRF, Eq. (6), is not consistent if localization with a smooth correlation function is used. Whitaker and Hamill (2002) reported that their study used the GC99 function despite the possible violation of Eq. (6), because it resulted in estimates with lower estimation errors compared to the case when a Heaviside step function was used, which would avoid the inconsistency.

The reason for the inconsistency lies in the used update equation for the covariance matrix. In the derivation of the Kalman filter one obtains

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{K}\mathbf{H})\,\mathbf{P}^f\,(\mathbf{I} - \mathbf{K}\mathbf{H})^T + \mathbf{K}\mathbf{R}\mathbf{K}^T. \tag{22}$$

If the same $\mathbf{P}^f$ and $\mathbf{R}$ are used in Eq. (22) and in the Kalman gain $\mathbf{K} = \mathbf{P}^f\mathbf{H}^T\left(\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R}\right)^{-1}$, Eq. (22) simplifies to

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{K}\mathbf{H})\,\mathbf{P}^f. \tag{23}$$

Equation (23) is used to update the covariance matrix in all ensemble Kalman filters, except the classical EnKF with perturbed observations (Evensen 1994; Burgers et al. 1998). The localization methods CL and OL only modify the Kalman gain, but not $\mathbf{P}^f$ and $\mathbf{R}$ in Eq. (22). Hence, Eqns. (22) and (23) are no longer equivalent if localization is applied. When Eq. (23) is directly used with a localized gain $\mathbf{K}$ one can even obtain a non-symmetric matrix $\mathbf{P}^a$. This, however, will not occur in the ensemble-based Kalman filters as these update the covariance matrix implicitly by updating the state ensemble.

Over all, the inconsistency of the covariance matrix update does occur in all filter algorithms that base on the simplified single-sided update equation (23). The difference between synchronous observation assimilation (as in the LETKF) and serial observation processing (as in the EnSRF) is, however, that the former method computes a single update of the matrix $\mathbf{P}^f$ because it assimilates all observations at a given time at once, while the EnSRF computes an update of $\mathbf{P}^f$ for each single observation. In the LETKF, the ensemble members representing $\mathbf{P}^a$ are immediately propagated by the model after the ensemble transformation. In contrast, in the serial observation processing of the EnSRF, each intermediately computed $\mathbf{P}_{(i)}$ (represented by the ensemble states) is used to assimilate the next observation. In the repeated update of the covariance matrix, the inconsistencies can accumulate. This effect will result in the observed dependence of the assimilation result on the order in which the observations are processed and in an inferior assimilation result compared to filter algorithms that assimilate all observation synchronously.

For the EnSRF, the covariance matrix update is derived from Eq. (23). For the i's observation it follows from Eq. (6) as

$$\mathbf{P}^a_{(i)} = \left(\mathbf{I} - \tilde{\mathbf{K}}_{(i)}\mathbf{H}_{(i)}\right)\mathbf{P}^f_{(i)}\left(\mathbf{I} - \tilde{\mathbf{K}}_{(i)}\mathbf{H}_{(i)}\right)^T \tag{24}$$

with $\tilde{\mathbf{K}}_{(i)}$ defined by Eq. (7). Even though the matrix update in Eq. (24) is symmetric it is inconsistent with Eq. (22) when $\mathbf{P}^f_{(i)}$ is localized in $\tilde{\mathbf{K}}_{(i)}$. One can check that it is not possible to re-derive the single-observation update of Potter (see Maybeck 1979, Sec. 7.3) when the localization is taken into account. Thus, it is not possible to derive an alternative factor $\tilde{\alpha}_{(i)}$ that ensures the equality of $\mathbf{P}^a$ in Eqns. (22) and (24), because there is in general no solution for $\tilde{\alpha}_{(i)}$ that ensures the equality. However, even if the symmetric update Eq. (22) could be used, the analysis result of the serial observation processing would still depend on the order in which the observations are assimilated unless one localizes $\mathbf{P}^f_{(i)}$ in Eq. (22). The Appendix provides a simple 2-dimensional example for applying the three equations (22) to (24) with serial and bulk processing of observations.

3. Configuration of numerical experiments

To assess the assimilation performances of the EnSRF and LETKF, identical twin experiments are conducted using the Lorenz-96 model (Lorenz 1996; Lorenz and Emanuel 1998). This nonlinear model has been used in several studies to examine the behavior of different ensemble-based Kalman filters (e.g. Anderson 2001; Whitaker and Hamill 2002; Ott et al. 2004; Lawson and Hansen 2004; Sakov and Oke 2008; Janjić et al. 2011). The same configuration as in Nerger et al. (2012a) is used and the results can be directly compared with their results.

The Lorenz-96 model uses the non-dimensional equations

$$\frac{dx_j}{dt} = \left(x_{j+1} - x_{j-2}\right) x_{j-1} - x_j + F \tag{25}$$

where $j = 1, \ldots, 40$ is the grid point index with cyclic boundary conditions and $F = 8$ is a forcing parameter. The time stepping is performed using a fourth-order Runge-Kutta scheme with a non-dimensional time step size of 0.05. The model and the filter algorithms have been implemented within the Parallel Data Assimilation Framework (PDAF, Nerger et al. 2005; Nerger and Hiller 2013, http://pdaf.awi.de).

A trajectory representing the "truth" is computed over 60000 time steps from the initial state of constant value of 8.0 but x_{20} = 8.008, following Lorenz and Emanuel (1998). Synthetic observations of the full state are generated by disturbing the true trajectory by uncorrelated random normal noise. Three cases will be examined in which the standard deviation σ_R of the observation error will be 1, 0.5, and 0.1. The strength of the assimilation impact increases when the observation errors shrink. The initial error estimate from the ensemble used in the experiments is 2.5. Thus, the largest σ_R is 40% of the error estimate, while the smallest values is only 4% of it.

Second-order exact sampling from the true trajectory Pham (2001) is used to generate the initial ensemble. To assess the assimilation performance over a long assimilation experiment, the assimilation is performed at each time step over 50000 time steps with an ensemble of 10 states. For the observations, an offset of 1000 time steps of the true trajectory is used to avoid the spin-up phase of the model. The localization is applied with a fixed support radius. All experiments are repeated ten times with varying random numbers for the generation of the initial ensemble. The assimilation performance will be assessed by the root mean square error of each experiment averaged over each set of ten experiments. The random numbers used to perturb the observations are not varied. It would have a similar effect to varying the initial ensemble.

4. Mean assimilation performance

The effect of the serial observation processing can be demonstrated in a full-length experiment with the Lorenz-96 model. Figure 1 shows the averaged RMS errors for a range of forgetting factors and support radii of the localization function and three different observations errors. The filters diverge when the time-mean RMS error is larger than the observation error. If at least one of the 10 repetitions of each experiment diverges, the rectangle for this parameter pair is left white. The overall shape of the RMS error distribution, namely a minimum error region that is surrounded by larger errors, shows that the parameter ranges chosen for the experiments cover the optimal parameter values.

The first two rows of Fig. 1 show the average RMS errors for the serial EnSRF and LETKF, respectively. As discussed by Nerger et al. (2012a), the regulated localization as used here in the LETKF should make the filter results with OL very similar to those with CL. However, there are significant differences, which are most pronounced for the smallest observation error of $\sigma = 0.1$ (right panels of Fig. 1). In this case, the LETKF converges in a much larger parameter region than the EnSRF. Further, the LETKF yields significantly smaller mean RMS errors than the EnSRF. When the assimilation strength is reduced by increasing the observation error, the error differences become smaller. For $\sigma_R = 0.5$ (middle column of Fig. 1), the minimum RMS errors obtained with the EnSRF are slightly larger than for the LETKF. In addition, there is a parameter range (forgetting factors 0.95 and 0.96, localization radii 18 and 20), where the EnSRF yields larger errors than the LETKF. This effect is unusual as one typically obtains a closed area of minimal errors (see, e.g. Janjić et al. 2011) as is visible for the LETKF. For the largest observation error of $\sigma_R = 1.0$ (left panels of Fig. 1), the RMS error in dependence of the forgetting factor and the support radius are very similar for the EnSRF and LETKF.

The EnSRF-bulk update scheme discussed in Section 2a avoids the serial observation processing, but applies CL. Hence, comparing the serial EnSRF with EnSRF-bulk allows to directly see the influence of serial observation processing. The averaged RMS errors for EnSRF-bulk are shown in the third row of Fig. 1. In the stable assimilation regime, e.g. for $\sigma_R = 0.1$ with a support radius below 18 grid points, the serial EnSRF shows up to about 2% smaller RMS errors than the EnSRF-bulk. This behavior is probably due to the fact that the serial observation processing avoids matrix inversions. For larger support radii and smaller inflation the EnSRF-bulk shows smaller RMS errors and less tendency to diverge compared to the serial EnSRF. The parameter region in which the EnSRF-bulk converges is larger than for the serial EnSRF and similar to the convergence region of the LETKF. However, in the case of $\sigma_R = 0.1$ the EnSRF-bulk diverges for support radii above 28 grid points. This divergence can be attributed to a large condition number of the matrix $\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R}$, which needs to be inverted in the EnSRF-bulk. Overall, the LETKF shows the largest convergence region and the smallest RMS errors. This behavior is influenced by the OL with regulated localization function which is used by the LETKF.

5. Stability of the EnSRF analysis with localization

To examine the reasons for the differences in the RMS errors obtained with the EnSRF, EnSRF-bulk and LETKF, the first analysis step of the experiments discussed above is examined in more detail. While obviously the first analysis step is not necessarily representative for the whole assimilation experiment it nonetheless allows to study the different behaviors of the filters. At the first analysis step, the experiments start with a 'climatological' state estimate with an RMS error of about 3.5. The initial ensemble estimate of the error is slightly lower with about 2.5. The error of the analysis state after the first analysis step depends on the observation error. It is larger than the asymptotic error level, which is reached only after several forecast-analysis cycles. The advantage of examining the first analysis step is that it shows the instability in a very clear way. Further, the results are practically uninfluenced by the model nonlinearity as only a single time step was computed.

The parameters considered in this section are a forgetting factor of 0.95 and a support radius of 20 grid points. For these parameters, all three filter formulations converge and the averaged RMS errors discussed in Section 4 are close to their minimum.

The EnSRF is configured to assimilate each observation in a loop starting from the observation at the grid point with index 1 and then ordered with increasing index. Thus, when the state of size 40 is fully observed, the state estimate and the ensemble are modified 40 times in each analysis step. The panels of Fig. 2 show the true and estimated RMS errors of the state for the sequence of assimilating 1 to 40 observations. To be able to directly examine one assimilation series, only one ensemble realization is shown here. The exact shape of the curves shown in Fig. 2 is specific for the set of random numbers used to generate the ensemble and those used to generate the observations. However, using other random numbers does not change the overall conclusions. Fig. 2 also shows the RMS errors from the analogous experiments with the LETKF and the EnSRF-bulk. Here, all observations are assimilated at once. To be able to study the dependence of the RMS error on the number of observations, 40 experiments are performed for each filter and each observation error in which between 1 and 40 observations are assimilated. In contrast to the EnSRF, the intermediate results would not be realized in an experiment with 40 observations.

For $\sigma_R = 1.0$ the upper panel of figure 2 shows that with a growing number of observations, the true and estimated RMS errors generally decrease. However, when about half of the observations are assimilated the true RMS errors (solid lines) increase, but finally decrease again when more observations are assimilated. This interim increase is larger for the EnSRF and EnSRF-bulk than for the LETKF. Overall, it is visible that the estimated errors (dashed lines) of the EnSRF and EnSRF-bulk are smaller than those of the LETKF. In addition, when 40 observations are assimilated, the true error of the EnSRF is 2.02 and hence slightly larger than the true error of 1.86 of the LETKF, while the true error of the EnSRF-bulk is 1.8. The difference between EnSRF and LETKF for 40 observations is statistically significant, when repeating the experiment with different random numbers, while it is not significant for LETKF and EnSRF-bulk.

For smaller observation errors, the interim increase of the true errors for the EnSRF and EnSRF-bulk is larger. When 3, 27, or 28 observations are assimilated for $\sigma_R = 0.5$, the true error for the EnSRF is larger than without assimilating any observations. In contrast, the LETKF reduces the RMS error for 28 observations by about 40% compared to assimilating no observations.

For $\sigma_R = 0.1$ the true error in the EnSRF for assimilating between 23 and 30 observations is up to about twice as large than without assimilation. The error estimate of the EnSRF misses this error increase and strongly underestimates the true error. The EnSRF-bulk shows a similar behavior, but with smaller peak values and a smaller error when 40 observations are assimilated. In contrast, the estimated error of the LETKF is much closer to the true error. The comparison of the RMS errors of the LETKF with those of the EnSRF and EnSRF-bulk show that the different localization methods lead to state estimates of significantly different quality, in particular when not all available observations are assimilated. However, for 40 observations the serial processing of the EnSRF, in which the ensemble states for each number of assimilated observations are explicitly computed, leads to larger errors compared to the synchronous analysis of the EnSRF-bulk.

The effect that leads to the large increase of the RMS error for the EnSRF and EnSRF-bulk is further demonstrated in Fig. 3. Here, the state estimates for the EnSRF, EnSRF-bulk and LETKF are shown when different numbers of observations are assimilated in the case of $\sigma_R = 0.1$. For 20 observations, the estimates of all three filters are very similar. In particular, the state estimate is very close to the truth in the left half of the domain, where the observations were already assimilated. For 25 observations, where the mean RMS error of the EnSRF jumped to a value of 8.0, an unrealistically large amplitude of the wave is visible for the EnSRF in the part of the domain, where no observations have been assimilated yet. The behavior is similar for the EnSRF-bulk, but the RMS error remains smaller than for the serial EnSRF. In contrast, the LETKF estimates a wave of realistic amplitude. When the number of observations is further increased, the EnSRF and EnSRF-bulk continue to estimate a state with a large wave amplitude in the part where the observations haven't yet been assimilated. The large amplitude persists up to about 30 assimilated observations. Finally, the amplitude is reduced and for 40 observations the state estimates of all three filters are realistic but the error in the estimated state is larger for the EnSRF than for the LETKF and EnSRF-bulk.

The differences between the serial EnSRF and the EnSRF-bulk are only caused by the serial observation processing. From Fig. 2 it is visible that the difference between both filters accumulates with a growing number of assimilated observations. The repeated inconsistent covariance updates of the serial EnSRF do not always result in larger errors of the state estimate. E.g., if only observations in the first half of the model domain are assimilated, the serial EnSRF shows smaller errors compared to the EnSRF-bulk. However, for more than 30 observations, the RMS errors from EnSRF-bulk are smaller than those from the serial EnSRF for all experiments. The estimated RMS errors are almost identical for the EnSRF and EnSRF-bulk. However, the serial observation processing of the EnSRF results in covariance matrices that are distinct from those obtained with the EnSRF-bulk as is also demonstrated in the Appendix. The different variance and covariance estimates are tapered by the localization matrix and result in state updates that are different in both filters. The differences are most pronounced for the smallest observation error of $\sigma_R = 0.1$.

The differences between the EnSRF-bulk and the LETKF are mainly caused by the different localization schemes. While for a single observation, the regulated OL used for the LETKF results in a localization effect that is identical to the CL in the EnSRF and EnSRF-bulk, this is no longer the case if multiple observations are assimilated at the same time (see Nerger et al. 2012a, for a detailed discussion of the regulated OL). However, the regulated OL results in much better state estimates in particular if the observations are incomplete as is visible from Figures 2 and 3. For the Kalman gain, the regulated OL results in a different localization function that improves the state estimates without reducing the support radius of the localization. For the EnSRF, one would need to strongly reduce the localization support radius for CL (e.g. to 8 grid points for $\sigma_R = 0.1$) to obtain a similarly stable analysis as for the LETKF at the first analysis time. However, as Fig. 1 shows, the RMS error for an experiment over 50000 time steps would be significantly larger for this smaller support radius.

As pointed out in section 2c, the inconsistent update of the state error covariance matrix should not only appear in the EnSRF, but also in other filters that process observations serially. The LETKF method can be easily modified to perform a loop of analysis steps with single observations. For consistency, the forgetting factor has to be removed from Eq. (19). Instead, the ensemble perturbations are inflated once before the analysis step by the square-root of the inverse forgetting factor as done in the EnSRF. The lowermost panel of Fig. 2 shows also the RMS error for the LETKF with serial observation processing. Similar to the EnSRF, the RMS error shows a peak for 3 observations and the instability around 25 observations. The true RMS errors are lower than for the EnSRF and the estimated RMS errors are slightly larger. This shows that the influence of the localization on the update of the covariance matrix in the serial variant of the LETKF is not identical to that in the EnSRF. However, the general instability of the analysis also occurs for the LETKF when it is applied with serial observation processing.

Note, that the change of the EnSRF behavior that is demonstrated here for different observation errors is not an effect of model nonlinearity. Only a single model time step has been computed before the first analysis time, which does not have much influence on the ensemble distribution. Actually, the behavior shown in Figs. 2 and 3 would look very similar when the analysis would be performed at the initial time without any time stepping. Thus, one could perform this experiment even without the Lorenz-96 model. That is, one only needs a covariance matrix, and initial state estimate and a set of observations together with their error estimate. By sampling the covariance matrix and state estimate with a small ensemble of 10 members one could compute the analysis step. The larger differences in the state update for decreasing observation errors are due to the fact that the effect of the inconsistently updated covariances grows with the influence of the observations on the state estimate. However, the effect of the differences can sometimes average out, as is visible from the nearly identical RMS errors for $\sigma_R = 0.5$ for about 20 to 28 assimilated observations (middle panel of Fig. 2).

6. Influence of the observation order

The analysis result in case of serial observation processing depends on the order in which the observations are assimilated. Hence, one might wonder whether one can improve the analysis results obtained with the serial EnSRF by changing the order in which the observations are assimilated. Accordingly, the influence of the order is examined here for the application with the Lorenz-96 model. Only the case $\sigma_R = 0.1$ is considered, which showed the largest influence of the serial observation processing before. Further, only the serial EnSRF is examined and compared to the LETKF.

The lowermost panel of Fig. 2 shows that the true RMS was largest when observations at the grid points 25 to 30 were assimilated. This is far from grid point 1 where the assimilation series started. Thus, a first test is whether one can stabilize the analysis by using a more uniform sorting of the observations. To this end, the observation order is revised so that the grid point indices of the assimilated observations are chosen like 1, 21, 11, 31, 6, 26, 16, 36 and continued so that the remaining gaps are filled in an approximately uniform way. The upper panel of Fig. 4 shows the RMS error over the number of assimilated observations for this observation order. For comparison, also the LETKF assimilated the same observations. Using the revised observation order, the large peak in the RMS error of the EnSRF at around 25 assimilated observations (Fig. 2, bottom panel) has actually disappeared. In this respect, the re-ordering of the observations is successful. However, up to about 20 assimilated observations, the RMS errors are now very close to the error without assimilation. Also, there are smaller peaks where the true RMS error exceeds the error without assimilation with values up to about 4.5. Further, the final RMS error after assimilating all 40 observations in the revised order is 0.91 and hence almost identical to the error of 0.94 without reordering. Fig. 4 shows that also for the LETKF more than 20 observations need to be assimilated to significantly reduce the RMS error. However, the RMS error remains smaller than that of the EnSRF and reaches a value of 0.13 when 40 observation are assimilated.

The upper panel of Fig. 5 shows the mean RMS error for the full experiment in which the EnSRF with the reordered observations is applied over 50,000 analysis steps. Compared to the case $\sigma_R = 0.1$ in Fig. 1, the mean RMS errors are identical, except for some parameter choices at the edge to filter divergence. Even, if the observation order is randomized and a different order is used at each analysis time, a very similar distribution of the errors would be found (not shown).

Thus, the state estimate of the EnSRF with 40 observations is not significantly influenced by the observation order.

An alternative to the series of global state updates in the EnSRF was introduced by Whitaker et al. (2008). This variant of the EnSRF, denoted below the L-EnSRF, performs individual local analysis updates for each grid point with the observations ordered by their influence on the state at the grid point. For this method one computes for each grid point the variance reduction in the analysis update induced by a single observation. Then, the observations are assimilated individually at each grid point in decreasing order according to the variance reduction.

The lower panel of Fig. 4 shows the RMS error for the L-EnSRF as a function of the number of assimilated observations. The RMS error remains close to the RMS error without assimilation, or even above it, until about 29 observation are assimilated. Thus, the individual sorting of the observations in the L-EnSRF also avoids the instability peak around grid points 25 to 30 in the original EnSRF without re-ordering. For more than 29 observation, the RMS error decreases strongly. The final error for 40 assimilated observations is reduced to 0.51. Hence it is significantly smaller than the error of the EnSRF with the original order, but larger than that of the LETKF. The reduction of the RMS error is also visible in the full experiment over 50000 analysis steps as is shown in the lower panel of Fig. 5. The minimum mean RMS error is reduced from 0.0193 to 0.0190. This change is small, but statistically significant. Further, the filter is stabilized and the parameter region in which the assimilation converges is increased. However, the RMS errors obtained with the L-EnSRF are still larger than those of the LETKF. In addition, the region of filter convergence is larger for the LETKF than for the L-EnSRF.

7. Practical relevance of the EnSRF instability

The numerical experiments conducted in the sections above clearly show the effect of the instability in the EnSRF analysis. However, these experiments are highly idealized. In particular, the Lorenz-96 model simulates only a single model field. Further, the dynamics of the model are homogenous and hence also the distribution of the errors in the state estimate and the ensemble perturbations is rather uniform. Also, the full model state was observed. The observation errors were varied by one order of magnitude in the experiments. This allowed to vary the strength of the assimilation impact. The largest influence of the serial observation processing in case of the EnSRF and of the regulated OL in case of the LETKF occurred for the smallest observation error which was only 4% of the error of the initial state estimate.

For real-world cases (e.g. Whitaker et al. 2008; Sakov et al. 2012; Losa et al. 2014), the mean RMS error estimated by the ensemble is typically of the same order as the observation errors. In this respect, these applications should operate in the regime of the largest observation errors used in the idealized experiments. In this case, no significant differences between the LETKF and EnSRF are to be expected. However, in realistic cases the estimated errors will show spatial variations and larger error estimates can occur locally. E.g., if eddies appear in a high-resolution ocean model, the ensemble spread could become large due to varying locations of the eddies or when only some ensemble members simulate the eddies while other miss them. In the atmosphere a situation might appear with convective scale models, when some ensemble members estimate convection while others don't. When in this situation accurate observations are assimilated, the effect of serial observation processing might deteriorate the assimilation performance. However, in this case also the spatial extent of the region with large state error estimates and small observation errors will influence the effect of the serial observation processing. It is unclear which spatial extent is necessary to make the effect visible.

The experiments with the Lorenz-96 model showed only a negligible effect of reordering or randomizing the observation sequence, unless one sorts the observations explicitly with decreasing influence and performs local analyses. However, in atmospheric data assimilation also the location of the observations can vary nearly randomly between successive analysis times. This kind of randomization might also influence the effect of the serial observation processing.

8. Conclusion

This study examined the influence of localization in ensemble-based Kalman filter formulations that perform the assimilation of an observation vector as a series over single observations. Filter algorithms of this type are the ensemble adjustment Kalman filter (EAKF) and the Ensemble Square-root Filter (EnSRF).

Most ensemble Kalman filters update in the analysis step the state error covariance matrix, which is represented by the ensemble of model states, using the non-symmetric update equation of the Kalman filter. This equation is cheaper to evaluate than the more general symmetric update equation, but only valid when the Kalman gain is computed with the same forecast state error covariance matrix as used in the update equation. Using a localized covariance matrix in the gain while using the non-localized matrix in the update equation, results in an inconsistent analysis state error covariance matrix. To some extent this inconsistency is inherent to all ensemble-based Kalman filters because they approximate the state error covariance matrix by the low-rank ensemble covariance matrix, but they increase the rank for the analysis step by applying localization. Filter algorithms that assimilate a whole observation vector simultaneously, update the covariance matrix only once during an analysis step. In contrast, in filters with serial observation processing the size of the observation vector defines how often the covariance matrix is updated.

The assimilation performance of the EnSRF was compared with that of the local ensemble transform Kalman filter (LETKF) with regulated observation localization using twin experiments with the Lorenz-96 model. When the observation errors were of a similar magnitude as the initial errors of the state estimate, both filter methods showed a similar behavior. When the observation errors were decreased, the EnSRF showed a stronger tendency to diverge and larger minimum RMS errors than the LETKF and a variant of the EnSRF that assimilates all observations at once.

Changing the observation order resulted in an improvement of the assimilation performance of the EnSRF. For this, each single grid point needed to be updated with an individual order of the observations. As proposed by Whitaker et al. (2008), ordering the observation with decreasing influence to reduce the estimate variance resulted in the best assimilation performance. However, in the twin experiments the EnSRF with localized update and individually ordered observations still exhibited larger minimum errors and a stronger tendency to diverge than the LETKF.

The idealized experiments used the Lorenz-96 model. However, the repeated inconsistent update of the covariance matrix and hence the ensemble states is a general property with serial observation processing. Thus, the instability of the analysis with serial observation processing should also occur with other models. However, for practical applications the deterioration of the filter performance of the EnSRF will often not be relevant. Overall, the experiments indicate, that the inconsistent ensemble update does only deteriorate the filter performance of the EnSRF in cases when the observations have a strong influence, i.e. when the observation error is small compared to the estimated error of the state. In most real-world applications, the observation and state errors have a similar magnitude and the serial observation processing should be stable. This finding is consistent with the fact that the EnSRF or EAKF algorithms have been successfully applied for a wide range of data assimilation problems. However, one should be careful that the observation errors do not become significantly smaller than the estimated state errors and hence induce a strong assimilation influence.

The LETKF method performed better than the EnSRF with smaller estimation errors and better stability. This difference was caused by the different localization schemes and the application of regulated observation localization for the LETKF. However, it is obvious that also the LETKF – as all other ensemble Kalman filters – performs an inconsistent update of the state error covariance matrix when it is are applied with localization. Thus, while the localization methods are empiric schemes that have been demonstrated to improve the state estimates and the stability of ensemble Kalman filters, their influence on the error estimates is still unclear. For example, Janjić et al. (2011) examined a localization variant of the SEIK filter in which the covariance matrix is updated using a Heaviside step function and using the smooth weighting function only for the update of the state estimate. While the update of the ensemble perturbations is also not fully consistent in this formulation, it exhibited very good assimilation performance with the Lorenz-96 model. Further research into localization is required to ensure consistent corrections of both the state estimate and the ensemble perturbations in the analysis steps of the ensemble-based Kalman filters.

Acknowledgment. The author likes to thank the two anonymous reviewers as well Dr. J. Whitaker and the Editor, Dr. A. Aksoy, for their helpful comments.

APPENDIX

A1. 2D Example of the serial observation assimilation

This appendix shows a simple example of the influence of serial observation processing with localization and of the application of a single-sided update of the covariance matrix. Let the forecast state and covariance matrix be

$$\mathbf{x}^f = \left\{ \begin{matrix} 1 \\ 1 \end{matrix} \right\}; \quad \mathbf{P}^f = \left\{ \begin{matrix} 1 & 0.8 \\ 0.8 & 1 \end{matrix} \right\} \tag{A1}$$

Two observations are available, which are defined by

$$\mathbf{y} = \left\{ \begin{matrix} 0 \\ 0 \end{matrix} \right\}; \quad \mathbf{R} = \left\{ \begin{matrix} 0.1 & 0 \\ 0 & 0.1 \end{matrix} \right\}; \quad \mathbf{H} = \left\{ \begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix} \right\} \tag{A2}$$

The localization matrix is

$$\mathbf{D} = \left\{ \begin{matrix} 1 & 0.25 \\ 0.25 & 1 \end{matrix} \right\} \tag{A3}$$

Now, compute the analysis covariance matrices, applying the covariance localization only in the Kalman gain. When all observations are assimilated at once, one obtains

$$\mathbf{P}^a_{(Eq.\ 22)} = \left\{ \begin{matrix} 0.089 & 0.007 \\ 0.007 & 0.089 \end{matrix} \right\}; \quad \mathbf{P}^a_{(Eq.\ 23)} = \left\{ \begin{matrix} 0.080 & 0.058 \\ 0.058 & 0.080 \end{matrix} \right\} \tag{A4}$$

Using the serial observation processing, assimilating first the observation defined by the first row of $\mathbf{H}$, followed by the second row, one obtains

$$\mathbf{P}^a_{(Eq.\ 22,serial)} = \left\{ \begin{matrix} 0.088 & 0.009 \\ 0.009 & 0.088 \end{matrix} \right\}; \quad \mathbf{P}^a_{(Eq.\ 24,serial)} = \left\{ \begin{matrix} 0.089 & 0.055 \\ 0.055 & 0.076 \end{matrix} \right\} \tag{A5}$$

The analysis state estimates after assimilating both observations are

$$\mathbf{x}^a_{(bulk)} = \left\{ \begin{matrix} 0.077 \\ 0.077 \end{matrix} \right\}; \quad \mathbf{x}^a_{(Eq.\ 22,serial)} = \left\{ \begin{matrix} 0.097 \\ 0.073 \end{matrix} \right\}; \quad \mathbf{x}^a_{(Eq.\ 24,serial)} = \left\{ \begin{matrix} 0.091 \\ 0.046 \end{matrix} \right\} \tag{A6}$$

The correct state estimate is $\mathbf{x}^a_{(bulk)}$ with the same value in both elements. With serial observation processing, both state estimates show significant errors. However, the second element of $\mathbf{x}^a_{(Eq.\ 22,serial)}$, which results from applying the symmetric update Eq. (22), is close to the true value. For the covariance matrices, the single-sided update Eqns. (23, 24) result in much larger covariances than the symmetric update equation. This effect is similar for both the bulk and the serial updates. However, when the update equation (24) of the EnSRF is used, also the variance estimate for the second state element is significantly underestimated.

References

Anderson, J., T. Hoar, K. Raeder, H. Liu, N. Collings, R. Torn, and A. Arellano, 2009: The data assimilation research testbed: A community facility. *Bull. Am. Meteorol. Soc.*, **90**, 1283–1296.

Anderson, J. L., 2001: An Ensemble Adjustment Kalman Filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884–2903.

Anderson, J. L., 2003: A local least squares framework for ensemble filtering. *Mon. Wea. Rev.*, **131**, 634–642.

Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical aspects. *Mon. Wea. Rev.*, **129**, 420–436.

Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998: On the analysis scheme in the Ensemble Kalman Filter. *Mon. Wea. Rev.*, **126**, 1719–1724.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99 (C5)**, 10 143–10 162.

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Q. J. Roy. Meteor. Soc.*, **125**, 723–757.

Greybush, S. J., E. Kalnay, T. Miyoshi, K. Ide, and B. R. Hunt, 2011: Balance and ensemble Kalman filter localization techniques. *Mon. Wea. Rev.*, **139**, 511–522.

Holland, B., and X. Wang, 2013: Effects of sequential or simultaneous assimilation of observations and localization methods on the performance of the ensemble Kalman filter. *Q. J. R. Meteorol. Soc.*, **139**, 758–770.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796–811.

Houtekamer, P. L., and H. L. Mitchell, 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., L. Nerger, A. Albertella, J. Schröter, and S. Skachko, 2011: On domain localization in ensemble based Kalman filter algorithms. *Mon. Wea. Rev.*, **139**, 2046–2060.

Lawson, W. G., and J. A. Hansen, 2004: Implications of stochastic and deterministic filters as ensemble-based data assimilation methods in varying regimes of error growth. *Mon. Wea. Rev.*, **132**, 1966–1981.

Lorenz, E. N., 1996: Predictability - a problem partly solved. *Proceedings Seminar on Predictability*, ECMWF, Reading, UK, 1–18.

Lorenz, E. N., and K. A. Emanuel, 1998: Optimal sites for supplementary weather observations: Simulation with a small model. *J. Atm. Sci.*, **55**, 399–414.

Losa, S. N., S. Danilov, J. Schröter, T. Janjić, L. Nerger, and F. Janssen, 2014: Assimilating NOAA SST data into BSH operational circulation model for the North and Baltic Seas: Part 2. Sensitivity of the forecast's skill to the prior model error statistics. *J. Mar. Syst.*, **129**, 259–270.

Maybeck, P. S., 1979: *Stochastic models, estimation, and control*, Vol. 1. Academic Press.

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. Hiller, 2013: Software for ensemble-based data assimilation systems - implementation strategies and scalability. *Computers & Geosciences*, **55**, 110–118.

Nerger, L., 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 11. ECMWF Workshop*, W. Zwieflhofer, and G. Mozdzynski, Eds., World Scientific, 63–83.

Nerger, L., T. Janjić, J. Schröter, and W. Hiller, 2012a: A regulated localization scheme for ensemble-based Kalman filters. *Q. J. Roy. Meteor. Soc.*, **138**, 802–812.

Nerger, L., T. Janjić, J. Schröter, and W. Hiller, 2012b: A unification of ensemble square root Kalman filters. *Mon. Wea. Rev.*, **140**, 2335–2345.

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

Pham, D. T., 2001: Stochastic methods for sequential data assimilation in strongly nonlinear systems. *Mon. Wea. Rev.*, **129**, 1194–1207.

Pham, D. T., J. Verron, and L. Gourdeau, 1998a: Singular evolutive Kalman filters for data assimilation in oceanography. *C. R. Acad. Sci., Ser. II*, **326 (4)**, 255–260.

Pham, D. T., J. Verron, and M. C. Roubaud, 1998b: A singular evolutive extended Kalman filter for data assimilation in oceanography. *J. Mar. Syst.*, **16**, 323–340.

Sakov, P., F. Counillon, L. Bertino, K. A. Lisaeter, P. R. Oke, and A. Korablev, 2012: TOPAZ4: an ocean-sea ice data assimilation system for the North Atlantic and Arctic. *Ocean Sci.*, **8**, 633–656.

Sakov, P., G. Evensen, and L. Bertino, 2010: Asynchronous data assimilation with the EnKF. *Tellus*, **62A**, 24–29.

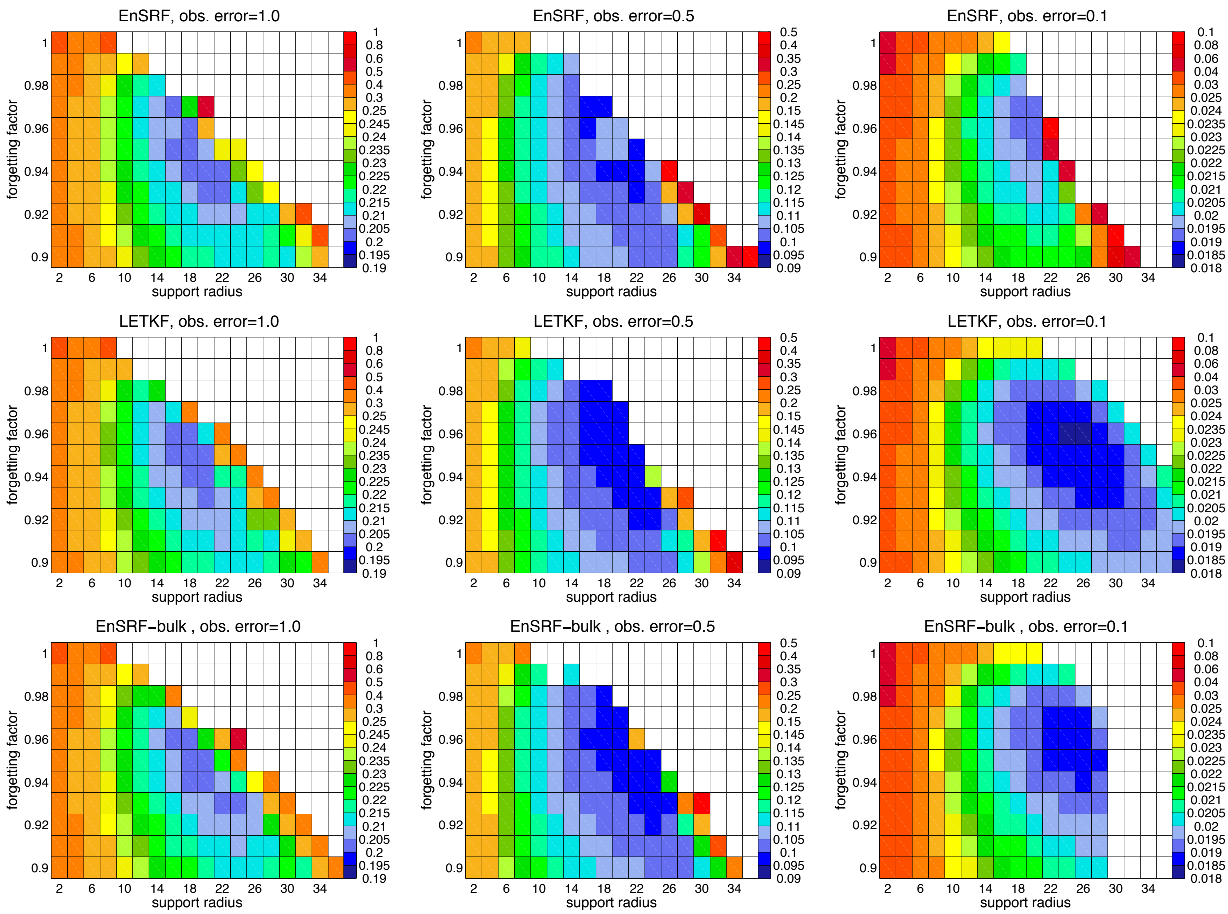
Sakov, P., and P. R. Oke, 2008: Implications of the form of the ensemble transformation in the ensemble square root filters. *Mon. Wea. Rev.*, **136**, 1042–1053.

Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble square root filters. *Mon. Wea. Rev.*, **131**, 1485–1490.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1927.

Whitaker, J. S., T. M. Hamill, X. Wei, Y. Song, and Z. Toth, 2008: Ensemble data assimilation with the NCEP global forecast system. *Mon. Wea. Rev.*, **136**, 463–482.

LIST OF FIGURES



683 FIG. 1. Average RMS errors for the EnSRF (top), LETKF (middle) and EnSRF-bulk (bottom) for three different observational errors: 1.0 (left), 0.5 (center), 0.1 (right). White fields denote filter divergence, which is defined here as the case that the averaged RMS error is larger than the observational error.

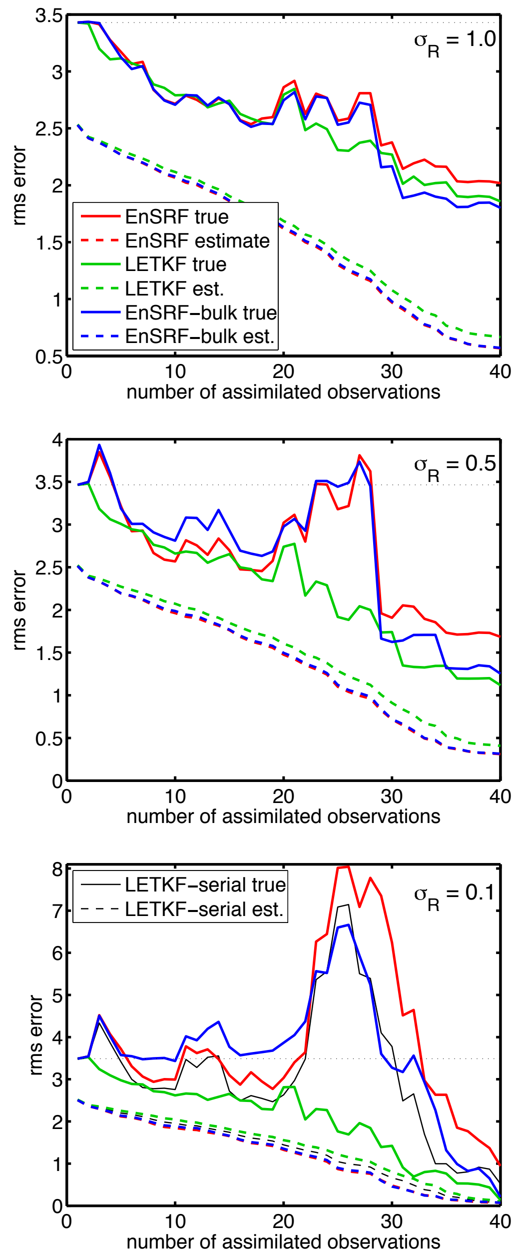


FIG. 2. True and estimated RMS errors for the first analysis step as a function of the number of assimilated observations for observation errors $\sigma_R = 1.0$ (top), 0.5 (middle), and 0.1 (bottom) for the case of $\rho = 0.95$ and a support radius of 20 grid points. Shown are errors for the cases EnSRF (red), LETKF (green), and EnSRF-bulk (blue). The solid lines represent the true RMS errors, while the dashed lines are estimate errors. The black dotted line marks the RMS error before the assimilation of observations. The lowermost panel also shows the RMS errors for the case that the LETKF performs serial observation processing (blue). The error increase for serial observation processing is caused by the inconsistent covariance update induced by the localization and by different localization influences of OL and CL.

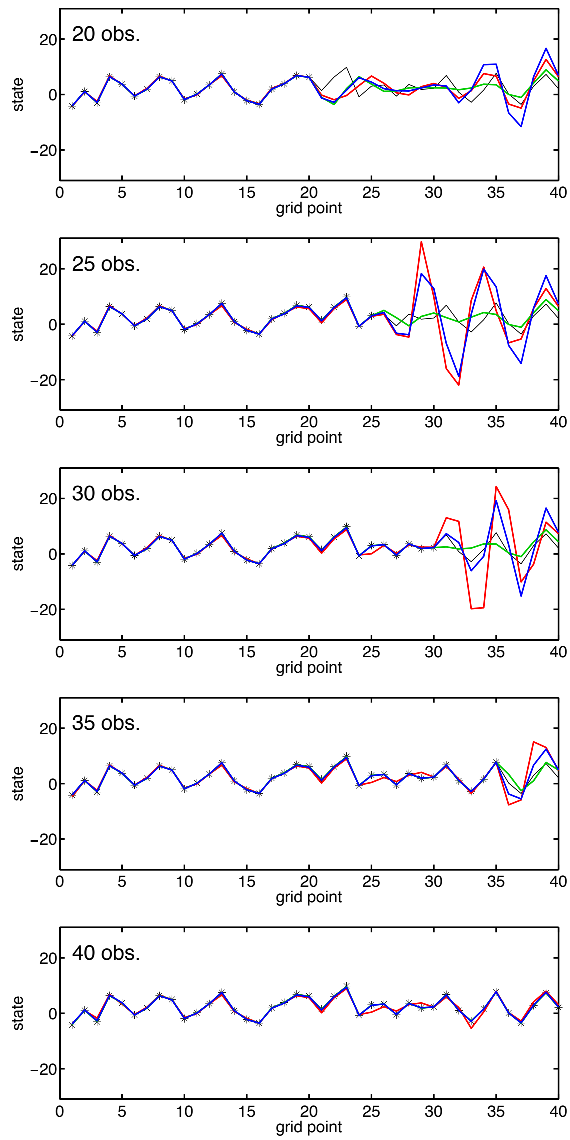


FIG. 3. Sequence of state estimates from EnSRF (red), LETKF (green), and EnSRF-bulk (blue) for different numbers of assimilated observations for $\sigma_R = 0.1$ (bottom), $\rho = 0.95$ and a support radius of 20 grid points. Shown are also the true state (black) and the observations (stars).

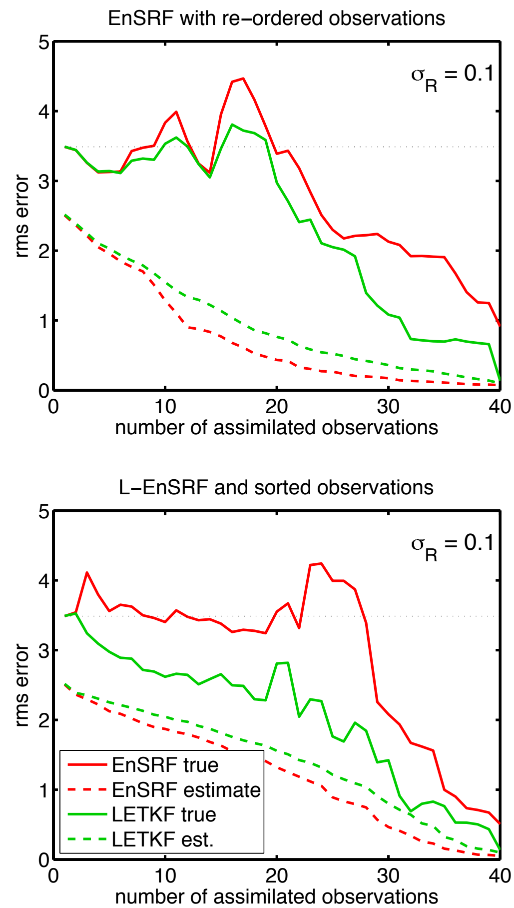


FIG. 4. True and estimated RMS errors for the first analysis step as a function of the number of assimilated observations for $\sigma_R = 0.1$ for the case of $\rho = 0.95$ and a support radius of 20 grid points. Shown are the errors for EnSRF with observations ordered for maximum distance (top) and error for the EnSRF with local analysis and observations sorted for decreasing influence (bottom).

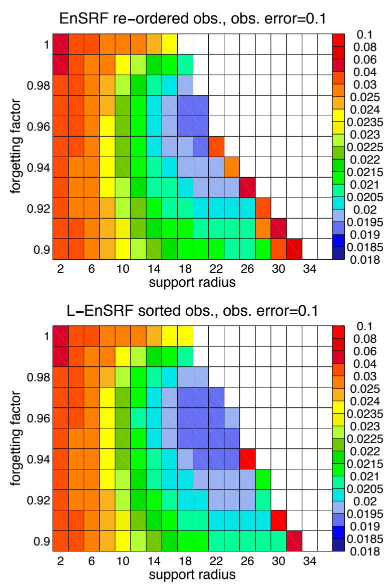


FIG. 5. Average RMS errors for $\sigma_R = 0.1$. Shown are the errors for the EnSRF with observations ordered for maximum distance (top) and the EnSRF with local analysis and observations sorted for decreasing influence (bottom).