

A comparison of assimilation results from the ensemble Kalman Filter and a reduced-rank extended Kalman Filter

X. Zang and P. Malanotte-Rizzoli

Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, USA

Received: 26 August 2002 – Revised: 31 July 2003 – Accepted: 5 November 2003

Abstract. The goal of this study is to compare the performances of the ensemble Kalman filter and a reduced-rank extended Kalman filter when applied to different dynamic regimes. Data assimilation experiments are performed using an eddy-resolving quasi-geostrophic model of the wind-driven ocean circulation. By changing eddy viscosity, this model exhibits two qualitatively distinct behaviors: strongly chaotic for the low viscosity case and quasi-periodic for the high viscosity case.

In the reduced-rank extended Kalman filter algorithm, the model is linearized with respect to the time-mean from a long model run without assimilation, a reduced state space is obtained from a small number (100 for the low viscosity case and 20 for the high viscosity case) of leading empirical orthogonal functions (EOFs) derived from the long model run without assimilation. Corrections to the forecasts are only made in the reduced state space at the analysis time, and it is assumed that a steady state filter exists so that a faster filter algorithm is obtained. The ensemble Kalman filter is based on estimating the state-dependent forecast error statistics using Monte Carlo methods. The ensemble Kalman filter is computationally more expensive than the reduced-rank extended Kalman filter.

The results show that for strongly nonlinear case, chaotic regime, about 32 ensemble members are sufficient to accurately describe the non-stationary, inhomogeneous, and anisotropic structure of the forecast error covariance and the performance of the reduced-rank extended Kalman filter is very similar to simple optimal interpolation and the ensemble Kalman filter greatly outperforms the reduced-rank extended Kalman filter. For the high viscosity case, both the reduced-rank extended Kalman filter and the ensemble Kalman filter are able to significantly reduce the analysis error and their performances are similar. For the high viscosity case, the model has three preferred regimes, each with distinct energy levels. Therefore, the probability density of the system has a multi-modal distribution and the error of the ensemble mean

for the ensemble Kalman filter using larger ensembles can be larger than with smaller ensembles.

1 Introduction

The dynamics governing oceanic processes vary strongly with location. In regions close to the mid-latitude western boundary jets, the ocean circulation is dominated by highly nonlinear, energetic mesoscale eddies, and the covariance model is inhomogeneous and anisotropic. By the expression of “inhomogeneous”, we mean that variances of the ocean circulation depend on location and by the expression of “anisotropic”, we mean that global correlation fields with respect to one point depend not only on the separation distance but also on the direction relative to that point. The variability of the ocean circulation makes it non-stationary. As an example, in a simple reduced gravity quasi-geostrophic model of the double-gyre circulation, McCalpin and Haidgovel (1996) noted three preferred regimes, each with distinct total energy levels, including a high energy state with a jet penetrating deeply into the basin, a low energy state with a short meandering jet, and an intermediate energy state with intermediate jet penetration. Analogous behaviors of the single and double-gyre wind-driven circulation were found by Meacham (2000) and Primeau (2002).

Estimation of the present state of the ocean and forecasting its future evolution have become increasingly important because the ocean plays a crucial role in climate change. Numerous approaches have been proposed for the state estimation and prediction problems. This study is aimed at assessing and comparing two sophisticated filtering techniques: the ensemble Kalman filter and a reduced-rank extended Kalman filter, each applied to two different flow regimes. Evensen (1994) pointed out that the results from ensemble Kalman filter was better than the full extended Kalman filter when applied to a nonlinear quasi-geostrophic ocean circulation model. In contrast to the model used by Evensen (1994), it is computationally infeasible to apply the full Kalman filter

to the model we used. Therefore, a reduced-rank approach must be used for our study. Another important difference is that in our study the ensemble size is far less than the model size, while in the study by Evensen (1994), the ensemble size is comparable to the model size. Furthermore, we will compare the performances of ensemble Kalman filter and reduced-rank steady-state Kalman filter under different flow regimes, from quasi-linear to fully chaotic and nobody has previously investigated the issue of filter performances under different dynamical flow regimes.

The paper is organized as follows. Section 2 provides background information on the ocean model used, the data assimilation algorithms and how to generate proper ensemble initial conditions. By the expression of “proper”, we mean that the sample covariance of ensemble initial conditions is consistent with the error covariance in the initial condition. In Sect. 3, the fully nonlinear filter, the ensemble Kalman filter, the reduced-rank extended Kalman filter and optimal interpolation algorithms are discussed. The assimilation results are given in Sect. 4. Section 5 presents the conclusions and discussions.

2 Experiment environment

2.1 Ocean model

The model used in this study is a reduced gravity, eddy-resolving, wind-driven, quasi-geostrophic ocean model, which can be written as

$$\frac{\partial q}{\partial t} + J(\psi, q) + \beta\psi_x = \nu\nabla^4\psi + \frac{fW_E}{H}, \quad (1)$$

where f and β are the planetary vorticity and its gradient, ν is a uniform eddy viscosity, W_E is the Ekman suction velocity, H is the upper layer depth, ψ is the stream function and q is the potential vorticity given by

$$q = \nabla^2\psi - \frac{1}{R_d^2}\psi. \quad (2)$$

Here R_d is the Rossby radius of deformation.

The model uses an idealized rectangular ocean basin that is 2048 km in the zonal direction and 4096 km in the meridional direction with a horizontal resolution of 16 km. A no-slip lateral boundary condition is used on the sides of the basin. The potential vorticity equation is discretized in space with a second-order finite difference scheme and advanced in time with a second-order Runge-Kutta scheme and the time step is 0.089 day. This same model was used by Mahadevan et al. (2001) to study the predictability of large-scale wind-driven ocean flows and by Buehner and Malanotte-Rizzoli (2003) to evaluate the performance of a reduced-rank Kalman filter.

The degree of nonlinearity of the model increases as the eddy viscosity decreases and the model exhibits qualitatively distinct circulation pattern with different eddy viscosity coefficients. As shown in Fig. 1, the model is quasi-linear and

exhibits a quasi-periodic behavior with a period of approximately 9 years for the high viscosity case ($\nu = 500$). The circulation is characterized by an oscillation between states of high and low energy, and the circulation pattern is simple and well-organized. As the eddy viscosity is reduced to 100, the quasi-periodic behavior disappears; the flow pattern becomes fully chaotic; and the circulation is characterized by strong meanders and energetic mesoscale eddies. Figure 2 shows that for the low viscosity case, the covariance model of the ocean circulation is anisotropic and inhomogeneous: the correlation length scale varies strongly with location and direction. The variability in the quiescent eastern basin is dominated by large-scale Rossby waves with very large correlation length scales. The variations near the western boundary, however, have very small correlation length scales. The radiation pattern shown in Fig. 2d implies that the western boundary current is the energy source of mesoscale eddies radiating away from the jet.

2.2 Assimilation scheme

In this study, the “identical twin” approach (assimilation performed using “a perfect model”) is used to compare the performances of the ensemble Kalman filter and a reduced-rank extended Kalman filter. A schematic diagram of assimilation experiments is shown in Fig. 3. The model is first spun up from rest using a steady meridionally antisymmetric wind forcing with an addition of small random perturbations to break the meridional symmetry of the flow. After spin-up, a steady meridionally antisymmetric wind forcing is used. Starting from a fully spun-up state at t_1 , a long integration of the model is produced: about 37 years for the low viscosity case and 146 years for the high viscosity case. The state is sampled every 44 days for the low viscosity case and every 88 days for the high viscosity case. This sampling frequency is high enough to capture the dominant natural variability of the model. EOFs are calculated from the output of these unconstrained model runs. For the low viscosity case, the first 100 leading EOFs explain 94.8% of the variability of the unconstrained model run. For the high viscosity, the first 20 leading EOFs account for 96.8% of the total variability.

As shown in Fig. 3, a “false” ocean and a “true” ocean are defined to be the simulation between time t_1 and t_2 and between t_3 and t_4 , respectively. In this study, the model is perfect: the model used for the assimilation run is identical to the model used to generate the “true” ocean. The “false” ocean differs from the “true” ocean only in the initial condition. The states at time t_3 and t_1 of the same model run are chosen to be the initial conditions for the “true” ocean and our best guess of the initial condition for the assimilation experiment, respectively.

The observation network utilized for this study is shown in Fig. 4. A set of measurements of potential vorticity are made at 30 model grid-points. Most of them are distributed near the western boundary where mesoscale variability is highest. The observations are generated from the true state using

$$\mathbf{y}(t) = \mathbf{H}\mathbf{x}^t(t) + \boldsymbol{\epsilon}. \quad (3)$$

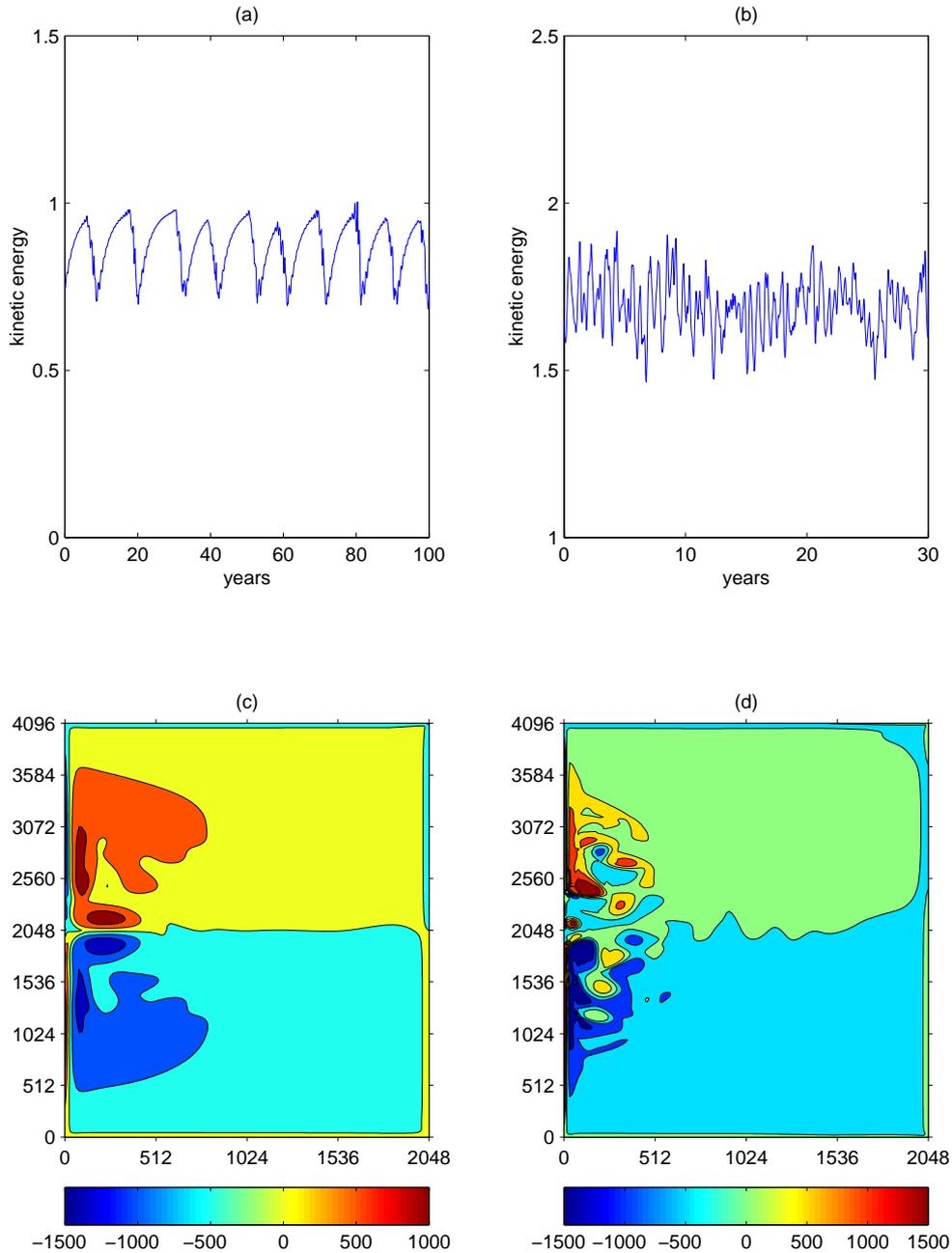


Fig. 1. Evolution of kinetic energy in a long model run without assimilation for (a) high and (b) low viscosity cases. Snapshots of the potential vorticity field in the long model run without assimilation for (c) high and (d) low viscosity cases.

Here \mathbf{H} is the measurement matrix which interpolates the true state points \mathbf{x}^f to the observation points \mathbf{y} and \mathbf{H} is assumed to be linear in this study; ϵ is the measurement error. Define N_s and N_o to be the number of state variables and observations, then \mathbf{H} is an $N_o \times N_s$ matrix and ϵ is an N_o -dimensional vector. In this study, the measurement errors are assumed to be stationary and uncorrelated with each other. The magnitude of the measurement error at each location is 10% of the natural variability of the model. For the data assimilation experiment, the observations are available every 9 days for the low

viscosity case and every 44 days for the high viscosity case. The higher sampling frequency for the low viscosity case is necessary because of the shorter correlation time scale.

2.3 Generation of ensemble initial conditions

As described in Sect. 2.2, the best guess and true initial conditions are chosen to be the model states at two different times from an unconstrained long model run. Therefore, the uncertainty of the best guess initial condition is represented

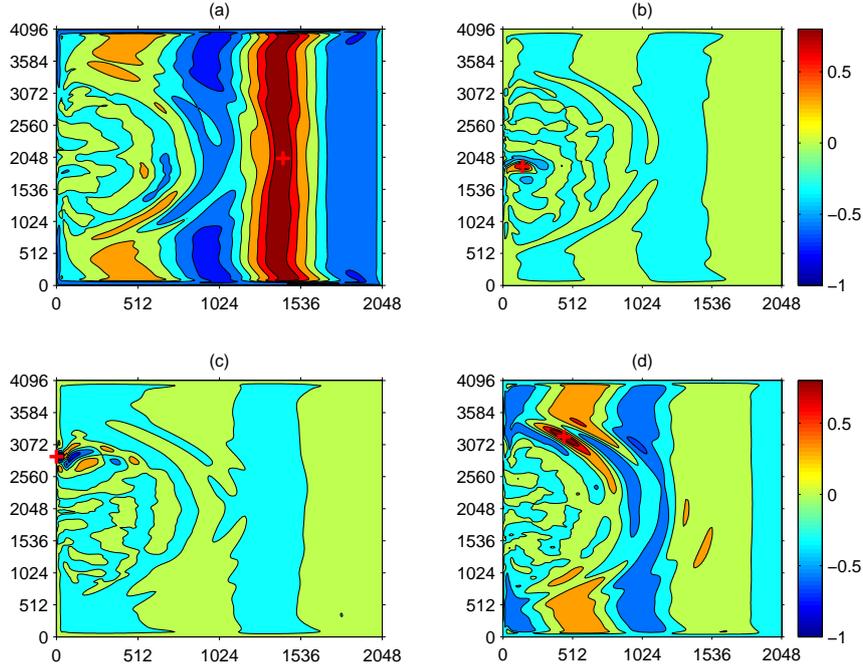


Fig. 2. Global correlation fields for the low viscosity case with respect to a point (a) in the quiescent eastern basin, (b) close to the western boundary, (c) at the western boundary, and (d) in the north-western part of the basin. The position of the point is denoted by a plus. This figure is generated from the time series data of a long model run without data assimilation.

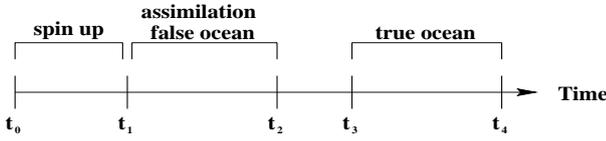


Fig. 3. Schematic diagram of the assimilation experiments. The model is spun-up from rest at t_0 and a fully spun-up state is obtained at t_1 which is then used as initial condition to produce a long model run. The states between t_2 and t_3 are defined to be “true ocean” and the state at t_1 is treated as the best guess of the initial condition for the assimilation.

by the variability of the unconstrained model run. As shown in Fig. 2, the sample covariance of the unconstrained model run is very complicated: the decorrelation length scale varies strongly with location and the correlation field is anisotropic. It is very difficult to directly generate an ensemble of initial states with such a complicated covariance structure for a high dimensional system. In this study, we generate the proper ensemble initial condition indirectly. It is assumed *a priori* that the stochastic dynamic process is covariance-ergodic so that the sample covariance calculated from a long model run converges to the covariance estimated from an ensemble forecast. We first generate an ensemble of small perturbations in potential vorticity with zero mean and a simple isotropic covariance given by

$$\text{cov}(\mathbf{r}_1, \mathbf{r}_2) = e(\mathbf{r}_1)e(\mathbf{r}_2)\exp\left(-\frac{|\mathbf{r}_1 - \mathbf{r}_2|^2}{r_d^2}\right), \quad (4)$$

where the horizontal decorrelation length is $r_d = 80$ km, the magnitude of the error is $e(\mathbf{r}) = 7 \text{ s}^{-1}$ (10% of the globally averaged standard deviation of the natural variability) in the middle of the basin decreasing to zero when approaching the boundary for the low viscosity case. For the high viscosity case, the decorrelation length is the same but the error magnitude is $e(\mathbf{r}) = 3.8 \text{ s}^{-1}$. Note that although the variance varies from location to location, the covariance function is assumed to have a universal decorrelation length scale. As in Evensen (1994), the fast Fourier transform (FFT) approach is used to create the random fields with the covariance structure given by Eq. (4).

The ensemble of small perturbations is added to the best guess of the initial condition to produce an ensemble of initial states which is called ensemble A. The variability of the unconstrained model run is much higher than the spread among the ensemble states of A. Ensemble A is integrated using the model to produce ensemble forecasts for about 9.7 years for the low viscosity case and 73 years for the high viscosity case. Figure 5 shows the evolution of the rms spread of the ensemble forecast which is given by

$$\sqrt{\frac{1}{N_s(N_e - 1)} \sum_{j=1}^{N_e} |\mathbf{x}_j^f - \bar{\mathbf{x}}^f|^2}$$

where $\bar{\mathbf{x}}^f$ is the mean of the ensemble forecast and N_e is the ensemble size. For the low viscosity case, the spread increases slowly in the first 8 months, it then increases exponentially until a statistically steady state is reached at the end

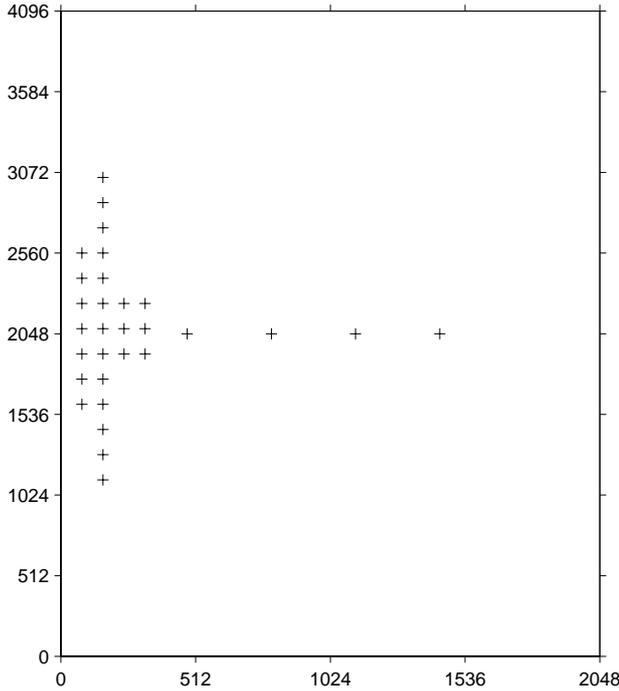


Fig. 4. The 30 locations where measurements of potential vorticity are made.

of third year. For the high viscosity case, it takes about 25 years to reach a statistically steady state. The error-doubling time of the model can be approximately inferred from the rms spread-doubling in the ensemble forecast shown in Fig. 5. As expected, Fig. 5 shows that the error-doubling time for the low viscosity case is shorter than that for the high viscosity case because the low viscosity case is highly chaotic.

2.4 Fully nonlinear filter

The discussion in this section is primarily drawn from the excellent book on stochastic processes and filtering theory by Jazwinski (1970). The general approach of filtering and prediction is probabilistic or Bayesian which does not require the model to be linear or the stochastic processes involved to be Gaussian. Dynamical systems such as motions of the atmosphere or ocean can be modeled by Markov processes:

$$d\mathbf{x} = \mathbf{F}(\mathbf{x}, t)dt + \mathbf{G}(\mathbf{x}, t)d\boldsymbol{\beta}(t), \quad (5)$$

where \mathbf{x} is an N_s -dimensional state vector of the dynamical system at time t ; \mathbf{F} is the N_s -dimensional vector for computing the state vector at time t from knowledge of its value at time $t - 1$ and externally imposed forces; \mathbf{G} is an $N_s \times N_b$ matrix which represents the relationship between the N_b -dimensional Brownian motions and the N_s -dimensional state vector; and $\boldsymbol{\beta}(t)$ is an N_b -dimensional Brownian motions with the covariance matrix of $\langle d\boldsymbol{\beta}(t)d\boldsymbol{\beta}(t)^T \rangle = \mathbf{Q}(t)dt$. Here the angle bracket denotes the expected value.

The time evolution of the probability density of the state vector \mathbf{x} is described by the Komogorov's forward equa-

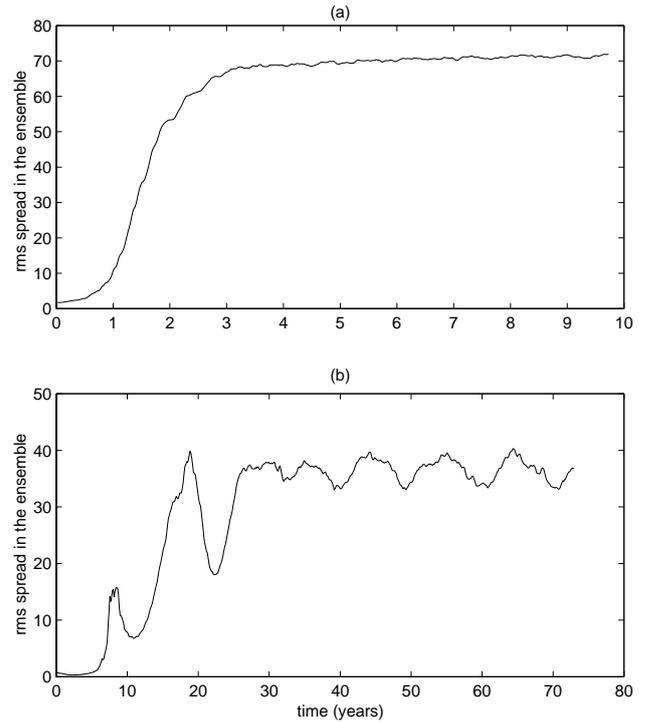


Fig. 5. The evolution of the rms spread of the ensemble forecast with 64 ensemble members for (a) low and (b) high viscosity cases.

tion or the Fokker-Planck equation (Jazwinski, 1970; Gardiner, 2002)

$$\frac{\partial p(\mathbf{x}, t)}{\partial t} = - \sum_{i=1}^{N_s} \frac{\partial [p(\mathbf{x}, t)\mathbf{F}_i]}{\partial \mathbf{x}_i} + \frac{1}{2} \sum_{i=1}^{N_s} \sum_{j=1}^{N_s} \frac{\partial^2 [p(\mathbf{x}, t)(\mathbf{G}\mathbf{Q}\mathbf{G})_{ij}]}{\partial \mathbf{x}_i \partial \mathbf{x}_j}. \quad (6)$$

As shown in Figs. 6 and 7, the variance and correlation structure of the final states of the ensemble forecast are very similar to those calculated from the unconstrained long model run. This confirms our covariance-ergodic assumption. The mean of the final states of the ensemble forecast is replaced with the best guess initial condition and the resulting ensemble state is called ensemble B. Note that the sample covariance of ensemble B is consistent with the error covariance in the initial condition.

3 Assimilation algorithms

Although this study is primarily concerned with the ensemble Kalman filter and a reduced-rank extended Kalman filter, we give a short review of fully nonlinear filter, because the Kalman filter, the extended Kalman filter and the ensemble Kalman filter are all special forms of the fully nonlinear filter.

Given the initial condition of the probability of the state vector, Eq. (6) can be used to forecast the probability of the state vector at any time t , which is usually referred to as the *a priori density* $p(\mathbf{x})$. At the time when observations are available, a conditional density $p(\mathbf{y}|\mathbf{x})$ can be inferred from the

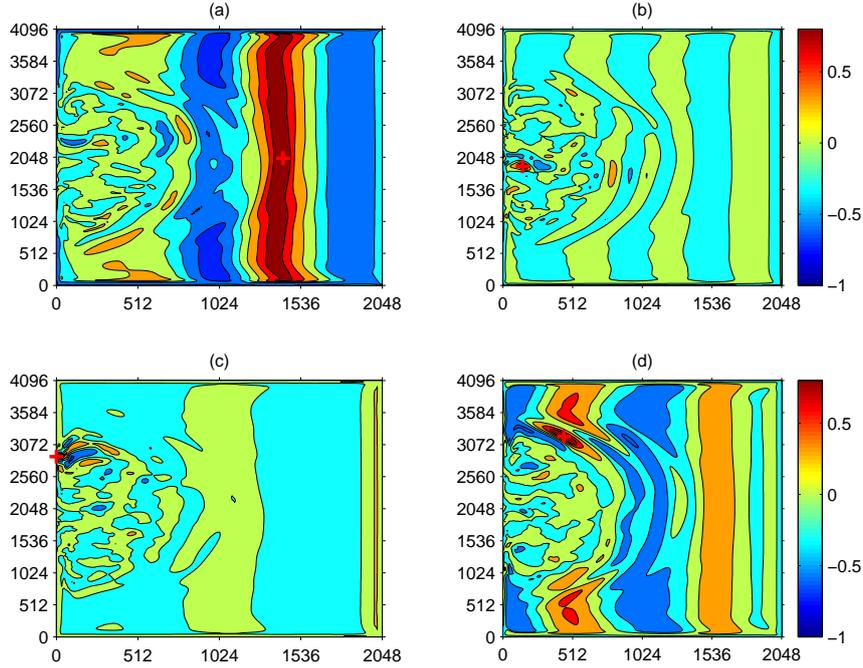


Fig. 6. Same as Fig. 2 except that the correlations are calculated from the final states of the ensemble forecast rather than from the unconstrained long model run.

measurement model. This condition density fully characterizes the relationship between the state vector and the observations. Applying Bayes's rule yields the *posterior* density

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{x})p(\mathbf{y}|\mathbf{x})}{p(\mathbf{y})}. \quad (7)$$

The denominator in Eq. (7) is a normalization factor given by

$$p(\mathbf{y}) = \int_{-\infty}^{+\infty} p(\mathbf{x})p(\mathbf{y}|\mathbf{x})d\mathbf{x}. \quad (8)$$

The *posterior* density $p(\mathbf{x}|\mathbf{y})$ is the filtering solution which is used as initial conditions for the Fokker-Planck equation to produce the forecast of the probability density of the state vector for the next assimilation cycle. So far, all studies on Fokker-Planck equation are limited to very low-dimensional systems (with dimension less than 4), because it is computationally infeasible to directly integrate the Fokker-Planck equation for high-dimensional system. An alternative to predict the evolution of probability density is to use Monte Carlo method to solve the stochastic partial differential Eq. (5). With large enough ensemble size, the solution from Monte Carlo method is a good approximation for the solution to the Fokker-Planck equation.

3.1 Ensemble Kalman Filter

The ensemble Kalman filter is introduced by Evensen (1994) as an approximation to nonlinear filtering. In this approach, the nonlinear dynamical model is used directly to produce the forecast of the probability density evolution through Monte

Carlo method. It is then assumed that both the forecast errors and measurement noises are normally distributed. Under this assumption, the first two moments – mean and covariance – fully characterize the probability density; the maximum likelihood (Bayesian) estimate is the same as the minimum variance estimate; and the general filtering Eq. (7) becomes the analysis equation in the Kalman filter:

$$\mathbf{x}^a(t) = \mathbf{x}^f(t) + \mathbf{K}(t) [\mathbf{y}(t) - \mathbf{H}(t)\mathbf{x}^f(t)], \quad (9)$$

where $\mathbf{x}^f(t)$ and $\mathbf{x}^a(t)$ are the forecast and analysis vector, respectively; and $\mathbf{K}(t)$ is the Kalman gain matrix.

In the ensemble Kalman filter algorithm, an ensemble of perturbed observations instead of a single unperturbed observation are used to update the ensemble forecasts. Otherwise, it will result in an updated ensemble with a variance that is too low (Burgers et al., 1998). The analysis equation of the ensemble Kalman filter is

$$\mathbf{x}_j^a(t) = \mathbf{x}_j^f(t) + \mathbf{K}(t) [\mathbf{y}_j(t) - \mathbf{H}(t)\mathbf{x}_j^f(t)], \quad (10)$$

where j counts from 1 to N_e ; $\mathbf{x}_j^f(t)$ and $\mathbf{x}_j^a(t)$ are the forecast and analysis vector of ensemble member j ; $\mathbf{K}(t)$ is the Kalman gain matrix; and $\mathbf{y}_j(t)$ is a set of perturbed observations given by

$$\mathbf{y}_j(t) = \mathbf{y}(t) + \mathbf{v}_j, \quad (11)$$

where \mathbf{v}_j is drawn from a multivariate normal distribution with zero mean and covariance equal to the error covariance of the measurements, \mathbf{R}_m . The statistical error of the Monte Carlo method converge very slowly with the ensemble size:

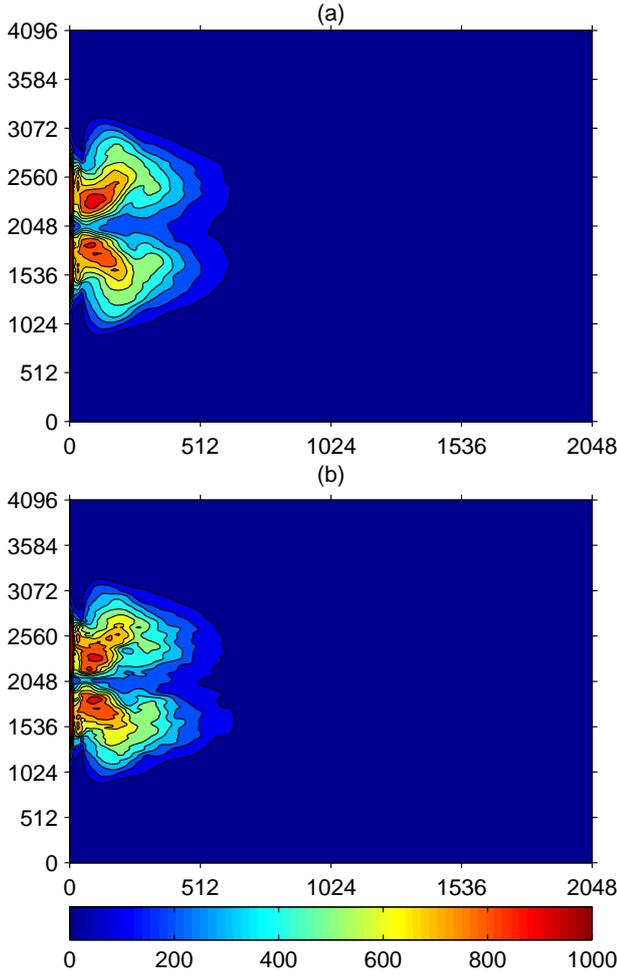


Fig. 7. Standard deviations of potential vorticity for the low viscosity case from (a) the unconstrained long model run and (b) the final state of the ensemble forecast.

the convergence rate is proportional to $1/\sqrt{N_e}$. The convergence rate is also affected by many other factors such as the model, observation errors, system noise, ensemble characteristics like localization factors, etc. In the case of high-dimensional state space, the ensemble size is limited by the capability of the computer. Therefore, the sample mean and covariance from random drawing could be far from the desired mean and covariance. In this study, the observation perturbations \mathbf{v}_j are constructed in such a way that its ensemble mean is zero exactly. This eliminates the statistical error in the ensemble mean and the filter is found to perform better under this condition. The ensemble members could also be constructed in such a way that both the sample mean and covariance matrix match exactly the desired ones (e.g. Pham, 2001). Whitaker and Hamill (2002) proposed another ensemble filter algorithm that eliminates the use of perturbed observations and they claimed that this new algorithm was likely to further increase the ensemble accuracy by 20–30%.

The Kalman gain matrix is given by

$$\mathbf{K}(t) = \mathbf{P}^f(t)\mathbf{H}^T(t) \left[\mathbf{H}(t)\mathbf{P}^f(t)\mathbf{H}^T(t) + \mathbf{R}(t) \right]^{-1}. \quad (12)$$

In the ensemble Kalman filter, the forecast error covariance matrix, $\mathbf{P}^f(t)$, is assumed to be identical to the sample covariance from the ensemble forecast. Houtekamer and Mitchell (1998, 2001) noted that when a limited size of ensemble were used to estimate the forecast error covariance, spuriously large correlation between greatly separated grid points was produced. They proposed to perform an element-wise multiplication of the sample covariance of the ensemble forecast with a correlation with a local support to remove the influence of those observations far away from the grid point being analyzed and they found that the analysis were substantially improved under this condition. In this study, we adopt their approach.

Let ρ denote a correlation function with local support and $\rho \circ \mathbf{P}^f(t)$ denote the Schur (element-wise) product of this correlation function and the sample covariance, $\mathbf{P}^f(t)$, calculated from the ensemble forecast. When the correlation function with local support is incorporated, the Kalman gain can be written as (Houtekamer and Mitchell, 2001)

$$\mathbf{K}(t) = \left[\rho \circ \left(\mathbf{P}^f(t)\mathbf{H}^T(t) \right) \right] \left[\rho \circ \left(\mathbf{H}(t)\mathbf{P}^f(t)\mathbf{H}^T(t) \right) + \mathbf{R}(t) \right]^{-1}. \quad (13)$$

In this study, a simple correlation function with local support is used

$$\rho(r) = \begin{cases} 1.0 - \frac{r^2}{r_0^2} & \text{if } r \leq r_0 \\ 0 & \text{if } r > r_0, \end{cases} \quad (14)$$

where r is the distance between the analyzed grid point and the observation location and r_0 is the “influence” radius beyond which the observations will not be used. This function is isotropic and decreases monotonically with distance with a rate that depends on only r_0 . Note that $\rho = 0$ when the separation distance exceeds the “influence” radius r_0 . The performance of the filter depends on the choice of r_0 . If r_0 is too small, the observations will not be fully utilized. On the hand hand, if r_0 is too large, the problem associated with “remote” observations will arise. In this study, a sensitivity study is performed to evaluate how the performance of the ensemble Kalman filter depends on the “influence” radius r_0 .

The ensemble forecast covariance matrix $\mathbf{P}^f(t)$ has such a large dimension that it cannot be handled directly. Fortunately, the number of observations and the number of ensemble members are usually much smaller than the number of state variables. In this case, one does not need to compute or store the full forecast covariance matrix $\mathbf{P}^f(t)$ as suggested by Evensen and van Leeuwen (1996). Suppose the ensemble of ocean states is stored in an $N_s \times N_e$ matrix \mathbf{C} . First, the ensemble mean of the ocean states is calculated and removed from \mathbf{C} . The resulting ensemble of the ocean state deviations is stored in the $N_s \times N_e$ matrix \mathbf{D} . Second, we calculate the

$N_o \times N_e$ matrix $\mathbf{S} = \mathbf{H}\mathbf{D}$. Third, the $N_s \times N_o$ matrix $\mathbf{P}^f \mathbf{H}^T$ is calculated using

$$\mathbf{P}^f \mathbf{H}^T = \frac{1}{N_e - 1} \mathbf{D} \mathbf{S}^T, \quad (15)$$

and the $N_o \times N_o$ matrix $\mathbf{H} \mathbf{P}^f \mathbf{H}^T$ is calculated using

$$\mathbf{H} \mathbf{P}^f \mathbf{H}^T = \frac{1}{N_e - 1} \mathbf{S} \mathbf{S}^T. \quad (16)$$

Then, we calculate $\rho \circ (\mathbf{P}^f \mathbf{H}^T)$ and $\rho \circ (\mathbf{H} \mathbf{P}^f \mathbf{H}^T)$. Finally, the Kalman gain matrix is calculated using Eq. (13).

3.2 Reduced-rank extended Kalman Filter

The Kalman filter is a linear filter and it can be extended to the nonlinear model through a linearization procedure. The resulting filter is referred to as the extended Kalman filter (Miller et al., 1994; Ghil, 1997). The model used in this study consists of about 3.3×10^4 prognostic variables. It is computationally infeasible to store and propagate the full error covariance matrix of the model. A number of techniques have been proposed to reduce the computational burden associated with the model state error covariance matrix (e.g. Fukumori and Malanotte-Rizzoli, 1995; Cane et al., 1996; Farrell and Ioannou, 2002). In this study, we adopt the approach used by Buehner and Malanotte-Rizzoli (2003): the empirical orthogonal functions (EOFs) calculated from a long model integration without assimilation are used as basis functions to reduce the dimension of the model state.

In the case where the system and measurement dynamics are linear, the state transition matrix, the spatial distribution of the observations, and the error statistics of the model and observations are stationary, the filtering process will reach a unique steady state: the forecast error statistics and Kalman gain matrix are constant if the system is both observable and controllable (Gelb, 1974). In this study, it is assumed that a steady-state filter exists so that a faster assimilation algorithm is obtained. The procedure of implementing the steady-state reduced-rank extended Kalman filter algorithm is summarized as follows.

First, we calculate the EOFs from the output of a model run without assimilation. The first few dominant EOFs, which are used as basis functions to reduce the model state, are stored in an $N_s \times N_r$ matrix \mathbf{E}_r , where N_r is the number of retained EOFs. In this study, the first 100 leading EOFs are retained for the low viscosity case and the first 20 leading EOFs are kept for the high viscosity case. The subspace spanned by these retained EOFs is referred to as resolved subspace and its complement is called unresolved subspace.

Second, the linearized model dynamics in the resolved subspace, \mathbf{M}_r , which is an $N_r \times N_r$ matrix, is obtained. The model is linearized with respect to the time-mean state from the model run without assimilation and the model is perturbed in the direction of each retained EOF. The i th column of \mathbf{M}_r is given by

$$\frac{1}{\alpha} \mathbf{E}_r^T [\mathbf{F}(\bar{\mathbf{x}} + \alpha \mathbf{e}_i) - \mathbf{F}(\bar{\mathbf{x}})], \quad (17)$$

where \mathbf{F} is the nonlinear ocean model, the N_s -dimensional vector $\bar{\mathbf{x}}$ is the time-mean of a long model run without assimilation, the N_s -dimensional vector \mathbf{e}_i is the i th retained EOF, and α is the size of the perturbation. In this study, the value of α is 10^{-3} . In Eq. (17), the model is integrated forward over a period equal to the time between analysis.

The Kalman filter equations in the resolved subspace are given by (Buehner and Malanotte-Rizzoli, 2003)

$$\mathbf{K}_r(t) = \mathbf{P}_r^f(t) \mathbf{H}_r^T(t) [\mathbf{H}_r(t) \mathbf{P}_r^f(t) \mathbf{H}_r^T(t) + \mathbf{R}(t)]^{-1}, \quad (18)$$

$$\mathbf{P}_r^a(t) = [\mathbf{I} - \mathbf{K}_r(t) \mathbf{H}_r] \mathbf{P}_r^f, \quad (19)$$

$$\mathbf{P}_r^f(t+1) = \mathbf{M}_r(t) \mathbf{P}_r^a(t) \mathbf{M}_r^T(t) + \mathbf{Q}_r(t), \quad (20)$$

where \mathbf{P}_r^f , \mathbf{P}_r^a , and \mathbf{H}_r are the forecast error covariance matrix, the analysis error covariance matrix, and the measurement model in the resolved subspace, respectively. \mathbf{Q}_r is the ‘‘pseudo-model error’’ covariance in the resolved subspace. There is dynamic coupling between the resolved and unresolved subspace, the errors in the unresolved subspace can affect the errors in the resolved subspace. The errors due to the use of a reduced space in the analysis are treated as model errors though the model is perfect. We call these errors ‘‘pseudo-model errors’’.

Third, observation error covariance and model error covariance are specified. Observation error is stationary in this study and consists of two components: the actual measurement error and representativeness error (Buehner and Malanotte-Rizzoli, 2003), that is,

$$\mathbf{R} = \mathbf{R}_m + \mathbf{H}_u \mathbf{P}_u^f \mathbf{H}_u. \quad (21)$$

where \mathbf{H}_u and \mathbf{P}_u^f are the measurement model and forecast error covariance matrix in the unresolved space, respectively. The characteristic of the actual measurement error is described in Sect. 2. Representativeness error accounts for the difference between the resolved subspace and the full space. As in Buehner and Malanotte-Rizzoli (2003), we assume that the resolved and unresolved subspace are dynamically uncoupled and the second term in Eq. (21) is calculated from the long model run without assimilation by assuming the forecast error in the unresolved subspace is equal to the variability in the long model run projected into the unresolved subspace.

The ‘‘pseudo-model errors’’ related to the use of a reduced space are very difficult to explicitly calculate. As in Buehner and Malanotte-Rizzoli (2003), these ‘‘pseudo-model errors’’ are estimated through parameterization. It is assumed that the ‘‘pseudo-model error’’ covariance is proportional to the sample covariance in the resolved subspace computed from the unconstrained long model run:

$$\mathbf{Q}_r = \gamma \Theta, \quad (22)$$

where Θ is the sample covariance in the resolved subspace. The proportionality constant, γ , is chosen in such a way that the filter performance (in terms of minimizing the analysis

error) is best. The exact value of γ used in each experiment is described in Sect. 4. Alternatively, this parameter can be estimated adaptively through the method of covariance matching (Maybeck, 1982): given γ , the theoretical forecast error covariance can be estimated through Eqs. (18) to (20) and the sample covariance of the forecast error in the resolved subspace can be calculated using the results from data assimilation:

$$\Omega = \frac{1}{N_t} \sum_{i=1}^{N_t} \mathbf{d}_i \mathbf{d}_i^T, \quad (23)$$

where \mathbf{d}_i is the forecast error vector in the resolved subspace at time step i and N_t is the total number of time steps for the data assimilation experiment. The parameter value of γ is chosen so that the theoretical forecast error covariance \mathbf{Q}_r is close to its sample value Ω .

Finally, given the model dynamics, the measurement model, the model and observation error covariances in the resolved subspace, the steady solutions to Eqs. (18) to (20) can be obtained efficiently using the doubling algorithm (Anderson and Moore, 1979). Multiplication of the stationary Kalman gain matrix with the innovation vector produces the corrections in the resolved subspace, which are mapped into the full model space before being added to the forecast to produce the analysis. The full nonlinear model is used to produce the forecast.

In this study, the basis functions used to reduce the model state are fixed in time. The SEEK filter introduced by Pham et al. (1998) uses dynamically evolving basis functions. Veron et al. (1999) applied the SEEK filter to the tropical Pacific and found no significant improvement, as compared to fixed basis functions. Ballabrea-Poy et al. (2001) applied the SEEK filter to study mid-latitude mesoscale eddies and found that using the dynamically evolving basis functions improved the long-term filter performance only when the basis functions were correctly initialized.

3.3 Optimal interpolation

Optimal interpolation, the traditional approach of data assimilation, is used as a benchmark to evaluate the performance of the ensemble Kalman filter and reduced-rank extended Kalman filter. Optimal interpolation is a simple form of the Kalman filter. In this algorithm, it is assumed that the forecast error covariance is temporally stationary, has a spatially uniform variance and a spatially uniform isotropic Gaussian correlation function with a decorrelation length approximately equal to the decorrelation length in the region of strong variability near the western boundary, which is 48 km for the low viscosity case and 80 km for the high viscosity case. The observation errors are assumed to be temporally stationary, uncorrelated and have a spatially uniform variance. It is found that if the forecast error covariance matrix is tuned so that its variance is 1/5 of the observation error variance, the filter performance is optimal in terms of minimum analysis error.

3.4 Expectations of assimilation results

In terms of accuracy, the ensemble Kalman filter is expected to outperform both the reduced-rank extended Kalman filter and optimal interpolation, because the fully nonlinear model is used directly in the ensemble Kalman filter to estimate the probability density evolution of the forecast error. In the reduced-rank extended Kalman filter, the nonlinear model is linearized first, the error covariance is propagated using the linearized model and all higher-order statistical moments are neglected. The reduced-rank stationary Kalman filter is very similar to the optimal interpolation: both of them are stationary. The major difference between them is how the forecast error covariance is obtained. In the optimal interpolation, the forecast error is assumed to have a spatially uniform covariance structure and the decorrelation length is specified empirically according to the statistics of an unconstrained long model run. In the reduced-rank extended Kalman filter, the approach used to determine the forecast error covariance is complex and has a rigorous theoretical justification. However, the “pseudo-model errors” in the reduced-rank stationary Kalman filter due to the dynamic coupling between the resolved and unresolved subspace are hard to explicitly calculate and they are estimated through parameterization.

In terms of computational cost, the optimal interpolation and the reduced-rank stationary Kalman filter are significantly less expensive than the ensemble Kalman filter. For the reduced-rank stationary Kalman filter, we have to make as many nonlinear forecasts as the number of retained EOFs using Eq. (17) to obtain the linearized model dynamics in the resolved subspace. However, these forecasts are made only for 9 days for the low viscosity case, which is equal to the sampling period of the observations or the period of one assimilation cycle. For the ensemble Kalman filter, the nonlinear forecasts are made over the entire assimilation period which is about 5 years for the low viscosity case. In addition, the Kalman gain needs to be recalculated at each analysis step in the ensemble Kalman filter. For the reduced-rank stationary Kalman filter, the Kalman gain only needs to be calculated once. Therefore, the ensemble Kalman filter is computationally more expensive than the reduced-rank stationary Kalman filter. The stationary filter requires only a slightly greater computational effort than a simple model integration without assimilation. For the ensemble Kalman filter with an ensemble size of N_e , the computational effort is at least N_e times as big as the effort required for a simple model integration without assimilation. Compared to the ensemble Kalman filter, the reduced-rank extended Kalman filter is more sophisticated and more difficult to implement. The reduced-rank Kalman filter and the ensemble Kalman filter can be combined to reduce the statistical errors of the Monte Carlo approach (Heemink et al., 2001).

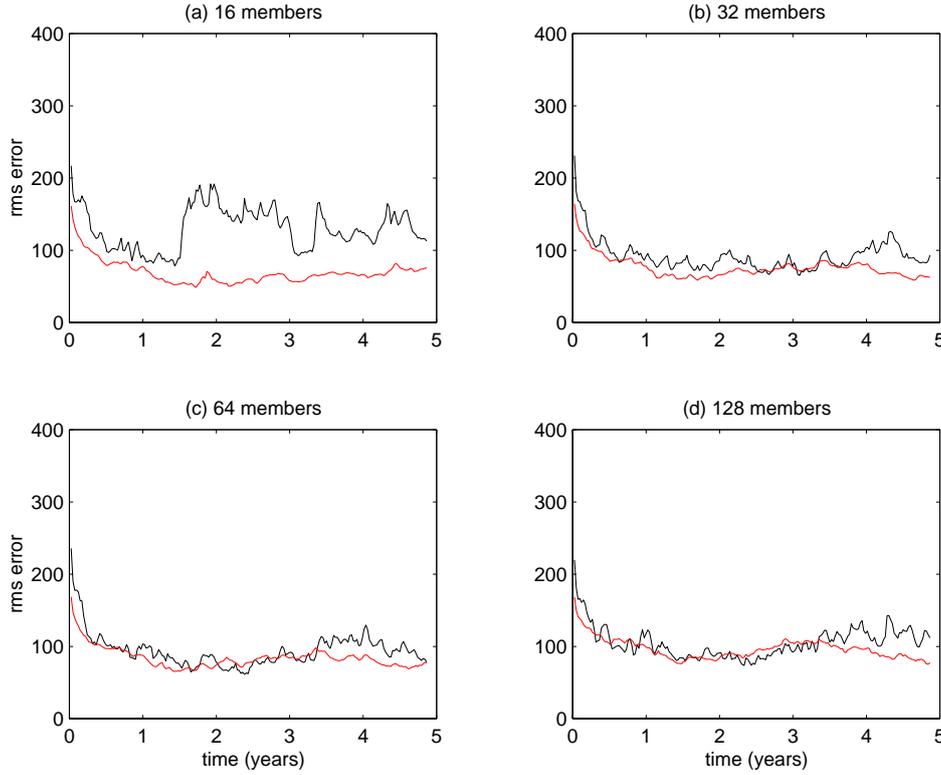


Fig. 8. Analysis error every 9 days during 5-year assimilation cycle for the low viscosity case under different ensemble sizes: **(a)** 16 members, **(b)** 32 members, **(c)** 64 members, and **(d)** 128 members. The black line is the rms error of the ensemble mean and the red line is the rms spread in the ensemble. In all cases the “influence” radius $r_0 = 128$ km.

4 Results

Figure 8 shows how the performance of the ensemble Kalman filter depends on the ensemble size for the low viscosity case. As discussed in Sect. 3.2, the performance of the ensemble Kalman filter depends on the choice of the “influence” radius r_0 . In this study, the value of r_0 is chosen to be 128 km for both low viscosity and high viscosity cases and the filter performance is nearly optimal (in terms of minimum analysis error). The choice of r_0 is not unique: it varies with location and ensemble size. The filter performance probably can be further improved through using smaller value of r_0 in the western part of the basin compared with the eastern part and using larger value of r_0 with more ensemble members. In Fig. 8, the red line is the rms spread in the ensemble, which is given by

$$\sqrt{\frac{1}{N_s(N_e - 1)} \sum_{j=1}^{N_e} |\mathbf{x}_j^a - \bar{\mathbf{x}}^a|^2}$$

where $\bar{\mathbf{x}}^a$ is the mean of the ensemble analysis. The black line in Fig. 8 is the rms difference between the ensemble mean and the true state, which is given by

$$\sqrt{\frac{1}{N_s} |\bar{\mathbf{x}}^a - \mathbf{x}^t|^2}$$

and it is called error in the ensemble mean. As seen in Fig. 8, the ensemble spread is less sensitive to the ensemble size than the error in the ensemble mean. With the ensemble size of 16, both the ensemble spread and the error in the ensemble mean decrease slowly during the first year. The ensemble spread remains nearly constant after the first year. The error in the ensemble mean, however, increases dramatically at the 18th month and remains at a higher level than the ensemble spread afterwards. Upon doubling the ensemble size, a substantial decrease of the error in the ensemble mean is observed and the ensemble spread is very close to the error in the ensemble mean from the beginning to the end. When the ensemble size is further quadrupled to 128, the changes in both the ensemble spread and the error in the ensemble mean are negligible. Figure 8 shows that 32 ensemble members are sufficient to accurately describe the forecast error statistics for the low viscosity case.

Figure 9 shows the results of the ensemble Kalman filter under two different sets of ensemble initial conditions: one corresponds to the ensemble A as described in Sect. 2.3 and the other corresponds to the ensemble B. Ensemble B has a proper covariance structure, while ensemble A has an incorrect covariance structure: its ensemble spread is too low. As seen in Fig. 9, with proper ensemble initial condition, the error in the ensemble mean decreases very quickly during the first 6 months and remains at a nearly constant level af-

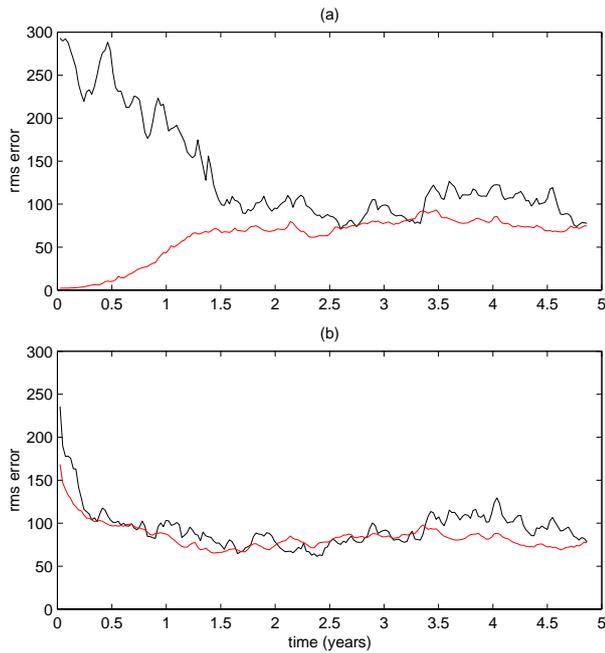


Fig. 9. Analysis error every 9 days during 5-year assimilation cycle for the low viscosity case under different ensemble initial conditions. The ensemble mean of the initial conditions for panel (a) and panel (b) is identical. The ensemble initial condition for panel (b) has correct covariance structure. The ensemble initial condition for panel (a), however, has a spread that is too low. The black line is the rms error of the ensemble mean and the red line is the rms spread in the ensemble. For both cases, the ensemble size is 64 and the “influence” radius $r_0 = 128$ km.

terwards. The ensemble spread is always indistinguishable from the error in the ensemble mean for the case with proper ensemble initial conditions. For the case in which the spread of the ensemble initial condition is too low, the error in the ensemble mean is much greater than the ensemble spread at the beginning. The ensemble spread increases and the error in the ensemble mean decreases rapidly during the first 18 months. After that, both the ensemble spread and the error in the ensemble mean change slowly with time. They intersect at the 30th month and remain nearly constant later.

As described in Sect. 3.3, the performance of the reduced-rank extended Kalman filter depends on the value of the tuning parameter γ which is the scaling factor applied to the covariance from the unconstrained model run to obtain the “pseudo-model error” covariance. It is found empirically that when the value of γ is chosen to be 0.0025 for the low viscosity case and 0.04 for the high viscosity case, the filter performance is best. The performance of the reduced-rank extended Kalman filter also depends on the number of retained EOFs, N_r . Buehner and Malanotte-Rizzoli (2003) found that the filter performance was nearly optimal when $N_r = 100$ for the low viscosity case and $N_r = 20$ for the high viscosity case and using more EOFs had very little effect on the filter performance.

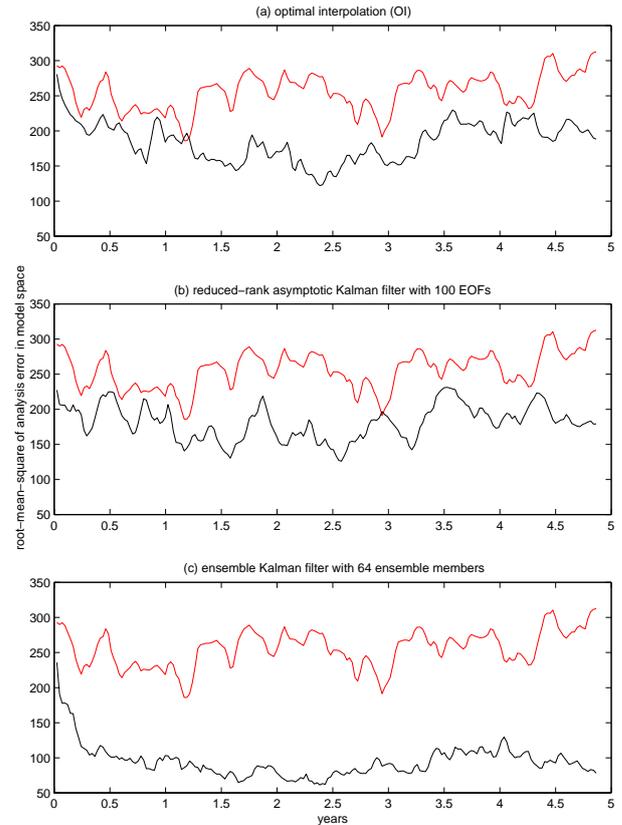


Fig. 10. Unconstrained model forecast error (red line) and analysis error (black line) every 9 days during 5-year assimilation cycle for the low viscosity case from the assimilation experiment of (a) optimal interpolation, (b) reduced-rank asymptotic Kalman filter with 100 EOFs and the tuning parameter $\gamma = 0.0025$, and (c) ensemble Kalman filter with 64 ensemble members and the “influence” radius $r_0 = 128$ km.

Figure 10 shows the rms error of the analysis from the optimal interpolation, the reduced-rank extended Kalman filter and the ensemble Kalman filter for the low viscosity case. The performance of each assimilation technique is evaluated by comparing the rms error from the assimilation experiment with the forecast error from the run without data assimilation. As seen there, every assimilation technique is able to reduce the analysis error to a certain level and the ensemble Kalman filter outperforms both the optimal interpolation and the reduced-rank extended Kalman filter. The reason why the ensemble Kalman filter is better than reduced-rank extended Kalman filter for the low viscosity case is that the system is strongly nonlinear over the time between assimilations.

The most striking feature in Fig. 10 is that the performance of the sophisticated reduced-rank extended Kalman filter is very similar to the simple optimal interpolation. The reason why the extended steady-state Kalman filter is no better than the statistical interpolation for the case in which the system is strongly nonlinear can be rationalized as follows. For both the extended steady-state Kalman filter and statistical interpolation, the Kalman gain is flow independent. The only dif-

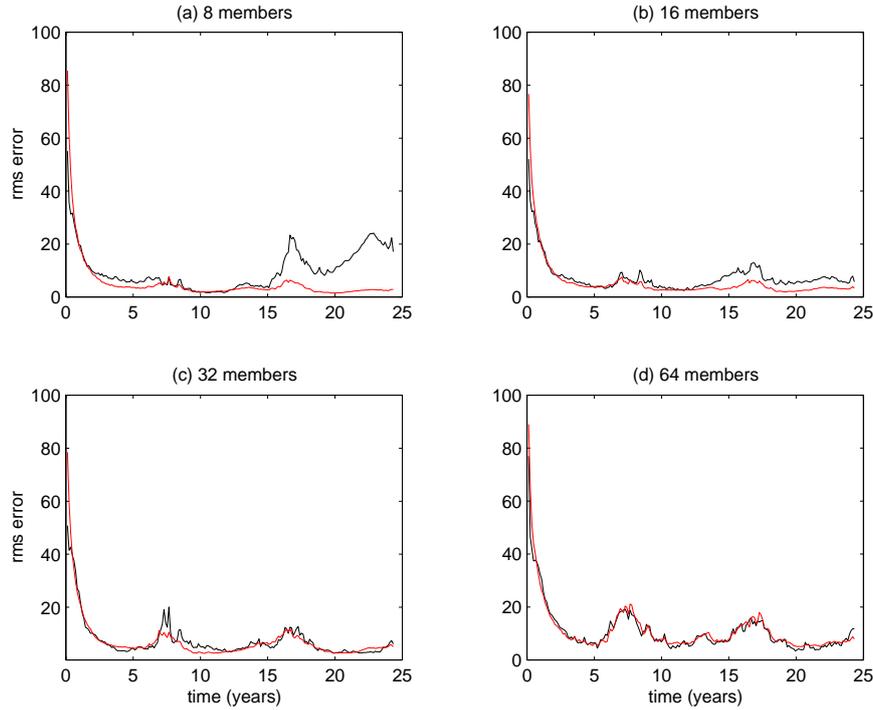


Fig. 11. Analysis error every 45 days during 25-year assimilation cycle for the high viscosity case under different ensemble sizes: **(a)** 8 members, **(b)** 16 members, **(c)** 32 members, and **(d)** 64 members. The black line is the rms error of the ensemble mean and the red line is the rms spread in the ensemble. In all cases the “influence” radius $r_0 = 128$ km.

ference between these two approaches is how the background error covariance is derived. In the statistical interpolation, the background error covariance is derived in an ad hoc way: it is assumed to be spatially homogeneous, the correlation length and variance level are tuned such that the assimilation performance is optimal in terms of minimum analysis error. In the extended steady-state Kalman filter, the model is linearized first to get the linear state transition operator and the background error covariance is derived for a limited subspace. Many assumptions are made in deriving the background error covariance for the reduced-rank extended steady-state Kalman filter and these assumptions are very hard to meet for highly nonlinear system. Therefore, it is of no surprise to see that the reduced-rank extended steady-state Kalman filter is no better than statistical interpolation for the strongly nonlinear system.

As shown in Fig. 10, the analysis errors of the optimal interpolation and the reduced-rank Kalman filter are on average 70% of the errors when no data are assimilated. The ensemble Kalman filter is able to reduce the analysis errors to about 30% of the errors in the run with no assimilation after 18 months.

Figure 11 shows how the performance of the ensemble Kalman filter depends on the ensemble size for the high viscosity case. The error in the ensemble mean and the ensemble spread decrease sharply for the first two years for all cases. With 8 ensemble members, the error in the ensemble mean is very similar to the ensemble spread during the

first 15 years and they diverge at the 15th year: the error in the ensemble mean suddenly increases, the ensemble spread, however, remains unchanged. With 16 ensemble members, the error in the ensemble mean and the ensemble spread also diverge at the 15th year. However, the difference between them is much smaller than that with 8 ensemble members. For the case with 32 ensemble members, the error in the ensemble mean and the ensemble spread remain always close to each other. Both of them undergo two slight increases in error: one at year 8 and the other at year 17. This is associated with the quasi-periodic behavior of the high viscosity case as the model shifts to a different climatology (higher or lower energy) at those times. The result from the ensemble Kalman filter with 64 ensemble members is very similar to that with 32 ensemble members. Figure 11 shows that with enough ensemble members (ensemble size ≥ 32), the rms spread in the ensemble is indistinguishable from the rms error of the ensemble mean. This implies that the covariances are not underestimated without covariance inflation. Covariance inflation was tried but did not reduce the error (figure not shown).

In Fig. 11, there is a spike in the error of the ensemble mean around year 6.6 and it seems odd that the spike is more pronounced with large ensemble members. As described in the introduction, for the high viscosity case, the model has three preferred regimes, each with distinct energy levels. Therefore, the probability density of the system will have a multi-modal distribution. Figure 12 shows the histogram of

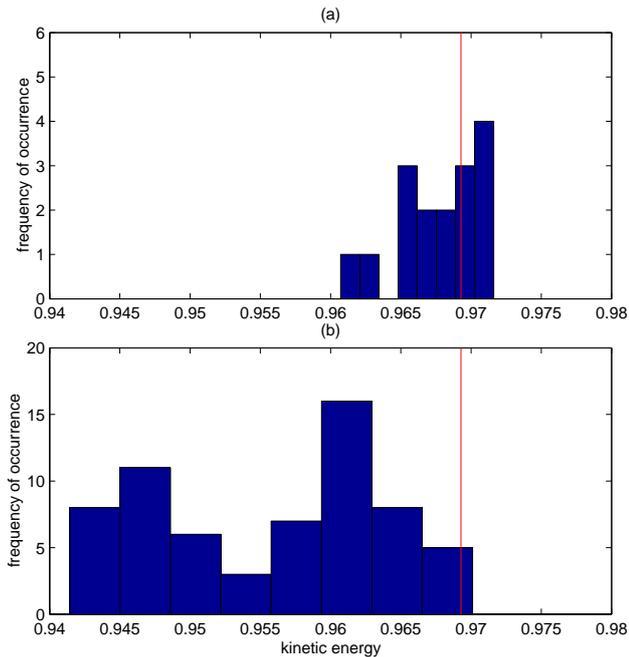


Fig. 12. The histogram of the kinetic energy of the ensemble analysis at year 6.6 with ensemble size of (a) 16 and (b) 64. The red line is the true state.

the kinetic energy of the ensemble analysis at year 6.6. As seen there, for the case with 16 ensemble members, the true state lies very close to the mean of the ensemble analysis. For the case with 64 ensemble members, the true state lies at the right end of the ensemble analysis. That's the reason why the error of the ensemble mean for the ensemble Kalman filter using larger ensembles can be larger than with smaller ensembles.

The rms errors of the analysis from the optimal interpolation, the reduced-rank extended Kalman filter and the ensemble Kalman filter for the high viscosity case are shown in Fig. 13. All filtering techniques are able to significantly reduce the analysis error relative to the run without data assimilation. The forecast error from the unconstrained model run exhibits a striking quasi-periodic behavior with a period of approximately 9 years. The analysis error from the optimal interpolation exhibits similar behavior but with a smaller amplitude. The quasi-periodic behavior severely diminishes after the reduced-rank extended Kalman filtering is applied and it is nearly invisible in the rms errors of the analysis from the ensemble Kalman filter. The errors of the analysis from the optimal interpolation, the reduced-rank extended Kalman filter and the ensemble Kalman filter decrease dramatically during the first two years. After two years, the analysis error of the optimal interpolation is on average 30% of the forecast error when no data are assimilated; the analysis error of the reduced-rank extended Kalman filter is on average 10% of the forecast error without data assimilation; and the analysis error of the ensemble Kalman filter is on average 6% of the forecast error without data assimilation. For the high vis-

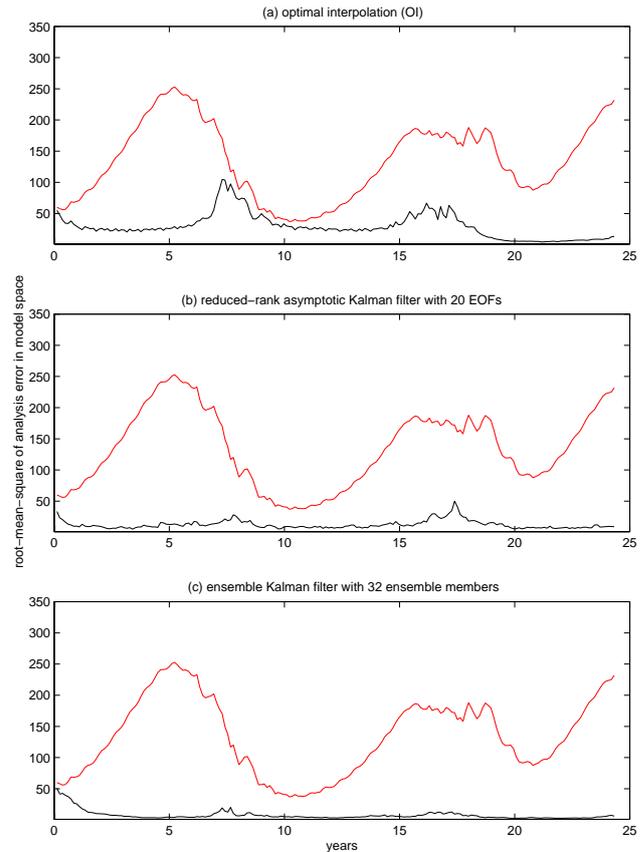


Fig. 13. Unconstrained model forecast error (red line) and analysis error (black line) every 45 days during 25-year assimilation cycle for the high viscosity case from the assimilation experiment of (a) optimal interpolation, (b) reduced-rank asymptotic Kalman filter with 20 EOFs and the tuning parameter $\gamma = 0.04$, and (c) ensemble Kalman filter with 32 ensemble members and the “influence” radius $r_0 = 128$ km.

cosity case, the reduced-rank extended Kalman filter greatly outperforms the optimal interpolation and is nearly as good as the ensemble Kalman filter.

5 Conclusions and discussions

The performances of the ensemble Kalman filter and the reduced-rank extended Kalman filter are compared using a quasigeostrophic, eddy-resolving model of the wind-driven ocean circulation. By changing the eddy viscosity, two qualitatively distinct types of behavior are obtained: chaotic state for the low viscosity case and quasi-periodic state for the high viscosity case.

For the low viscosity case in which the circulation is strongly chaotic and abounds with energetic mesoscale eddies, it is found that as expected, the error in the ensemble mean decreases as the ensemble size increases and 32 ensemble members are sufficient to accurately describe the temporally non-stationary, spatially inhomogeneous, and anisotropic structure of the forecast errors. In this case, the performance of the highly sophisticated stationary reduced-

rank extended Kalman filter is very similar to the simple optimal interpolation which is based on ad hoc assumptions about the forecast error covariance. The ensemble Kalman filter greatly outperforms the reduced-rank extended Kalman filter.

For the high viscosity case in which the circulation is quasi-periodic, simple, and well organized, all three assimilation techniques are able to significantly reduce the analysis error and the reduced-rank extended Kalman filter is nearly as good as the ensemble Kalman filter. For the high viscosity case, the model has three preferred regimes, each with distinct energy levels. Therefore, the probability density of the system has a multi-modal distribution and the error of the ensemble mean for the ensemble Kalman filter using larger ensembles can be larger than with smaller ensembles. The Bayesian approach as described in Sect. 3.1 should be used to obtain an optimal state estimate for this case. Attempts have been made to apply the Bayesian approach to very low dimensional nonlinear dynamic systems (Miller et al., 1999; Anderson and Anderson, 1999).

Verlaan and Heemink (2001) proposed a new method to quantify the nonlinearity of data assimilation problems and found that the nonlinearity depended not only the numerical model used, but also on the accuracy of the measurements, the sampling frequency, and the variance of the system noise. They used the well-known Lorenz model to compare the performance of several Kalman filter algorithms and found that different Kalman filter algorithms showed almost the same performance for weakly nonlinear problems. For strongly nonlinear problems, the more advanced methods were more accurate. Their results are consistent with ours. However, the nonlinearity in our data assimilation experiments depends only on the model used.

The steady-state reduced-rank extended Kalman filter has three major drawbacks. First, the extended Kalman filter has a closure problem: higher-order statistical moments are neglected in the error covariance evolution equation. In the ensemble Kalman filter, higher-order moments are retained in the forecast step. However, only the first two moments are used in the analysis step. Second, those approaches involving reduced-rank representation of the error covariance face the problem that there is dynamic coupling between the errors in the resolved and unresolved subspaces and it results in a “pseudo-model error” which is very hard to explicitly calculate. Third, the forecast error covariance and Kalman gain of the steady-state reduced-rank extended Kalman filter are time-independent.

Acknowledgements. This study was supported in part by the Office of Naval Research grant N00014-98-1-0881 and the National Science Foundation ITR project grant ACI-0121182. The computations were performed on the NCAR’s parallel machines. Special thanks go to M. Buehner and J. Hansen for discussions. Comments from reviewers helped to improve our presentation.

References

- Anderson, J. L. and Anderson, S. L.: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts, *Mon. Wea. Rev.*, 127, 2741–2758, 1999.
- Anderson, B. D. O. and Moore, J. B.: *Optimal filtering*, Prentice-Hall, 357, 1979.
- Ballabrera-Poy, J., Brasseur, P., and Verron, J.: Dynamical evolution of the error statistics with the SEEK filter to assimilate altimetric data in eddy-resolving ocean models, *Q. J. R. Meteorol. Soc.*, 127, 233–253, 2001.
- Buehner, M. and Malanotte-Rizzoli, P.: Reduced-Rank Kalman filter Applied to an Idealized Model of the Wind-Driven Ocean Circulation, *J. Geophys. Res.*, in press, 2003.
- Burgers, G., Van Leeuwen, P. J., and Evensen, G.: Analysis scheme in the ensemble Kalman filter, *Mon. Wea. Rev.*, 126, 1719–1724, 1998.
- Cane, M., Kaplan, A., Miller, R. N., Tang, B., Hackert, E. C., and Busalacchi, A. J.: Mapping tropical Pacific sea level: Data assimilation via a reduced state space Kalman filter, *J. Geophys. Res.*, 101, 22 599–22 617, 1996.
- Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte-Carlo methods to forecast error statistics, *J. Geophys. Res.*, 99, 10 143–10 162, 1994.
- Evensen, G. and van Leeuwen, P. J.: Assimilation of geosat altimeter data for the Agulhas Current using the ensemble Kalman filter with a quasigeostrophic model, *Mon. Wea. Rev.*, 124, 85–96, 1996.
- Farrell, B. F. and Ioannou, P. J.: State estimation using a reduced order Kalman filter, *J. Atmos. Sci.*, 58, 3666–3680, 2001.
- Fukumori, I. and Malanotte-Rizzoli, P.: An approximate Kalman filter for ocean data assimilation: An example with an idealized Gulf Stream model, *J. Geophys. Res.*, 100, 6777–6793, 1995.
- Gardiner, C. W.: *Handbook of stochastic methods: for physics, chemistry and natural sciences*, Springer-Verlag, Berlin, 444, 2002.
- Gelb, A.: *Applied optimal estimation*, M.I.T. Press, 374, 1974.
- Ghil, M.: Advances in sequential estimation for atmospheric and oceanic flows. *J. Meteor. Soc. Japan*, 75, 289–304, 1997.
- Heemink, A. W., Verlaan, M., and Segers, A. J.: Variance reduced ensemble Kalman filtering, *Mon. Wea. Rev.*, 129, 1718–1728, 2001.
- Houtekamer, P. L. and Mitchell, H. L.: Data assimilation using an ensemble Kalman filter technique, *Mon. Wea. Rev.*, 126, 796–811, 1998.
- Houtekamer, P. L. and Mitchell, H. L.: A sequential ensemble Kalman filter for atmospheric data assimilation, *Mon. Wea. Rev.*, 129, 123–137, 2001.
- Jazwinski, A. H.: *Stochastic processes and filtering theory*, Academic Press, New York, 376, 1970.
- Maybeck, P. S.: *Stochastic models, estimation and control*, Academic Press, 2, 423, 1982.
- Mahadevan, A., Lu, J., Meacham, S. P., and Malanotte-Rizzoli, P.: The predictability of large-scale wind-driven flows, *Nonlinear Processes in Geophysics*, 8, 449–465, 2001.
- McCalpin, J. D. and Haidvogel, D. B.: Phenomenology of the low-frequency variability in a reduced gravity, quasigeostrophic double-gyre model, *J. Phys. Oceanogr.*, 26, 739–752, 1996.
- Meacham, S. P.: Low-frequency variability in the wind-driven circulation, *J. Phys. Oceanogr.*, 30, 269–293, 2000.
- Miller, R. N., Ghil, M., and Gauthiez, F.: Advanced data assimila-

- tion in strongly nonlinear dynamical systems. *J. Atmos. Sci.*, 51, 1037–1056, 1994.
- Miller, R. N., Carter, E. F., and Blue, S. T.: Data assimilation into nonlinear stochastic models, *Tellus*, 51A, 167–194, 1999.
- Pham, D. T.: Stochastic methods for sequential data assimilation in strongly nonlinear systems, *Mon. Wea. Rev.*, 129, 1194–1207, 2001.
- Pham, D. T., Verron, J., and Roubaud, M. C.: A singular evolutive extended Kalman filters for data assimilation in oceanography, *J. Mar. Syst.*, 16, 323–340, 1998.
- Primeau, F.: Multiple equilibria and low-frequency variability of wind-driven ocean currents, *J. Phys. Oceanogr.*, 32, 2236–2256, 2002.
- Verlaan, M. and Heemink, A. W.: Nonlinearity in data assimilation applications: a practical method for analysis, *Mon. Wea. Rev.*, 129, 1578–1589, 2001.
- Verron, J., Gourdeau, L., Pham, D. T., Murtugudde, R., and Busalacchi, A. J.: An extended Kalman filter to assimilate satellite altimeter data into a nonlinear numerical model of the tropical Pacific Ocean: Method and validation, *J. Geophys. Res.*, 104, 5441–5458, 1999.
- Whitaker, J. S. and Hamill, T. M.: Ensemble data assimilation without perturbed observations, *Mon. Wea. Rev.*, 130, 1913–1924, 2002.