Available online at www.sciencedirect.com
SCIENCE DIRECT
Ocean Modelling 12 (2006) 378–400
Ocean Modelling
www.elsevier.com/locate/ocemod

# The representer method, the ensemble Kalman filter and the ensemble Kalman smoother: A comparison study using a nonlinear reduced gravity ocean model

Hans E. Ngodock [a,*], Gregg A. Jacobs [b], Mingshi Chen [a,1]

[a] Department of Marine Science, University of Southern Mississippi, Stennis Space Center, Mississippi, United States
[b] Naval Research Laboratory, Code 7320, Stennis Space Center, Mississippi, United States

Received 3 February 2005; received in revised form 4 August 2005; accepted 4 August 2005
Available online 8 September 2005

**Abstract**

This paper compares contending advanced data assimilation algorithms using the same dynamical model and measurements. Assimilation experiments use the ensemble Kalman filter (EnKF), the ensemble Kalman smoother (EnKS) and the representer method involving a nonlinear model and synthetic measurements of a mesoscale eddy. Twin model experiments provide the "truth" and assimilated state. The difference between truth and assimilation state is a mispositioning of an eddy in the initial state affected by a temporal shift. The systems are constructed to represent the dynamics, error covariances and data density as similarly as possible, though because of the differing assumptions in the system derivations subtle differences do occur. The results reflect some of these differences in the tangent linear assumption made in the representer adjoint and the temporal covariance of the EnKF, which does not correct initial condition errors. These differences are assessed through the accuracy of each method as a function of measurement density. Results indicate that these methods are comparably accurate for sufficiently dense measurement networks; and each is able to correct the position of a purposefully misplaced mesoscale eddy. As measurement density is decreased, the EnKS and the representer method retain accuracy longer than the EnKF. While the representer method is more accurate than the sequential methods within the time period covered by the observations (particularly during the first part of the assimilation time), the representer method is less accurate during later times and during the forecast time period for sparse networks as the tangent linear assumption becomes less accurate. Furthermore, the representer method proves to be significantly more costly (2–4 times) than the EnKS and EnKF even with only a few outer iterations of the iterated indirect representer method.


*Keywords:* Data assimilation; Representer method; Ensemble Kalman filter; Ensemble Kalman smoother

* Corresponding author.
*E-mail address:* Hans.Ngodock@nrlssc.navy.mil (H.E. Ngodock).
[1] Present address: Department of Mathematics, Colorado University, Denver, CO, USA.

1463-5003/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ocemod.2005.08.001

## 1. Introduction

There is an abundance of data assimilation algorithms, each falling into one of two categories: sequential or variational. The sequential methods are built on the application of the Kalman filter (Kalman and Bucy, 1961), and include the extended Kalman filter (EKF) (e.g. Miller et al., 1994), and some reduced-rank EKF such as the reduced-order extended Kalman filter (ROEK) of Cane et al. (1996), the sequential evolutive extended Kalman filter (SEEK) of Pham et al. (1997), the error subspace statistical estimation (ESSE) of Lermusiaux and Robinson (1999), the sequential evolutive interpolated Kalman filter (SEIK) of Pham et al. (1998), and the reduced-order information filter (ROIF) of Chin et al. (2002). Other implementations of the KF to nonlinear dynamics use Monte-Carlo techniques for the computation of the forecast error covariance matrix. Examples include the ensemble Kalman filter or EnKF (Evensen, 1994), and its many variants (see Tippett et al., 2003). The list of the sequential methods above is not exhaustive. It may be supplemented by the smoother version or implementation (e.g. the Kalman smoother (KS), the ensemble Kalman smoother (EnKS) ... etc.) and by the optimal interpolation (OI) based methods.

The variational methods are derived from the optimal control theory (Le Dimet and Talagrand, 1986) and objective analysis (Sasaki, 1958, 1970). They have been implemented in meteorology and oceanography as three- and four-dimensional variational assimilation methods, better known as 3D-Var (Sasaki, 1970) and 4D-Var (Lewis and Derber, 1985; Derber, 1987) respectively. The 4D-Var algorithm is computationally expensive. Efforts have been made to make it affordable, e.g. the incremental 4D-Var of Courtier et al. (1994). The 4D-Var may be implemented in its dual formulation, where the minimization of the cost function is done in the data subspace, e.g. the representer algorithm (Bennett, 1992, 2002).

Each of these sequential and variational methods is being used (or being considered to be used) at many operational meteorological and oceanographic centers, academic and research institutions. Considering the diverse assimilation community, a growing interest in the field and limited computer resources, one fundamental question arises: which method is more accurate at less cost and under which circumstances? The answer to this question should be inferred from a careful and rigorous study that compares the assimilation methods within the same context, which includes model dynamics and associated external forcing, initial and boundary conditions, data, data error covariance, model error covariance or its representation, and of course, the same estimator. The study should be conducted with a model of high complexity as currently used in the atmospheric and oceanographic communities. Also, the comparison should be carried out in various assimilation settings (e.g. different circulation events or dynamical cases, different methods for generating or prescribing the model error covariance, different datasets and data types) to assess the impact of different input factors on the outcome of the assimilation. A sensitivity study with respect to each of these control parameters would imply an extensive multidimensional problem. Such an extensive study presents major difficulties: (1) the large number of assimilation methods cannot be considered for one comparison study; (2) there is not a method to reconcile the representation of the model error covariance for all the sequential methods; (3) it would require a considerable amount of time to develop the adjoint of a complex model, and should the adjoint exist, the assimilation experiments would be computationally costly. Moreover, the use of a complex model may hamper the insights that we can gain from a simple model. All these factors place the comparison study beyond the scope of one study and would require a collaborative effort of many assimilation practitioners using the same model with different assimilation techniques. The alternative is to narrow the scope of an individual study, and some progress has been made in this direction.

Some recent studies have compared assimilation methods. Brusdal et al. (2003) compared the EnKF, SEEK and the EnKS using the Miami Isopycnal coordinate ocean model. Their study was a demonstration of the potential of these methods for operational ocean forecasting systems. That comparison was between sequential methods only. Caya et al. (2004) compared a strong constraint 4D-Var with the ensemble square root filter (EnSRF) Tippett et al. (2003), using a cloud model and simulated radar observations. The rms error of model variables was computed during the assimilation window, and for the forecast. The 4D-Var yielded the lowest rms error. However, with the assimilation window confined to the first fifth of the entire time interval of the experiment (assimilation and forecast), the sequential method was at disadvantage because of its inability to correct the initial condition.

The data assimilation and model evaluation experiment (DAMEE) project consisted of a comparison of assimilation results from different ocean models using different assimilation schemes on the same region (Willems et al., 1994). In such a study, it is difficult to draw consistent conclusions since many input parameters were different across the experiments. Kurapov et al. (2002) also conducted a comparison study between the generalized inverse or GIM (the name adopted for the representer method), the EnKF and the OI, using a linear baroclinic coastal ocean model. They reported that both the representer and the EnKF methods could fit the data satisfactorily, and that the representer method was able to compensate for poorly specified statistics for the model and data errors. Another comparison between the representer method and the EnKF was carried out by Reichle et al. (2002) with a land surface hydrology model and synthetic microwave radiobrightness data. They found that the EnKF could be less accurate than the representer method. However, the EnKF could be very inexpensive (in terms of computational cost) when a reasonable analysis was achieved with a moderate ensemble size. In all these studies, the number of assimilation methods to be compared was reduced to two or three. The criteria for comparison were (1) the fit to the data, measured by the rms error of the assimilated solution, (2) the forecast error associated with each method's estimate of the model state at the final time (e.g. Caya et al., 2004) , and (3) the computational cost of each method involved.

The study here is different from the DAMEE experiment, and from the study by Kurapov et al. (2002) in the following sense: we use a nonlinear model and implement both the EnKF and the representer method for the same model and datasets. Our approach is similar to that of Reichle et al. (2002) and Caya et al. (2004), comparing a sequential method with a variational, and not sequential with sequential, or variational with variational. We consider a weak-constraint 4D-Var, a change from Caya et al. (2004) who used a strong constraint 4D-Var. The strong constraint algorithm may become problematic when the assimilation time interval is longer than the predictability time of the model, i.e. the time range of influence of the initial condition. The comparison within this work follows the work of Reichle et al. (2002), i.e. a comparison between the representer method and the EnKF. Here an ocean model with mesoscale activity is employed. In addition, the EnKS solution is computed. It allows not only an assessment of the improvement in the EnKF solution by using the smoother version of the sequential method, but also to compare a sequential smoother (EnKS) to a variational smoother (representer method).

The variational method may be formulated as either strong constraint (i.e. assuming a perfect dynamical model), or weak constraint (i.e. allowing errors in the model). Note that the errors in the dynamical model can arise from many sources: bad parameterization, neglected and/or unresolved processes, erroneous external forcing ... etc. In the weak constraint formulation, the search for the minimum of the cost function can be carried out in the state space, which is prohibitively costly for models of high dimensions, or in the data subspace, which is generally much smaller. The representer method is a weak constraint variational method that minimizes a cost function in the data space. It will be applied as the variational method in this study in the form of the iterated indirect representer method (Bennett, 2002; Chua and Bennett, 2001; Ngodock et al., 2000; Muccino and Bennett, 2002). A variational method is implemented on the premise that the tangent linear model (TLM) of the nonlinear model is available and stable for the length of the assimilation window. The adjoint is based on the TLM. In this study, we used the full TLM of the ocean model under consideration, and the adjoint thereof. The iterated indirect representer method is an attempt to address the nonlinear properties of the system within the context of the linear minimization. For the first iteration, a solution of the full nonlinear model is used as the background state within the TLM. The representer method provides an optimal estimate of the correction to the background state. The next iteration uses the improved background state, and so on. The successive corrections within each iteration are expected to become smaller, and the TLM approximation more accurate.

The original Kalman filter (KF) algorithm is designed for linear models. For nonlinear models, the extended Kalman filter (EKF) is used, where the propagation of the model error covariance matrix is done with a linearized version of the model. Within these frameworks, the large dimension error covariance matrix has to be computed, and the linearization of the model might not work for strongly nonlinear models, e.g. Evensen (1992) and Miller et al. (1994). This makes the KF and EKF unlikely candidates for sequential data assimilation using nonlinear models with large dimensions. The EnKF avoids the large dimensional matrices of the KF and EKF and the oversimplifications of the reduced-order filters by generating the model error covariance matrix based on ensemble integrations with the nonlinear model dynamics. Thus, the EnKF does

not require any linearization of the model. Miller et al. (1999) and Evensen (1994) show that the ensemble approach resolves the major difficulties associated with the EKF.

Many factors influence the assimilation solution accuracy: data density, representation of model error covariance, properly assigning errors to the appropriate sources and the ensemble generation and size are a few. This study does not pursue each of these factors individually. Rather it focuses on the data density since it can easily be made common to both the sequential and variational methods. The model error covariance can be reconciled between the methods in the sense that the statistical properties (variance and decorrelation scales) used in the ensemble generation are the same for the model error covariance prescribed in the representer method. The selected methods are expected to perform satisfactorily accurately with dense data coverage. As data density decreases, the solution accuracy of each method is expected to degrade. An indication of which method is more capable is measured by retaining accuracy (in the analysis and forecast) as data density decreases.

It should be mentioned that at any final analysis time, the EnKS and the EnKF should always provide the same solution. Also, both the representer and the EnKS should yield the same solution at the analysis times, provided that: (1) the model is linear, and (2) the EnKS uses an infinitely large ensemble. Such consistency results are not expected when a nonlinear model is used as is the case here, even if a large ensemble were used, which would be impractical. Other reasons why we should expect differences follow. The representer method solves for the modal trajectory, i.e. the mode of the conditional joint pdf, while the general filter (ensemble or not) solves for the marginal conditional pdf. These will differ for nonlinear dynamics. This can be seen from the fact that with no observations the representer method results in the central forecast as the estimator, while the EnKF results in the ensemble mean (Evensen, 2004, personal communication). Furthermore, the EnKF cannot correct the initial conditions. The method has to progressively adjust toward the data until it starts fitting the data. The ensemble generates a good spatial covariance matrix that takes into account the nonlinear evolution of the model errors.

The EnKS uses the EnKF solution at the latest analysis time as first guess. The method then propagates this information backwards through a time–space covariance matrix generated by an ensemble of model trajectories, and is able to update the solution at all previous analysis times. This backward propagation is determined by the time correlation estimated from the ensemble of trajectories. Thus, the accuracy of the previous analyses depends on the accuracy of the EnKF analysis in the first place, and the time correlation in the EnKS covariance matrix. Nonetheless, the EnKS has a better covariance matrix than the EnKF because the trajectories that make up the ensemble comprise analyzed states at all previous analysis times. In the representer method, the covariance functions are prescribed. The potential of being wrong, i.e. prescribing covariance functions that do not reflect the 'reality' of the 'state' errors is greater. Specifying model errors is equally difficult in either method. The state error variance is more correctly represented in the ensemble methods because of the ability to take into account the nonlinearities. However, adding model error covariances to the ensemble is just as difficult as specifying them in the variational method. The TLM and the adjoint will miss the nonlinear interactions and propagation of the model errors. This can be compensated by the ability of the method to propagate the information backward and forward from any data to the entire model space–time domain. However, the method is clearly limited by the accuracy of the TLM. This is especially crucial when strong nonlinearities are present.

Owing to all these potential sources of disagreement between the selected methods, this study will not focus on such aspects like choosing the right ensemble size or prescribing the right covariance functions for the representer method. Since a large ensemble is impractical and since the literature abounds with applications using limited-size ensembles, this paper focuses on what accuracy is achieved with the representers and a limited-size ensemble for EnKF and EnKS as one solves the assimilation problem. Thus, this paper is not an attempt to cross-validate the assimilation methods, justify the choice of covariance functions of ensemble size.

In the next section, we describe the dynamical model. Assimilation experiment setup and evaluations are discussed in Section 3, and results of the experiments are examined in detail in Section 4. In Section 5, we discuss the cost of implementing each of these methods, and concluding remarks follow in Section 6.

## 2. The dynamical model

A nonlinear reduced gravity (primitive equation) model is used to simulate an idealized eddy shedding off the loop current (hereafter LC) in the Gulf of Mexico (hereafter GOM). It is the same as the $1\frac{1}{2}$ layer

version of the reduced gravity model introduced by Hurlburt and Thompson (1980). The dynamical equations are:

$$\begin{aligned}
&\frac{\partial hu}{\partial t} - fhv + g'h\frac{\partial h}{\partial x} = A_M\left(\frac{\partial^2 hu}{\partial x^2} + \frac{\partial^2 hu}{\partial y^2}\right) + \tau^x - \mathrm{drag}_x,\\
&\frac{\partial hv}{\partial t} + fhu + g'h\frac{\partial h}{\partial y} = A_M\left(\frac{\partial^2 hv}{\partial x^2} + \frac{\partial^2 hv}{\partial y^2}\right) + \tau^y - \mathrm{drag}_y,\\
&\frac{\partial h}{\partial t} + \frac{\partial hu}{\partial x} + \frac{\partial hv}{\partial y} = 0,
\end{aligned} \tag{1}$$

where $u$ and $v$ are the zonal and meridional components of velocity, $h$ is the layer thickness, $f$ is the Coriolis parameter (here a $\beta$-plane is adopted), $g$ is the acceleration due to gravity, $g'$ is the reduced gravity, $A_M$ is the horizontal eddy diffusivity, computed based on the prescribed Reynolds number $Re$, the maximum inflow velocity and half the width of the inflow port. The model parameters are listed in Table 1.

Hurlburt and Thompson (1980) showed that it is possible to simulate the eddy shedding by specifying time-invariant transport at the inflow and outflow open boundaries (see the model domain in Fig. 1). In this case the wind stress and the bottom drag are neglected. With a transport of 35Sv at inflow and outflow ports, we can simulate an eddy shedding with a period of about 4 months. The computations are carried out on an Arakawa C-grid (Mesinger and Arakawa, 1976), with uniform space steps $\Delta x = 20$ km and $\Delta y = 18.75$ km and a time step of 20 min. The equations are solved for the transports $hu$ and $hv$ and for the layer thickness $h$, and the velocities are computed by dividing the transports by the instantaneous layer thickness. Thus, the nonlinearities are present only in the pressure gradient terms and in retrieving the velocities from the transport, making the tangent linearization easier. The domain is 1600 km × 900 km, with the inflow port 160 km wide, centered 400 km from the eastern boundary, and the outflow port 150 km wide, centered

Table 1
Table of model parameters

| $\beta$ | $f_0$ | $g$ | $g'$ | $Re$ |
|---|---|---|---|---|
| $2\times10^{-11}$ m$^{-1}$s$^{-1}$ | $5\times10^{-5}$ s$^{-2}$ | 9.806 m/s$^2$ | 0.03 m/s$^2$ | 50.2 |

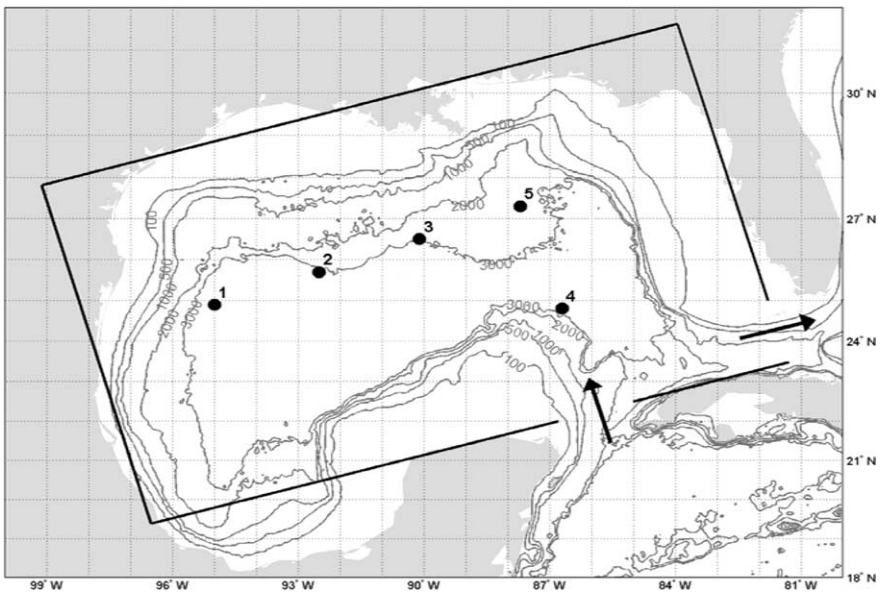


Fig. 1. The model domain is an idealized Gulf of Mexico representation with inflow and outflow ports. The selected diagnostic locations of the assimilation solution are marked with bullets.

75 km from the southern boundary. A schematic representation of the GOM is a rectangular domain with inflow and outflow ports, and the model grid is 20° from direct alignment with east–west and north–south (see Fig. 1).

In the reduced gravity model the initial upper-layer thickness is 300 m with the flow at rest. From the reduced gravity layer thickness, the corresponding sea surface height (SSH) is given as:

$$\mathrm{SSH} = (h - 300)/(1 + g/g'). \tag{2}$$

This linear relationship yields a linear measurement functional when the SSH data are assimilated into the model. The reduced gravity is $g' = g\Delta\rho/\rho$, where $\rho$ is the reference density, $\Delta\rho$ is the density difference between the active layer and the motionless bottom layer, and $g$ is the acceleration due to gravity.

## 3. Experiment setup and evaluation metrics

### 3.1. Experiment initialization, generating initial conditions and ensemble members

The assimilation experiments are set up to test the ability to correct the position of the mesoscale eddy field. This is accomplished through a twin model experiment in which one model run provides an initial condition for the assimilation as well as the data for the assimilation. In order to provide an eddy position error, the observations are offset in time.

The dynamical model of Section 2 is spun up from rest for 4 years, which is sufficient time to reach a statistical equilibrium (Hurlburt and Thompson, 1980), and provides the initial condition for the subsequent setup. The reference or truth solution is constructed by integrating the nonlinear model from the initial condition for an additional 8 months and storing only the last 4 months from which the observations are sampled. Thus the assimilation time window is 4 months. The background solution, which the assimilation experiments are supposed to correct, differs from the reference solution by being lagged in time. The SSH measurements are obtained from the layer thickness $h$ through Eq. (2). We have assigned a standard deviation of 5 cm to the SSH and 5 cm/s to the velocity measurements in all the assimilation experiments. For all assimilation experiments, the data error covariance matrix is assumed diagonal, the variance being the square of the respective measurement standard deviation.

To generate ensemble members, the model is integrated from the initial condition for 2 months and forced by a random wind stress. The random wind stress is computed as a standard deviation multiplying a random field with unity variance. Each realization of the random field is generated from the inverse Fourier transform of the prescribed wave number spectrum with a random phase at each wave number (Evensen, 1994). The wave number spectrum is a Gaussian function with e-folding length of 100 km in both $x$ and $y$ directions. The standard deviation ($10^{-4}$ m$^2$/s$^2$) of the wind stress field represents a typical stress of 0.1 N/m$^2$ plus uncertainty in the dynamical equations. The perturbation wind stress applied to each ensemble is recomputed every observation interval, which is 5 days or 10 days depending on the sampling network. This wind stress perturbation represents not only errors in the external momentum flux, but also errors in the system dynamics (Eq. (1)). A different random wind stress is provided to each ensemble member to generate as many members as are needed in the ensemble. The 2-month time period in the ensemble spin-up ensures two things: first, each ensemble covers a large portion of physically realizable solutions. That is, the ensemble properly represents the model pdf. Second, the 2-month offset provides a significant deviation from the observations. This constructs the initial ensemble (i.e. after the 2-month spin-up with random wind stress).

The model error covariance functions used in the representer method are given by

$$Q(\mathbf{x}, t, \mathbf{x}', t') = V(\mathbf{x}, \mathbf{x}') \exp\left(-\frac{(\mathbf{x} - \mathbf{x}')^2}{2L^2}\right) \exp\left(-\frac{|t - t'|}{\tau}\right) \tag{3}$$

where $L = 100$ km, i.e. the same decorrelation length as in the random fields used in the wind stress perturbations, and $\tau = 10$ day is the time decorrelation scale. The variance $V(\mathbf{x}, \mathbf{x}')$ is constant in space and time and is equal to the square of the standard deviation of the wind stress perturbations. Finally, the dynamical errors in the representer method are applied only to the momentum equations (Jacobs and Ngodock, 2003), and there is

no cross-correlation between the model error components. The continuity equation is considered strong constraint.

The initial ensemble mean is also taken as the initial condition for the representer experiments. The background used in the variational solution for the tangent linearization of the dynamical equations (for the first iteration of the indirect representer method) is obtained from integrating the nonlinear model for 4 months from the initial ensemble mean.

Fig. 2a shows the difference between the reference and the background initial conditions. The reference initial condition has a loop current eddy (hereafter LCE) in the center of the model domain, and the remnant of a LCE at the northwestern corner of the model domain. In the next 3 months, the LCE in the middle of the domain will quasi-linearly propagate westward, crash against the western boundary and slowly move northward while dissipating. In the meantime, the loop current intrudes further into the domain and by the wake of the fourth month another LCE is about to shed from the LC. In the background initial condition, the LCE is just about to shed from the LC, and a previously shed LCE has reached the western coast of the GOM. Fig. 2b shows that there is a substantial spread in the ensemble. Although there is more spatial structure in the spread with 25 members, the magnitude of the spread did not increase with 100 members. The values in Fig. 2b indicate that for SSH, the difference between members varies from −10 cm to 10 cm, where a change of 1 cm in SSH corresponds to a change of 3 m in the layer thickness according to Eq. (2).

A standard set of diagnostic stations is used throughout the evaluations in this study. The station locations are shown in Fig. 1. They are selected in such a way that they are common to all the sampling networks; locations 1–3 are distributed along the path of the LCE, location 4 is in the region where the LCE detaches, and

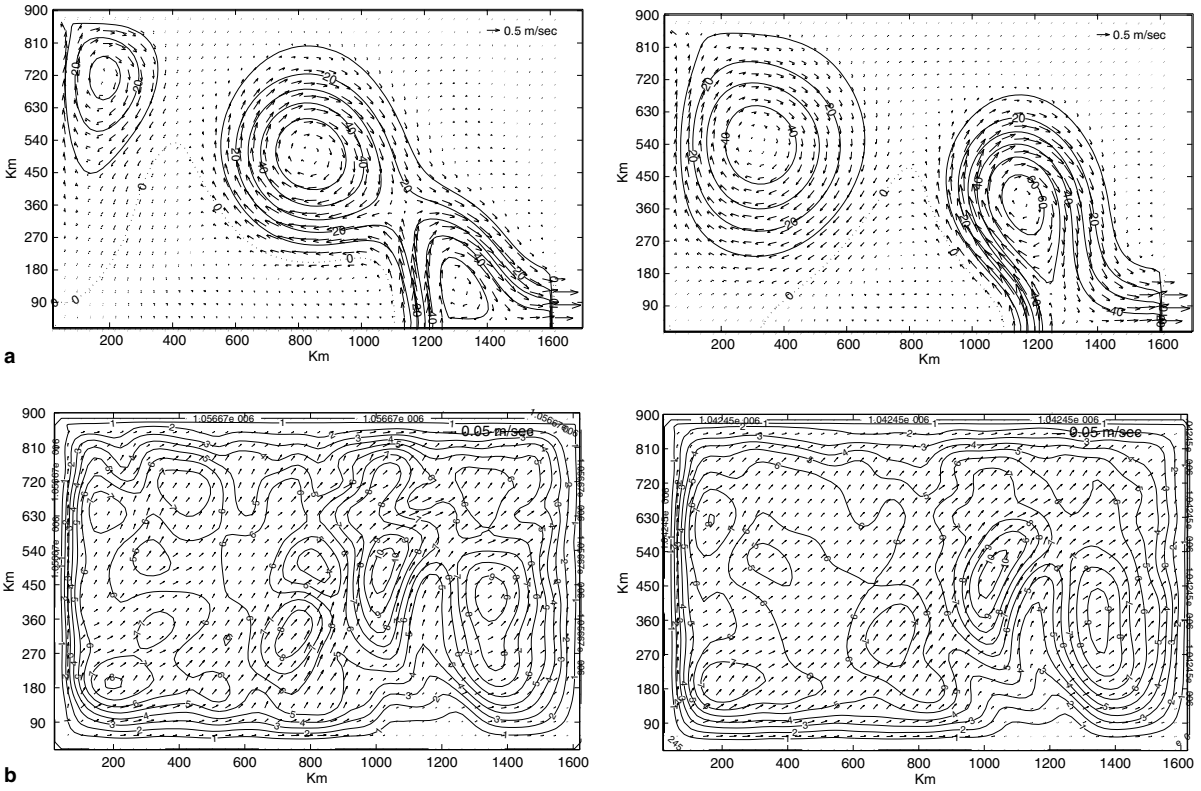


Fig. 2. (a) Initial condition of the reference solution (left) from which observations are taken and the initial ensemble mean (right) which is also the initial condition of the background for the variation assimilation. The contour lines are for SSH, and the contour interval is 10 cm. The correction that the assimilation systems must affect is to change the position of the loop current eddy. (b) The square root of the ensemble spread around the initial ensemble mean using 25 members (left) and 100 members (right).

location 5 is north of the LCE detachment region. The background displays a phase lag of about 2 months from the 'truth' solution as shown on sample time series in Fig. 3.

The rms errors between the background and the 'truth' solution, computed at the diagnostic stations over the 4-month and the last 2 months of the assimilation are shown in Table 2. The computation of the rms is restricted to the diagnostic stations for the following reason: the variability in the model domain is due to the LC intrusion and the LCE shedding and westward propagation. Elsewhere there is little variability, if at all. That is the reason we have selected diagnostic stations along the LCE path and in the region of the LCE

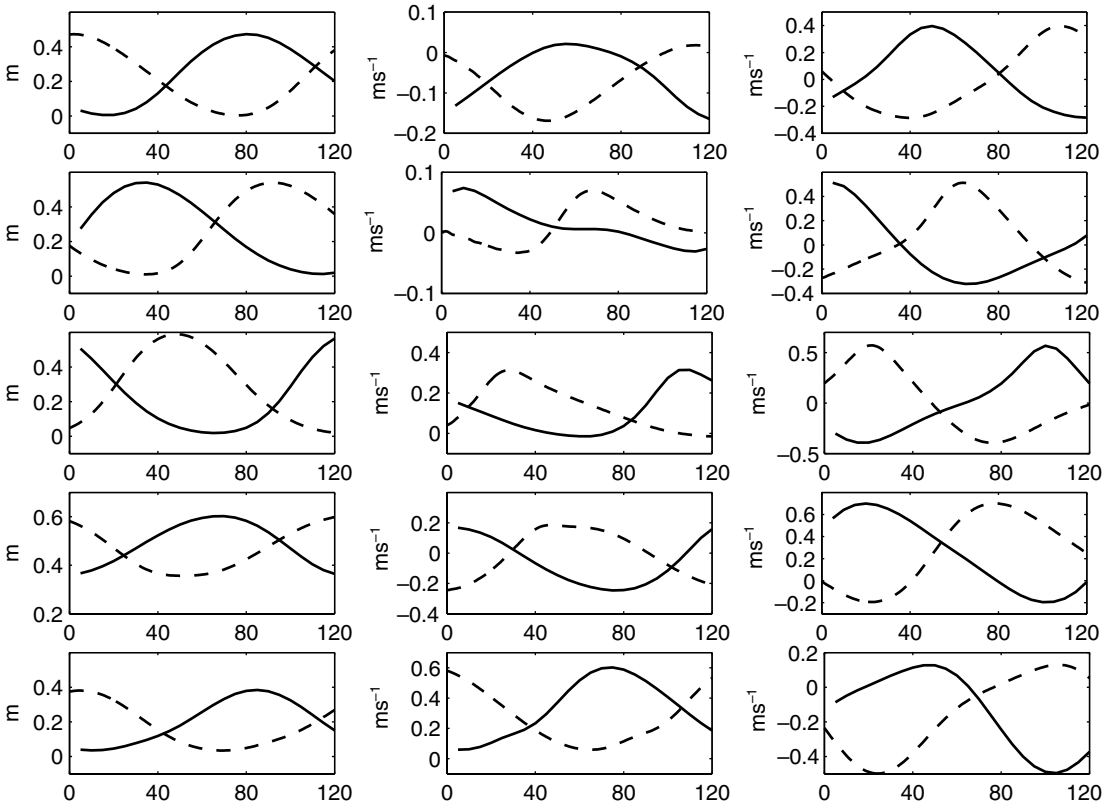


Fig. 3. Time series of the reference solution from which data is taken (solid line) and the background that must be corrected (dashed line) at diagnostic locations 1–5 (top to bottom rows respectively). The columns from left to right are for SSH, u velocity and v velocity respectively, and the $x$-axis is the number of days.

Table 2
Prior rms errors (between the reference and the background solutions) at the selected diagnostic locations computed for the 4-month and the last 2-month of the assimilation time window

| | 4-month | | | Last 2-month | | |
|---|---|---|---|---|---|---|
| | SSH | u | v | SSH | u | v |
| Location 1 | 0.323 | 0.1224 | 0.4552 | 0.3318 | 0.1144 | 0.4527 |
| Location 2 | 0.39 | 0.0544 | 0.4945 | 0.3750 | 0.0462 | 0.4808 |
| Location 3 | 0.3725 | 0.2038 | 0.6036 | 0.3494 | 0.2101 | 0.5766 |
| Location 4 | 0.1629 | 0.2851 | 0.6344 | 0.1652 | 0.2849 | 0.6224 |
| Location 5 | 0.2327 | 0.3419 | 0.4353 | 0.2426 | 0.366 | 0.4424 |

The units are meters (m) for SSH and meters per second (m/s) for both components of velocity.

shedding. Often, the rms misfits are computed based on a spatial average. However, that would lower the rms (by the contribution of the regions with no variability common to the model and the data) and cloud the impact of the assimilation.

The assimilation problem at hand can arguably be presented as an initialization problem. However, it is known that the sequential filter cannot estimate the initial condition from future observations. While correcting an error in the initial condition (a misplaced eddy) is important, it is equally important to consider the fit to the future observations and accuracy of a subsequent forecast beyond the observation interval.

### 3.2. Nonlinearities

Nonlinearities in the model are represented by strong advection when the LCE sheds from the LC. Also, as the LCE propagates westward, nonlinear interactions between advection, latitudinal variations in Coriolis force and the pressure gradient terms are important. Nonlinearities influence the performance of the EnKF, in the sense that the probability distribution function of the model error may not be Gaussian, when propagated by nonlinear dynamics. On the other hand, a nonlinear model must be preferably tangent linearized, with a stable TLM and the adjoint thereof, prior to using the representer method. Thus, the significant difference between the background and the reference solution (and thus the data, see Table 2 and Figs. 2 and 3), and the nonlinearity of the model provide a challenging framework for all methods, and hence the comparison. In these experiments, the TLM is stable for the assimilation window, and we have implemented the adjoint of the full TLM. The stability of the TLM was tested by the growth of small perturbations using the first-order Taylor's expansion. For several measurement types (u, v, and ssh) and sites in the space–time domain the representer matrix is symmetric, which insures that the adjoint–covariance–TLM system is consistent.

### 3.3. Ensemble size

The EnKF (and EnKS) solution is computed with increasing ensemble size, using 25, 50, 100, 150 and 200 members respectively. Results not reported here show that there is a substantial improvement to the solution of both the EnKF and EnKS as the ensemble size is increased from 25 to 50 to 100. Beyond 100 members, there is little change in the solutions. Therefore, for the comparison experiments below, we will consider the EnKF and EnKS solutions that used the ensemble of 100 members.

### 3.4. Observation networks and evaluation metrics

It is expected all assimilation methods perform well with an accurate observation network that covers most of the state space. In fact, if all the state space is observed with satisfactory small errors, we do not need to assimilate at all. However, the observations seldom cover all the state space (all model state variables observed throughout space and time), and the accuracy of the assimilation is affected by the observation density. Therefore, the experiments here have eight data networks ranging from sparse (network 1) to dense (network 8) to investigate the accuracy of the assimilation methods. Table 3 describes the networks, the total number of measurements from each network.

The accuracy of each method is measured by the rms error (data minus analyzed solution) of each model state variable computed at the measurement sites. This rms is square root of the mean of squared errors over of the assimilation time. A successful assimilation experiment should fit the observations and the dynamics to within respective errors. This implies that the rms should be less than the measurement standard deviation. This is the definition of 'accuracy' we will be using in the assimilation experiments.

It is known that the EnKF, being a sequential filter, does not correct the initial condition. Since the specified initial condition is lagged in time to represent a misplaced eddy, some time is required for the EnKF to adjust the model state to match the data. Thus, it should not be expected that the EnKF typically perform well early in the assimilation time window. The EnKS is implemented sequentially following Evensen and van Leeuwen (2000). However, being a smoother, it uses all the data within the time interval to adjust the model trajectory, which can include the earliest observation time. The representer method also has the ability to

Table 3
The number of observations provided by each sampling network

| Network | Assimilation time covered | 1 300 km 10 days | 2 300 km 5 days | 3 200 km 10 days | 4 200 km 5 days | 5 100 km 10 days | 6 100 km 5 days | 7 60 km 10 days | 8 60 km 5 days |
|---|---|---|---|---|---|---|---|---|---|
| Experiment 1 | 4 month | 540 | 1080 | 1152 | 2304 | 5184 | 10,368 | 14,976 | 29,952 |
| Experiment 2 | 6 month | | 1620 | 1728 | 3456 | | | | |
| Experiment 3 | 4 month Only SSH | | | 384 | 768 | 1738 | 3456 | 4992 | 9984 |

The observations are uniformly distributed in the space and time domain covered by each experiment. The second column provides some detail on the differences between the experiments. The abbreviations on the first row provide the spatial and temporal resolution of the observations. For example, 60 km and 5 days stand for a data network sampling every 60 km (in both the $x$ and $y$ directions) and every 5 days respectively.

adjust the model trajectory. As a consequence, these methods will be compared using three metrics: first, we compare the rms error of the three solutions (i.e. representer, EnKS and EnKF) over the last 2 months of assimilation, so that the EnKF is not directly put at disadvantage. Secondly, the study compares the rms of the representer and the EnKS solution errors over the entire time interval. Thirdly, we compare the forecast rms error from the three methods.

## 4. Experiments

### 4.1. Analyzed errors and relative assimilation system accuracies

The first set of experiments consists of assimilating the data from each of the eight networks individually using the representer, the EnKS and the EnKF methods. The data consist of SSH, u and v observations. The rms error at the five diagnostic stations for each network is computed over the last 2 months of the assimilation window, and plotted against the network number (Fig. 4) rather than the total number of measurements from the network, in order to maintain an equidistant $x$-axis in the plots. The correspondence between the network number and the number of measurements from the network is given in Table 3. We will use the same legend in Figs. 4–11, where the dotted line will represent the observation standard deviation and the assimilated solution and forecast rms will be plotted shown for the representer (solid line), the EnKS (dashed lined) and the EnKF (dashed-dotted line).

For the EnKF, there are only a few locations and state variables for which the rms is below or close to the data standard deviation for networks 4–8 (the denser networks, Fig. 4). The rms nonuniformly decreases as the network number increases. The EnKF errors have large peaks for networks 5 and 7 (using the longer 10 days sampling), particularly at diagnostic locations 3, 4 and 5, which are in the vicinity of the LCE shedding area. These networks sample data every 10 days, indicating that the method is more accurate with data sampled more frequently in time (every 5 days in this case), even though the spatial sampling may be more sparse. The rms errors associated with the EnKF solution show that the method is the least capable of driving the solution toward the reference. Nevertheless, considering the prior rms errors in Table 2, Fig. 4 shows that the EnKF is able to reduce the prior rms by as much as 90% (e.g. network 8 and location 1 for SSH and v, also location 2 for SSH), albeit for prior rms reductions lower than 40% at many locations: SSH and u at location 4 and network 5; u at location 1 and networks 1–3, location 2 and networks 3 and 5.

For the EnKS, the solution is significantly more accurate than that of the EnKF. The rms errors associated with the EnKS are generally lower that those of the EnKF with only one exception: u at location 2 and networks 3–5. This unexpected behavior may be solely due to the prior rms being already lower than the measurement standard deviation. Other than this one anomaly, the EnKS displays improvements to the EnKF that exceed one standard deviation of the observation noise level (even for sparse networks), which translate into a higher reduction of the prior rms errors.

The representer rms errors are generally lower than those of the EnKF and EnKS. There are only a few locations where the representer is less accurate that the other two methods: SSH at location 4 and for networks 1–3, v at locations 3–5 and for networks 1–3.

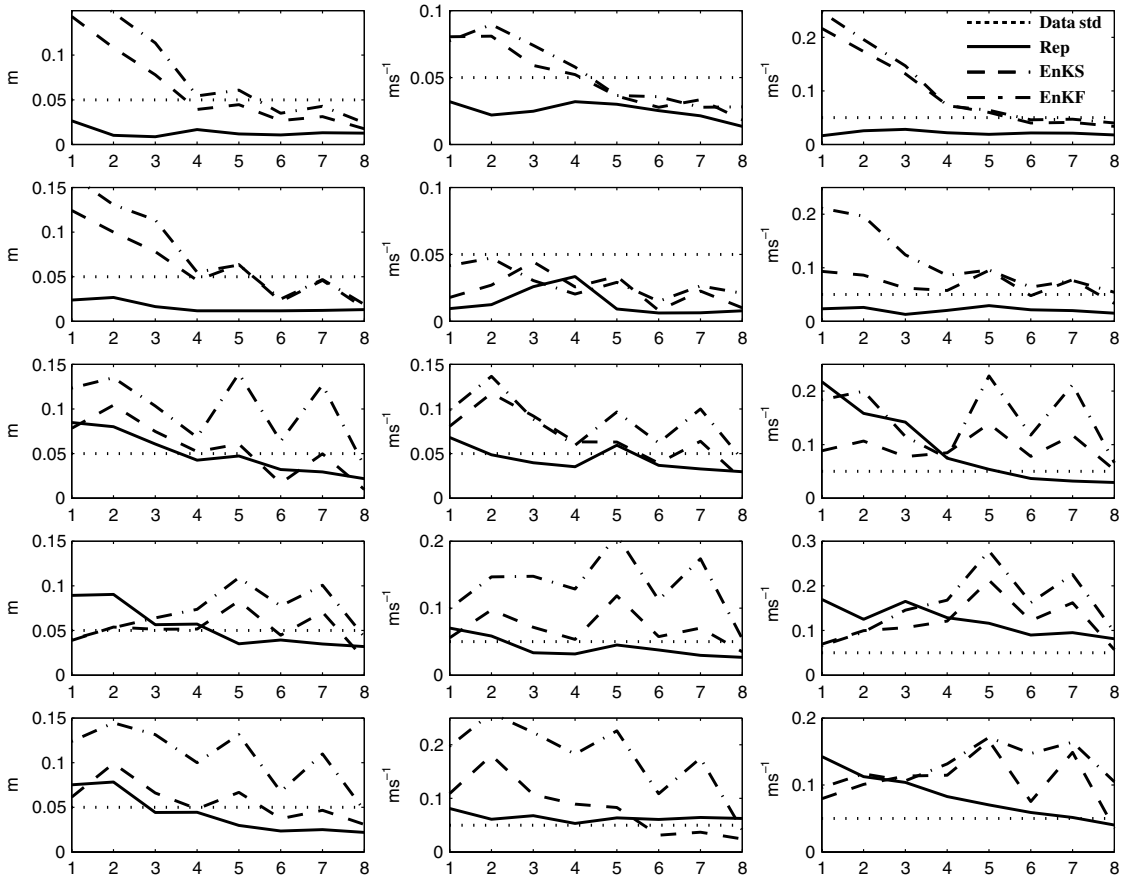


Fig. 4. The rms of all state variables: SSH (left column), u (center column) and v (right column) at all five diagnostic locations (from location 1 on top row to location 5 on bottom row). The *x*-axis is the network number. The association of the network number and the total number of data is given in Table 3. The rms averaging covers the last 2 months of the 4 month assimilation time interval. Solid lines are the representer rms errors, dashed lines are the EnKS errors, dashed-dotted lines are the EnKF errors, and dotted lines are the observation noise level.

### 4.2. EnKS and representer comparison

Because the EnKF does not correct the initial condition, comparisons in Fig. 4 are limited to the final 2 months of the assimilation window. However, both the EnKS and representer methods correct the initial condition, and a fair comparison may be conducted over the entire assimilation window. In this second comparison, the rms error of the EnKS and the representer solutions is computed over the entire assimilation window. Results are plotted in Fig. 5, and show that the representer is more accurate than the EnKS as it retains its accuracy with more sparse networks. The larger differences between the two methods in Fig. 5, as compared to Fig. 4, indicate that representer method fits the data much better in the first half of the assimilation window than its does in the second half. The converse seems to be true for the EnKS, although not uniformly across the networks and locations, and that may be due to initial condition inaccuracy in the EnKS solution.

### 4.3. Initial condition effects

The EnKF analysis time series of the assimilated solution (not shown), requires between 20 and 60 days (depending on model variables) to begin to match the data. This is a direct effect of the purposeful wrong initial condition. Eventually, the influence of the initial condition will decay over time, resulting in a better

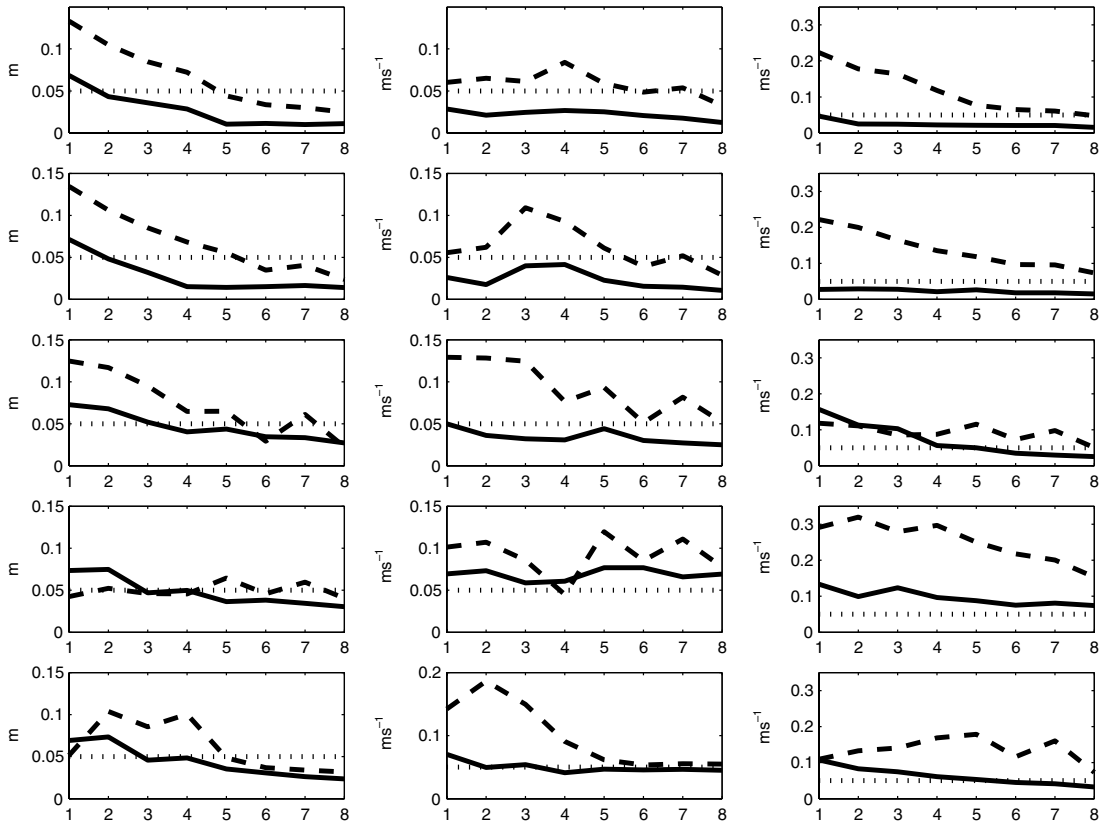


Fig. 5. This analysis is the same as Fig. 4 except the rms errors are computed over the entire assimilation window. Only the representer (solid line) and EnKS (dashed line) results are shown. The EnKF does not correct the initial condition, and thus is not comparable to the EnKS and representer solutions over the entire assimilation window.

fit to the data by the EnKF at times that are distant enough from the initial time. Therefore, a second set of experiments (Experiment 2 in Table 3) is designed where the assimilation period is extended to 6 months and the rms computation covers the last 3 months. The total number of data involved in these experiments is listed in Table 3. The assimilation experiments are conducted using data only for networks 2, 3 and 4. These are the (sparse) networks for which the EnKF and EnKS accuracy sharply degrades after satisfactory performance for the denser networks (see Figs. 4 and 5).

In comparison with Fig. 4, the results in Fig. 6 indicate that the sequential methods have a much lower rms (across the networks and for almost all state variables) than in the previous experiment. This suggests that sequential methods benefit from a longer adjustment time, and therefore are more accurate in the second half of the assimilation window. The representer method on the other hand yields mixed results. In most cases (e.g. zonal velocity at all locations, meridional velocity at locations 1–2, SSH at locations 1–2) the rms slightly increased, especially using network 3, which samples every 10 days. Some improvements in the representer solution rms are noted for the meridional velocity at locations 3–5 for all networks involved and the SSH at locations 4–5 for network 2, and location 3 for network 4. The main point here is that in going from 4 to 6 months the rms did improve for the sequential methods in the last 3 months. The implication is that when the sequential methods are given enough time to adjust, the analysis accuracy improves. However, at most locations and networks, the representer is more accurate than the EnKS and EnKF, with some exceptions.

The representer and the EnKS rms errors are also computed over the entire 6-month assimilation window. Results (Fig. 7) show that the representer outperforms the EnKS. Note here that the EnKS rms increases while the representer rms is slightly lower than that of the second half window (Fig. 6). This again indicates that while the representer is fitting the data slightly better in the first half of the assimilation window than it does

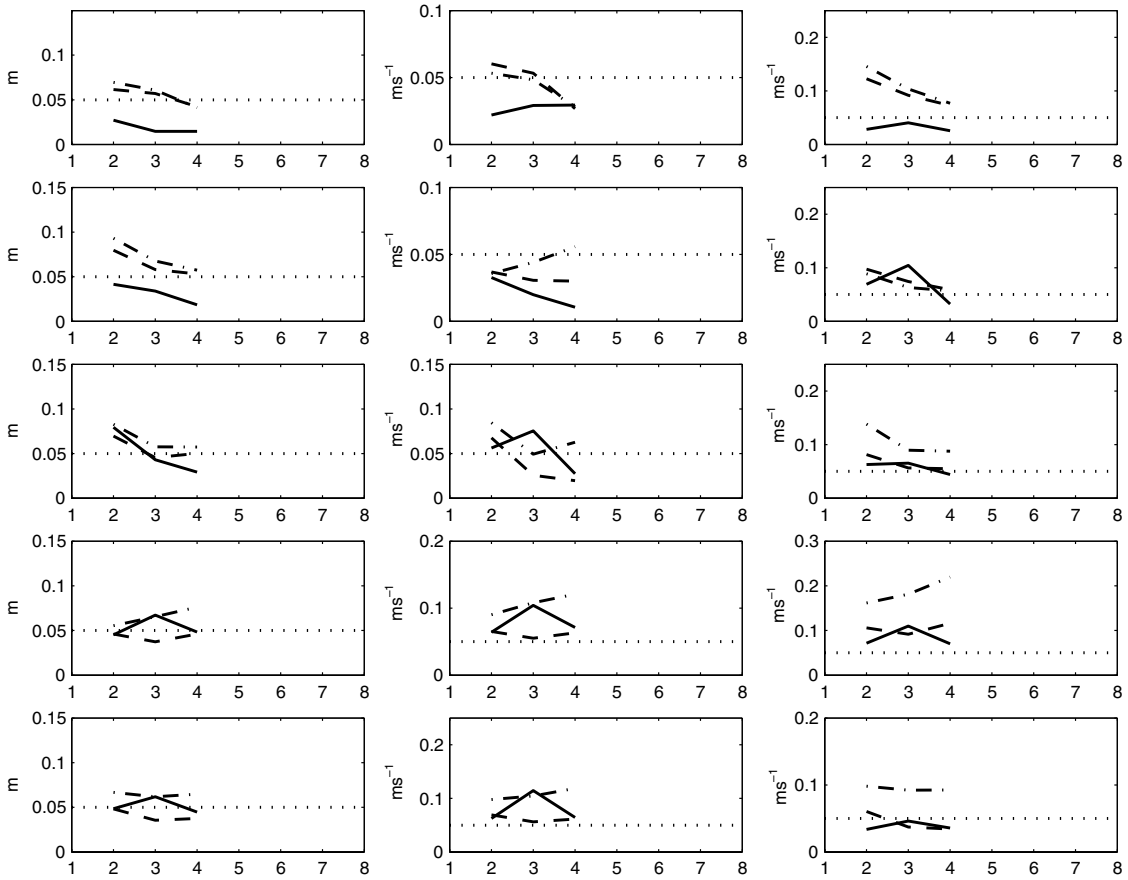


Fig. 6. The results here are similar to those in Fig. 4 except the assimilation window is 6 months, and the averaging of the rms covers the last 3 months, and the experiments are carried out only for networks 2, 3 and 4.

in the second, the EnKS is fitting the data better in the second half than it does in the first. This finding becomes critical in the forecast comparison.

### 4.4. SSH only

The assimilation of only SSH data is motivated by the potential use of satellite altimeter data. Experiment 3 covers 4 months as in Experiment 1, but only SSH is assimilated. The assimilation is carried out for networks 3–8. The corresponding number of data is listed in Table 3.

The rms errors are computed for the last 2 months. The results (Fig. 8) indicate that the accuracy in this experiment follows the pattern of Experiment 1. The representer method is accurate for SSH at locations 1, 2 and 5 for all networks, location 3 for networks 4, and 6–8. At location 4 the accuracy decreases for networks 3–6. However, contrary to the other methods, the representer SSH rms error never exceeds 2 standard deviations of the data error. For the zonal velocity the accuracy is achieved at location 1 networks 5–8, location 2 for all networks, location 3 for networks 7 and 8, and location 4 for networks 5–8. Elsewhere, the solution accuracy decreases with an rms that sometimes exceeds 3 standard deviations, e.g. location 3 for networks 3-4, location 5 for network 3. In the meridional velocity, accuracy is achieved only at locations 1 and 2 for all networks. For locations 3–5, the rms rapidly increases from below 2 standard deviations (networks 7 and 8) to more than 6, as the data density decreases to network 3.

The challenging stations for the v-component are located around the eddy shedding area, where stronger advective nonlinearities occur. This may indicate inadequacies in the model error covariance functions: the

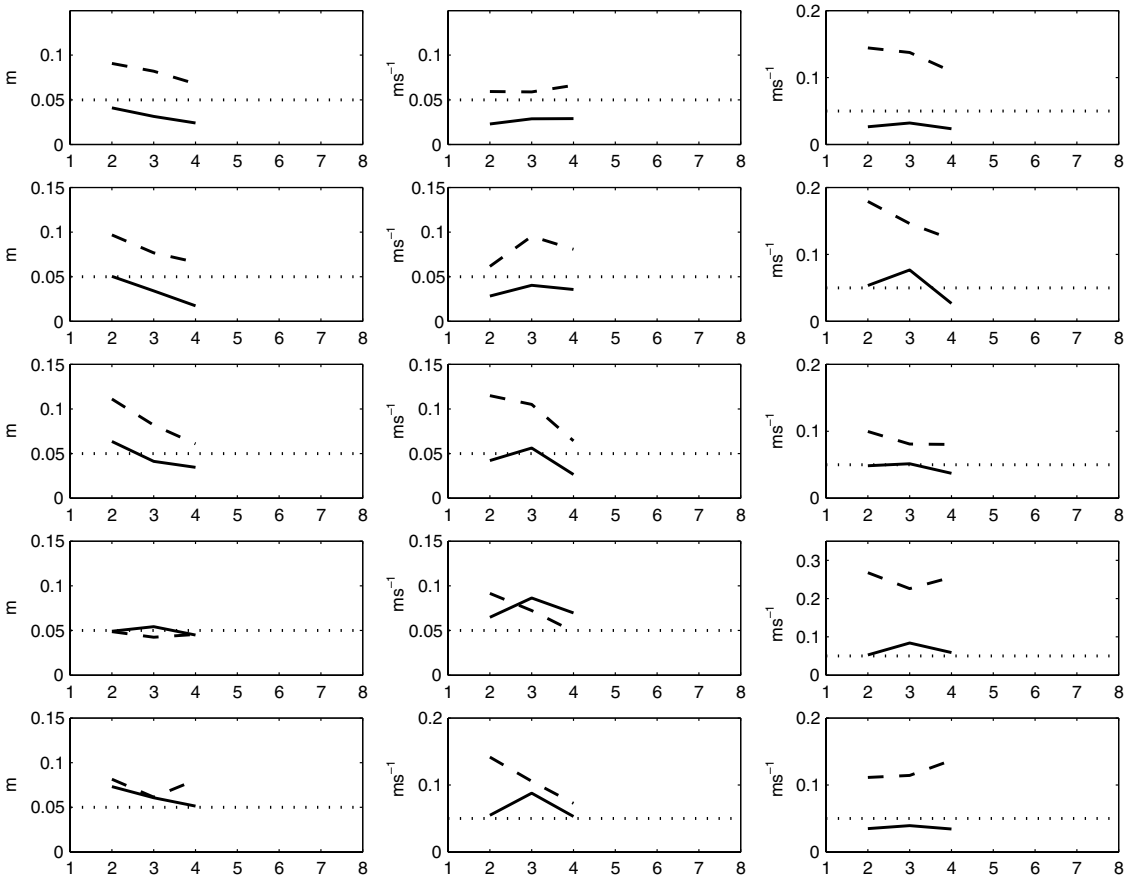


Fig. 7. Comparison of the representer and EnKS for the 6-month assimilation experiment. The rms is computed over the entire assimilation window.

covariance prescribed for the representer lacks cross-correlation between variables, while this cross-correlation is inherently generated in the covariance computed from the ensemble. On the other hand, the covariance in the variational method is not designed to represent the nonlinear interaction of the model errors.

The EnKF solution SSH is accurate only at location 4 for networks 5–8, having an rms error less or equal to the observation standard deviation. At all other locations, the SSH rms is always above a standard deviation. As the data density decreases, the rms gradually increases to more than 4 standard deviations, e.g. network 3 at locations 1 and 2. For the zonal velocity, accuracy is achieved at location 2 for all networks, location 3 for networks 6–8, and location 4 for networks 5–6. Elsewhere the solution rms error is greater than the observation standard deviation. The rms for the meridional velocity is almost always above the observation standard deviation, except at locations 4–5 for networks 5–6. In general, except for a few locations and variables, the EnKF rms error is always greater than that of the representer.

As we have observed in previous experiments, the EnKS rms error is significantly lower than the EnKF. There are several locations and networks where the EnKS rms is lower than the observation standard deviation whereas the EnKF rms is well above. Also, there are locations (and networks) where the EnKS rms is lower than the representer rms, especially for the velocity components.

The rms errors computed over the 4-month assimilation window of Experiment 3 are compared for representer and the EnKS. Results in Fig. 9 show that the rms error slightly increases for the representer method, whereas the increase is quite significant for the EnKS, particularly for the velocity. This increase in the rms indicates that, contrary to the previous experiments, both methods are fitting the data with a lower accuracy in the first half of the assimilation window than in the second. The EnKS rms error is lower or equal to the

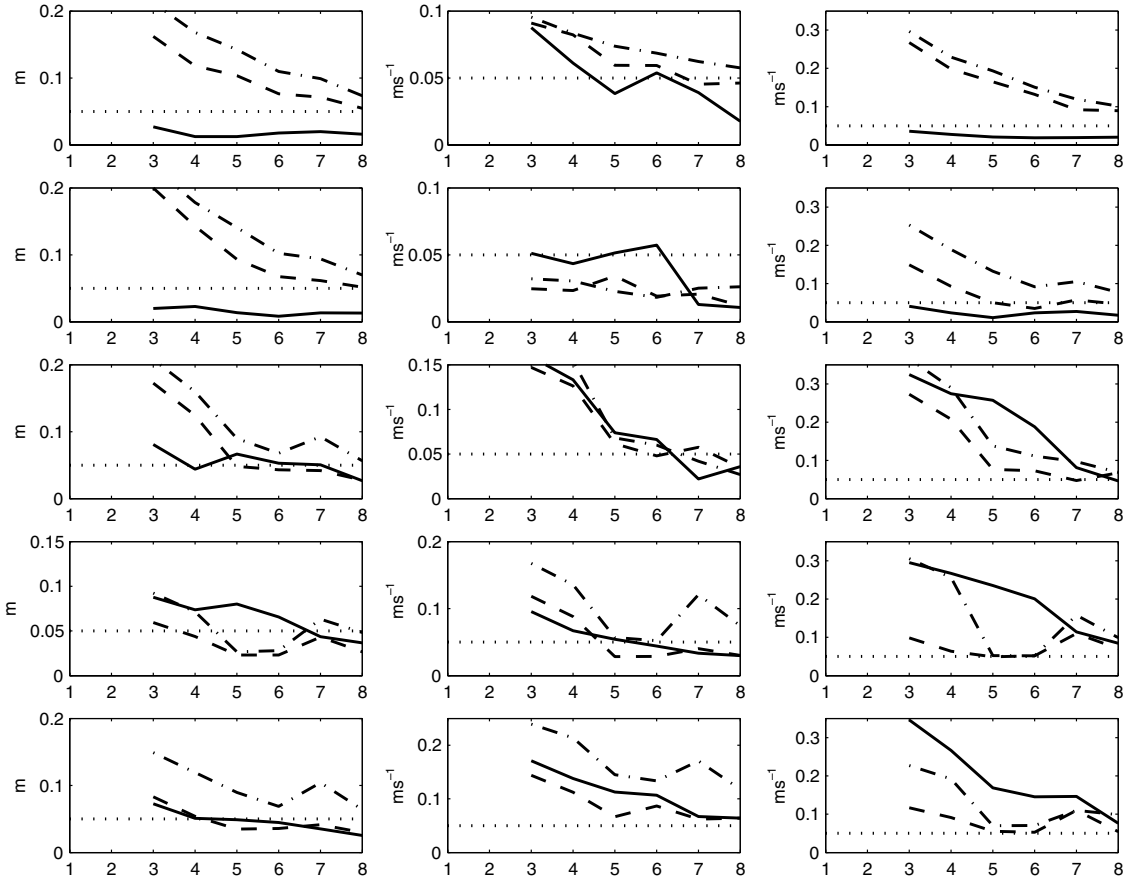


Fig. 8. Same as in Fig. 4, except only SSH measurements are assimilated.

observation standard deviation only for SSH at location 4 (all networks), location 5 (networks 5–8), for the zonal velocity at location 1 (networks 5 and 7), location 2 (network 6) and location 3 (network 8). The rms error for the meridional velocity is always greater than the standard deviation. The representer method on the other hand is more accurate. The SSH rms error is less than the standard deviation at all locations and networks 4–8, except at location 4. Also for both the velocity components, the method is accurate at locations 1–3 for networks 7–8.

### 4.5. Forecast accuracy

The comparison experiments above were all carried out for the hindcast problem. Since these methods are potential candidates for analysis-forecast cycling systems, they must be compared also on performance in the forecast after the analysis. It should be noted that theoretically, the EnKF and the EnKS produce the same solution at the final analysis (Evensen and van Leeuwen, 2000), and thus the same forecast. This is confirmed in our experiments. Therefore, we will compare the forecast from the representer and the EnKS solutions. The forecast is computed as the nonlinear model integrated forward from the analysis at the final time of the assimilation. It is carried out for 4 months, and the comparison is based on the rms error between the reference forecast and the forecast from the assimilation experiments. These forecast rms errors are computed for Experiment 1 (all data assimilated at each network) and Experiment 3 (only SSH data assimilated).

The first forecast comparison results in Fig. 10 show an almost identical scenario across the diagnostic locations and variables: the representer rms error is lower than that of the EnKS/F for the dense networks. As the data density decreases, the representer rms increases almost uniformly and becomes greater than the EnKS/F

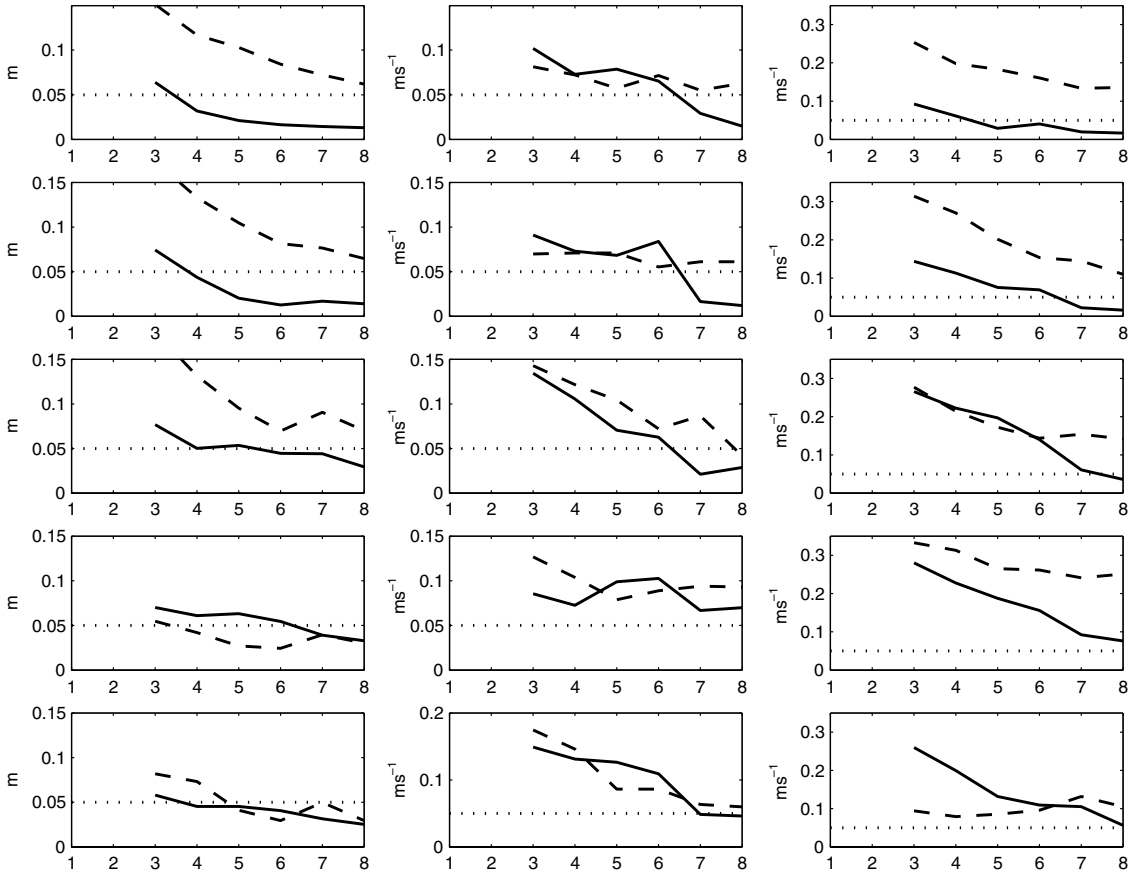


Fig. 9. Same as in Fig. 5, except only SSH measurements are assimilated.

rms. The forecast rms error depends solely on the accuracy attained in the estimation of the final state of the solution. Thus, Fig. 10 illustrates that the representer estimate of the final model state is better than that of the EnKS/F for the denser networks, and the converse is true for the sparser networks.

This result seems to contradict the better accuracy achieved by the representer method in the hindcast problem. One should remember that the hindcast rms depends on the accuracy of the assimilated solution at all the analysis times, whereas the forecast rms depends only on the accuracy of the assimilated solution at the final analysis. Also, we have seen in the hindcast experiments that the representer method is more accurate in the first half of the assimilation window than in the second. Therefore, the lesser accuracy in the estimate of the final condition (for the sparse networks) is not a contradiction or surprise.

The performance of the representer forecast rms degrades when only SSH data are assimilated (Experiment 3) as shown in Fig. 11. The rms error grows rapidly with decreasing data density, except for a few locations where it appears steady. The EnKS/F forecast rms, on the other hand, is mostly level and substantially lower than that of the representer, with a very few exceptions.

The EnKF better accuracy at the final time can be explained by its representation of the covariance matrix: the computation of the covariance matrix uses the full nonlinear model, and the analysis may be improving at each observation time, i.e. getting closer to the data. In the EnKF, the covariance matrix $P^{\mathrm{f}}$ at any time is approximated by ensemble integrations using the full nonlinear dynamics. The counterpart of $P^{\mathrm{f}}$ in the representer method is $LCL^{\mathrm{T}}$, where $L$ is the TLM, $L^{\mathrm{T}}$ is the adjoint and $C$ is the prescribed model error covariance. It is clear that unlike $P^{\mathrm{f}}$, $LCL^{\mathrm{T}}$ misses the nonlinear interactions of the dynamics in its representation of the covariance (this is especially true in the presence on strong nonlinearities as during the eddy shedding), and $C$ may not represent the nonlinear dynamical evolution of the model errors. These covariance flaws in the

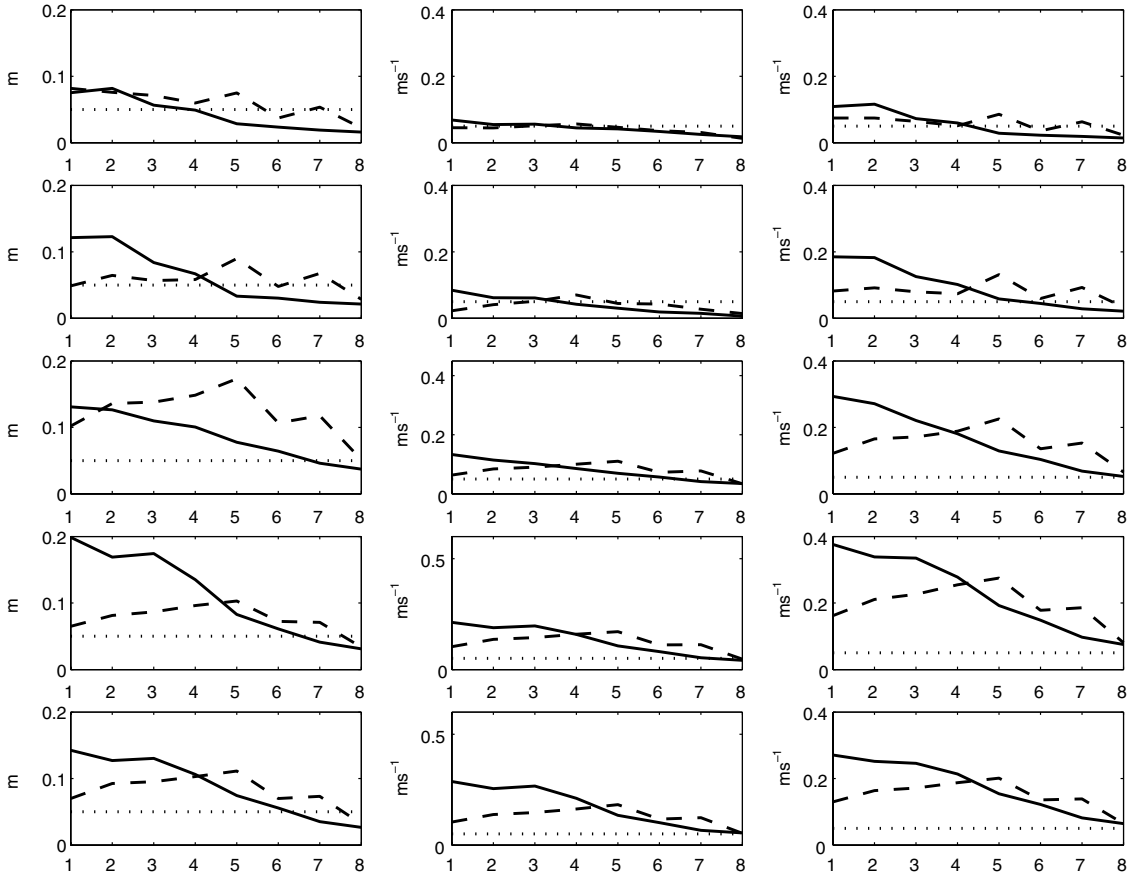


Fig. 10. The forecast rms error for the representer (solid line) and the EnKS/EnKF (dashed line), with all data assimilated. The dotted line is the measurement noise level, and the $x$-axis is the network number.

representer method may be compensated either by the amount of data for dense observation networks, or, in the case of sparse networks, by the contribution of past and future observation for observation times well within the assimilation window.

As mentioned above, the dynamics represented by the data are characterized by a strong nonlinear event in the last month of the assimilation window, namely an eddy shedding. In the first three months of the data, the dynamics are characterized by a westward propagation of an LCE while the LC is intruding into the domain. In the wake of the fourth month, the LC further intrudes and a new LCE starts shedding, a result of stronger advective nonlinearities. The representer method is expected to be less accurate in fitting the data in this last segment because the TLM is no longer as accurate as in the first 3 months, and the covariance functions have become inadequate: they are missing the nonlinear interactions and cross-correlation of the model errors. The ensemble-based methods on the other hand generate a covariance matrix based on these nonlinearities, and are therefore more able to fit the data. To elicit the role of these strong nonlinearities and inaccuracy of the TLM toward the end of the assimilation window, we made a 5-month forecast from the representer solution 1 month (30 days) before the end of the assimilation, and a 6-month forecast from the representer solution 2 months (60 days) before the final time. These forecasts are compared to the EnKF/EnKS forecast and the representer forecast from the end time for 4 months, for network 3.

The results in Fig. 12 show that the representer forecasts 30 and 60 days prior to the end time are more accurate than that from the end time and the one from EnKF. Thus the strong nonlinear event during the last month has a big impact on the representer solution. However, this is not to say that the representer method is unable to assimilate data accurately in the presence of strong nonlinearities. The assimilation experiment

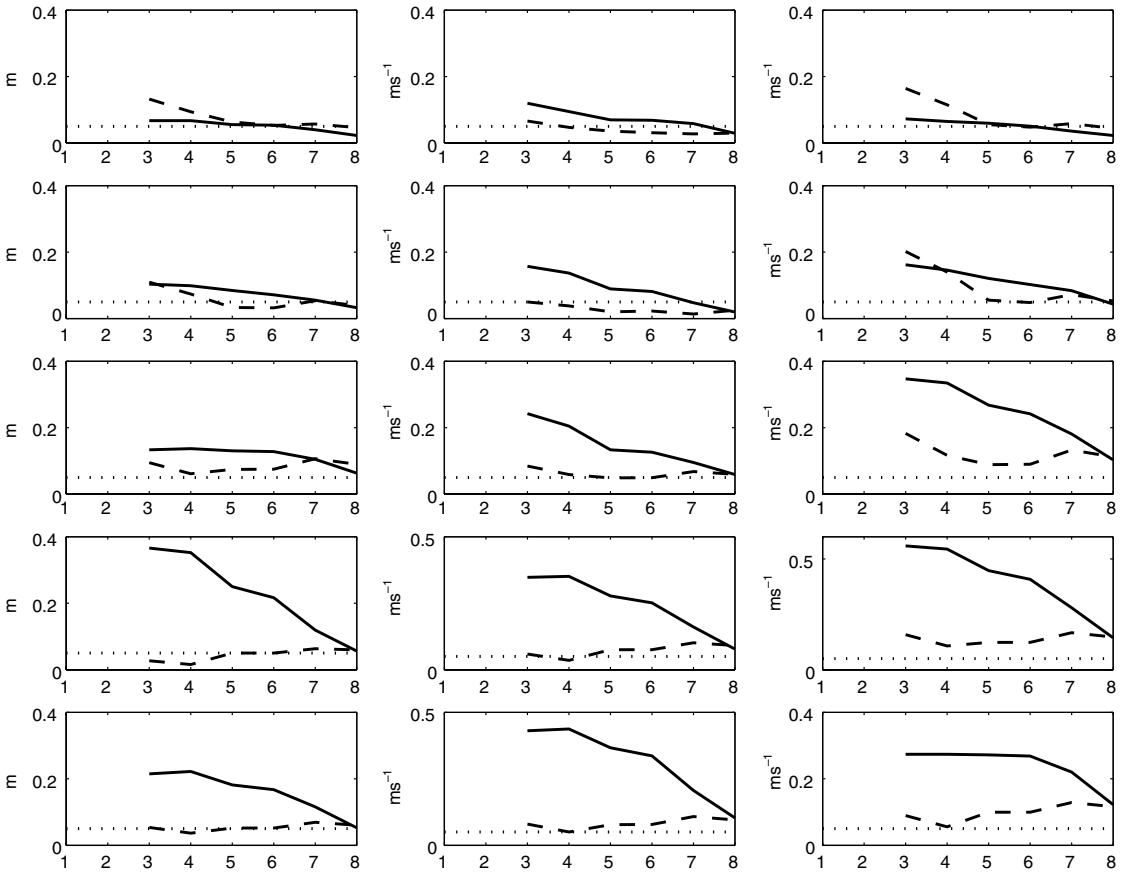


Fig. 11. Same as in Fig. 10, except only SSH data are assimilated.

using a 6-month time window (see Fig. 6) fitted the data comparably well during this strong nonlinear event. The difficulty therefore may also be arising from a poor representation of the time correlation within the prescribed covariance functions, in the sense that the contribution of past and future measurements is limited. This combines with strong nonlinearities to render the method inaccurate toward the end of the smoothing interval.

## 5. What is the cost?

In this section, the cost of each method is evaluated in terms of the number of integrations of the forward model. Given a first guess field, a rough estimate of the cost to implement the iterated indirect representer method is $L(F+A)(I+1)$ where $L$ is the number of iterations over the linearizations of the model. A (linear) iteration computes a new background that will be used for the tangent linearization in the next iteration and so on. The successive iterations are intended to make smaller corrections and thus allow the TLM to be a more accurate approximation. $F$ is the cost of the forward model, $A$ is the cost of the adjoint, $I$ is the number of iterations it takes to invert the representer matrix (added to the data error covariance matrix) on the innovation vector through a conjugate gradient method, and one additional backward and a forward integration completes the analysis by computing the best estimate as demonstrated by Amodei (1995), Egbert et al. (1994), Bennett et al. (1996) and Chua and Bennett (2001). The cost above does not include the covariance multiplications or the actual conjugate gradient matrix operations.

Assuming that the ensemble has been initialized, the cost of implementing the EnKF is $MF$ (where $M$ is the number of samples), plus the cost of inverting the representer matrix (added to the data error covariance

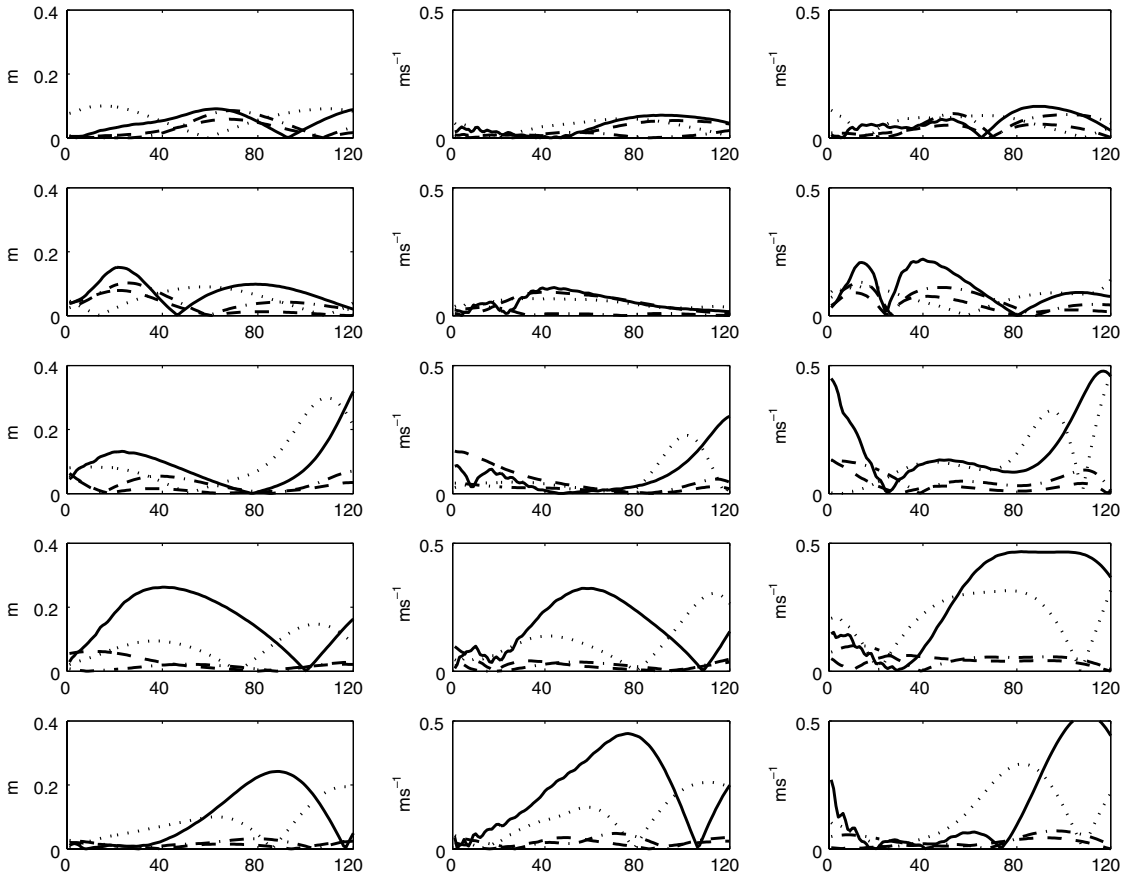


Fig. 12. Time series of the absolute difference between the forecast and the reference solutions for the representer solution from the final time (solid line), 30 days (dashed line) and 60 days (dashed-dotted lines) before the final time, and the EnKF/EnKS from the final time (dotted line) at all diagnostic stations. The $x$-axis is the number of days.

matrix) on the innovation vector. Other matrix operations should be added to this cost, such as the computation of the forecast error covariance matrix and the analyzed covariance matrix. The largest unknown in the EnKF is $M$, the number of ensemble members required to provide an accurate solution.

The EnKS uses the same number of model integrations as the EnKF, but the analysis step is more expensive (in core memory and CPU time) than in the EnKF, as a larger forecast error covariance matrix is involved in exactly the same way. Remember that in the EnKS the forecast error covariance is computed using the ensemble trajectory.

A typical representer experiment for the system used in this study converges in about three outer iterations ($L = 3$), each of which required about 50 iterations of the conjugate gradient for the dense networks (about 30 or less for the sparse). It should be emphasized that these numbers of conjugate gradient iterations are excellent, given the sizes of the datasets involved. The adjoint model at hand is about 1.5 times more computationally expensive than the forward model. Even by neglecting the computational cost of the covariance multiplications, the cost of the representer is about 382 times the cost of the forward model for dense networks (5–8) and 232 times the cost of the forward for the sparse (1–4). The EnKF and EnKS results reported here used an ensemble of 100 members, which costs 100 forward model integrations.

## 6. Conclusions

We have conducted a few assimilation experiments in an effort to compare the representer method, the EnKF and EnKS. In Experiment 1, the EnKF solution is generally less accurate than the other two solutions

during the analysis time period. The EnKS significantly improves accuracy over the EnKF, but the representer is still the more accurate during the analysis interval. We also conducted a smoother-to-smoother (i.e. representer–EnKS) comparison over the entire assimilation window for the same experiment. Here also, the representer is generally more accurate than the EnKS for the denser networks, and retains its accuracy longer with the sparser networks. Due to the wrong initial condition, the sequential methods need some adjustment time (20–60 days for the EnKF, and much less for the EnKS) to start matching the data. For this reason, we conducted a 6-month assimilation experiment and obtained much lower rms for both the EnKF and EnKS over the second half of the assimilation window. The representer rms for that experiment is still lower than those of the sequential methods. Another assimilation experiment where only SSH data are assimilated is carried out, with rms results almost following the pattern of the previous experiments.

The difference between the representer rms and the EnKS rms is larger in the 4-month comparison than in the 2-month, due to two factors: (1) the representer is more accurate in the first half of the assimilation window and tends to be less accurate in the second half, and (2) the EnKS is less accurate in the first half and becomes more accurate in the second half of the assimilation window. This behavior is true for all the experiments conducted and is related to the dynamics that are represented by the data and the algorithmic design. In the first three months of the data, the dynamics are characterized by a westward propagation of an LCE while the LC is intruding into the domain. In the wake of the fourth month, the LC further intrudes and a new LCE starts shedding, a result of stronger advective nonlinearities. The representer method is expected to be less accurate in fitting the data in this last segment because the TLM is no longer as accurate as in the first three months. Only an abundance of data or more outer iterations (which exacerbates the cost) will prevent the method from losing accuracy. The ensemble-based methods on the other hand generate an error covariance matrix that includes the effects of these nonlinearities and are therefore more able to fit the data.

The decreasing accuracy of the representer towards the end of the assimilation window negatively affects the associated forecast. On the other hand, the increasing accuracy of the EnKS (and EnKF) towards the end of the assimilation window positively affects the forecast. The representer forecast rms increases rapidly with decreasing data density, revealing inaccuracies in the final condition estimate and the subsequent forecast. This raises the question of the suitability or the potential implementation of the proper method for cycles of analysis and forecasts. The results here indicate that a variational method provides a more accurate analysis solution. For providing forecasts, the sequential methods are indicated to provide improved accuracy. The results seem to be dependent on the TLM accuracy within the assimilation time interval, and for the case here, a strong nonlinear event dominates toward the end of the assimilation interval. It may be that with a shorter assimilation interval the TLM inaccuracies are not a factor, and this should be the subject of future research. In general, and particularly for a small number of data, the representer is more accurate in the analysis time interval, but the sequential methods are more accurate in the assimilation-forecast cycles if there is sufficient time to adjust and fit the data. That is, if the initial condition errors decay. Finally, the sequential methods are more straightforward in implementation (they do not require the adjoint), and are less costly. The comparison exercise of this paper used a simplified dynamical model. The concept of this study should be applied to a much more sophisticated model, if the computational resources that such a model would require are available, before a decision of any kind is made.

A new implementation of the EnKF and EnKS was recently proposed by Evensen (2004), for improved ensemble generation and analysis scheme. This implementation was not used in this study. We are aware that the new version of the EnKF and EnKS would improve their accuracy while reducing the cost, according Evensen (2004). Although it is not possible to quantify how the performance of these methods will improve in comparison to the representer at this time, we believe that their implementation in the present framework would not adversely change the conclusions of this study.

## Acknowledgments

This work was sponsored by the Office of Naval Research (Program Element number 0601153N) as part of the project 'Shelf to Slope Energetics and Exchange Dynamics'. This paper is NRL paper contribution number JA/7320-05-5112. We are grateful to Geir Evensen for providing the EnKF/EnKS computer code and for

remarks that helped improve the manuscript. We also thank the anonymous reviewers for their constructive comments, and Scott Smith for positive feedback.

## Appendix A. Implementation of the assimilation methods

For the sake of clarity, the dynamical model is written as

$$\mathbf{u}^{\mathrm{f}}_{k+1} = M(\mathbf{u}^{\mathrm{a}}_{k}) + \eta_{k+1}, \tag{A.1}$$

where $\mathbf{u}$ represents the state vector, $M$ is the nonlinear dynamical model operator, $\eta$ is the model error. The superscripts f and a denote the forecast and analysis respectively, and the subscript $k$ the time index. If a set of observations is available at time level $k+1$, i.e.

$$\mathbf{y}_{k+1} = \mathbf{H}\mathbf{u}_{k+1} + \varepsilon_{k+1}, \tag{A.2}$$

where $\mathbf{H}$ is the observation operator, $\mathbf{y}$ is the observation vector and $\varepsilon$ is vector of measurement errors, whose covariance matrix $\mathbf{R}$, the EnKF analysis at time $k+1$ is given by the formula:

$$\mathbf{u}^{\mathrm{a}}_{k+1} = \mathbf{u}^{\mathrm{f}}_{k+1} + \mathbf{P}^{\mathrm{f}}_{k+1}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}^{\mathrm{f}}_{k+1}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}(\mathbf{y}_{k+1} - \mathbf{H}\mathbf{u}^{\mathrm{f}}_{k+1}), \tag{A.3}$$

where the forecast model error covariance matrix $\mathbf{P}^{\mathrm{f}}_{k+1}$ is ensemble approximated by

$$\mathbf{P}^{\mathrm{f}}_{k+1} = \frac{1}{n_{\mathrm{ens}} - 1}(\mathbf{A}^{\mathrm{f}} - \overline{\mathbf{A}}^{\mathrm{f}})(\mathbf{A}^{\mathrm{f}} - \overline{\mathbf{A}}^{\mathrm{f}})^{\mathrm{T}}. \tag{A.4}$$

Each forecasted model state has been stored in one column of a big matrix $\mathbf{A}^{\mathrm{f}}$, and the mean of the ensemble is stored in $\overline{\mathbf{A}}^{\mathrm{f}}$, i.e. each column of $\overline{\mathbf{A}}^{\mathrm{f}}$ contains the ensemble mean. In our implementation, the analysis equation (A.3) is performed for the entire ensemble $\mathbf{A}^{\mathrm{f}}$, and the analysis is taken as the mean of the analyzed ensemble (Evensen, 2003). The implementation of the EnKS uses the same update equation applying to the ensemble of trajectories. So each column of $\mathbf{A}^{\mathrm{f}}$ in the EnKS contains an ensemble member at all previous analysis times and the forecast of that same member at the current analysis time.

For the variational method, the assimilation problem consists of finding the minimum of the penalty function

$$J = \int\int\int\int \eta(\mathbf{x},t)Q^{-1}(\mathbf{x},t,\mathbf{x}',t')\eta(\mathbf{x}',t')\,\mathrm{d}\mathbf{x}\,\mathrm{d}t\,\mathrm{d}\mathbf{x}'\,\mathrm{d}t' + \boldsymbol{\varepsilon}^{\mathrm{T}}\mathbf{R}^{-1}\boldsymbol{\varepsilon} \tag{A.5}$$

by solving the Euler–Lagrange (EL) system associated to the vanishing of the first variation or gradient of the penalty function with respect to the controls. The contribution of initial and boundary condition errors has been omitted from (A.5) for the sake of simplicity. The EL system is

$$\boldsymbol{\lambda}_k = M^*\boldsymbol{\lambda}_{k+1} + [\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\hat{\mathbf{u}})]_k \tag{A.6}$$

$$\hat{\mathbf{u}}_{k+1} = M(\hat{\mathbf{u}}_k) + (Q\cdot\boldsymbol{\lambda})_{k+1} \tag{A.7}$$

where $\boldsymbol{\lambda}(\mathbf{x},t) = \iint Q^{-1}(\mathbf{x},t,\mathbf{x}',t')\eta(\mathbf{x}',t')\,\mathrm{d}\mathbf{x}'\,\mathrm{d}t'$ is the weighted residual, $M^*$ is the adjoint model, and $Q\cdot\boldsymbol{\lambda}(\mathbf{x},t) = \iint Q(\mathbf{x},t,\mathbf{x}',t')\lambda(\mathbf{x}',t')\,\mathrm{d}\mathbf{x}'\,\mathrm{d}t'$.

The representer expansion

$$\hat{u}(\mathbf{x},t) = u^{\mathrm{b}}(\mathbf{x},t) + \sum_{m=1}^{K}\beta_m r^m(\mathbf{x},t) \tag{A.8}$$

allows to uncouple the (A.6) and (A.7), where $u^{\mathrm{b}}$ is the first guess solution (it also serves as the background solution for linearization purposes), $r^k(\mathbf{x},t)$, $m = 1,\ldots,K$ are the representer functions and the expansion (A.8) requires that the dynamical model be linear. In applications using nonlinear dynamics, the model is linearized in (A.7) and one can invoke (A.8). In that case the representer functions are computed by

$$\mathbf{r}^m_{k+1} = M\mathbf{r}^m_k + (Q\cdot\alpha_m)_{k+1} \tag{A.9}$$

$$\alpha^m_k = M^*\alpha^m_{k+1} + \mathbf{H}^{\mathrm{T}}\delta(\mathbf{x} - \mathbf{x}_m)\delta(t - t_m) \tag{A.10}$$

The representer coefficients $\beta_m$ are obtained by solving the linear system:

$$(\mathbf{R}_e + \mathbf{R})\boldsymbol{\beta} = \mathbf{y} - \mathbf{H}\mathbf{u}^{\mathrm{b}}. \tag{A.11}$$

The representer matrix $\mathbf{R}_e$ is obtained by evaluating the representer functions $\mathbf{r}^m(\mathbf{x}, t)$ at the measurement sites. In practice, solving (A.11) does not require the computation of the entire representer matrix. An iterative method such as the conjugate gradient may be invoked, as long as the matrix–vector product on the left-hand side of (A.11) can be computed for any vector. This is made possible through the indirect representer algorithm, Amodei (1995) and Egbert et al. (1994). Thus our implementation is the iterated indirect representer algorithm, Chua and Bennett (2001), in which the best estimate at the current iteration becomes the first guess and background for the next iteration.

## References


Amodei, L., 1995. Solution approchée pour un problème d'assimilation de données avec prise en compte de l'erreur du modèle. Comptes Rendus de l'Académie des Sciences 321 (Série IIa), 1087–1094.

Bennett, A.F., 1992. Inverse Methods in Physical Oceanography. Cambridge University Press, New York, 347pp.

Bennett, A.F., 2002. Inverse Modeling of the Ocean and Atmosphere. Cambridge University Press.

Bennett, A.F., Chua, B.S., Leslie, L.M., 1996. Generalized inversion of a global numerical weather prediction model. Meteor. Atmos. Phys. 60, 165–178.

Brusdal, K., Brankart, J.M., Halberstadt, G., Evensen, G., Brasseur, P., van Leeuwen, P.J., Dombrowsky, E., Verron, J., 2003. A demonstration of ensemble-based assimilation methods with a layered OGCM from the perspective of operational ocean forecasting systems. J. Mar. Syst. 40–41, 253–289.

Cane, M., Kaplan, A., Miller, R.N., Tang, B., Hackert, E., Busalacchi, A., 1996. Mapping tropical Pacific sea level: data assimilation via e reduced state space Kalman filter. J. Geophys. Res. 101 (C10), 22,599–22,617.

Caya, A., Sun, J., Snyder, C., 2004. A comparison between the 4D-Var and the ensemble filter techniques for radar data assimilation. In: Proceedings of the 19th Conference on Weather Analysis and Forecasting, 11–15 January 2004, Seattle (Washington).

Chin, T.M., Haza, A.C., Mariano, A.J., 2002. A reduced-order information filter for multilayer shallow-water models: profiling and assimilation of sea surface height. J. Atmos. Oceanic. Technol. 19, 517–533.

Chua, B.S., Bennett, A.F., 2001. An inverse ocean modeling system. Ocean Modeling 3, 137–165.

Courtier, P., Thépaut, J.-N., Hollingsworth, A., 1994. A strategy for operational implementation of 4D-Var, using an incremental approach. Quart. J. Roy. Meteor. Soc. 120, 1367–1387.

Derber, J.C., 1987. Variational four-dimensional analysis using quasi-geostrophic constraints. Mon. Wea. Rev. 115, 998–1008.

Egbert, G.D., Bennett, A.F., Foreman, M.G.G., 1994. TOPEX/POSEIDON tides estimated using a global inverse method. J. Geophys. Res. 99, 24821–24852.

Evensen, G., 1992. Using the extended Kalman filter with a multilayer quasi-geostrophic ocean model. J. Geophys. Res. 97, 17,905–17,924.

Evensen, G., 1994. Sequential data assimilation with nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J. Geophys. Res. 99 (C5), 10143–10162.

Evensen, G., 2003. The ensemble Kalman filter: theoretical formulation and practical implementation. Ocean Dyn. 53, 367–443.

Evensen, G., 2004. Sampling strategies and square root analysis schemes for the EnKF. Ocean Dyn. 54, 539–560.

Evensen, G., van Leeuwen, P.J., 2000. An ensemble Kalman smoother for nonlinear dynamics. Mon. Wea. Rev. 128, 1852–1867.

Hurlburt, H.E., Thompson, J.D., 1980. A numerical study of the loop current intrusions and eddy shedding. J. Phys. Oceanogr. 10 (10), 1611–1651.

Jacobs, A.G., Ngodock, H.E., 2003. The maintenance of conservative physical laws within data assimilation systems. Mon. Wea. Rev. 131 (11), 2595–2607.

Kalman, R., Bucy, R., 1961. New results in linear prediction and filtering theory. Trans. AMSE J. Basic Eng. 83D, 95–108.

Kurapov, A.L., Egbert, G.D., Miller, R.N., Allen, J.S., 2002. Data assimilation in a baroclinic coastal ocean model: ensemble statistics and comparison of methods. Mon. Wea. Rev. 130, 1009–1025.

Le Dimet, F., Talagrand, O., 1986. Variational algorithm for analysis and assimilation of meteorological observations: theoretical aspects. Tellus 38A, 97–110.

Lermusiaux, P.F.J., Robinson, A.R., 1999. Data assimilation via error subspace statistical estimation. Part I. Theory and schemes. Mon. Wea. Rev. 127, 1385–1407.

Lewis, J.M., Derber, J.C., 1985. The use of adjoint equations to solve a variational adjustment problem with advective constraints. Tellus 37A, 309–322.

Mesinger, F., Arakawa, A., 1976. Numerical methods used in atmospheric models. GARP Publication Series No. 14, WMO/ICSU Joint Organizing Committee, 64 pp.

Miller, R.N., Ghil, M., Gauthiez, F., 1994. Advanced data assimilation in strongly nonlinear systems. J. Atmos. Sci. 51 (8), 1037–1056.

Miller, R.N., Carter Jr., E.F., Blue, S.T., 1999. Data assimilation into nonlinear stochastic models. Tellus 51A, 167–194.

Muccino, J.C., Bennett, A.F., 2002. Generalized inversion of the Korteweg–de Vries equation. Dyn. Atmos. Oceans 35 (3), 227–263.

Ngodock, H.E., Chua, B.S., Bennett, A.F., 2000. Generalized inversion of a reduced gravity primitive equation ocean model and tropical atmosphere ocean data. Mon. Wea. Rev. 128, 1757–1777.

Pham, D.T., Verron, J., Roubaud, M.C., 1997. Singular evolutive Kalman filter with EOF initialization for data assimilation in oceanography. J. Mar. Syst. 16, 323–340.

Pham, D.T., Verron, J., Gourdeau, L., 1998. Singular evolutive Kalman filter for data assimilation in oceanography. C. R. Acad. Sci. Paris 326, 255–260.

Reichle, R.H., McLaughlin, D.B., Entekhabi, D., 2002. Hydrologic data assimilation with the ensemble Kalman filter. Mon. Wea. Rev. 130, 103–114.

Sasaki, Y., 1958. An objective analysis based on the variational method. J. Meteor. Soc. Jpn. 36, 77–88.

Sasaki, Y., 1970. Some basic formalisms in numerical variational analysis. Mon. Wea. Rev. 98, 875–883.

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

Willems, R.C., Glenn, S.M., Crowley, M.F., Malanotte-Rizzoli, P., Young, R.E., Ezer, T., Mellor, G.L., Arango, H.G., Robinson, A.R., Lai, C.-C.A., 1994. Experiment evaluates ocean models and data assimilation in the Gulf Stream. EOS 75, 385, 391, 394.