Original Article

https://doi.org/10.1590/s2675-28242020068309

OCEAN AND COASTAL RESEARCH

Covariance localization in the ensemble transform Kalman filter based on an augmented ensemble

Jichao Wang[1,2]*, Jiacheng Li[1], Shaodong Zang[1], Jungang Yang[2], Guoli Wu[1]

[1] College of Science, China University of Petroleum, Qingdao, 266580, China
[2] First Institute of Oceanography, Ministry of Natural Resources, Qingdao, 266061, China

*Corresponding author: wangjc@upc.edu.cn

Abstract

With the increased density of available observation data, data assimilation has become an increasingly important tool in marine research. However, the success of the ensemble Kalman filter is highly dependent on the size of the ensemble. A small ensemble used in data assimilation could cause filter divergence, undersampling and spurious correlations. The primary method to alleviate these problems is localization. It can eliminate some spurious correlations and increase the rank of the forecast error covariance matrix. The ensemble transform Kalman filter has been widely used in various studies as a deterministic filter. Unfortunately, the covariance localization cannot be directly applied to ensemble transform Kalman filter. The new covariance localization needs to be presented to adapt the ensemble transform Kalman filter. Based on the method of expanded ensemble and eigenvalue decomposition, this study describes a variation of covariance localization that takes advantage of an unbiased covariance matrix from the expanded ensemble. Experiments described herein show that the new method outperforms the localization methods proposed by others when used in the ensemble transform Kalman filter. The new method yields an analysis estimate that is closer to the true state under different experimental conditions.

Descriptors: Data Assimilation, Ensemble Transform Kalman Filter, Covariance Localization, Augmented Ensemble.

INTRODUCTION

Data assimilation in oceanography and meteorology seeks to provide a current analysis of the state of the atmosphere and ocean, to be used as an initial condition for a forecast. The use of Monte Carlo experiments and ensemble data assimilation also have long roots in Numerical Weather Prediction (NWP), in particular in the fields of ensemble forecasting and observing system simulation experiments (OSSEs). Ensemble data assimilation methods are being continuously improved and have become a viable choice in operational numerical prediction. The Ensemble Kalman filter (EnKF) was originally introduced by Evensen (1994), an outstanding ensemble data assimilation method that allows for nonlinear models being employed in a formulation based on the Kalman filter. In ensemble-based data assimilation, observation error and forecast error must be resolved. For the study of observation error, please see the article by Zang and Wang (2019). The statistical accuracy of forecast error is extremely important for any ensemble data assimilation scheme, since the forecast error covariance matrix (the P^f matrix, hereinafter) is often estimated from ensemble members.

However, the computational cost of producing an ensemble of forecasts large enough to estimate

Submitted on: 3/February/2020
Approved on: 24/October/2020
Associate Editor: Edmo J. D. Campos

Editor: Rubens M. Lopes

sufficiently accurate and high-rank covariance matrices is prohibitive. Current NWP models have state spaces of O(10^7) (UK Met Office, 2008) and thus require a large ensemble to adequately represent the statistics. Typically, an ensemble filter uses a smaller number of ensemble members than the size of the state. If it is so small that the ensemble is not statistically representative, then the system is said to be undersampled.

In general, undersampling would lead to filter divergence, underestimation of forecast error covariance and spurious correlations. Therefore, the EnKF could generate a non-optimal analysis due to these problems. To overcome these obstacles, various methods have been presented in recent years.

The most famous methods are covariance inflation, local analysis and covariance localization in the current ensemble-based data assimilation research. Covariance inflation is a method of correcting an underestimation in the forecast error covariance, which was introduced by Anderson and Anderson (1999). The main idea of covariance inflation is to increase the forecast error covariance by inflating the difference between ensemble mean and ensemble members. It can be applied before the analysis calculations, $x_i^f \leftarrow q\left(x_i^f - \overline{x}^f\right) + \overline{x}^f$, where the $\leftarrow$ represents the replacement of the previous value and q is an inflation factor. Although the covariance inflation overcomes the underestimation of forecast error covariance, the inflation factor does not help to correct the problem of long-range spurious correlations. Therefore, more sophisticated methods are required to settle spurious correlations in the forecast error covariance matrix. There are currently two approaches including the covariance localization and local analysis. Covariance localization (Houtekamer and Mitchell, 2001; Whitaker and Hamill, 2002) is a method that cuts off longer range correlations existing in the forecast error covariance matrix at a specified distance. It works by applying a Schur product (Horn, 1990) to the forecast error covariance matrix P^{f}. An important and significant benefit of the Schur product is that it could increase the rank of forecast error covariance matrix P^{f} (Oke et al., 2007). Local analysis (Hunt et al., 2007; Ott et al., 2004) is presented to perform an analysis to update the local state variables by using local observations. In general, one could update a variable state at a specific location or grid point by assimilating only the observations within a fixed range. The length of the fixed range determines the number of observations during assimilation.

Compared with the covariance localization, the local analysis is an independent approach. It could be used in any data assimilation scheme. Because the Schur product is only applied in the forecast error covariance matrix P^f that is never computed in ensemble transform Kalman filter (ETKF) (Bishop et al., 2001), covariance localization cannot be combined with the ETKF. See Janjić et al. (2011) for a specific discussion of this problem.

Currently, many research studies have focused on covariance localization and local analysis (Greybush et al., 2010; Janjić et al., 2011; Sakov and Bertino, 2011; Jiang and Gorell, 2019; Luo et al., 2019). Although local analysis can be applied in any data assimilation scheme, it does not have the same performance as the covariance localization. Miyoshi and Yamane (2007) pointed out that local analysis is similar to covariance localization, but local analysis usually results in a weak localization effect. Because local analysis is performed independently, smoothness of the analysis fields across the sub-domain boundaries becomes an issue of concern (Janjić et al., 2011). The experiments of Janjić et al. (2011) show that the minimum RMS error obtained for the local analysis (domain localization) is larger than that for covariance localization. Overall, it seems that at strong levels of assimilation, covariance localization produces a more robust analysis (Sakov and Bertino, 2011). Therefore, as can be seen from the aforementioned studies, covariance localization may be a relatively good method.

To overcome the obstacle that covariance localization cannot be applied to the ETKF, Petrie (2008) presented a new approximation of localization in the ETKF. It achieved by applying a Schur product between the square root of the correlation function ρ (an $n \times n$-dimensional matrix, where n denotes the number of variables) and a new forecast ensemble perturbation matrix. The square root of correlation function is calculated by singular value decomposition (SVD) and this new forecast ensemble perturbation matrix is extended by $n-N$ (where N denotes the number of ensembles, $N \ll n$); columns of zeros are added to the raw forecast ensemble perturbation matrix to correct the dimensions and then a new forecast ensemble perturbation matrix obtained. The purpose of this correction is to unify the dimensions of the square root of the correlation function and the new forecast ensemble perturbation matrix. However, this new covariance localization is a poor approximation and the effect of assimilation poor.

Based on sample correlations and matrix factorization, several adapted methods have been developed to compute the localization function, which is dynamically applied as the covariance localization to the ETKF. Bishop and Hodyss (2009a) presented a new method for generating localization functions that depend on the true error correlation functions and that also adapt to the width of the true error correlation function. The method uses ensemble correlations raised to a power (ECO-RAP). The calculation of this method would be too expensive, so Bishop and Hodyss (2009b) used a factorization property for a Covariance Adaptively Localized with ECO-RAP (CALECO) forecast error covariance matrix that, together with other simplifications, reduces the cost. In this new presented method, a CALECO ensemble is provided and each member of the CALECO ensemble is an element-wise product between one raw ensemble member and one column of the square root of the ECO-RAP matrix.

Although the new covariance localization is presented and could be used in the ETKF, there are still two problems in the actual implementation. First, it gives *M* posterior ensemble members but there are only enough computational resources to propagate *N* members ($N < M$) in a practical application leaving the question of how to obtain the *M* posterior ensemble members. Second, how the square root of the ECO-RAP matrix could be rapidly computed must be resolved. Here, we choose a simple method to obtain the *N* posterior ensemble members by using a stochastic subsampling approach, previously recognized as a poor method (Bishop et al., 2017). This method has been adapted, however, using a constant inflation factor to obtain an analysis with sufficient accuracy. Then, instead of the ECO-RAP matrix, the Gaspari and Cohn localization function (GC function) (Gaspari and Cohn, 1999) is used. The covariance localization method presented here was then used in two models to exam its performance. Although the *N* analysis ensemble members are obtained randomly, the effect of localization is satisfactory in subsequent experiments.

Section 2 of this study briefly describes the new covariance localization in the ETKF. Section 3 describes the Lorenz-96 model and Kuramoto–Sivashinsky (KS) equation used in the specific experimental designs. Section 4 summarizes the main results of the paper. Conclusions are in Section 5.

METHOD

Model space covariance localization was chosen as the localization method; Campbell et al., (2010) compared the performances of EnKFs that use model space covariance localization and observation space covariance localization and found that model space covariance localization was superior. Bishop and Hodyss (2009b) showed a simple method of incorporating model space localization with an ETKF-type data assimilation scheme (explained in detail above). Based on that idea, we adapted their method to propose a new covariance localization method.

The expansion of ensemble size

In the ETKF method, the difficultly of applying covariance localization is that an $n \times n$-dimensional forecast error covariance matrix P^f is not explicitly expressed and it is implied through the forecast ensemble perturbation matrix X'^f (See the following paragraph for the definition of matrix X'^f). This causes problems when we wish to introduce a Schur product which acts on the forecast error covariance matrix P^f and not on the perturbation matrix X'^f. Therefore, the current purpose is to achieve a Schur product on the perturbation matrix X'^f, but not the forecast error covariance matrix P^f.

To achieve a Schur product on the perturbation matrix X'^f and the square root of localization function ρ, we introduce the following notation and definitions. Let $W = \left[w_1, \ldots, w_L\right]$ denote the square root of the localization function ρ, so that $\rho = WW^T$. The parameter *L* represents the number of columns of the matrix *W*, the number of selected eigenvalues from the localization function ρ. For more details, please see section 2.2. Let $P^f = X'^f X'^{fT}$, where $X'^f = \left[x_1^f - \overline{x}^f, x_2^f - \overline{x}^f, \ldots, x_N^f - \overline{x}^f\right] \Big/ \sqrt{N-1} = \left[u_1, u_2, \ldots, u_N\right]$ gives the square root of the sample error covariance matrix P^f of the raw *N* ensemble members. $\{x_i^{\mathrm{f}}, i = 1, \ldots, N\}$ represents different ensemble members and $\bar{x}^f$ represents the ensemble mean. Then, the localized ensemble covariance matrix can be given by

$$P^{\mathrm{f}}_{\mathrm{loc}} = P^{\mathrm{f}} \circ \rho = Z^{\mathrm{f}} Z^{\mathrm{f}T} \tag{1}$$

Here *n* is the model space dimension (the number of variables). The square root Z^f of localized forecast error covariance matrix $P^{\mathrm{f}}_{\mathrm{loc}}$ is a $n \times NL$ matrix given by the modulation product (Δ) of X'^f by W defined by

$$Z^{\mathrm{f}} = W \Delta X'^{\mathrm{f}} = [w_1 \circ u_1, w_1 \circ u_2, \ldots, w_1 \circ u_N, \ldots, w_L \circ u_1, w_L \circ u_2, \ldots, w_L \circ u_N] \tag{2}$$

Note that is a square root of the localized forecast error covariance matrix $P^{\mathrm{f}}_{\mathrm{loc}}$. It will replace X'^f as the new perturbation matrix in the subsequent process of assimilation. The expansion of ensemble size is evident from Eq. (2): Z^f has $M = N \times L$ ensemble members whereas X'^f only has *N* ensemble members. This method solves the problem of fewer ensemble members in the assimilation system. Meanwhile, the Schur product has been successfully used in the perturbation matrix X'^f.

GC localization function decomposition

The localization function is normally defined to be a correlation function with local support. Here, the GC localization function ρ is used as a correlation function in the new model space covariance localization. It can be given by,

$$\rho = \begin{cases} 1 - \frac{5}{3}\left(\frac{|z|}{c}\right)^2 + \frac{5}{8}\left(\frac{|z|}{c}\right)^2 + \frac{1}{2}\left(\frac{|z|}{c}\right)^2 - \frac{1}{4}\left(\frac{|z|}{c}\right)^2, 0 \le |z| \le c \\ -\frac{2}{3}\left(\frac{|z|}{c}\right)^2 + 4 - 5\left(\frac{|z|}{c}\right) + \frac{5}{3}\left(\frac{|z|}{c}\right)^2 + \frac{5}{8}\left(\frac{|z|}{c}\right)^3 - \frac{1}{2}\left(\frac{|z|}{c}\right)^4 + \\ \quad \frac{1}{12}\left(\frac{|z|}{c}\right)^5, c \le |z| \le 2c \\ 0, 2c < |z| \end{cases} \tag{3}$$

where z is the Euclidean distance between the two grid points in physical space. When this localization function is applied to specified grid points in the model discrete domain, it becomes a matrix. The length scale c is defined by $c = \sqrt{10/3}r$, where the r is a chosen cutoff length scale. Note that the factor $\sqrt{10/3}$ could tune the localization function to be optimal as determined by Lorenc (2003).

To get the square root matrix of the GC localization function ρ, we used a simple and effective method different than that used in Petrie (2008). First, the eigenvectors and eigenvalues pairs of ρ were computed and ordered from the largest eigenvalue to smallest eigenvalue. Bishop et al. (2017) then used the top 10 largest eigenvalues and their corresponding eigenvectors to construct a square root matrix (*L*=10). Obviously, when dealing with a large number of state variables, the top 10 largest eigenvalues and their corresponding eigenvectors cannot fully represent the error variation of the system variables, so the choice of L is redefined in this paper. Because the top 10% of eigenvalues could account for 70% of the sum of all the eigenvalues, the selected interval of the parameter L was defined as follows,

$$[\min(10, eig_{10\%}), \max(10, eig_{10\%})],$$

where $eig_{10\%}$ represents the number of the first 10% of eigenvalues in descending order. We can determine the optimal value of *L* from the interval by conducting multiple experiments. The value of L can thus be dynamically selected to better represent the error of the variable. Then the square root W of the localization function ρ can be obtained by using the following equation,

$$W = eigenvectors_L \, eigenvalue_L^{1/2} \tag{4}$$

where the $eigenvectors_L$ represents the matrix of the first *L* eigenvectors and $eigenvalue_L^{1/2}$ represents the square root matrix of the matrix of corresponding eigenvalues.

New covariance localization used in ETKF

When the square root $z^f = [z^f_1, \ldots, z^f_M]$ of localized forecast error covariance matrix $P^{\mathrm{f}}_{\mathrm{loc}}$ is obtained from Eq. (2) in section 2.1, then one can use it to create an *M*-member ensemble $v^f = [v^f_1, v^f_2, \ldots, v^f_M]$, the specific calculation is as follows,

$$v^{\mathrm{f}}_k = \bar{v}^{\mathrm{f}} + \sqrt{M} - 1 z^{\mathrm{f}}_k,$$

$$\bar{v}^{\mathrm{f}} = \frac{1}{N} \sum_{j=1}^{N} x^{\mathrm{f}}_j,$$

$$Z^{\mathrm{f}} = \frac{1}{\sqrt{M-1}} [v^{\mathrm{f}}_1 - \bar{v}^{\mathrm{f}}, \ldots, v^{\mathrm{f}}_M - \bar{v}^{\mathrm{f}}], \tag{5}$$

where the constant factor $\sqrt{M-1}$ is different from that of Bishop et al. (2017). This treatment is to ensure the unbiasedness of the covariance estimate. From Eq. (5), the new forecast ensemble has the same mean as the raw ensemble. Then, except for the changes in the forecast ensemble, the next operation is exactly the same as the operations of standard ETKF. The scaled forecast observation ensemble perturbation matrix $\widehat{Y}^f$ is introduced.

$$\widehat{Y}^f = R - \tfrac{1}{2} Y'^f, Y'^f = HZ^f, \tag{6}$$

where R is a $p \times p$ dimension observation error covariance matrix (p denotes the number of observations) and H is the observation operator. Matrix $\widehat{Y}^f$ is the forecast ensemble perturbation matrix projected into the observational space and scaled by $R^{-1/2}$. The effect of scaling the observations by the square root of observation error covariance is to normalize the observation of physical quantities with their respective error standard deviations (Petrie 2008). The singular value decomposition (SVD) of the scaled forecast observation ensemble perturbation matrix $\widehat{Y}^f$ is given as follows,

$$\widehat{Y}^{fT} = U\Sigma V^T, \tag{7}$$

where the matrix U and V are orthogonal matrices, the singular values of $\hat{Y}^{fT}$ are given by Σ. U has dimensions $M \times M$, Σ has dimensions $M \times p$ and V has dimensions $p \times p$. Next, the square root Z^a (analysis perturbation) of analysis error covariance matrix P^a is given by

$$Z^a = Z^f\left[U(\Sigma\Sigma^T + I)^{-1/2} U^T\right] \tag{8}$$

The posterior mean of the distribution of truth is given by

$$\bar{v}^a = \bar{v}^f + Z^f U \Sigma(\Sigma^T\Sigma + I)^{-1} V^T R^{-1/2}(y - H\bar{v}^f). \tag{9}$$

Obviously, if one uses the analysis perturbation matrix Z^a and posterior mean $\bar{v}^a$ to calculate the analysis ensemble V^a, the size of the analysis ensemble V^a would be $n \times M$. But there are only enough computational resources to propagate N members, so the analysis ensemble V^a cannot be used in the next forecast equation.

A STOCHASTIC SUBSAMPLING APPROACH

To solve the problem mentioned in section 2.3, an adapted stochastic subsampling approach method is presented. The specific operation is given by

$$X^a = \bar{v}^a 1_N^T + \sqrt{N-1}\left[Z^a\, randn(M,N)\right] \tag{10}$$

where the matrix X^a is an $n \times N$ analysis ensemble and $1_N^T = \underbrace{[1,\ldots,1]}_{N\text{ elements}}$. Note that the matrix $randn(M,N)$ is an $M \times N$ matrix whose elements are independent random draws from a normal distribution, so it can transform the matrix Z^a to be an $n \times N$ matrix as a square root of an analysis error covariance matrix P^a. The $\sqrt{N-1}$ used in Eq. (9) could be recognized as a constant inflation factor. In the next section, we use these above-mentioned methods in two different models to examine their feasibility.

The pseudo-codes of the new method proposed in this article are as follows:

1. Initialization

• Initial system state x_0^{f}, initial forecast error covariance matrix P_0^{f} and initial forecast ensemble $\{x_i^{\mathrm{f}}\}_{i=1}^{N}$.

Note: in the initialization process of this paper, the forecast error covariance matrix is assumed in advance. The initial forecast ensemble is constructed by using the true value at the initial time and the forecast error covariance matrix. Then the average of the forecast ensemble is taken as the initial system state.

2. For $t_k = 1, 2, \ldots$

(a) Augmented ensemble

• Compute the square root of GC localization function ρ

$$\rho = WW^T, W = [w_1, \ldots, w_L],$$

Selected interval of parameter L:

$$\left[\min(10, eig_{10\%}), \max(10, eig_{10\%})\right]$$

• Compute the forecast ensemble perturbation matrix

$$X'^f = \frac{1}{\sqrt{N-1}}\left[x_1^f - \bar{x}^f - x_2^f - \bar{x}^f, \ldots x_N^f - \bar{x}^f\right] = [u_1, u_2, \ldots, u_N]$$

• A new Schur product between matrix W and matrix X'^f (equivalent to the augmented ensemble)

$$Z^f = W\Delta X'^f = [(w_1 \circ u_1, w_1 \circ u_2, ..., w_1 \circ u_N), ..., (w_L \circ u_1, w_L \circ u_2, ..., w_L \circ u_N)]$$

• Compute new forecast ensemble

$$v_k^f = \bar{v}^f + (\sqrt{M-1}) Z_k, k = 1, ...M$$
$$\bar{v}^f = \frac{1}{N}\sum_{j=1}^{N} x_j^f$$

• Compute new forecast ensemble perturbation matrix

$$Z^f = \frac{1}{\sqrt{M-1}}[(v_1^f - \bar{v}^f), ..., (v_M^f - \bar{v}^f)]$$

(b) Analysis

• The new forecast ensemble perturbation matrix Z^f and new forecast ensemble $\{v_i^f\}_{i=1}^{M}$ used in standard ETKF

• Compute the analysis ensemble perturbation matrix Z^a and associated mean

$$\bar{v}^a = \bar{v}^f + Z^f U \Sigma (\Sigma^T \Sigma + I)^{-1} V^T R^{-1/2} (y - H\bar{v}^f)$$
$$Z^a = Z^f [U(\Sigma\Sigma^T + I)^{-1/2} U^T]$$

• Stochastic subsampling: $[x_1^a, ..., x_N^a] = X^a$
$= \bar{v}^a 1_N^T + \sqrt{N-1}[Z^a \, randn(M,N)]$

(c) Forecast

• Compute the forecast ensemble for $i = 1, ..., N, x_i^f = M(x_i^a)$

and the mean $\bar{x}^f = \frac{1}{N}\sum_{i=1}^{N} x_i^f$

• Compute the forecast ensemble perturbation matrix

$$X'^f = \frac{1}{\sqrt{N-1}}(x_1^f - \bar{x}, ..., x_N^f - \bar{x}^f)$$

MODEL AND EXPERIMENTAL DESIGN

This section describes the experimental designs and the models that are implemented within the ETKF. They are the Lorenz-96 model (Lorenz and Emanuel, 1998) and Kuramoto-Sivashinsky equation (KS) (Sivashinsky, 1977), both used previously in studies of state estimation problems.

Lorenz-96 model

The Lorenz-96 model is a one-dimensional toy-model and has been widely used to test the performance of a data assimilation method. It is assumed to contain n variables $\{x_i\}, i = 1, 2, \ldots, n$. The dynamic system is represented by the following ordinary differential equations,

$$\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})x_{i-1} - x_i + F, i = 1, ..., n \tag{11}$$

Note that the domain in which the n variables are defined is circle-like, so that $x_{-1} = x_{n-1}, x_0 = x_n, x_1 = x_{n+1}$. The constant forcing term F is set as 8 and can cause chaotic behavior. In this study, the Lorenz-96 model is solved by the fourth-order Runge-Kutta scheme with a time step Δt of 0.01 units (the time unit equal to 5 days) until a final time of $T = 20$.

KS equation

The KS equation is expressed as follows,

$$\frac{\partial u}{\partial t} = -u\frac{\partial u}{\partial x} - \frac{\partial^2 u}{\partial x^2} - \frac{\partial^4 u}{\partial x^4} \tag{12}$$

It is a non-linear partial differential equation; complex and chaotic behavior can be produced due to the presence of second and fourth order terms. Here, the KS equation is solved by using an exponential time differentiating Runge-Kutta 4 numerical scheme with a time step of $\Delta t = 0.25$ until a final time of $T = 250$.

Experimental design

To analyze the ETKF with new covariance localization (referred to as NCL) and to demonstrate the comparisons of NCL with the methods presented in Petrie (2008) (referred to as the P_method) and Bishop et al. (2017) (referred to as GETKF), time series values using multiple methods were calculated. The specific experimental operations are demonstrated in this section.

In general, the true trajectory x^t is determined by evolving the perfect model equations from known initial conditions, where the superscript t denotes true. Then the observations *y* can be created by using the true trajectory x^t and observation error covariance matrix *R*,

$$y = Hx^t + v, v \sim N(0,R). \tag{13}$$

where H denotes an observation operator that maps the true state vector into the observation space. Here, the uncorrelated observation errors are considered so that the observation error covariance matrix R is a diagonal matrix in which all diagonal elements are one. As for ensemble members, all the members are evolved using the perfect model equations but beginning from perturbed initial conditions. The details for the experiments are provided as follows.

To generate the true trajectory, the Lorenz-96 equations are started from initial conditions where $x_i = 8, i = 1,\ldots,40$ with a small perturbation of 0.2 added to x_{20}. Then to provide the initial ensemble members for different experiments, the *N* = 5 and *N* = 10 pseudo-random samples from the normal distribution $N\left(0,\sigma_b^2 I\right)$ are added to the true initial condition, where the matrix I demotes an identity matrix and σ_b^2 is the forecast error variance, respectively. The observations are calculated by using Eq. (13) with the number of observations of $p = 35$ (partial observation) or $p = 40$ (full observation) in different experiments, respectively. The frequency varies between experiments, with the chosen frequencies being observations available every 5 or 10 time steps, respectively.

The experimental design for KS equation are similar to the above experimental designs. The true trajectory is defined by the solution to the KS equation the periodic domain $0 \le x \le 32\pi$ from initial conditions $u = \cos\left(\frac{x}{16}\right)\left(1+\sin\left(\frac{x}{16}\right)\right)$ using *n* = 256 spatial points. The initial ensemble members for different experiments are calculated by adding the and pseudo-random samples from the normal distribution $N\left(0,\sigma_b^2 I\right)$ to the true initial condition, respectively. Then, the observations are calculated by using Eq. (12) with the number of observations, $p = 235$ (partial observation) or $p = 256$ (full observation) in experiments, respectively. The frequency varies between experiments, with the chosen frequencies being observations available every 5 or 10 time steps, respectively.

RESULTS AND DISCUSSION

Several experiments were conducted to test the performances of the NCL, P_method and GETKF combined with the ETKF, respectively. All experimental conditions were the same for each of the three methods. Meanwhile, to demonstrate the potential of the proposed method, we performed multiple experiments without covariance localization as a control. In these runs, the other settings are identical to the experiments with covariance localization. Meanwhile, the covariance localization was tested with different frequencies of observations and number of ensemble members in the above-mentioned models.

We first consider the performance of NCL compared with the P_method and GETKF. To evaluate method performance for a particular variable, the root mean square error (RMSE) between the true trajectory and that of the analysis was calculated to assess analysis accuracy. The definition of the RMSE is given by,

$$RMSE = \left(\frac{1}{l}\sum_{j=1}^{l}(x_j^a - x_j^t)^2\right)^{1/2} \tag{14}$$

Here, l represents total assimilation time steps. The parameter settings for different experiments are shown in Table 1. The RMSEs were then calculated for different models and parameters, and the comparisons of the true trajectory and the analysis estimates are shown in Figures 1 to 4. From the different figures of RMSE in two models, there is a clear significant positive impact on the result from the NCL. It not only successfully uses covariance localization in ETKF, but also decreases the RMSE of analysis estimates compared with the method without covariance localization (NOCL), the P_method and GETKF. However, as for the NCL, the RMSEs of the last several model variables have a small increase in the conditions of partial observations (such as the conditions of C, G and H). The reason for this situation may be that these variables have no observations in data assimilation, and the assimilation process is less affected. The NCL had an obvious disadvantage compared with the GETKF for the condition of D, but it still outperforms the method without covariance localization, P_method. Overall, the results of the NCL are satisfactory.

Table 1. The parameter settings for different experiments, here 'all' denotes 40 and 256 for the Lorenz-96 model and KS model, respectively; 'partial' denotes 35 and 235 for the Lorenz-96 model and KS model, respectively.

Condition	Obs. Freq. (time steps)	Obs. Num. (p)	Ens. Num. (N)
A	5	all	5
B	5	all	10
C	5	partial	5
D	5	partial	10
E	10	all	5
F	10	all	10
G	10	partial	5
H	10	partial	10

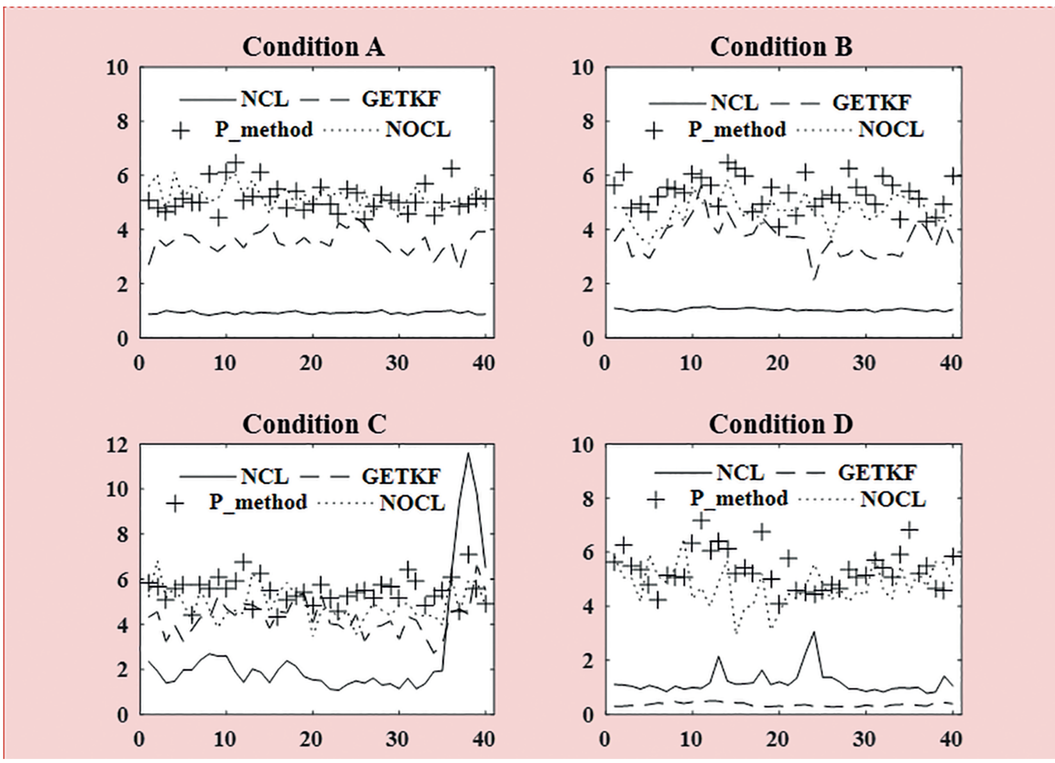

Figure 1. Comparison of the RMSE with different experimental conditions (Lorenz-96 model).

Compared with covariance inflation, the major advantage of covariance localization is that it masks spurious correlations between distant state vector elements. Because the solution of the model used in experiments satisfies periodic boundary conditions, the distance between the state vector elements is determined by the index of two grids. An example of its effect on the Lorenz-96 model with a one-dimensional periodic boundary condition (NCL under experimental condition A), the GC localization function ρ, forecast error covariance matrix P^{f}, Schur product $\rho \circ P^{\mathrm{f}}$ and new Schur product $\left(W\Delta X'^{\mathrm{f}}\right)\left(W\Delta X'^{\mathrm{f}}\right)^{T}$ is displayed from left to right in Fig. 5, which shows that this approximation $\left(W\Delta X'^{\mathrm{f}}\right)\left(W\Delta X'^{\mathrm{f}}\right)^{T} \approx \rho \circ P^{\mathrm{f}}$ effectively removed spurious correlations. The matrix $\left(W\Delta X'^{\mathrm{f}}\right)\left(W\Delta X'^{\mathrm{f}}\right)^{T}$ approximates a diagonal matrix due to the fact that spurious correlations are removed. More importantly, P_method not only reduces the value of the forecast variance but also does not remove the spurious correlations; this result with specific figures has been presented in Petrie (2008).

Filter divergence occurs when an incorrectly specified analysis state is unable to be adjusted

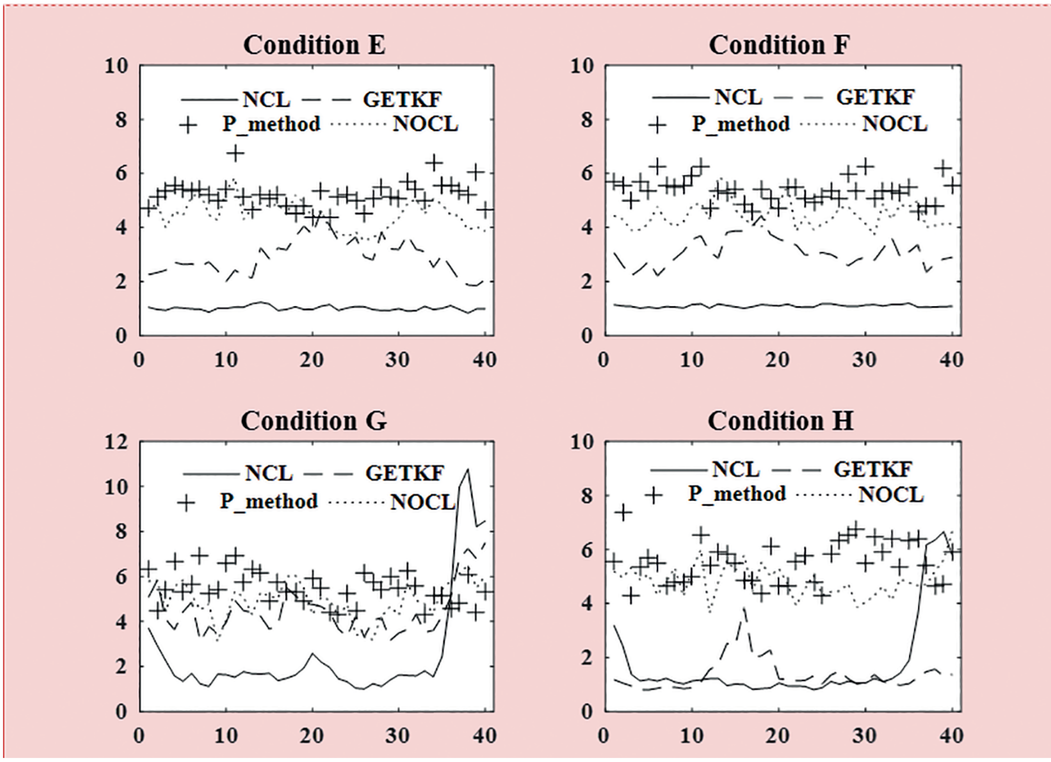

Figure 2. Comparison of the RMSE with different experimental conditions (Lorenz-96 model).

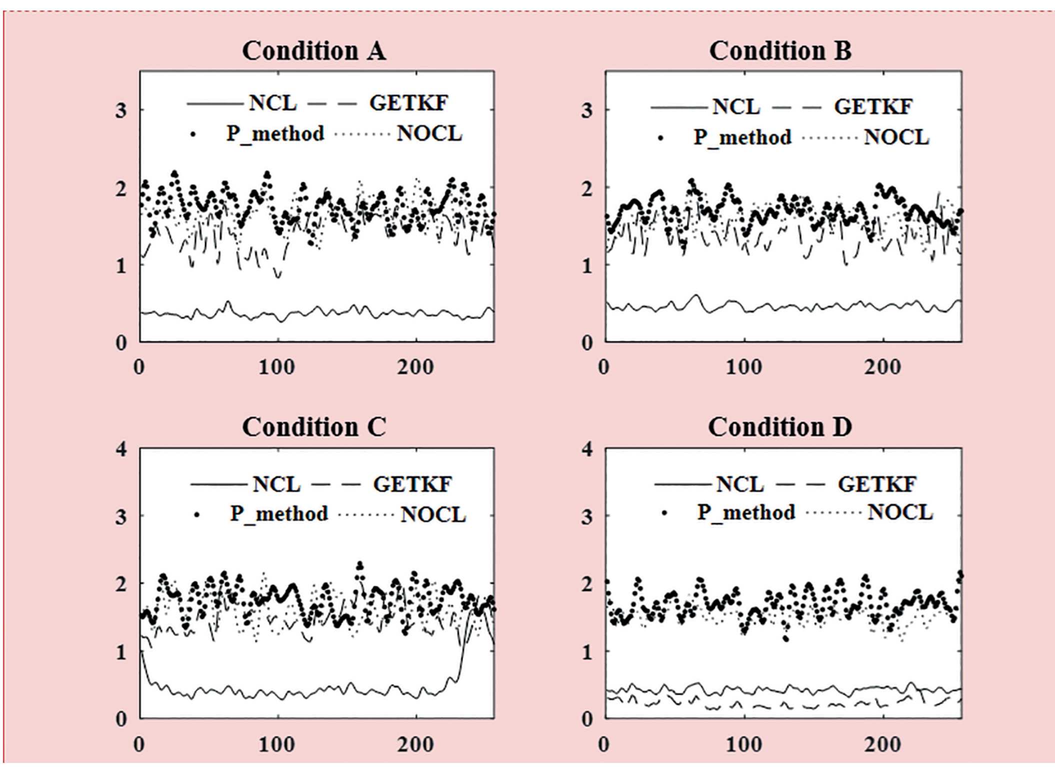

Figure 3. Comparison of the RMSE with different experimental conditions (KS model).

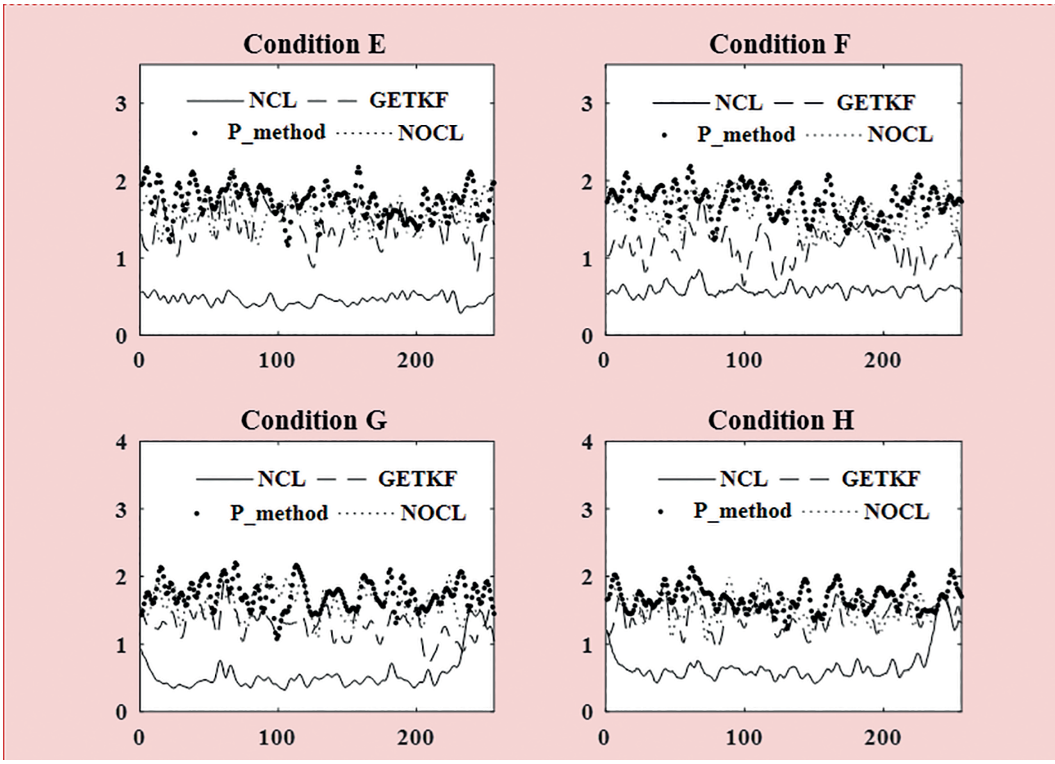

Figure 4. Comparison of the RMSE with different experimental conditions (KS model).

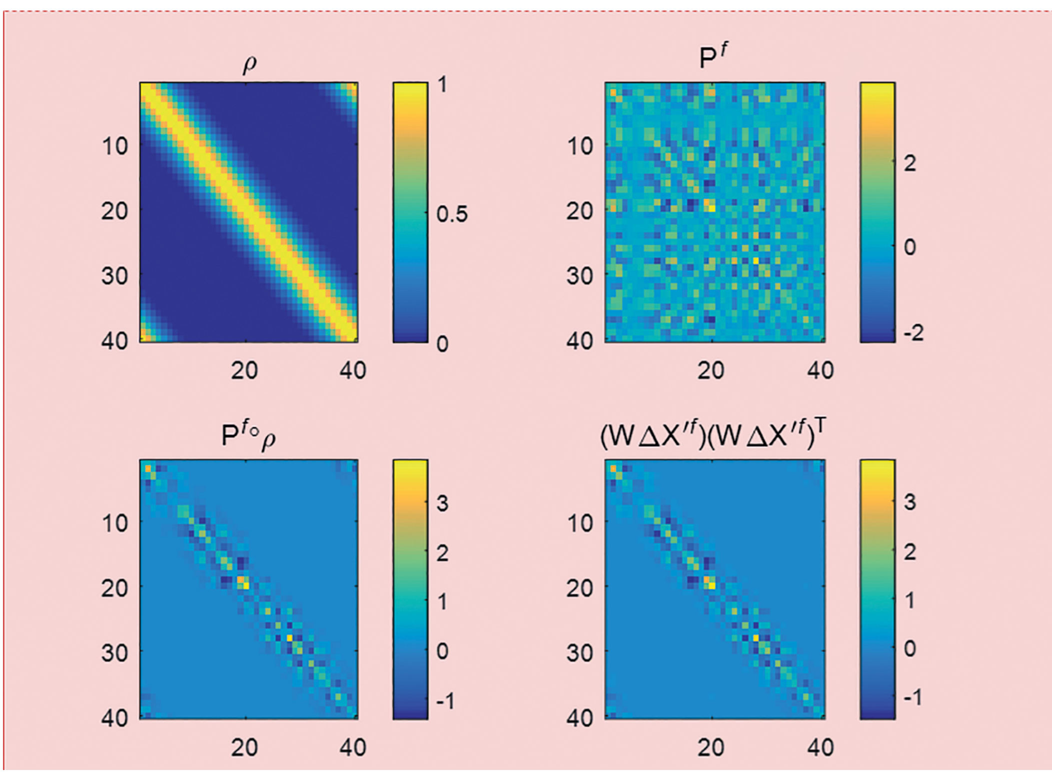

Figure 5. The effect of NCL on the forecast error covariance matrix (Lorenz-96 model, Condition A).

by observation assimilation to more accurately represent the true state of the system. One cause of filter divergence is undersampling; the ensemble member size may be important to the occurrence of filter divergence. Figure 6 and 7 show the difference between analysis states (NCL and GETKF) and the true trajectory in the Lorenz-96 model and the KS model under condition A, respectively. A large difference was observed between the analysis for NCL and GETKF. The analysis with NCL more accurately represented the true solution compared with the other covariance localization assessed. Hence, filter divergence may be well solved by the new method.

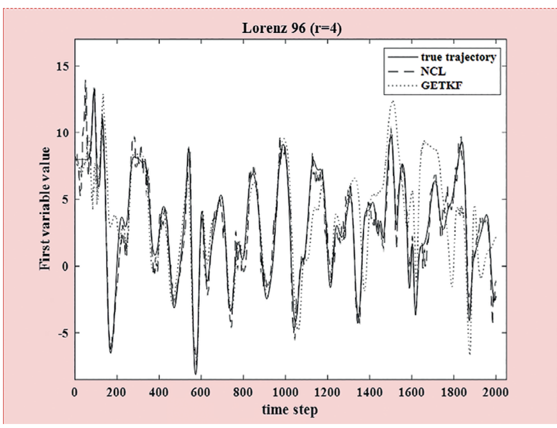

Figure 6. Value of the first variable according to NCL, GETKF and the truth over time (Condition A).

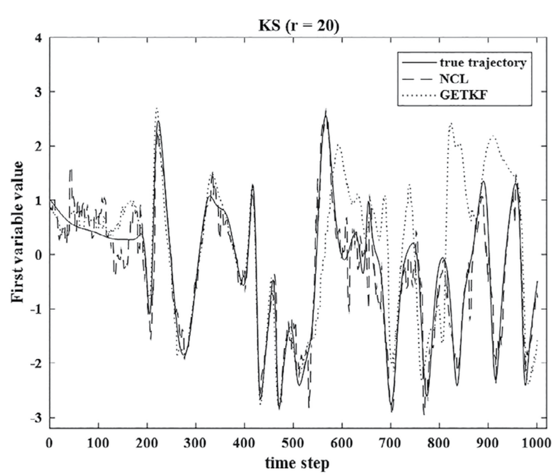

Figure 7. Value of the first variable according to NCL, GETKF and the truth over time (Condition A).

The above results show that the NCL is a good localization method in the models analyzed herein. We recommend the NCL method is implemented in future by other research groups.

CONCLUSIONS

This study proposes a variation of the localization method shown in Bishop et al. (2017) to realize the ETKF with a reduced number of ensemble members. To examine the effects of the method on resolving undersampling, a new method for ETKF that uses a GC localization function has been described. The square root of the localization function is composed of the top 10% of the largest eigenvalue and its corresponding eigenvector pair. From the results of assimilation and the Schur product, the effect of this processing is better. The NCL can remove the long range spurious correlations and the range of forecast error covariance when the Schur product is unchanged, an advantage over the original method. Meanwhile, the NCL method greatly reduces the RMSE in different experimental conditions. Note, however, that RSME would increase occasionally in experiments under partial observation conditions, so the new method still needs to be improved.

ACKNOWLEDGEMENTS

The authors thank the National Key Research and Development Program of China (No. 2017YFC1405600), the National Natural Science Foundation of China (No. 41406007) and the Fundamental Research Funds for the Central Universities (No.19CX05003A-5).

AUTHOR CONTRIBUTIONS

J.W.: Conceptualization; Investigation; Methodology; Writing original draft; Writing - review & editing; Supervision; J.L.: Funding acquisition; Visualization;

S.Z.: Conceptualization; Writing - review & editing; Resources; Visualization; J.Y.: Software; Funding acquisition; G.W.: Visualization; Validation.

REFERENCES

ANDERSON, J. L. & ANDERSON, S. L. 1999. A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Monthly Weather Review*, 127(12), 2741-2758, DOI: https://doi.org/10.1175/1520-0493(1999)127<2741:AMCIOT>2.0.CO;2

BISHOP, C. H., ETHERTON, B. J. & MAJUMDAR, S. J. 2001. Adaptive sampling with the ensemble transform kalman filter. Part I: Theoretical aspects. *Monthly Weather Review*, 129(3), 420-436, DOI: https://doi.org/10.1175/1520-0493(2001)129<0420:ASWTET>2.0.CO;2

BISHOP, C. H. & HODYSS, D. 2009a. Ensemble covariances adaptively localized with ECO-RAP. Part 1: Tests on simple error models. *Tellus A: Dynamic Meteorology and Oceanography*, 61(1), 84-96, DOI: https://doi.org/10.1111/j.1600-0870.2008.00371.x

BISHOP, C. H. & HODYSS, D. 2009b. Ensemble covariances adaptively localized with ECO-RAP. Part 2: A strategy for the atmosphere. *Tellus A: Dynamic Meteorology and Oceanography*, 61(1), 97-111, DOI: https://doi.org/10.1111/j.1600-0870.2008.00372.x

BISHOP, C. H., WHITAKER, J. S. & LEI, L. 2017. Gain form of the ensemble transform Kalman filter and its relevance to satellite data assimilation with model space ensemble covariance localization. *Monthly Weather Review*, 145(11), 4575-4592, DOI: https://doi.org/10.1175/MWR-D-17-0102.1

CAMPBELL, W. F., BISHOP, C. H. & HODYSS, D. 2010. Vertical covariance localization for satellite radiances in ensemble kalman filters. *Monthly Weather Review*, 138(1), 282-290, DOI: https://doi.org/10.1175/2009MWR3017.1

EVENSEN, G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans*, 99(C5), 10143-10162, DOI: https://doi.org/10.1029/94JC00572

GASPARI, G. & COHN, S. E. 1999. Construction of correlation functions in two and three dimensions. Quarterly Journal of the Royal Meteorological Society, 125(554), 723-757, DOI: https://doi.org/10.1002/qj.49712555417

GREYBUSH, S. J., KALNAY, E., MIYOSHI, T., IDE, K. & HUNT, B. R. 2011. Balance and ensemble kalman filter localization techniques. *Monthly Weather Review*, 139(2), 511-522, DOI: https://doi.org/10.1175/2010MWR3328.1

HORN, R. A. 1990. The Hadamard product. In: JOHNSON, C. R. (ed.). *Matrix theory and applications*. Providence: Proceedings of Symposia in Applied Mathematics; American Mathematical Society, 40, 87-169.

HOUTEKAMER, P. L. & MITCHELL, H. L. 2001. A sequential ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review*, 129(1), 123-137.

HUNT, B. R., KOSTELICH, E. J. & SZUNYOGH, I. 2007. Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter. *Physica D: Nonlinear Phenomena*, 230(1-2), 112-126, DOI: https://doi.org/10.1016/j.physd.2006.11.008

JANJIĆ, T., NERGER, L., ALBERTELLA, A., SCHROTER, J. & SKACHKO, S. 2011. On domain localization in ensemble-based Kalman filter algorithms. *Monthly Weather Review*, 139(7), 2046-2060, DOI: https://doi.org/10.1175/2011MWR3552.1

JIANG, J. & GORELL, S. B. 2019. Analysis of localization methods in ensemble Kalman filter with SVD analytical solution. *Journal of Petroleum Science and Engineering*, 175, 919-931, DOI: https://doi.org/10.1016/j.petrol.2019.01.030

LORENC, A. C. 2003. The potential of the ensemble Kalman filter for NWP—a comparison with 4D-Var. *Quarterly Journal of the Royal Meteorological Society*, 129(595), 3183-3203, DOI: https://doi.org/10.1256/qj.02.132

LORENZ, E. N. & EMANUEL, K. A. 1998. Optimal sites for supplementary observations: simulation with a small model. *Journal of the Atmospheric Sciences*, 55(4), 399-414, DOI: https://doi.org/10.1175/1520-0469(1998)055<0399:OSFSWO>2.0.CO;2

LUO, X., LORENTZEN, R. J., VALESTRAND, R. & EVENSEN, G. 2019. Correlation-based adaptive localization for ensemble-based history matching: applied to the Norne field case study. *Society of Petroleum Engineers - Reservoir Evaluation & Engineering*, 22(3), 1-26, DOI: https://doi.org/10.2118/191305-MS

MET OFFICE (UK - Met Office College). 2008. *UK observations* [online]. London: Met Office College. Available at: http://www.metoffice.gov.uk/research/nwp/observations/.

MIYOSHI, T. & YAMANE, S. 2007. Local ensemble transform Kalman filtering with an AGCM at a T159/L48 resolution. *Monthly Weather Review*, 135(11), 3841-3861, DOI: https://doi.org/10.1175/2007MWR1873.1

OKE, P. R., SAKOV, P. & CORNEY, S. P. 2007. Impacts of localisation in the EnKF and EnOI: experiments with a small model. *Ocean Dynamics*, 57, 32-45, DOI: https://doi.org/10.1007/s10236-006-0088-8

OTT, E., HUNT, B. R., SZUNYOGH, I., ZIMIN, A. V., KOSTELICH, E. J., CORAZZA, M., KALNAY, E., PATIL, D. J. & YORKE, J. A. 2004. A local ensemble Kalman filter for atmospheric data assimilation. *Tellus A: Dynamic Meteorology and Oceanography*, 56(5), 415-428, DOI: https://doi.org/10.3402/tellusa.v56i5.14462

PETRIE, R. 2008. *Localization in the ensemble Kalman filter*. MSc. Reading: Atmosphere, Ocean and Climate University of Reading.

SAKOV, P. & BERTINO, L. 2011. Relation between two common localisation methods for the EnKF. *Computational Geosciences*, 15, 225-237, DOI: https://doi.org/10.1007/s10596-010-9202-6

SIVASHINSKY, G. I. 1977. Nonlinear analysis of hydrodynamic instability in laminar flames-I. Derivation of basic equations. *Acta Astronautica*, 4(11-12), 1177-1206, DOI: https://doi.org/10.1016/0094-5765(77)90096-0

WHITAKER, J. S. & HAMILL, T. M. 2002. Ensemble data assimilation without perturbed observations. *Monthly Weather Review*, 130(7), 1913-1924, DOI: https://doi.org/10.1175/1520-0493(2002)130<1913:edawpo>2.0.co;2

ZANG, S. & WANG, J. 2019. Considering the impact of observation error correlation in ensemble square-root Kalman filter. *Brazilian Journal of Oceanography*, 67, e19261, DOI: https://doi.org/10.1590/s1679-87592019026106717