



**Multivariate  
localization methods  
for ensemble Kalman  
filtering**

S. Roh et al.

# Multivariate localization methods for ensemble Kalman filtering

S. Roh<sup>1</sup>, M. Jun<sup>1</sup>, I. Szunyogh<sup>2</sup>, and M. G. Genton<sup>3</sup>

<sup>1</sup>Department of Statistics, Texas A&M University, College Station, TX, USA

<sup>2</sup>Department of Atmospheric Sciences, Texas A&M University, College Station, TX, USA

<sup>3</sup>CEMSE Division, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

Received: 14 April 2015 – Accepted: 17 April 2015 – Published: 8 May 2015

Correspondence to: M. Jun (mjun@stat.tamu.edu)

Published by Copernicus Publications on behalf of the European Geosciences Union & the American Geophysical Union.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





## Multivariate localization methods for ensemble Kalman filtering

S. Roh et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



of a background (prior) distribution based on the latest observations. The background distribution is represented by an ensemble of short-term forecasts from the previous analysis time. This ensemble is called the background ensemble.

Because the number of background ensemble members that is feasible to use in a realistic atmospheric model is small, the estimates of weak covariances (the entries with small absolute values in the background covariance matrix) tend to have large relative estimation errors. These large relative errors have a strong negative effect on the accuracy of an EnKF estimate of the analysis mean. The standard approach to alleviating this problem is to apply a physical-distance-dependent localization to the sample background covariances before their use in the state update step of the EnKF. In essence, localization is a method to introduce the empirical understanding that the true background covariances tend to rapidly decrease with distance into the state estimation process.

Data assimilation schemes treat the spatially discretized state vector,  $\mathbf{x}$ , as a multivariate random variable. We use the conventional notation  $\mathbf{x}^b$  and  $\mathbf{x}^a$  for the background and the analysis state vectors, respectively. We also use the notation  $\mathbf{y}^o$  for the vector of observations. In an EnKF scheme, the analysis mean,  $\bar{\mathbf{x}}^a$ , is computed from the background mean,  $\bar{\mathbf{x}}^b$ , by the update equation

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^b + \mathbf{K} \left( \mathbf{y}^o - \overline{h(\mathbf{x}^b)} \right). \quad (1)$$

The function  $h(\cdot)$  is the observation function, which maps the finite-dimensional state vector into observables. Thus,  $\overline{h(\mathbf{x}^b)}$  is the ensemble mean of the prediction of the observations by the background. The matrix

$$\mathbf{K} = \mathbf{P}^b \mathbf{H}^T \left( \mathbf{H} \mathbf{P}^b \mathbf{H}^T + \mathbf{R} \right)^{-1} \quad (2)$$

is the Kalman gain matrix, where  $\mathbf{P}^b$  is the background covariance matrix,  $\mathbf{H}$  is the linearization of  $h$  about  $\bar{\mathbf{x}}^b$  and  $\mathbf{R}$  is the observation error covariance matrix. The entry  $K_{ij}$

## Multivariate localization methods for ensemble Kalman filtering

S. Roh et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



of  $\mathbf{K}$  determines the effect of the  $j$ th observation on the  $i$ th component of the analysis mean,  $\bar{\mathbf{x}}^a$ . Under the standard assumption that the observation errors are uncorrelated, the matrix,  $\mathbf{R}$ , is diagonal. Hence, the way the effect of the observations is spread from the observations to the different locations and state variables is determined by  $\mathbf{P}^b$  and  $\mathbf{H}$ . The sampling variability in the estimates of  $\mathbf{P}^b$  affects the accuracy of the information propagated in space and between the different state variables through the matrix products,  $\mathbf{P}^b \mathbf{H}^T$  and  $\mathbf{H} \mathbf{P}^b \mathbf{H}^T$ . The goal of localization is to reduce the related effects of sampling variability on the estimates of  $\mathbf{K}$ .

Over the years, many different localization methods have been proposed. Hamill et al. (2001); Houtekamer and Mitchell (1998, 2001); Hunt et al. (2007); Ott et al. (2004), and Whitaker and Hamill (2002) used localization functions which set the covariance to zero beyond a certain distance (localization radius). Jun et al. (2011) proposed a nonparametric statistical method to estimate the covariance. Anderson (2007) used a hierarchical ensemble filter which estimates the covariance using an ensemble of ensemble filters. Bishop and Hodyss (2007, 2009a, b) adaptively determined the width of localization by computing powers of the sample correlations. Buehner and Charron (2007) examined the spectral and spatial localization of error covariance. Anderson and Lei (2013) and Lei and Anderson (2014) proposed an empirical localization function based on the output of an observing system simulation experiment.

The focus of the present paper is on the family of schemes that localize the covariances by taking the Schur (Hadamard) product of the sample background covariance matrix and a correlation matrix of the same size, whose entries are obtained by the discretization of a distance-dependent correlation function with local (compact) support (e.g., Hamill et al., 2001; Houtekamer and Mitchell, 2001; Whitaker and Hamill, 2002). Such a correlation function is usually called a *localization* or *taper function*. The commonly used localization functions were introduced by Gaspari and Cohn (1999). Beyond a certain distance, all localization functions become zero, forcing the filtered estimates of the background covariance between state variables at locations that are

far apart in space to zero. This property of the filtered background covariances can also be exploited to increase the computational efficiency of the EnKF schemes.

Equations (1) and (2) provide the solution of a formulation of the data assimilation problem that assumes that  $\mathbf{P}^b$  is invertible (e.g., Sects. 4.2.1 and 4.2.3 of Szunyogh, 2014). Because  $\mathbf{P}^b$  is symmetric, its eigenvalues are real and non-negative, which implies that it is invertible, only if it is also positive-definite. (An  $n \times n$  symmetric matrix  $\mathbf{A}$  is positive-definite if  $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$  for all non-zero vectors  $\mathbf{x} \in \mathbb{R}^n$ .) Thus, Eqs. (1) and (2) provide a solution of the formal data assimilation problem for a localized estimate of  $\mathbf{P}^b$ , only if that localized estimate is symmetric and positive-definite.

Because the computation of the right-hand-side of Eq. (2) does not require the invertibility of  $\mathbf{P}^b$ , singularity of the localized  $\mathbf{P}^b$  usually does not lead to a breakdown of the computations in practice. This does not change the fact, however, that the solution obtained by Eqs. (1) and (2) can no longer be the solution of the problem that they were designed to solve. In addition, an ill-conditioned estimate of  $\mathbf{P}^b$  can degrade the conditioning (increase the condition number) of  $\mathbf{H} \mathbf{P}^b \mathbf{H}^T + \mathbf{R}$ , making the numerical computation of the right-hand side of Eq. (2) less stable.

A realistic atmospheric model has multiple scalar state variables (e.g., temperature, coordinates of the wind vector, surface pressure, humidity). If a univariate localization function, such as that described by Gaspari and Cohn (1999), is applied directly to a multivariate state vector, the resulting localized background covariance matrix may not be positive-definite. This motivates us to seek rigorously-derived multivariate localization functions for ensemble Kalman filtering. As will be demonstrated, such rigorously-derived multivariate localization functions often produce more accurate analyses than those that apply the same univariate localization functions to each scalar component of the state vector. Kang et al. (2011) also introduced a multivariate localization method that zeros out covariances between physically unrelated variables. Their motivation to zero out cross-covariances, however, was to filter apparent spurious covariances rather than to preserve the positive-definiteness of the background error covariance matrix.

Multivariate localization methods for ensemble Kalman filtering

S. Roh et al.

Title Page

Abstract Introduction

Conclusions References

Tables Figures

◀ ▶

◀ ▶

Back Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper

## Multivariate localization methods for ensemble Kalman filtering

S. Roh et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



In our search for proper multivariate localization functions, we take advantage of recent developments in the statistics literature. In particular, we use the localization functions developed in Porcu et al. (2012), who studied the radial basis functions to construct multivariate correlation functions with compact support. Note that Sect. 5 in Zhang and Du (2008) described general methodology for covariance tapering in the case of multiple state variables. Du and Ma (2013) used a convolution approach and a mixture approach to derive covariance matrix functions with compactly supported marginal and cross-covariances. Kleiber and Porcu (2015) constructed nonstationary correlation functions with compact support for multivariate random fields. Genton and Kleiber (2015) reviewed approaches to building cross-covariance models such as compactly supported correlation functions for multivariate Gaussian random fields.

The rest of the paper is organized as follows. Section 2 briefly describes EnKF and localization for the special case of two state variables. Section 3 describes the bivariate Lorenz-95 model we use to test our ideas. Section 4 summarizes the main results of the paper.

## 2 Methodology

### 2.1 Univariate localization

In principle, localization can be implemented by using filtered estimates of the background covariances rather than the raw sample covariances to define the matrix,  $\mathbf{P}^b$ , used in the computation of  $\mathbf{K}$  by Eq. (2). The filtered (localized) version of covariance matrix,  $\tilde{\mathbf{P}}^b$ , is obtained by computing the Schur (entry-wise) product:

$$\tilde{\mathbf{P}}^b = \hat{\mathbf{P}}^b \circ \mathbf{C}, \quad (3)$$

where  $\mathbf{C}$  is a correlation matrix, which has the same dimensions as the sample covariance matrix,  $\hat{\mathbf{P}}^b$ . In practice, however, the localization is often done by taking advantage of the fact that localization affects the analysis through  $\mathbf{P}^b \mathbf{H}^T$  and  $\mathbf{H} \mathbf{P}^b \mathbf{H}^T$ , or, ultimately,

## Multivariate localization methods for ensemble Kalman filtering

S. Roh et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



through  $\mathbf{K}$ . In particular, because a distance,  $d$ , can be defined for each entry,  $K_{ij}$ , of  $\mathbf{K}$  by the distance between the  $i$ th analyzed variable and the  $j$ th observation, the simplest localization strategy is to set all entries,  $K_{ij}$ , that are associated with a distance longer than a prescribed localization radius,  $R$  ( $d > R$ ), to zero, while leaving the remaining entries unchanged (e.g., Houtekamer and Mitchell, 1998; Ott et al., 2004; Hunt et al., 2007).

Another approach is to localize  $\mathbf{P}^b \mathbf{H}^T$  and  $\mathbf{H} \mathbf{P}^b \mathbf{H}^T$  by a tapering function (e.g., Hamill et al., 2001; Houtekamer and Mitchell, 2001). The usual justification for this approach is that  $\mathbf{H}$  is typically the linearization of a local interpolation function,  $h(\cdot)$ , for which the localized matrix products provide good approximations of the products computed by using localized estimates of  $\mathbf{P}^b$ . Note that  $\mathbf{P}^b \mathbf{H}^T$  is the matrix of background covariances between the state variables at the model grid points and at the observation locations, while  $\mathbf{H} \mathbf{P}^b \mathbf{H}^T$  is the matrix of background covariances between the state variables at the observation locations. Thus, a distance can be associated with each entry of the two matrix products, which makes the distance-dependent localization of the two products possible. The approach becomes problematic, however, when  $h(\cdot)$  is not a local function, which is the typical case for remotely sensed observations (Campbell et al., 2010).

We consider the situation where localization is applied directly to the background error covariance matrix,  $\hat{\mathbf{P}}^b$ . Recall that the localized covariance matrix is expressed as in Eq. (3). In particular,  $\mathbf{C}$  is a positive-definite matrix with strictly positive eigenvalues, while the sample covariance matrix,  $\hat{\mathbf{P}}^b$ , may have zero eigenvalues (as it is only non-negative definite). The localization in Eq. (3) helps to eliminate those zero eigenvalues of  $\hat{\mathbf{P}}^b$  and alleviate the related large relative estimation errors. The positive-definiteness of  $\mathbf{C}$  ensures that localization does not introduce new zero eigenvalues in the process of eliminating the zero eigenvalues of  $\hat{\mathbf{P}}^b$ . The proper definition of the localization function that ensures that  $\mathbf{C}$  is positive-definite has been thoroughly investigated for the univariate case ( $N = 1$ ) in the literature (e.g. Gaspari and Cohn, 1999).

## 2.2 Multivariate localization

We now consider a model with  $N$  state variables. For instance, we take a simple model based on the hydrostatic primitive equations. This model solves the equations for the two horizontal components of wind, the surface pressure, the virtual temperature and a couple of atmospheric constituents. The state of the model is represented by the state vector,  $\mathbf{x} = (x_1, x_2, \dots, x_N)$ , where  $x_i, i = 1, 2, \dots, N$ , represents the spatially discretized state of the  $i$ th state variable in the model.

The sample background covariance matrix,  $\hat{\mathbf{P}}^b$ , can be partitioned as

$$\hat{\mathbf{P}}^b = \begin{pmatrix} \hat{\mathbf{P}}_{11}^b & \hat{\mathbf{P}}_{12}^b & \cdots & \hat{\mathbf{P}}_{1N}^b \\ \hat{\mathbf{P}}_{21}^b & \hat{\mathbf{P}}_{22}^b & \cdots & \hat{\mathbf{P}}_{2N}^b \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\mathbf{P}}_{N1}^b & \hat{\mathbf{P}}_{N2}^b & \cdots & \hat{\mathbf{P}}_{NN}^b \end{pmatrix}. \quad (4)$$

The entries of the submatrices,  $\hat{\mathbf{P}}_{ij}^b, i = 1, \dots, N$ , are the marginal-covariances for the  $i$ th state variable. In practical terms, if the  $i$ th state variable is the virtual temperature, for instance, each diagonal entry of  $\hat{\mathbf{P}}_{ii}^b$  represents the sample variance for the virtual temperature at a given model grid point, while each off-diagonal entry of  $\hat{\mathbf{P}}_{ij}^b$  represents the sample covariances between the virtual temperatures at a pair of grid points. Likewise, the entries of  $\hat{\mathbf{P}}_{ij}^b, i \neq j$ , are the sample cross-covariances between the grid point values of the  $i$ th and the  $j$ th state variables at pairs of locations, where the two locations for an entry can be the same grid point.

We thus consider matrix-valued localization functions,  $\boldsymbol{\rho}(d) = \{\rho_{ij}(d)\}_{i,j=1,\dots,N}$ , which are continuous functions of  $d$ . The component  $\rho_{ij}(d)$  of  $\boldsymbol{\rho}(d)$  is the localization function used for the calculation of the covariances included in the sub-matrix  $\mathbf{P}_{ij}^b$  of  $\mathbf{P}^b$ . Each entry of the localization matrix  $\mathbf{C}$  is computed by considering the value of the appropriate component of  $\boldsymbol{\rho}(d)$  for a particular pair of state variables and the distance,  $d$ , associated with the related entry of  $\hat{\mathbf{P}}^b$ .



## Multivariate localization methods for ensemble Kalman filtering

S. Roh et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



In order to get a proper matrix-valued localization function,  $\rho$ , a seemingly obvious approach to extend the results of Gaspari and Cohn (1999) would be to compute the entries of  $\mathbf{C}$  based on a univariate correlation function for a multivariate variable. That is, for the pair of state variables  $i$  and  $j$ , we localize the corresponding sample background covariance matrix,  $\hat{\mathbf{P}}_{ij}^b$ , by multiplying a localization matrix from the same correlation function for all  $i$  and  $j$ . Formally, this would be possible because the distance  $d$  is uniquely defined for each entry of  $\hat{\mathbf{P}}^b$  the same way in the multivariate case as in the univariate case. This approach, however, cannot guarantee the positive-definiteness of the resulting matrix,  $\mathbf{C}$ . As a simple illustrative example, consider the situation where the discretized state vector has only two components that are defined by two different scalar state variables at the same location (e.g., the temperature and the pressure). In this case, if  $n$  is the number of locations, the localization matrix for the two state variables together can be written as

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_0 & \mathbf{C}_0 \\ \mathbf{C}_0 & \mathbf{C}_0 \end{pmatrix} \quad (5)$$

independently of the particular choice of the localization function. Here  $\mathbf{C}_0$  is an  $n \times n$  localization matrix from a univariate localization function. From Eq. (5), it is clear that  $n$  eigenvalues of  $\mathbf{C}$  are zero and the rank of  $\mathbf{C}$  is  $n$ , while its dimension is  $2n \times 2n$ .

As in Eq. (2), although  $\mathbf{C}$  is rank-deficient and thus so is the localized covariance matrix  $\tilde{\mathbf{P}}^b$ , we may still be able to calculate the inverse of  $\mathbf{H}\tilde{\mathbf{P}}^b\mathbf{H}^T + \mathbf{R}$ , as  $\mathbf{R}$  is a diagonal matrix. The smallest eigenvalue of  $\mathbf{H}\tilde{\mathbf{P}}^b\mathbf{H}^T + \mathbf{R}$  is the smallest (positive) value of  $\mathbf{R}$ , and thus the matrix,  $\mathbf{H}\tilde{\mathbf{P}}^b\mathbf{H}^T + \mathbf{R}$ , is still invertible and has positive eigenvalues. However, unless the diagonal elements of  $\mathbf{R}$  are large (which implies large observation error variance), the matrix  $\mathbf{H}\tilde{\mathbf{P}}^b\mathbf{H}^T + \mathbf{R}$  is seriously ill-conditioned and the computation of its inverse may be numerically unstable. Therefore, the numerical stability of the computation of the inverse of the matrix heavily relies on the observation error variance, which is an undesirable property.



same for each pair of state variables, leaving no flexibility to account for the potential differences in the correlation lengths and decay rate for the different state vector components.

The second proposed method takes advantage of the availability of multivariate compactly supported functions from the spatial statistics literature. To the best of our knowledge, only a few papers have been published on this subject; one of them is Porcu et al. (2012). The function class they considered was essentially a multivariate extension of the Askey function (Askey, 1973),  $f(d; \nu, c) = \left(1 - \frac{d}{c}\right)_+^\nu$ , with  $c, \nu > 0$ . Here,  $x_+ = \max(x, 0)$  for  $x \in \mathbb{R}$ . For instance, a bivariate Askey function, which is a special case of the results of Porcu et al. (2012), is given by ( $i, j = 1, 2$ )

$$\rho_{ij}(d; \nu, c) = \beta_{ij} \left(1 - \frac{d}{c}\right)_+^{\nu + \mu_{ij}}, \quad (8)$$

where  $c > 0$ ,  $\mu_{12} = \mu_{21} \leq \frac{1}{2}(\mu_{11} + \mu_{22})$ ,  $\nu \geq [\frac{1}{2}s] + 2$ ,  $\beta_{ii} = 1$  ( $i = 1, 2$ ),  $\beta_{12} = \beta_{21}$ , and

$$|\beta_{12}| \leq \frac{\Gamma(1 + \mu_{12})}{\Gamma(1 + \nu + \mu_{12})} \sqrt{\frac{\Gamma(1 + \nu + \mu_{11})\Gamma(1 + \nu + \mu_{22})}{\Gamma(1 + \mu_{11})\Gamma(1 + \mu_{22})}}. \quad (9)$$

Here,  $\Gamma(\cdot)$  is the gamma function (e.g., Wilks, 2006),  $s$  is the dimension of the Euclidean space where the state variable is defined, and  $[x]$  is the largest integer that is equal to or smaller than  $x$ . The Askey function in Eq. (8) has the support  $c$  because it sets covariances beyond a distance  $c$  to zero. It can be seen from Eq. (9) that, if the scalars,  $\mu_{ij}$ , are chosen to be the same for all values of  $i$  and  $j$ , the condition on  $\beta_{12}$  for  $\rho$  to be valid is  $|\beta_{12}| \leq 1$ . Note that, for this choice, the second method is essentially the same as the first method with the Askey function set to  $\tilde{\rho}$ . The localization function given by Eq. (8) is more flexible than the functions of the first method with the Askey function set to  $\tilde{\rho}$  because  $\mu_{ij}$  can be chosen to be different for each pair of indexes,  $i$  and  $j$ . The localization length, however, is still the same for the different pairs of the state variables.

To illustrate the differences between the shape of the Gaspari–Cohn and the Askey functions, we show the Gaspari–Cohn function for  $c = 25$  and the univariate Askey

**Multivariate localization methods for ensemble Kalman filtering**

S. Roh et al.

Title Page	
Abstract	Introduction
Conclusions	References
Tables	Figures
◀	▶
◀	▶
Back	Close
Full Screen / Esc	
Printer-friendly Version	
Interactive Discussion	



function for  $c = 50$ ,  $v = 1, 2, 3$ , and  $\mu_{11} = 0$  (Fig. 1). This figure shows that for a given support, the Askey functions are narrower.

### 3 Experiments

#### 3.1 The EnKF scheme

5 There are many different formulations of the EnKF update equations, which produce not only an updated estimate of the mean, but also the ensemble of analysis perturbations that are added to the mean to obtain an ensemble of analyses. This ensemble of analyses serves as the ensemble of initial conditions for the model integration that produce the background ensemble. In our experiments, we use the method of perturbed observations. It obtains the analysis mean and the ensemble of analysis perturbations  
10 by the equations

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^b + \mathbf{K}(\mathbf{y} - \mathbf{H}\bar{\mathbf{x}}^b), \quad (10)$$

$$\mathbf{x}_k^{a'} = \mathbf{x}_k^{b'} + \mathbf{K}(\mathbf{y}_k^{o'} - \mathbf{H}\mathbf{x}_k^{b'}), \quad (11)$$

15 where  $\mathbf{x}_k^{a'}$ ,  $k = 1, 2, \dots, M$  are the ensemble perturbations and  $\mathbf{y}_k^{o'}$ ,  $k = 1, 2, \dots, M$  are random draws from the probability distribution of observation errors. As the notation suggests, we consider a linear observation function in our experiments. This choice is made for the sake of simplicity and limits the generality of our findings much less than the use of an idealized model of atmospheric dynamics.

20 For the case of multiple state variables, the ensemble members are considered to be in a single ensemble, that is, not being grouped into distinct sub-ensembles.

**Multivariate  
localization methods  
for ensemble Kalman  
filtering**

S. Roh et al.

Title Page	
Abstract	Introduction
Conclusions	References
Tables	Figures
⏪	⏩
◀	▶
Back	Close
Full Screen / Esc	
Printer-friendly Version	
Interactive Discussion	



### 3.2 The Bivariate Lorenz model

The idealized model we use is the bivariate Lorenz-95 model (Lorenz, 1995). The model mimics the nonlinear dynamics of two linearly coupled atmospheric state variables,  $X$  and  $Y$ , on a latitude circle. The variable,  $X$ , is a “slow” variable represented by  $K$  discrete values,  $X_k$ , and  $Y$  is a “fast” variable represented by  $J \times K$  discrete values. The governing equations are

$$\frac{dX_k}{dt} = -X_{k-1}(X_{k-2} - X_{k+1}) - X_k - (ha/b) \sum_{j=1}^J Y_{j,k} + F, \quad (12)$$

$$\frac{dY_{j,k}}{dt} = -abY_{j+1,k}(Y_{j+2,k} - Y_{j-1,k}) - aY_{j,k} + (ha/b)X_k, \quad (13)$$

where  $Y_{j-J,k} = Y_{j,k-1}$  and  $Y_{j+J,k} = Y_{j,k+1}$  for  $k = 1, \dots, K$  and  $j = 1, \dots, J$ . The “boundary condition” is periodic; that is,  $X_{k-K} = X_{k+K} = X_k$ , and  $Y_{j,k-K} = Y_{j,k+K} = Y_{j,k}$ . In our experiments,  $K = 36$  and  $J = 10$ . The parameter  $h$  controls the strength of the coupling between  $X$  and  $Y$ ,  $a$  is the ratio of the characteristic time scales of the slow motion of  $X$  to the fast motion of  $Y$ ,  $b$  is the ratio of the characteristic amplitudes of  $X$  to  $Y$ , and  $F$  is a forcing term. We choose the parameters to be  $a = 10$ ,  $b = 10$ ,  $F = 10$ , and  $h = 2$ . These values of the model parameters are equal to those originally suggested by Lorenz (1995), except for the value of the coupling coefficient  $h$ , which is twice as large in our case. We made this change in  $h$  to increase the covariances between the errors in the estimates of  $X$  and  $Y$ , which makes the model more sensitive to the choices of the localization parameters. We use a fourth-order Runge–Kutta time integration scheme with a time step of 0.005 non-dimensional units as Lorenz (1995) did. We define the physical distances between  $X_{k_1}$  and  $X_{k_2}$ , between  $Y_{j_1,k_1}$  and  $Y_{j_2,k_2}$ , and between  $X_{k_1}$  and  $Y_{j_1,k_2}$  by  $|10(k_1 - k_2)|$ ,  $|10(k_1 - k_2) + j_1 - j_2|$ , and  $|10(k_1 - k_2) - j_1|$ , respectively. Figure 2 shows a typical state of the model for the selected parameters. The figure shows that  $X$  tends to drive the evolution of  $Y$ : the hypothetical process represented by  $Y$  is more active (its variability is higher) with higher values of  $X$ .

### 3.3 Experimental design

Since the estimates of the cross-covariances play a particularly important role at locations where one of the variables is unobserved, we expect an improved treatment of the cross-covariances to lead to analysis improvements at locations where only one of the state variables is observed. This motivates us to consider an observation scenario in which  $X$  and  $Y$  are partially observed. The variable  $X$  is observed at randomly chosen 20% of all locations and  $Y$  is observed at randomly chosen 90% of those locations where  $X$  is not observed. Spatial locations of the partially observed  $X$  and  $Y$  are illustrated in Fig. 3. The results from this experiment are compared to those from a control experiment, in which both  $X$  and  $Y$  are fully observed.

We first generate a time series of “true” model states by a 2000-time-step integration of the model. We initialize an ensemble by adding the standard Gaussian noise to the true state; then, discarding the first 3000 time steps. We then generate simulated observations by adding random observation noise of mean zero and variance 0.02 to the the appropriate components of the “true” state of  $X$  at each time step. We use the same procedure to generate simulated observations of  $Y$ , except that the variance of the observation noise is 0.005. Observations are assimilated at every time step by first using a 20-member ensemble with a constant covariance inflation factor of 1.015. The error in the analysis at a given verification time is measured by the root-mean-square distance between the analysis mean and the true state. We refer to the resulting measure as the root-mean-square error (RMSE). The probability distribution of the RMSE for the last 1000 time steps of 50 different realizations of each experiment is shown by a boxplot. The boxplot is an effective way of displaying summary of the distribution of numbers. The lower and upper bounds of the box respectively give the 25th and 75th percentiles. The thick line going across the interior of the box gives the median. The whisker depends on the interquartile range (IQR) that is precisely equal to the vertical length of the box. The whiskers extend to the extreme values which are

## Multivariate localization methods for ensemble Kalman filtering

S. Roh et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



no more than 1.5 IQR from the box. Any values that fall outside of the end points of whiskers are considered outliers and they are displayed as circles.

In the boxplot figures in the next section, we compare the RMSE for four different localization schemes. We use the following notation to distinguish between them in the figures:

1. S1 – the bivariate sample background covariance is used without localization;
2. S2 – same as S1 except that the cross-covariance terms are replaced by zeros;
3. S3 – a univariate localization function is used to filter the marginal-covariances, while the cross-covariance terms are replaced by zeros;
4. S4 – one of the bivariate localization methods described in Sect. 2.2.2 is used to filter both the marginal- and the cross-covariances.

In the experiments identified by S4, we consider two different bivariate localization functions: The first one is  $\rho^{(1)}(\cdot) = \{\beta_{ij}\rho^{(1)}(\cdot)\}_{i,j=1,2}$  with  $\beta_{ii} = 1$  ( $i = 1, 2$ ),  $\beta_{ij} = \beta$  ( $i \neq j$ ), and  $|\beta| < 1$ . We use the fifth-order piecewise-rational function of Gaspari and Cohn (1999) to define the univariate correlation function,  $\rho^{(1)}$ , in the following form,

$$\rho^{(1)}(d; c) = \begin{cases} -\frac{1}{4}(|d|/c)^5 + \frac{1}{2}(d/c)^4 + \frac{5}{8}(|d|/c)^3 - \frac{5}{3}(d/c)^2 + 1, & 0 \leq |d| \leq c, \\ \frac{1}{12}(|d|/c)^5 - \frac{1}{2}(d/c)^4 + \frac{5}{8}(|d|/c)^3 + \frac{5}{3}(d/c)^2 - 5(|d|/c) + 4 - \frac{2}{3}c/|d|, & c \leq |d| \leq 2c, \\ 0, & 2c \leq |d|. \end{cases} \quad (14)$$

This correlation function attenuates the covariances with increasing distance, setting all the covariances to zero beyond distance  $2c$ . So this function has the support  $2c$ . If  $|\beta| < 1$  and  $c$  is the same for both the marginal- and the cross-covariances, the matrix-valued function,  $\rho^{(1)}$ , is positive-definite and of full rank. We test various values of the localization parameters  $c$  and  $\beta$ .

**Multivariate localization methods for ensemble Kalman filtering**

S. Roh et al.

Title Page	
Abstract	Introduction
Conclusions	References
Tables	Figures
◀	▶
◀	▶
Back	Close
Full Screen / Esc	
Printer-friendly Version	
Interactive Discussion	



The second multivariate correlation function we consider,  $\rho^{(2)}$ , is the bivariate Askey function described in Sect. 2.2. In particular, we use  $\mu_{11} = 0$ ,  $\mu_{22} = 2$ ,  $\mu_{12} = 1$ , and  $\nu = 3$ . According to Eq. (9), for these choices of parameters, the one remaining parameter,  $\beta_{12}$ , must be chosen such that  $|\beta_{12}| < 0.79$ .

### 3.4 Results

Figure 4 shows the distribution of RMSE for variable  $X$  for different configurations of the localization scheme in the case where the state is only partially observed. This figure compares the Askey function and Gaspari–Cohn function which have the same support (localization radius), so setting all the covariances to zero beyond the same distance. We recall that because  $X$  is much more sparsely observed than  $Y$ , we expect to see some sensitivity of the analyses of  $X$  to the treatment of the cross-covariance terms. The figure confirms this expectation. A comparison of the results for configurations S1 and S2 suggests that ignoring the cross-covariances is a better strategy than to use them without localization. This conclusion does not hold once a univariate localization is applied to the marginal covariances, as using configuration S3 produces worse results than applying no localization at all (S1).

Figure 4 also shows that the distribution of the state estimation error is less sensitive to the choice of localization strategy for the larger values of support. Of all localization schemes, S4 with  $\beta = 0.1$  performs best regardless of the localization radius: the distribution of the state estimation error is narrow with a mean value that is lower than those for the other configurations of the localization scheme. For this choice of localization scheme and  $\beta$ , the Askey function produces smaller errors than the Gaspari–Cohn function, particularly, for smaller localization radii.

Figure 5 is the same as Fig. 4, except for variable  $Y$  rather than for variable  $X$ . A striking feature of the results shown in this figure is that the Askey function clearly performs better than the Gaspari–Cohn function. Another obvious conclusion is that using a smaller localization radius (a lower value of support) is clearly advantageous

Multivariate localization methods for ensemble Kalman filtering

S. Roh et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





## Multivariate localization methods for ensemble Kalman filtering

S. Roh et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



for the estimation of  $Y$ . This result is not surprising, considering that  $Y$  is densely observed and its spatial variability is much higher than that of  $X$ . In contrast to the results for variable  $X$ , configuration S3 produces much more accurate estimates of variable  $Y$  than do configurations S1 or S2. In addition, configuration S4 performs only slightly better, and only for the lowest value of support, than does configuration S3. The latter observations indicate that the marginal covariances play a more important role than do the cross-covariances in the estimation of the densely observed  $Y$ . The proper filtering of the marginal covariances can thus greatly increase the accuracy of the estimates of  $Y$ . In other words, the densely observed  $Y$  is primarily estimated based on observations of  $Y$ . Hence, the low signal-to-noise ratio for the sample estimate of the marginal covariances for  $Y$  greatly limits the value of the observations of  $Y$  at longer distances.

Figure 6 is the same as Fig. 4, except for the case of a fully observed state. By comparing the two figures, we see that the analysis is far less sensitive to the localization radius in the fully observed case than in the partially observed case. As can be expected, the state estimates are also more accurate in the fully observed case. In the fully observed case, localization strategy S3 performs much better than do strategies S1 and S2 and similarly to S4. This result indicates that in the fully observed case,  $X$  is primarily analyzed based on observations of  $X$ , making the analysis of  $X$  more sensitive to the localization of the marginal covariances than to the localization of the cross-covariances. Similar to the partially observed case, the Askey function tends to perform better than the Gaspari–Cohn function, but the differences between the accuracy of the state estimates for the two filter functions are negligible, except for the shortest localization radius.

Finally, Fig. 7 shows the distribution of the errors for variable  $Y$  in the fully observed case. The best results are obtained by using a short localization radius with the Askey function, even though the variability of the error is relatively large in that case. The fact that localization strategies S3 and S4 perform similarly well shows that the estimates of the cross-covariances do not play an important role in this case; that is,  $X$  is pri-

marily estimated based on observations of  $X$  and  $Y$  is dominantly estimated based on observations of  $Y$ .

Figures 7–10 show the results for the 40-member ensemble. We use an inflation factor of 1.01, because the optimal value of the inflation factor is typically smaller for a larger ensemble. Figure 7 shows that the ensemble size has little effect on the estimates of  $X$  in the case of the partially observed state. For variable  $Y$  in the partially observed case (Fig. 8) and both variables  $X$  and  $Y$  in the fully observed case (Figs. 9 and 10), however, the best results are obtained with a larger localization radius than in the case of the 20-member ensemble. This behavior is expected, as a larger ensemble can more accurately estimate the weaker covariances associated with the longer distances. As for the 20-member ensemble, the localization schemes using the Askey function perform better than those using the Gaspari–Cohn function.

## 4 Discussion

The central argument of this paper is that using a univariate localization function for the localization of both the marginal and the cross-covariances in an EnKF scheme may lead to a rank deficient estimate of the background covariance matrix. We suggested two different approaches for the construction of positive-definite filtered estimates of the background covariance matrix. One of them takes advantage of the knowledge of a proper univariate localization function, whereas the other uses a multivariate extension of the Askey function. The results of our numerical experiments show that a mathematically proper localization function often leads to improved state estimates. The results of the numerical experiments also suggest that of the two approaches we introduced, the one based on the Askey function produces more accurate state estimates than that based on the Gaspari–Cohn function.

*Acknowledgement.* Mikyoung Jun’s research was supported by NSF grant DMS-1208421, while Istvan Szunyogh’s research was supported by ONR Grant N000140910589. This publica-

**Multivariate  
localization methods  
for ensemble Kalman  
filtering**

S. Roh et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## References

- Anderson, J. L.: Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter, *Physica D*, 230, 99–111, 2007. 836
- Anderson, J. and Lei, L.: Empirical localization of observation impact in ensemble kalman filters, *Mon. Weather Rev.*, 142, 739–754, 2013. 836
- Askey, R.: Radial Characteristic Functions, technical report no. 1262, Mathematical Research Center, University of Wisconsin-Madison, Madison, 1973. 843
- Bishop, C. H. and Hodyss, D.: Flow adaptive moderation of spurious ensemble correlations and its use in ensemble based data assimilation, *Q. J. Roy. Meteor. Soc.*, 133, 2029–2044, 2007. 836
- Bishop, C. H. and Hodyss, D.: Ensemble covariances adaptively localized with ECO-RAP. Part 1: Tests on simple error models, *Tellus A*, 61, 84–96, 2009a. 836
- Bishop, C. H. and Hodyss, D.: Ensemble covariances adaptively localized with ECO-RAP. Part 2: A strategy for the atmosphere, *Tellus A*, 61, 97–111, 2009b. 836
- Buehner, M. and Charron, M.: Spectral and spatial localization of background-error correlations for data assimilation, *Q. J. Roy. Meteor. Soc.*, 133, 615–630, 2007. 836
- Campbell, W. F., Bishop, C. H., and Hodyss, D.: Vertical covariance localization for satellite radiances in ensemble kalman filters, *Mon. Weather Rev.*, 138, 282–290, 2010. 839
- Du, J. and Ma, C.: Vector random fields with compactly supported covariance matrix functions, *J. Stat. Plan. Infer.*, 143, 457–467, 2013. 838
- Gaspari, G. and Cohn, S. E.: Construction of correlation functions in two and three dimensions, *Q. J. Roy. Meteor. Soc.*, 125, 723–757, 1999. 836, 837, 839, 841, 847
- Genton, M. G. and Kleiber, W.: Cross-covariance functions for multivariate geostatistics, *Statist. Sci.*, in press, 2015. 838
- Hamill, T. M., Whitaker, J. S., and Snyder, C.: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter, *Mon. Weather Rev.*, 129, 2776–2790, 2001. 836, 839

## Multivariate localization methods for ensemble Kalman filtering

S. Roh et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

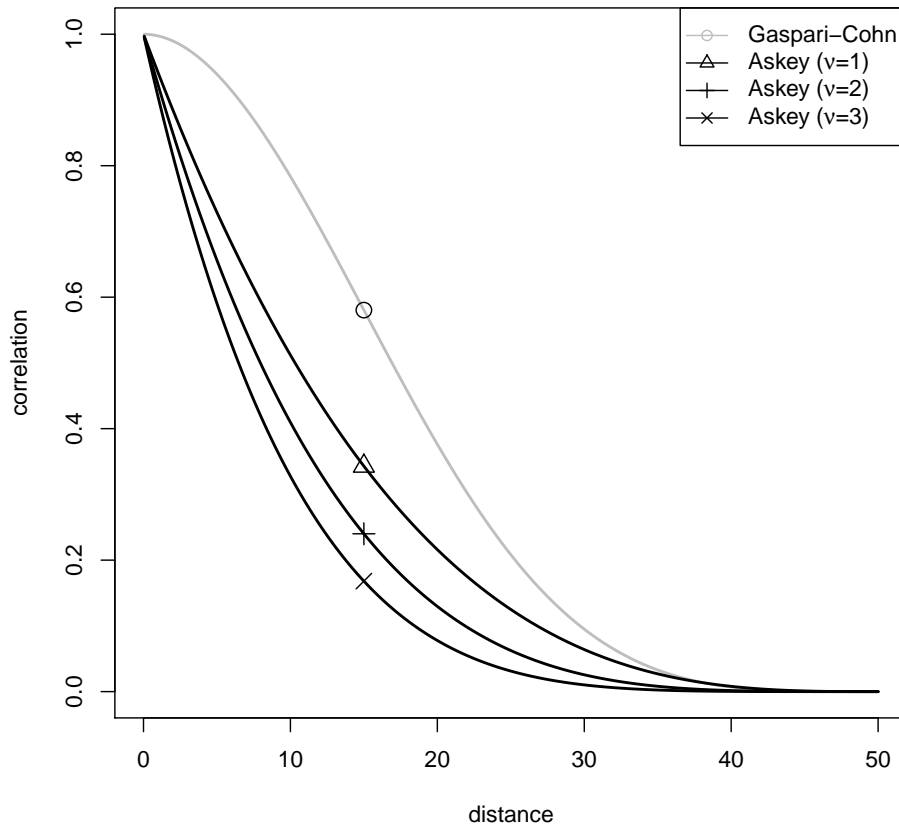
Full Screen / Esc

Printer-friendly Version

Interactive Discussion







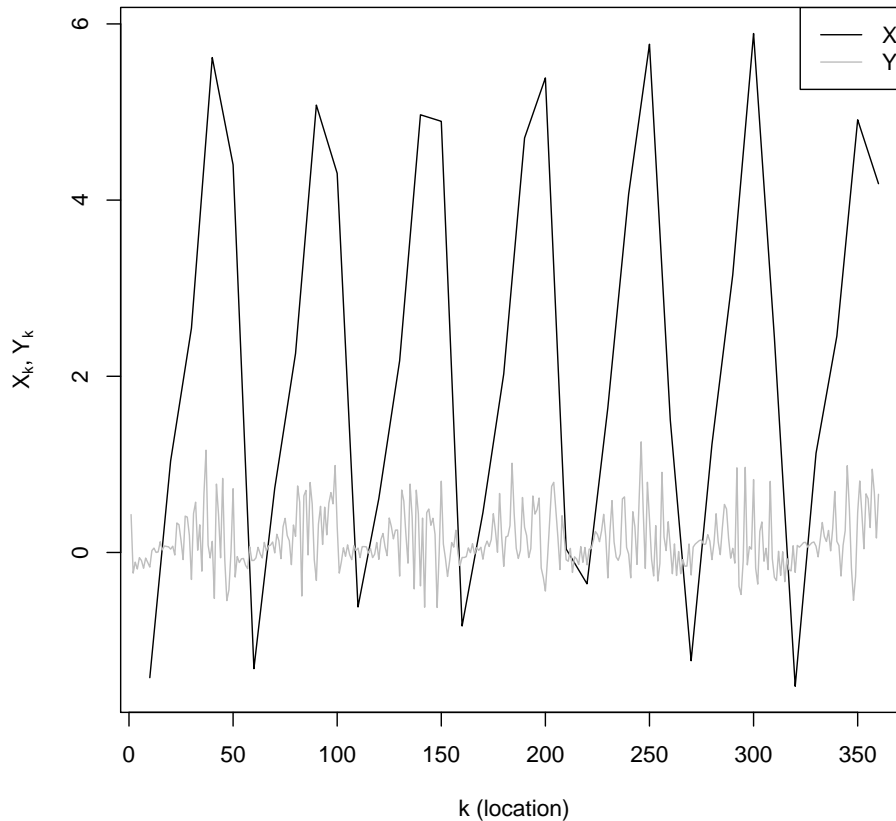
**Figure 1.** The Gaspari–Cohn covariance function with a localization constant  $c = 25$  (support of 50) and the Askey covariance function with a support parameter  $c = 50$  and various shape parameters.

## Multivariate localization methods for ensemble Kalman filtering

S. Roh et al.

<a href="#">Title Page</a>	
<a href="#">Abstract</a>	<a href="#">Introduction</a>
<a href="#">Conclusions</a>	<a href="#">References</a>
<a href="#">Tables</a>	<a href="#">Figures</a>
<a href="#">⏪</a>	<a href="#">⏩</a>
<a href="#">◀</a>	<a href="#">▶</a>
<a href="#">Back</a>	<a href="#">Close</a>
<a href="#">Full Screen / Esc</a>	
<a href="#">Printer-friendly Version</a>	
<a href="#">Interactive Discussion</a>	





**Figure 2.** A snapshot of the variables  $X$  and  $Y$  from a numerical integration of the system of Eqs. (12) and (13) with  $K = 36$ ,  $J = 10$ ,  $F = 10$ ,  $a = 10$ ,  $b = 10$ , and  $h = 2$ .

## Multivariate localization methods for ensemble Kalman filtering

S. Roh et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

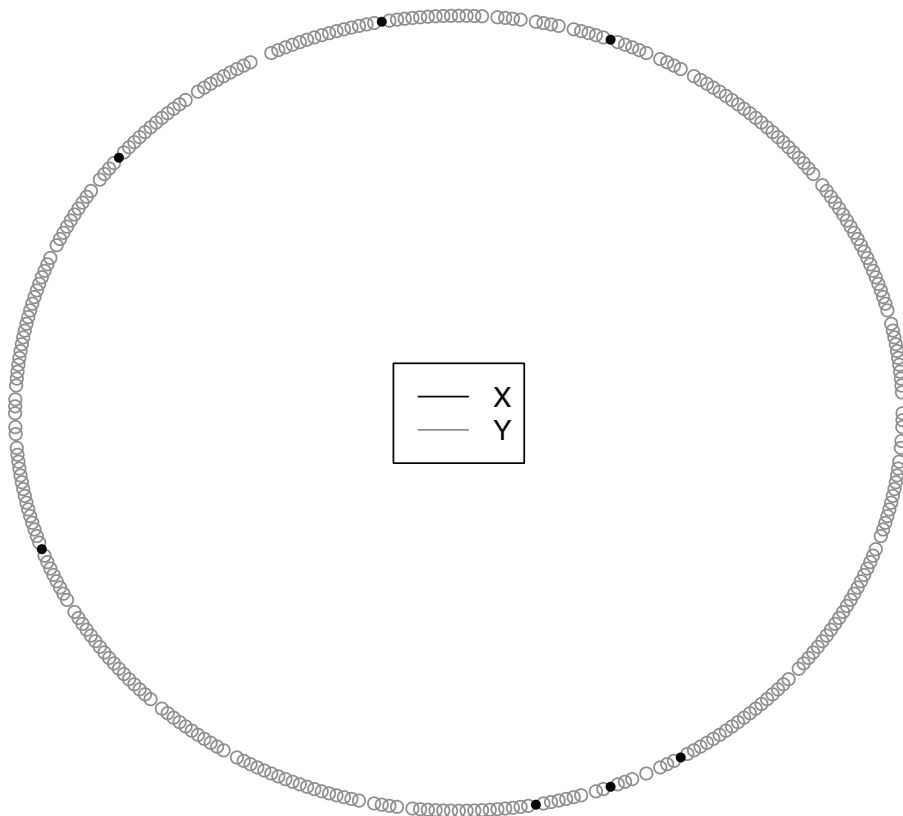
Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





**Figure 3.** Spatial locations of partial observation of  $X$  and  $Y$ .

**Multivariate  
localization methods  
for ensemble Kalman  
filtering**

S. Roh et al.

Title Page

Abstract Introduction

Conclusions References

Tables Figures

⏪ ⏩

◀ ▶

Back Close

Full Screen / Esc

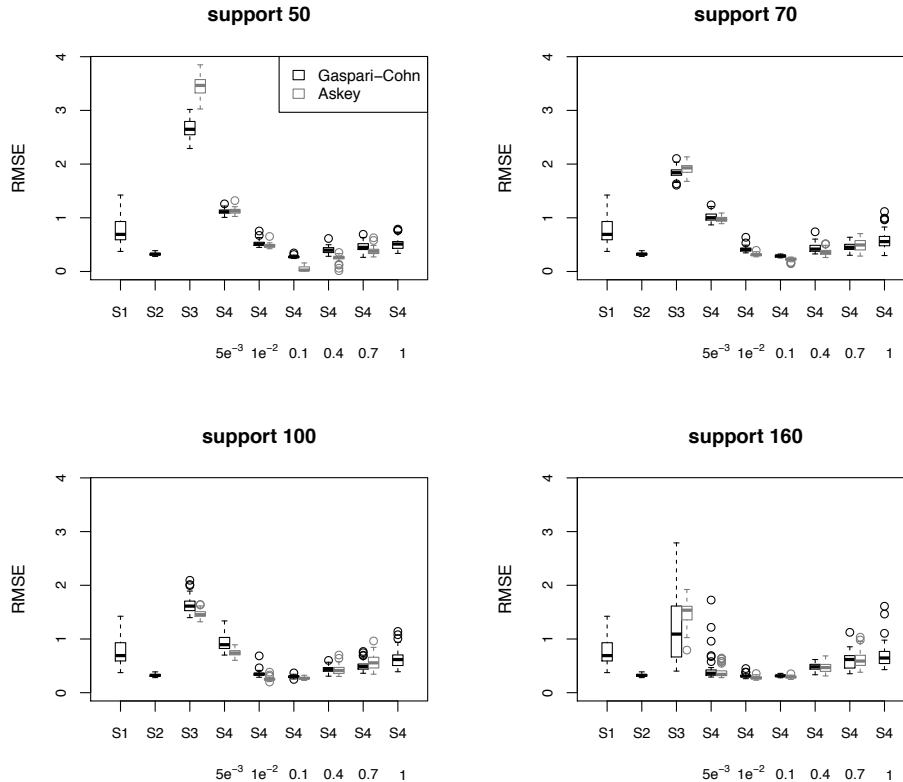
Printer-friendly Version

Interactive Discussion



## Multivariate localization methods for ensemble Kalman filtering

S. Roh et al.



**Figure 4.** The probability distribution of RMSE for variable  $X$  in the case when the system is only partially observed. Results are shown for different localization strategies. For the definitions of localization strategies S1, S2, S3 and S4, see the text. The title of each panel indicates the localization radius (length of support). The numbers below S4 indicate the value of  $\beta$ .

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

⏪

⏩

◀

▶

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

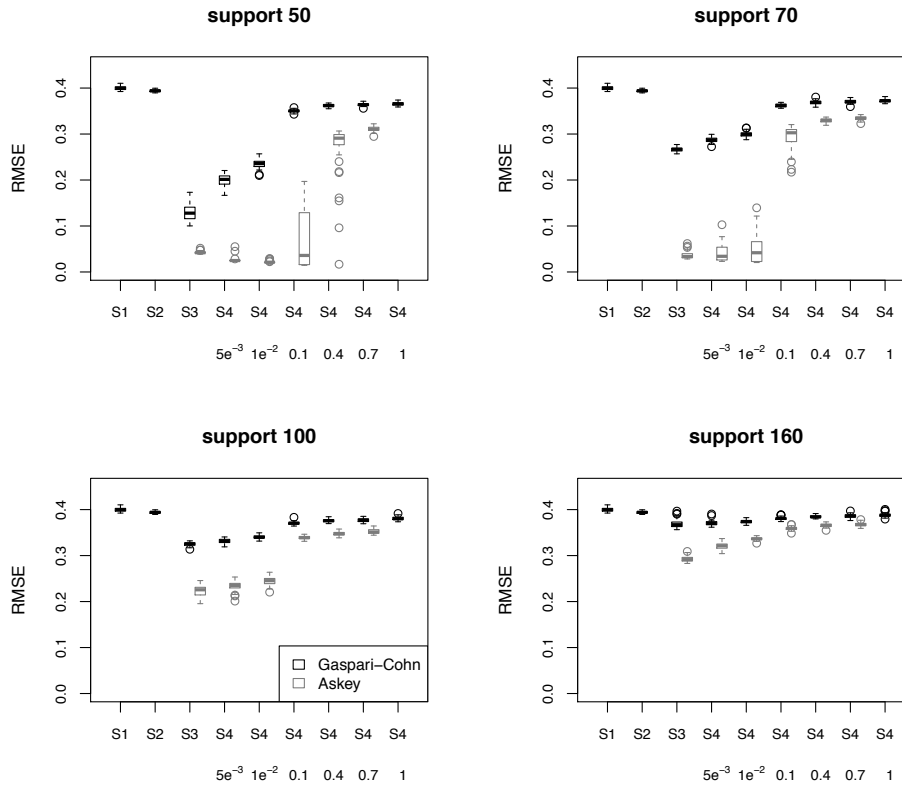
[Interactive Discussion](#)





## Multivariate localization methods for ensemble Kalman filtering

S. Roh et al.



**Figure 5.** Same as 4, except for variable  $Y$ .

Title Page

Abstract Introduction

Conclusions References

Tables Figures

◀ ▶

◀ ▶

Back Close

Full Screen / Esc

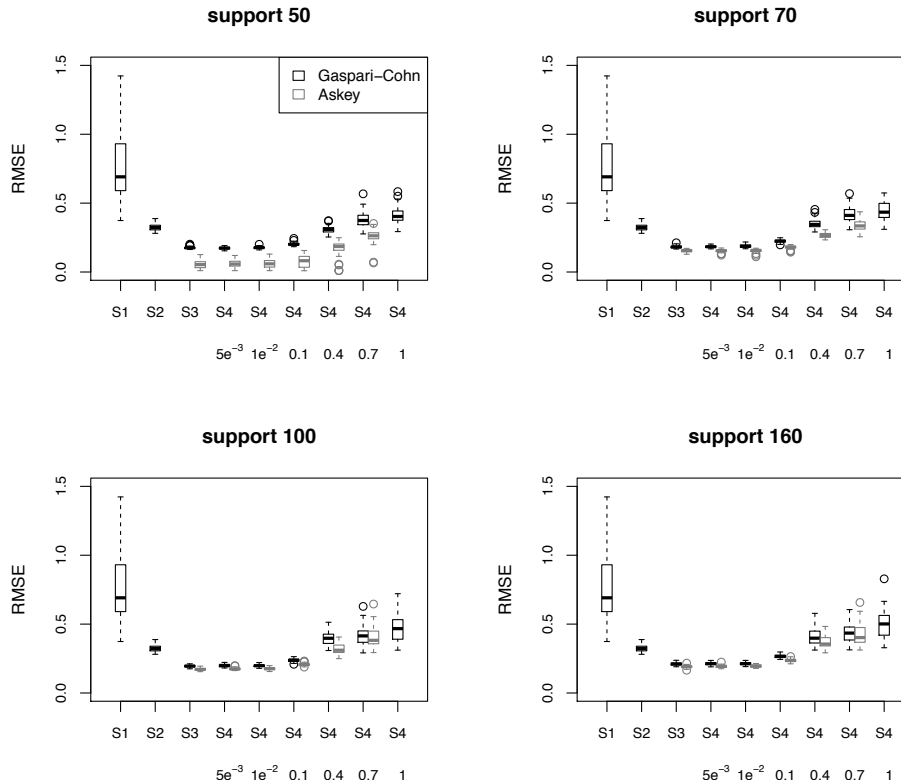
Printer-friendly Version

Interactive Discussion



## Multivariate localization methods for ensemble Kalman filtering

S. Roh et al.



**Figure 6.** Same as 4, except for the case when the system is fully observed.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Multivariate localization methods for ensemble Kalman filtering

S. Roh et al.

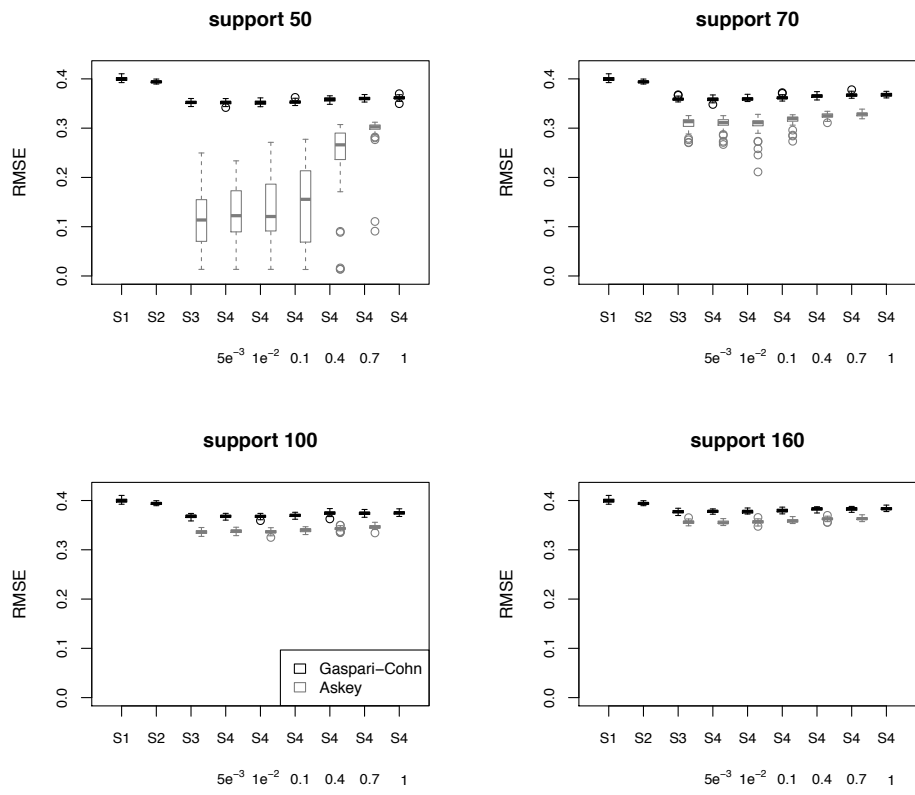


Figure 7. Same as 6, except for variable  $Y$ .

Title Page

Abstract Introduction

Conclusions References

Tables Figures

◀ ▶

◀ ▶

Back Close

Full Screen / Esc

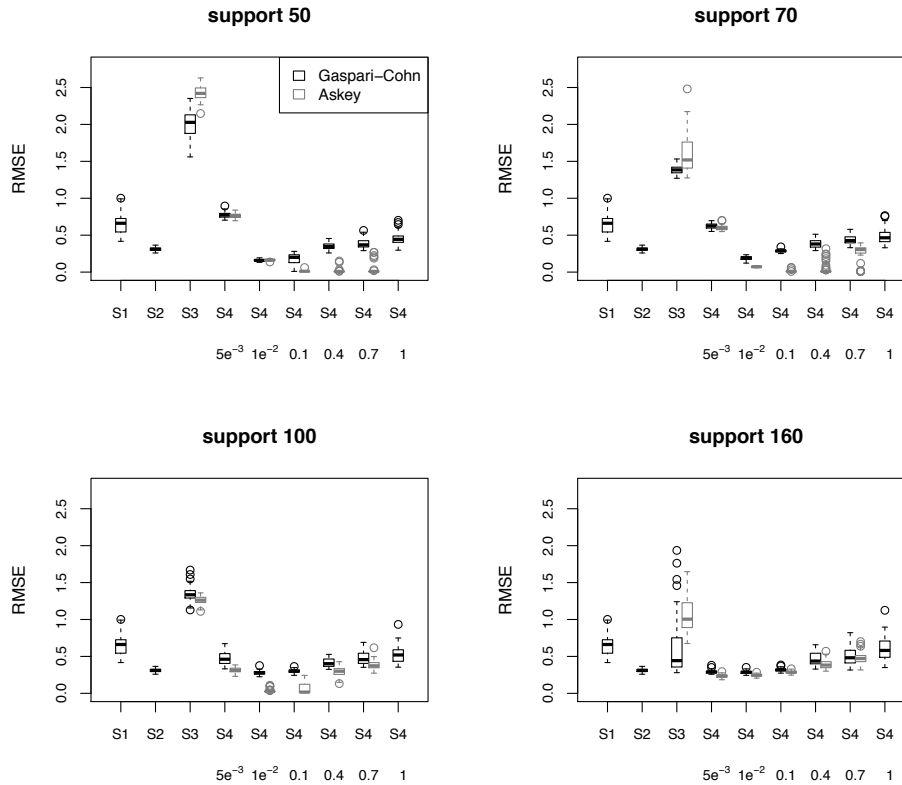
Printer-friendly Version

Interactive Discussion



## Multivariate localization methods for ensemble Kalman filtering

S. Roh et al.



**Figure 8.** Same as 4, except for 40 ensemble members.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

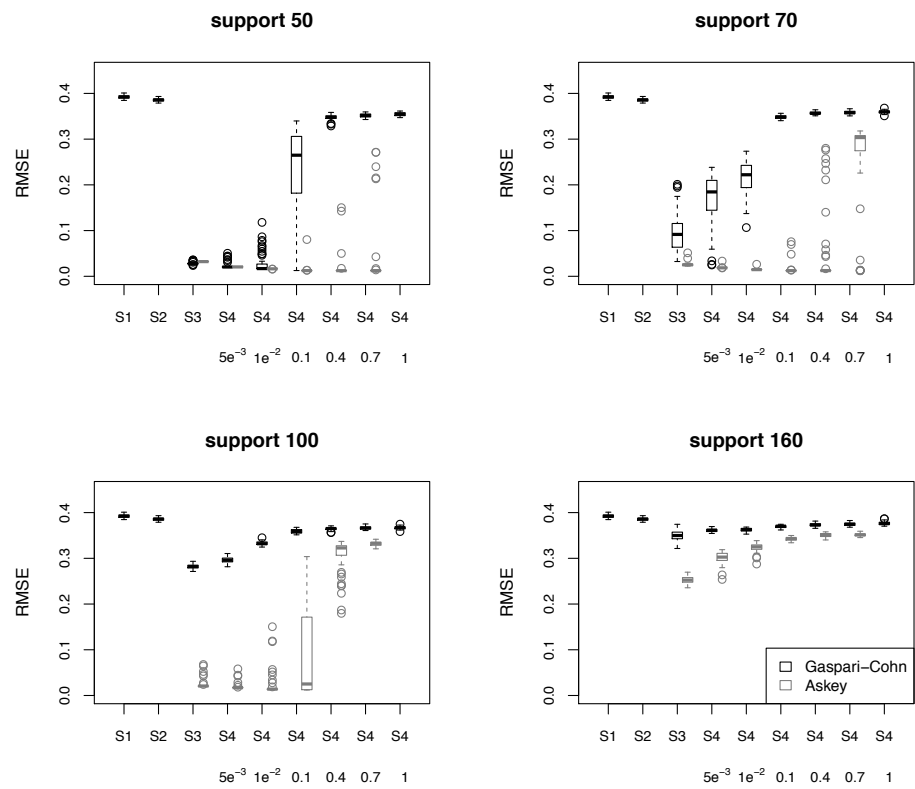
Printer-friendly Version

Interactive Discussion



## Multivariate localization methods for ensemble Kalman filtering

S. Roh et al.



**Figure 9.** Same as 5, except for 40 ensemble members.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Multivariate localization methods for ensemble Kalman filtering

S. Roh et al.

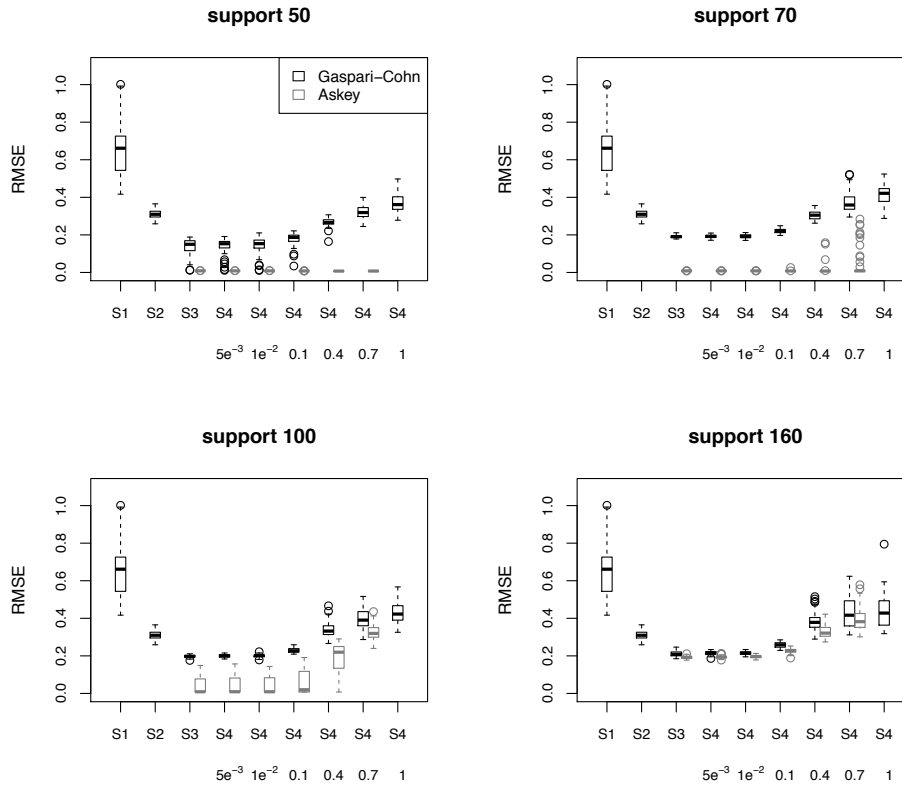


Figure 10. Same as 6, except for 40 ensemble members.

Title Page

Abstract Introduction

Conclusions References

Tables Figures

◀ ▶

◀ ▶

Back Close

Full Screen / Esc

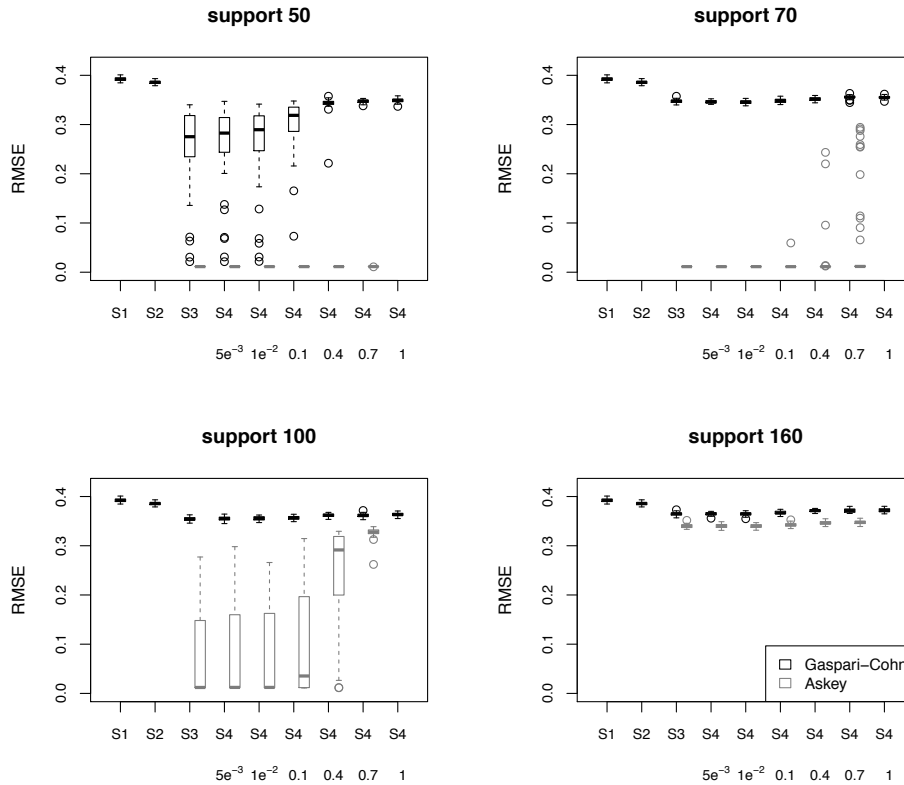
Printer-friendly Version

Interactive Discussion



## Multivariate localization methods for ensemble Kalman filtering

S. Roh et al.



**Figure 11.** Same as 7, except for 40 ensemble members.

Title Page

Abstract Introduction

Conclusions References

Tables Figures

◀ ▶

◀ ▶

Back Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

