



## Application and comparison of Kalman filters for coastal ocean problems: An experiment with FVCOM

Changsheng Chen,<sup>1,2</sup> Paola Malanotte-Rizzoli,<sup>3</sup> Jun Wei,<sup>3</sup> Robert C. Beardsley,<sup>4</sup> Zhigang Lai,<sup>1</sup> Pengfei Xue,<sup>1</sup> Sangjun Lyu,<sup>3</sup> Qichun Xu,<sup>1</sup> Jianhua Qi,<sup>1</sup> and Geoffrey W. Cowles<sup>1</sup>

Received 9 September 2007; revised 19 October 2008; accepted 16 February 2009; published 13 May 2009.

[1] Twin experiments were made to compare the reduced rank Kalman filter (RRKF), ensemble Kalman filter (EnKF), and ensemble square-root Kalman filter (EnSKF) for coastal ocean problems in three idealized regimes: a flat bottom circular shelf driven by tidal forcing at the open boundary; an linear slope continental shelf with river discharge; and a rectangular estuary with tidal flushing intertidal zones and freshwater discharge. The hydrodynamics model used in this study is the unstructured grid Finite-Volume Coastal Ocean Model (FVCOM). Comparison results show that the success of the data assimilation method depends on sampling location, assimilation methods (univariate or multivariate covariance approaches), and the nature of the dynamical system. In general, for these applications, EnKF and EnSKF work better than RRKF, especially for time-dependent cases with large perturbations. In EnKF and EnSKF, multivariate covariance approaches should be used in assimilation to avoid the appearance of unrealistic numerical oscillations. Because the coastal ocean features multiscale dynamics in time and space, a case-by-case approach should be used to determine the most effective and most reliable data assimilation method for different dynamical systems.

**Citation:** Chen, C., P. Malanotte-Rizzoli, J. Wei, R. C. Beardsley, Z. Lai, P. Xue, S. Lyu, Q. Xu, J. Qi, and G. W. Cowles (2009), Application and comparison of Kalman filters for coastal ocean problems: An experiment with FVCOM, *J. Geophys. Res.*, 114, C05011, doi:10.1029/2007JC004548.

### 1. Introduction

[2] One of the primary goals of numerical circulation model development is to simulate and predict hydrographic and flow fields in the ocean. The basic idea of four-dimensional (4-D) data assimilation is to make a synthesis of existing data and model dynamics to provide a systematically better model simulation consistent with the observed fields that are normally noisy and incomplete in space and time. Data assimilation systems consist of three components: a set of observed data, a dynamical model, and an assimilation scheme. Since the data have errors and models are imperfect, a well-constructed assimilation scheme should provide a better match between data and model within the bounds of observational and modeling errors [Ghil, 1989]. Therefore, once a dynamical model is fully developed, the appropriate assimilation scheme

becomes more critical in ensuring successful model applications.

[3] The most widely used data assimilation techniques are nudging, optimal interpolation (OI), adjoint-based methods, and Kalman filters (KFs). Nudging and OI are the simplest and least computationally intensive data assimilation methods. They work efficiently in applications with full data coverage of the model domain, but can lead to unrealistic density gradients and current shears at the data boundary when the data coverage is limited to a discrete portion of the model domain [Lorenc, 1981; Chen *et al.*, 2007, 2008]. The adjoint-based methods are based on control theory in which a cost function defined by the difference between model-derived and measured quantities is minimized in a least squares sense under the constraints of the model equations [Le Dimet and Talagrand, 1986; Thacker and Long, 1988; Tziperman and Thacker, 1989; Bergamasco *et al.*, 1993; Morrow and De Mey, 1995]. This method has been widely used in both linear and nonlinear systems for the estimation of model parameters, initial and boundary conditions, and surface forcing such as wind stress and heat/moisture fluxes. KFs are sequential data assimilation methods based on estimation theory [Kalman, 1960]. Since they provide prediction of both analysis and forecast error covariances, they can be used for nowcasting and forecasting with global- to basin-scale ocean and atmospheric models [Evensen, 1992,

<sup>1</sup>School for Marine Science and Technology, University of Massachusetts–Dartmouth, New Bedford, Massachusetts, USA.

<sup>2</sup>Also at Marine Ecosystem and Environmental Laboratory, Shanghai Ocean University, Shanghai, China.

<sup>3</sup>Department of Atmospheric, Earth and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.

<sup>4</sup>Department of Physical Oceanography, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA.

1993, 2003; *Blanchet et al.*, 1997; *Ghil and Malanotte-Rizzoli*, 1991].

[4] Malanotte-Rizzoli and collaborators have constructed a package of KFs, including a reduced rank Kalman filter (RRKF) [*Buehner and Malanotte-Rizzoli*, 2003; *Buehner et al.*, 2003], an ensemble Kalman filter (EnKF) [*Evensen*, 1994; *Zang and Malanotte-Rizzoli*, 2003], a deterministic and stochastic ensemble square-root Kalman filter (EnSKF), and an ensemble transform Kalman filter (EnTKF) [*Bishop et al.*, 2001]. This package was first applied to idealized cases as proof-of-concept tests and subsequently has been successively incorporated into fully realistic, primitive equation basin-scale ocean models. In a linear dynamical system, when the temporal variability of model forecast errors are characterized by the dominant modes, the error evolution can be resolved in a subspace defined by a few leading empirical orthogonal functions (EOFs). This RRKF based on EOF theory is not computationally intensive, and can effectively reduce model errors with better performance than OI [*Buehner and Malanotte-Rizzoli*, 2003]. *Evensen* [1994] first suggested that the error covariance relative to the mean of ensemble members could provide a better estimation of the error covariance defined in the classical Kalman filter. The EnKF is traditionally constructed by running a forecast model driven by an ensemble of initial conditions with random perturbations generated using a Monte Carlo approach and then the error covariance relative to the ensemble mean is estimated to determine the ensemble analysis values for the next forecast [*Evensen*, 1994]. Recently, the singular evolutive extended Kalman (SEIK) [*Pham et al.*, 1998a] and singular evolutive interpolated Kalman (SEIK) [*Pham et al.*, 1998b] filters were developed, in which the ensemble members are specified initially using either dominant eigenvectors (or EOFs) or reinitialized periodically in suberror space to improve the computational efficiency by optimally constructing the error covariance using lower ensemble numbers [*Nerger et al.*, 2005]. The EnSKF and EnTKF are derivatives of EnKF that are applied using deterministic observations [*Bishop et al.*, 2001; *Whitaker and Hamill*, 2002]. The EnTKF was introduced by *Bishop et al.* [2001] for optimally deploying adaptive observations. This filter was used by S. J. Lyu et al. (A comparison of data assimilation results from the deterministic and stochastic ensemble Kalman filters, unpublished manuscript, 2009; Optimal fixed and adaptive observation arrays in an idealized wind-driven ocean model, unpublished manuscript, 2009) to design optimal fixed and adaptive observational arrays in an idealized model of the wind-driven circulation in a double gyre ocean.

[5] Recently, a multi-institutional (UMASSD-MIT-WHOI) effort was made to implement KF methods into the unstructured grid Finite-Volume Coastal Ocean Model (FVCOM) for coastal and estuarine applications [*Chen et al.*, 2006a]. A series of proof-of-concept tests were conducted to compare the abilities of various KFs to restore a dynamical coastal or estuarine system after they had been subjected to random perturbations. Experiments were made to examine the sensitivity of the convergence rate to the sampling location and assimilation configuration (univariate or a multivariate covariance approach related to the number of state variables in assimilation). Three idealized coastal and estuarine problems were selected: (1) tidal oscillations

in a flat bottom circular basin, (2) a low-salinity plume over an idealized continental shelf, and (3) tidal flushing in addition to freshwater discharge in an idealized rectangular estuary with intertidal zones. Unlike global- and basin-scale ocean model systems, KFs have not been widely used in coastal ocean and estuarine models. The three idealized cases presented in this study represent fundamental processes that occur widely in the coastal ocean, and results obtained from these experiments can provide useful guidance for future application of KFs to realistic coastal and estuarine forecast or hindcast systems.

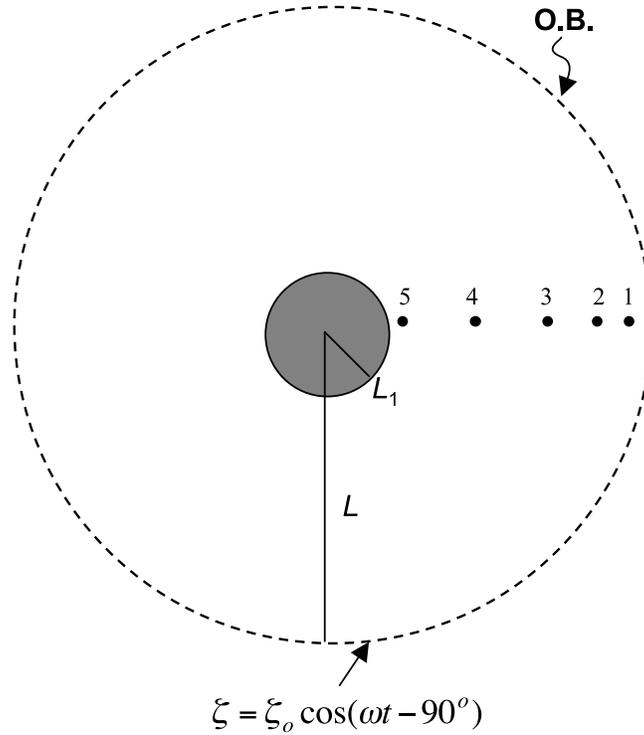
[6] This paper summarizes the validation experiment results with a focus on comparisons of RRKF and EnKF for the three idealized cases. Although the experiments were made using FVCOM, the results are applicable in general to any unstructured or structured grid ocean model. The paper is organized as follows: FVCOM and schematics of KF-FVCOM coupling are briefly described in section 2, the results for the three idealized cases are presented in sections 3, 4 and 5, and the discussion and conclusions are presented in section 6.

## 2. FVCOM With Implementation of KFs

[7] FVCOM is a state-of-the-art unstructured grid, finite volume, 3-D free surface primitive equation coastal ocean model developed originally by *Chen et al.* [2003] and modified and upgraded by a joint effort of UMASSD-WHOI scientists [*Chen et al.*, 2006a, 2006b]. The governing equations in this model are closed by default using the Mellor and Yamada level 2.5 turbulent closure scheme for vertical eddy viscosity [*Mellor and Yamada*, 1982] with the option to select from the suite of turbulence closure schemes available from the General Ocean Turbulence Model (GOTM) [*Burchard*, 2002]. A Smagorinsky scheme is employed for the parameterization of the horizontal eddy diffusivity [*Smagorinsky*, 1963]. FVCOM subdivides the computational domain into a set of nonoverlapping unstructured triangular meshes in the horizontal and discrete layers defined by the generalized terrain-following coordinate in the vertical [*Pietrzak et al.*, 2002]. It is solved numerically by flux calculation using an integral form of the discrete equations over the triangular grids. The flux calculation ensures both conservation of total mass over the entire computational domain and on individual stencils used to compute water properties. The finite volume numerical approach combines the advantages of finite element methods for geometric flexibility and finite difference methods for simple discrete code structure and computational efficiency. FVCOM is fully parallelized for efficient multiprocessor execution [*Cowles*, 2008].

[8] The KFs were implemented into FVCOM in an independent data assimilation module with optimal utilization of domain decomposition based parallelization. The RRKF and EnKF procedures in FVCOM follow the flowcharts illustrated in Figures S1 and S2 in the auxiliary material.<sup>1</sup> The RRKF is implemented on the basis of the stationary error covariance method and the EnKF is constructed following the Monte Carlo sampling of ensemble

<sup>1</sup>Auxiliary materials are available in the HTML. doi:10.1029/2007JC004548.



**Figure 1.** Idealized circular basin configuration. Dots indicate the observational sites where the currents and surface elevation are “measured.”

members. The model integration procedures of EnSKF and EnTKF in FVCOM are very similar to EnKF. The algorithms used in these filters are described by *Buehner and Malanotte-Rizzoli* [2003], *Evensen* [1994], *Houtekamer and Mitchell* [1998], *Bishop et al.* [2001], *Whitaker and Hamill* [2002], *Wang and Bishop* [2003], and *Zang and Malanotte-Rizzoli* [2003]. The mathematical equations and covariance operations used in RRKF, EnKF, EnSKF and EnTKF are briefly described in the auxiliary material.

### 3. Tidal Oscillation in a Flat Bottom Circular Basin

[9] Consider a two-dimensional, nonrotating tidal oscillation problem in a flat bottom circular basin with a circular island at the center (Figure 1). Let  $L_1$  and  $L$  be radii of the two concentric circles (island and basin) and  $H_o$  is the constant mean water depth. Assuming that the surface elevation and currents are uniform in the azimuthal direction, the inviscid, linear shallow water equations in polar coordinates are given as

$$\frac{\partial V_r}{\partial t} + g \frac{\partial \zeta}{\partial r} = 0; \frac{\partial \zeta}{\partial t} + \frac{\partial r V_r H_o}{r \partial r} = 0 \quad (1)$$

where  $r$  is the radial coordinate;  $V_r$  is the  $r$  component of velocity; and  $\zeta$  is the surface water elevation. We impose a wave solution in the form of  $\zeta = \text{Re}[\zeta_o(r) \cdot e^{-i(\omega t - 90^\circ)}]$  and  $V = \text{Re}[iV_o(r) \cdot e^{-i(\omega t - 90^\circ)}]$ , which is forced by periodic tidal forcing with amplitude  $\zeta_o|_{r=L} = A$  at the outer open boundary and satisfies the no-normal flux condition  $\frac{\partial \zeta_o}{\partial r}|_{r=L_1} = 0$  at the

boundary of the circular island. The analytical solutions of (1) for  $\zeta_o$  and  $V_o$  are

$$\zeta_o(r) = c_1 J_0\left(r \frac{\omega}{\sqrt{gH_o}}\right) + c_2 Y_0\left(r \frac{\omega}{\sqrt{gH_o}}\right) \quad (2)$$

$$V_o(r) = -\sqrt{\frac{g}{H_o}} \left[ c_1 J_0'\left(r \frac{\omega}{\sqrt{gH_o}}\right) + c_2 Y_0'\left(r \frac{\omega}{\sqrt{gH_o}}\right) \right] \quad (3)$$

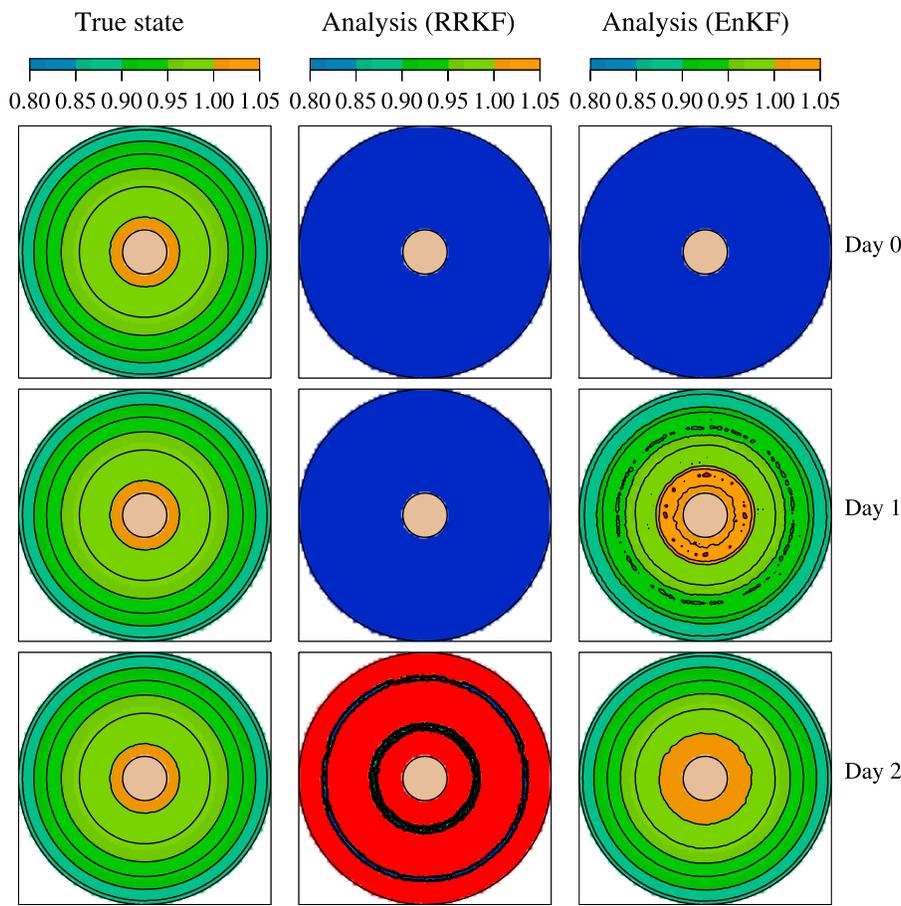
where

$$c_1 = A \cdot Y_0'\left(L_1 \frac{\omega}{\sqrt{gH_o}}\right) / \left[ J_0\left(L \frac{\omega}{\sqrt{gH_o}}\right) Y_0'\left(L_1 \frac{\omega}{\sqrt{gH_o}}\right) - J_0'\left(L_1 \frac{\omega}{\sqrt{gH_o}}\right) Y_0\left(L \frac{\omega}{\sqrt{gH_o}}\right) \right]$$

$$c_2 = -A \cdot J_0'\left(L_1 \frac{\omega}{\sqrt{gH_o}}\right) / \left[ J_0\left(L \frac{\omega}{\sqrt{gH_o}}\right) Y_0'\left(L_1 \frac{\omega}{\sqrt{gH_o}}\right) - J_0'\left(L_1 \frac{\omega}{\sqrt{gH_o}}\right) Y_0\left(L \frac{\omega}{\sqrt{gH_o}}\right) \right]$$

and  $J_0$  and  $Y_0$  are the zeroth-order Bessel functions of the first and second kind, respectively. The superscript “'” represents the first derivative with respect to  $r$ .

[10] The periodic tidal forcing specified in this study was the semidiurnal  $M_2$  tidal wave with frequency  $\omega = 2\pi / (12.42 \times 3600)$  s. This is a typical standing wave problem in which the amplitude of the oscillation is geometrically controlled for a given wave frequency. Two cases were



**Figure 2.** Comparison of the surface elevation between the true, analysis (RRKF), and analysis (EnKF) values at days 0, 1, and 2 for the “normal oscillation” case. Days used in this plot refer to the time when the perturbation is added. The unit for the surface elevation is meters.

studied: (1) small amplitude (hereafter referred to as “normal case”) and (2) large amplitude (hereafter referred to as “near-resonance case”) oscillations. In the first case,  $L_1 = 10$  km,  $L = 57$  km,  $H_o = 10$  m, and  $A = 1.0$  m. The maximum tidal elevation is about 1.2 m, appearing at the coast of the island. In the second case,  $L_1$  and  $L$  are the same as those in the first case, but  $H_o = 1.0$  m and  $A = 0.01$  m. In this case,  $\zeta$  is very close to a resonant condition, in which the maximum elevation exceeds 0.84 m near the island and the tidal oscillation is out of phase relative to a node at  $r = 56$  km.

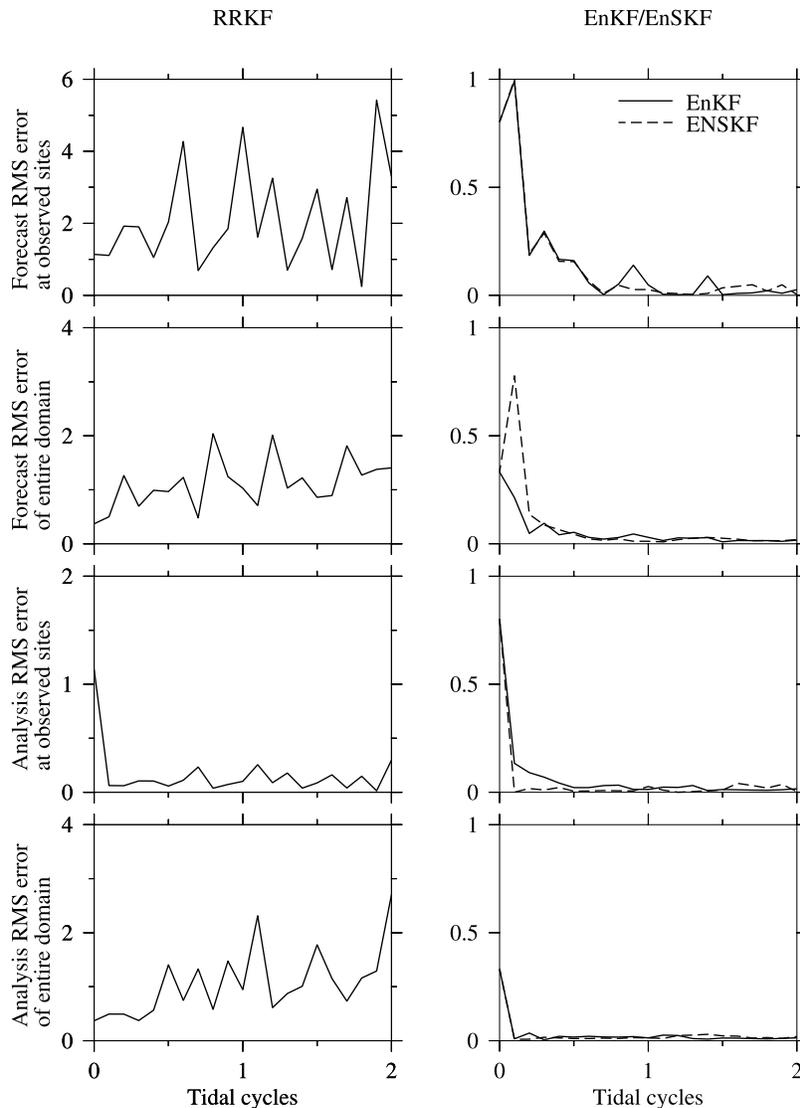
[11] In both “normal” and “near-resonance” cases, the FVCOM domain consists of an unstructured triangular mesh with a horizontal resolution of  $\sim 1.0$  km (defined as the shortest edge of a triangle). The total number of nodes and cells are 11,196 and 21,897, respectively, and the numerical solution converged accurately to the analytical solution in a few tidal cycles after an initial spin up.

[12] Numerical experiments using the Kalman filter were performed to examine the capability of RRKF, EnKF and EnSKF methods to restore the true solution after introduction of a perturbation. In the RRKF tests, the model was run prognostically for 15 days and then the state variables ( $\zeta$  and  $V_r$ ) output hourly over days 5–10 were used to calculate EOFs. In both RRKF and EnKF tests, the model “hot starts” from the beginning of day 11 with an “incorrect”

initial field. In the EnKF test, this initial field was constructed using

$$ENS(i) = 0.2\mathbf{x}(1) + 0.8 \left[ \mathbf{x}(i) - \frac{1}{20} \sum_{j=1}^{20} \mathbf{x}(j) \right]; \quad i = 1, 2, \dots, 20 \quad (4)$$

where  $i$  is the index of the ensemble member and  $\mathbf{x}(i)$  is the state variable matrix specified for the  $i$ th ensemble member. The ensemble member is constructed using the hourly model output field backward from the end of the 10th model day. In the RRKF test, the initial field was specified by the mean of 20 ensemble members constructed from (4). Therefore, both RRKF and EnKF tests were run using the same initial perturbation. The observations used for RRKF and EnKF were the values of the surface elevation or currents derived from the analytical solution (2)–(3) at the measurement sites shown in Figure 3. We assume that the observations are imperfect, even though the values are derived from the analytical solution. Random noise was added to the observational field, which was in general specified by forming the product of 1% of the standard deviation of the true solution relative to its mean and random numbers drawn from a Gaussian distribution with zero mean and unity standard deviation. The same



**Figure 3.** Forecast and analysis RMS errors (m) at measurement sites and over the entire domain over the two tidal cycles for the assimilation experiments with (left) RRKF and (right) EnKF/EnSKF.

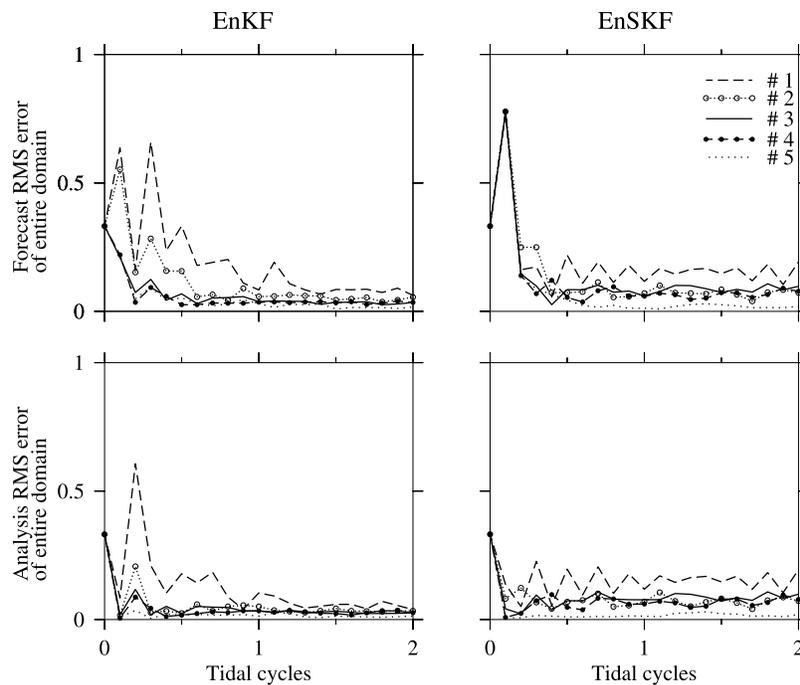
numerical approach was used in both the EnSKF and EnKF tests. The time interval between the two analysis steps was 1 h, which was determined from the decorrelation time scale of tidal currents estimated by an autocorrelation analysis. This time interval is applied for all twin experiments described in this paper.

### 3.1. Case 1: “Normal” Case

[13] The RRKF is incapable of restoring the numerical solution to the “true” values after perturbation (Figures 2 and 3). The dynamical system for this case is dependent on both boundary forcing and initial conditions. Initialized with the analytical solution, the model reaches the “true” state within two tidal cycles and the EOFs calculated with hourly model output from days 5–10 shows that the oscillation is dominated by the first four EOF modes with relative variances of 97.1, 1.0, 0.8, and 0.5%. In this case, the difference in maximum amplitude of the surface elevation between the open boundary and the coast is 0.2 m, which is 20% of the amplitude of the boundary forcing. When this

system is started with a perturbed initial condition, the interaction of perturbed and boundary-driven motions produces large oscillations with dominant periods around 2.5 and 7.1 h. These oscillations override the true  $M_2$  tidal motion and produce considerably different EOFs from the true solution. We also ran this case using the initialized field specified by a random selection of the model output at a time step before the end of 10th day. The results show that in such a frictionless system, RRKF fails once the perturbation produced by the wrong initial field is large enough to change the structure of the “true” EOFs.

[14] With an ensemble of perturbation fields having the same order of magnitude as those used in the RRKF tests, the EnKF and EnSKF succeeded in forcing the numerical solution toward the “true” solution (Figures 2 and 3). An examination of surface elevation at a select site near the coast (site 5 in Figure 1) with tests using the EnKF shows that both analysis and forecast RMS errors estimated either at this site or over the entire domain decrease rapidly over a few hours after the perturbation was introduced and the



**Figure 4.** Forecast and analysis RMS errors (m) of (left) EnKF and (right) EnSKF over the entire domain over the two tidal cycles for the assimilation experiment with selection of the surface elevation measurement at sites 1, 2, 3, 4, and 5.

surface elevation returns to the true state after one day. In this case, the difference between EnKF and EnSKF is insignificant, although the change of the forecast RMS error over the entire domain for these two methods differs during the first few hours after the filters are applied. For this 2-D problem, the elevation is determined by the convergence or divergence of water transport. Therefore, the velocity returns to the true solution at the same rate as the elevation.

[15] In EnKF and EnSKF, the convergence rate to the true solution depends on the location of sampling (Figures 4 and 5). The fastest convergence always occurs at sites with the most significant variation in both elevation and currents. For the case with elevation sampling, both analysis and forecast errors decrease fastest at site 5 and gradually slow down as the sampling locations are moved to sites 4, 3 and 2. When the sampling is set up at site 1 near the open boundary, the RMS errors remain at a level that is one or two times larger than that found at the other sampling sites. The opposite result was found in the case with current sampling (Figure 5). These two results are consistent with each other, because the amplitude of the elevation increases toward the coast, while the magnitude of the current decreases toward the coast as a result of the no-normal flux boundary condition.

### 3.2. Case 2: “Near-Resonance” Case

[16] This is a geometrically controlled tidal system in which the mean depth has been reduced to 1.0 m to create near-resonance conditions. Forced by a tidal oscillation with amplitude of 0.01 m, the tidal elevation at the coast can reach 0.84 m. At site 5 near the coast, both RRKF and EnKF drive a rapid convergence of the elevation to the true solution after the perturbation (Figures 6 and 7). In RRKF, the oscillation is dominated by the first EOF mode which accounts for 98.1% of the variance. Even when the initial

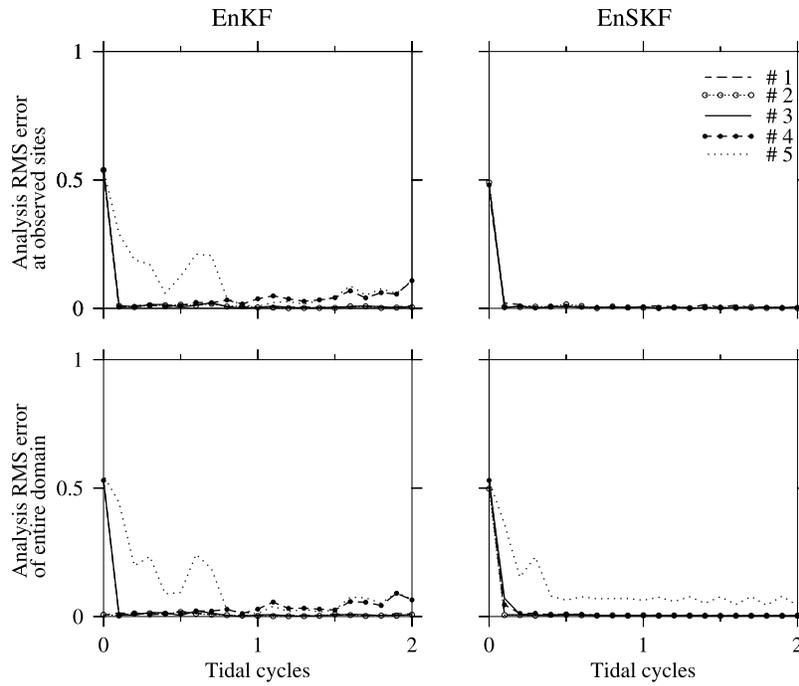
condition is reset back to zero at high tide, the EOF mode structure remains unchanged after perturbation. Therefore, RRKF succeeds in filtering the initial error and returning the numerical solution back to the true condition. This is not surprising because the near-resonance feature in this system is caused by the basin response to a small external forcing at the open boundary, with little influence from the initial condition. Success of EnKF in both “normal” and “near-resonance” cases suggests that this method is less sensitive to the physical setting for tidally driven flows. This is also true for EnSKF.

[17] Similar to the “normal” case, the rate of convergence to the true solution in RRKF, EnKF, and EnSKF is sensitive to the sampling location (Figure 7). In all these cases, the model predicts the fastest convergence at site 5 for elevation and site 1 for currents. When the elevation sampling is set to site 1 near the open boundary, analysis errors of all three filters display a large oscillation pattern and the numerical solutions never return to the true condition. In the current sampling case, although the analysis errors at site 5 eventually decrease, the convergence speed to the true solution is significantly slower.

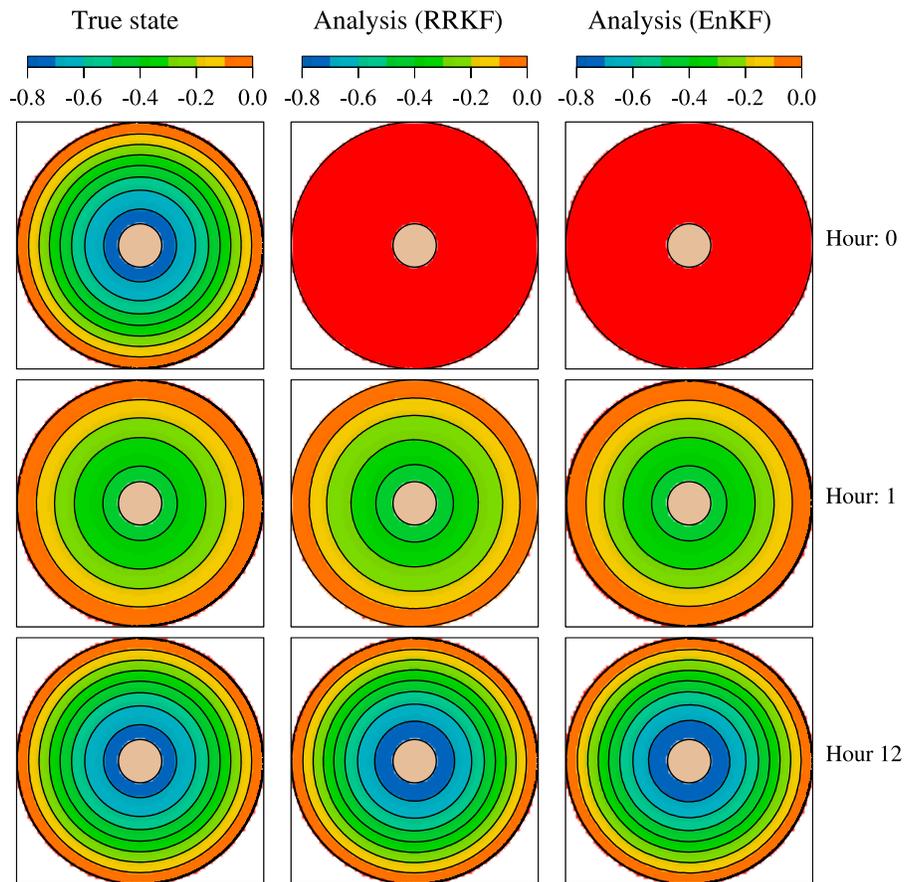
[18] Both “normal” and “near-resonance” cases are highly idealized, but the results indicate that sitting of instruments too near the coast may neglect more dynamically significant areas. These results may be useful in the design of a coastal observing system.

## 4. Low-Salinity Plume Over an Idealized Continental Shelf

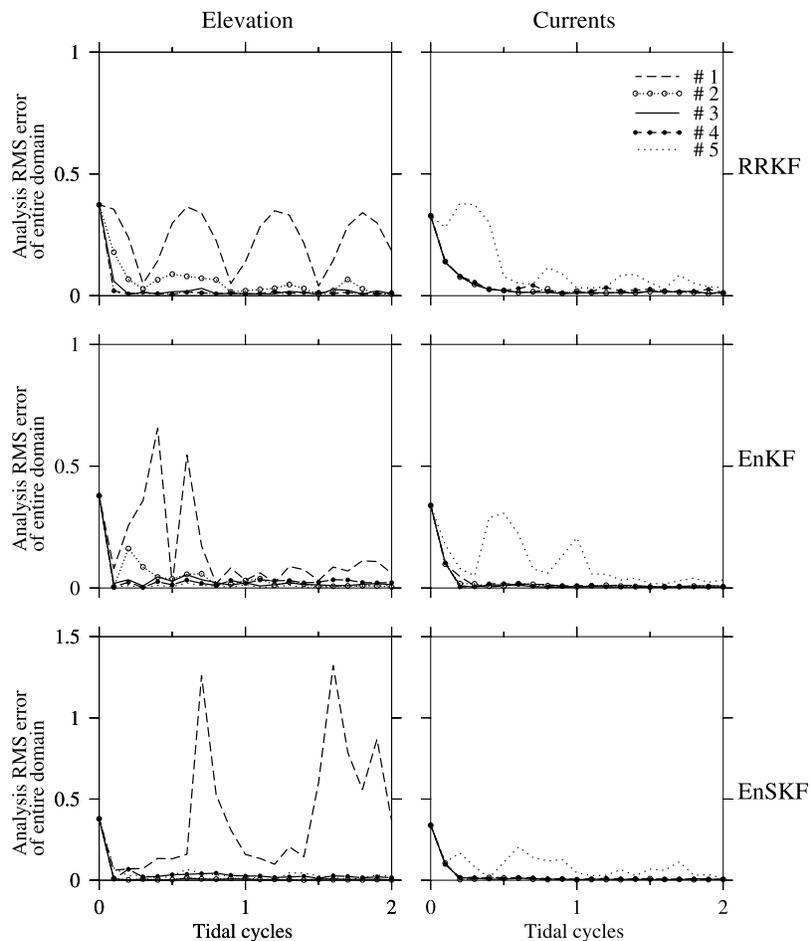
[19] Consider the idealized continental shelf shown in Figure 8, with water depth 20 m at the coast, a linear slope of  $6 \times 10^{-4}$  over the shelf, and a flat bottom region 80 m



**Figure 5.** Analysis RMS errors (m/s) of (left) EnKF and (right) EnSKF at the measurement sites and over the entire domain over the two tidal cycles for the assimilation experiment with selection of the velocity measurement at sites 1, 2, 3, 4, and 5. The definition of RMS errors is the same as that used in Figure 5.



**Figure 6.** Comparison of the surface elevation (m) between the true, analysis (RRKF), and analysis (EnKF) values at hours 0, 1, and 2 for the “near-resonance oscillation” case. The hours represented here refer to the time when the perturbation is added.



**Figure 7.** Analysis RMS errors of (top) RRFK, (middle) EnKF, and (bottom) EnSKF over the entire domain over the two tidal cycles for assimilation experiments with selection of the (left) elevation (m) and (right) velocity (m/s) measurements at sites 1, 2, 3, 4, and 5.

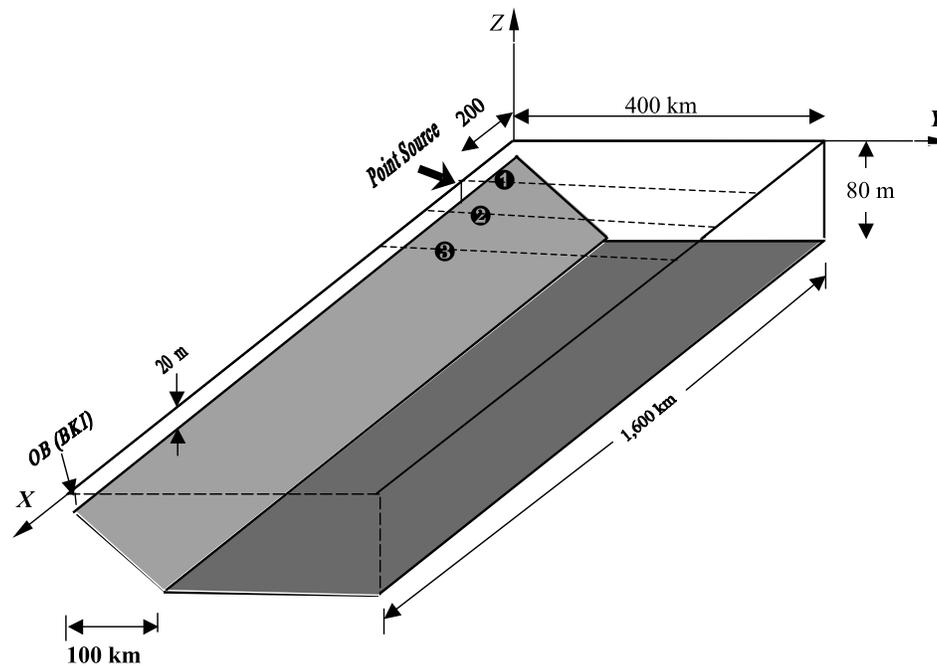
deep off shelf. A constant freshwater discharge of  $1000 \text{ m}^3/\text{s}$  is injected at the coast 200 km down-shelf from the origin at  $x = 0$ . The background salinity is set to a constant value of 30 PSU. At the open boundary, located 1400 km downstream from the river source, an implicit gravity wave radiation boundary condition (BKI [Blumberg and Kantha, 1985]) is applied to allow wave energy to propagate out of the computational domain with minimum reflection. The model domain features a nonoverlapping triangular grid in the horizontal and  $\sigma$  levels in the vertical. The horizontal resolution is 20 km everywhere except for cells connected to the river source, where the horizontal resolution is reduced to about 5 km for proper resolution of the plume [Chen *et al.*, 2007]. The vertical resolution is defined by 10  $\sigma$  layers with uniform thickness, which range from 2 m at the coast to 8 m off shelf.

[20] Driven by the freshwater river discharge, a surface-intensified low-salinity plume is established on the shelf, with a relatively strong clockwise circulation out of the river mouth and a coastal-trapped along-shelf current in the “downstream” (+ $x$ ) direction. This plume develops rapidly in the first 10 days and then gradually evolves as it propagates downstream. The fields of salinity and currents in the river mouth region vary slowly after 30 days.

[21] Numerical experiments were made to examine the robustness of RRFK, EnKF and EnSKF in assimilating the coastal low-salinity plume as it moved through a set of fixed location “mooring” sites. By using the stationary error covariance approach, RRFK requires a “steady state” condition to construct the EOFs. The salinity and currents near the river mouth reach an equilibrium state after 30 days, and thus EOFs were calculated using the hourly model fields for days 30–50. EnKF and EnSKF show no significant difference in the convergence rate and RMS errors, and thus EnSKF results are not included here for the plume case.

[22] Three moorings are “deployed” on the 44-m isobath along the coast within the plume, with M1 at the river mouth and M2 and M3 at 80 km and 160 km downstream from M1, respectively (Figure 8). At each mooring, the salinity and currents were “measured” at the midpoint of the first  $\sigma$  layer ( $\sim 1.5$  m) and these data were used as observations. Observations were assumed to be imperfect, and a random perturbation was added using the same method described in section 3.

[23] In both RRFK and EnKF experiments, we tested the rate of convergence of the perturbed numerical solution toward the true (unperturbed) solution using different



**Figure 8.** Idealized continental shelf configuration. Symbols 1–3 with filled dots represent the sites of the fixed location moorings.

perturbed initial fields. Four sets of these perturbed initial conditions were constructed, all of which were composed of 20 ensemble members constructed from the hourly model outputs. The first was sampled randomly from day 10 to 50, the second from day 20 to 50, the third from day 30 to 50, and the fourth between day 40 to 50 (hereafter referred to as “days 10, 20, 30 and 40” perturbations, respectively). These ensemble members were directly used as the initial perturbation field for EnKF, while the mean of ensemble members was used for RRKF.

[24] To examine the influence of selected assimilation methods on the convergence rate toward the true solution, we conducted five types of assimilation experiments using both univariate and multivariate covariance approaches in this plume study (Table 1). The first and second cases used the univariate assimilation approach in which the error covariance was estimated using only one state variable. In case 1 (hereafter called C1), only salinity measurements were assimilated. In case 2 (C2), only velocity measurements were assimilated. The remaining three cases utilized the multivariate covariance approach in which the error covariance was constructed using both salinity and velocity variables. The difference is that in case 3 (C3) only salinity measurements are assimilated; in case 4 (C4) only current measurements are assimilated; and in case 5 (C5) both salinity and current measurements are assimilated. In each case, two experiments were conducted, the first using data at only one “mooring” and the second using data from all three “moorings.”

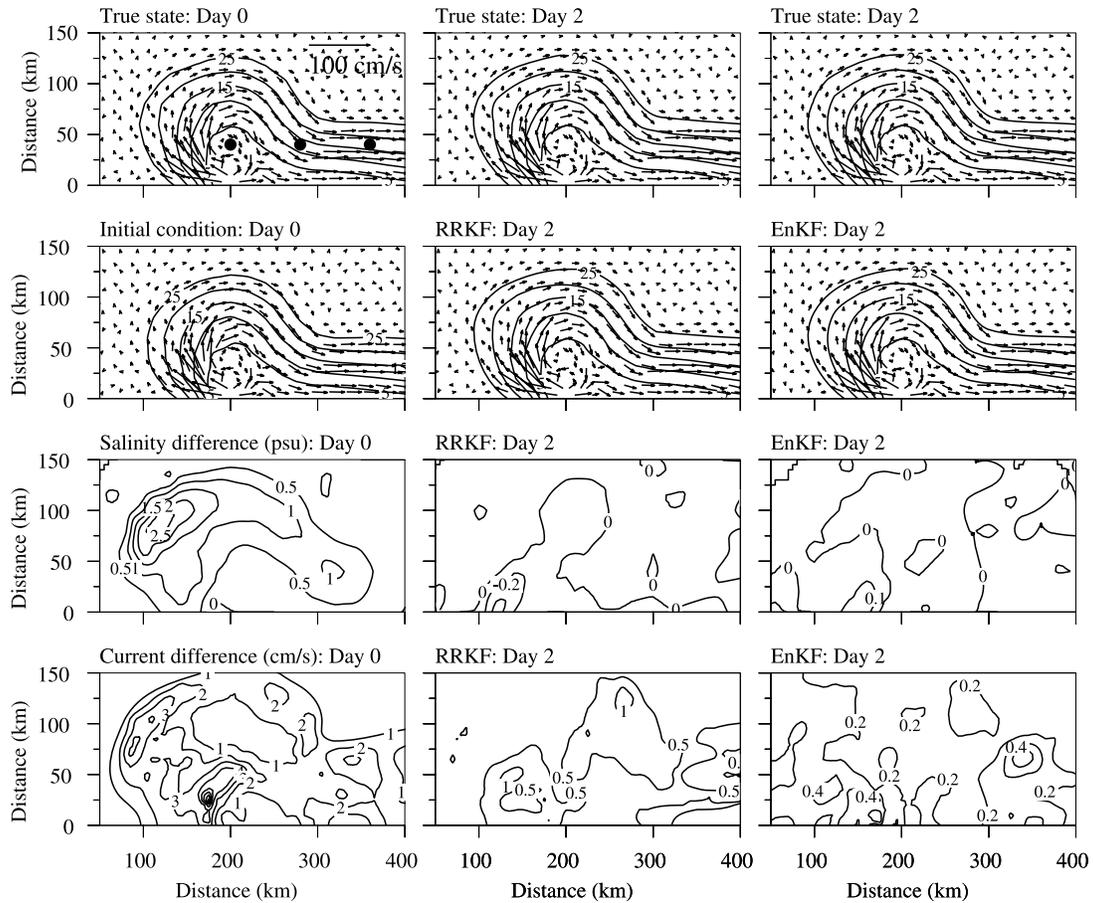
[25] Given the day 40 initial perturbation field, the RMS of the forecast and the true solutions obtained from running the model without assimilation diverges. With the same initial perturbed field, both RRKF and EnKF (for the

multivariate approach case C5 with measurements at all three “moorings”) succeeded in filtering the primary perturbation errors and enabled the numerical solution to quickly converge toward the true state. For example, Figure 9 shows the comparison between assimilated and true fields of the salinity and currents at the onset of the model run and 2 days after assimilation for the RRKF and EnKF runs using the day 40 model field to construct ensemble members. The difference between true and assimilated fields is  $\sim 0.5$ – $2.5$  PSU in salinity and  $\sim 2$ – $7$  cm/s in velocity at the beginning and decreases to  $\sim 0.2$  PSU or less in salinity and  $< 0.6$  cm/s in velocity in course of the first 2 days.

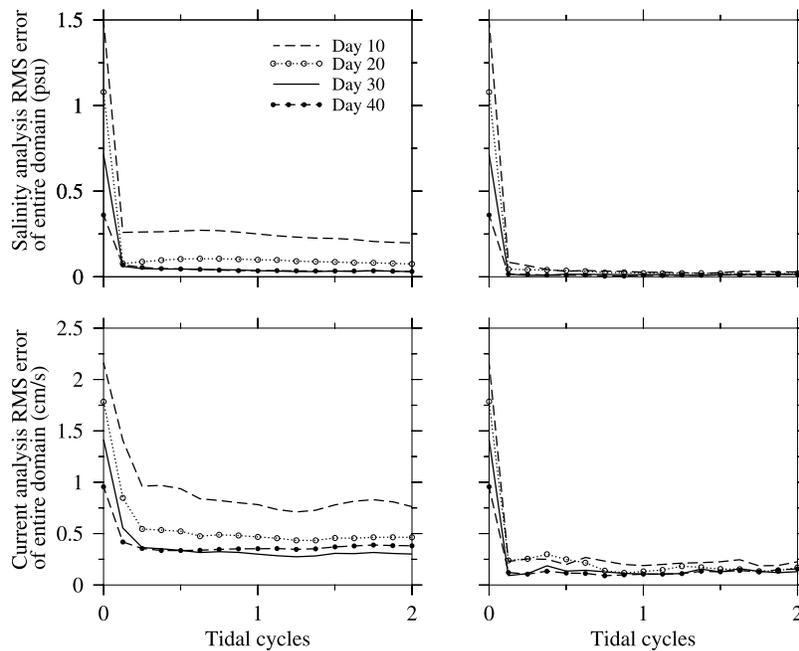
[26] Given the same perturbed initial field, RRKF and EnKF show a similar tendency to converge rapidly toward the true field in the first few hours, with the remaining RMS errors dependent on the initial perturbation. In the case with the day 10 initial perturbation field, RRKF shows larger RMS errors in both salinity and currents than EnKF (Figure 10). The errors reduce rapidly for the case with the day 20 initial perturbation field. Results for both RRKF and EnKF with the day 30 and day 40 initial perturbation fields are quite similar. In this experiment, RRKF uses a fixed forecast error

**Table 1.** Summary of Configurations of Five Assimilation Experiments

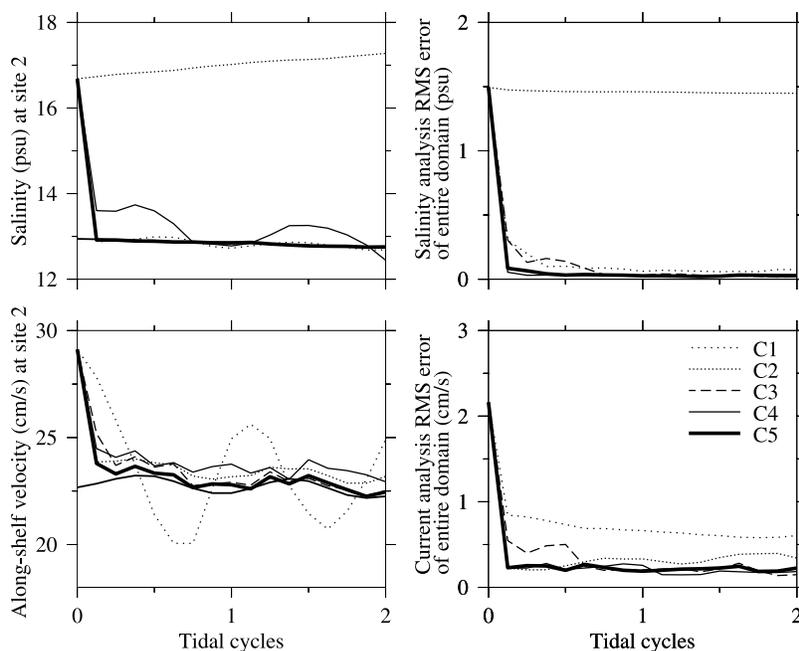
Experiment		Error Covariance	Measurements
Case	Case Type		
C1	univariate	salinity	salinity
C2	univariate	currents	currents
C3	multivariate	salinity and currents	salinity
C4	multivariate	salinity and currents	currents
C5	multivariate	salinity and currents	salinity and currents



**Figure 9.** Distributions of the true, RRKF, and EnKF analysis salinity and near-surface current vectors, the true analysis salinity difference, and the true analysis velocity difference at day 0 and day 2 after perturbation. Filled dots in the top left distribution are the sites of the fixed location moorings.



**Figure 10.** Analysis RMS errors in (top) salinity and (bottom) velocity for RRKF and EnKF experiments with the same initial perturbation fields constructed by the ensemble sets for days 10, 20, 30, and 40.



**Figure 11.** Comparison of the (top left) true salinity and (bottom left) along-shelf velocity values at site 2 with the analysis values calculated by the EnKF. Black line indicates the true value. Analysis RMS errors in (top right) salinity and (bottom right) velocity over the entire domain for the five difference cases (C1–C5).

covariance matrix, so the resulting EOFs work well for a quasi-linear system that reaches a steady state. An equilibrium state for salinity and currents near the river mouth is reached after 30 days, the day 30 and day 40 initial fields contain dominant modes similar to the EOFs calculated on the basis of the hourly model fields for days 30–50, and thus display similar results. The day 10 initial field shows slightly different mode structures. Although not large enough to override the dominant EOFs, this can lead to relatively large RMS errors.

[27] The Kalman gain calculated in the above case includes both salinity and velocity; that is, EnKF is applied to both currents and water properties. In other assimilation methods (e.g., nudging and OI), assimilation usually is performed only for a single state variable. Both RRKF and EnKF also allow assimilation of single model field. The experiments C1–C5 were made to examine the adjustment of the model to different assimilation methods based on univariate or multivariate covariance approaches. In these experiments, RRKF and EnKF display the same behavior. The EnKF results are provided here.

[28] The performance of EnKF varies across the five cases (Figures 11 and 12). In C1, the salinity error at the measurement sites drops rapidly in a few hours after the filter is applied. Because in this univariate configuration only salinity is directly influenced by the assimilation, near-inertial oscillations appear in the velocity as a result of the rapid dynamical adjustment of the flow field to the salinity field. Therefore, the RMS error for the entire domain drops rapidly in salinity but the RMS error in currents remains at a significant level after a rapid drop during the first 6 h.

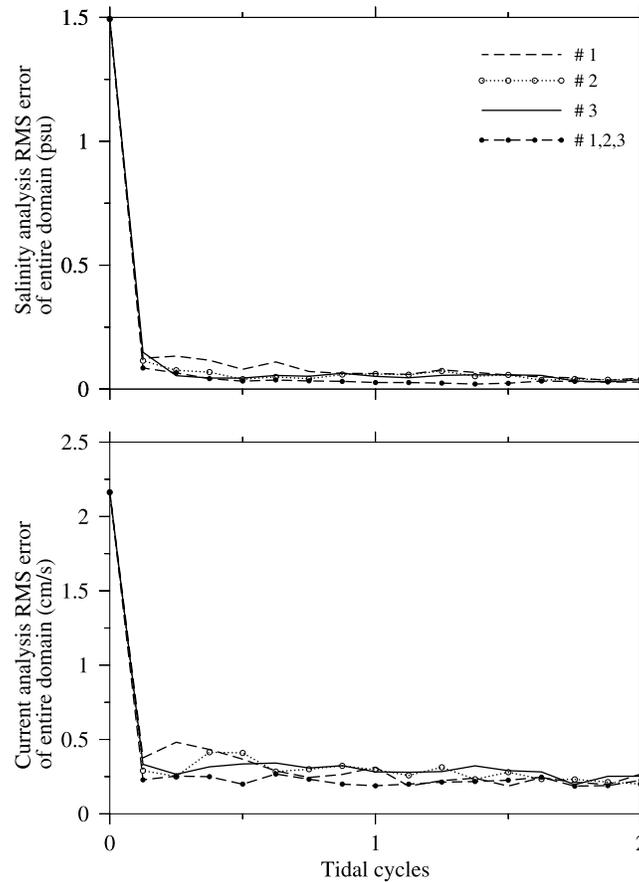
[29] In C2, there is only a nominal adjustment in salinity at the measurement sites and in the remainder of the

domain, even though the velocity error is observed to rapidly decrease in the first 6 h after perturbation. Correspondingly, the RMS error in the domain-averaged salinity does not change significantly with time. It is clear that if the EnKF is applied only to currents, this method is incapable of adjusting the salinity field back to the true solution. In fact, the salinity structure in the plume depends on the salt flux from the river. The assimilation approach used in this case provides no mechanism to restore the salt that is removed by substituting the perturbed initial salinity field.

[30] In C3, when both salinity and currents are used to construct the error covariance in the EnKF, the model-predicted salinities and currents at the measurement sites show rapid convergence toward the true solution. Similar to C1 (where only salinity data were assimilated), the rapid adjustment of the flow to the adjusted salinity produces near-inertial oscillations not present in the true solution. These oscillations, however, are much weaker than those found in C1. In this case, both RMS errors in salinity and velocity rapidly decrease to the random noise level in the first assimilation day.

[31] In C4, EnKF achieves a similar performance to C3. The primary difference is that in this case, the RMS errors in salinity and velocity drop quickly to the level of background noise within a few hours without the slowly decreasing period noted in C3. The results in C5 contain no significant differences from C4.

[32] The five experiments described above suggest that in order to accurately predict the low-salinity plume over the shelf, the EnKF should include both salinity and currents in assimilation. With the multivariate covariance approach, this filter is capable of removing the perturbation errors and restoring the numerical solution to the true state through



**Figure 12.** EnKF analysis RMS errors in salinity and velocity over the entire domain for experiments with selection of salinity and velocity measurements at sites 1, 2, and 3 and at all three sites.

assimilation using either salinity or current measurements at fixed locations.

[33] Neither RRKF or EnKF display strong sensitivity to the number of measurement sites in this test case. In C5, for example, the EnKF assimilations made with salinity and velocity data at mooring site 1, 2, or 3 show the same convergence speed in the first 3 h after perturbation, and the resulting RMS errors differ only slightly afterward (Figure 12). This suggests that if the coefficient of coherence scale in the filter is appropriately specified, the number of measurement sites within the region of coherence does not significantly influence the convergence rate in EnKF. This is also generally true for RRKF for cases with the same initial perturbation fields as EnKF.

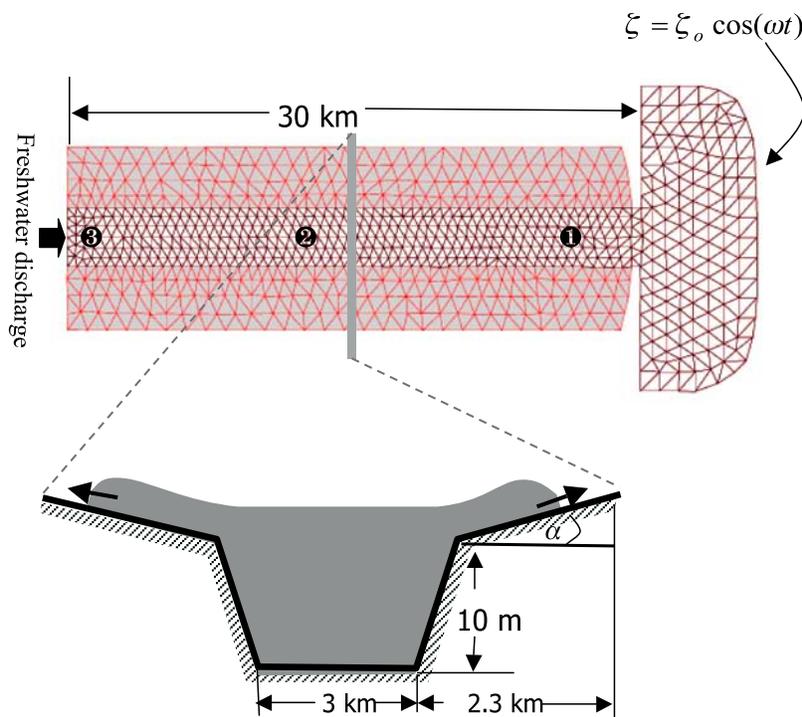
## 5. Tidal Flushing and River Plume in an Idealized Estuary

[34] Consider an idealized semienclosed rectangular-shaped channel with a width 3 km at the bottom of the river channel, a length of 30 km, a constant depth of 10 m and a lateral slope of about 0.033 (a change of water depth from 10 m to 0 m over a distance of 0.3 km (Figure 13). This channel is oriented east to west, and is connected to a relatively wide and flat bottom shelf to the east and intertidal zones along the northern and southern edges. The intertidal zones are distributed symmetrically to the

channel, with a constant slope of  $\alpha = 7.5 \times 10^{-4}$  (an increase in height of the intertidal zone from 0 m to 1.5 m over a distance of 2 km in the cross-channel direction).

[35] Numerical experiments were made for two cases: one is driven solely by the  $M_2$  tidal forcing at the open boundary at the shelf (case A), and the other by  $M_2$  tidal forcing in addition to freshwater discharge at the upstream end of the channel (case B). The computational domain (Figure 13) has a horizontal resolution of  $\sim 0.5$  km in the channel,  $\sim 0.6$  to 1.0 km over the shelf, and  $\sim 0.6$  to 0.9 km over the intertidal zones, and variable vertical resolution provided by 10 uniform  $\sigma$  levels. In both cases, the model was integrated for 20 days. In case A with the  $M_2$  tidal amplitude  $\zeta_o = 1.0$  m specified at the open boundary, the tidal motion and flooding/drying process over the river/intertidal zone complex reached an equilibrium state within 10 tidal cycles. In case B, both tidal forcing and freshwater discharge were activated at the start of model integration. Although the tidal motion reached an equilibrium state in a few days, the river plume resulting from mixing between freshwater and salt water gradually moved seaward and the residual and buoyancy-driven low-frequency currents did not reach an equilibrium state by the end of model integration on day 20.

[36] The twin experiments with RRKF, EnKF and EnSKF were made to examine the convergence tendency of the model to the true solution after the initial perturbation for case A but with EnKF and EnSKF only for case B. RRKF



**Figure 13.** Idealized estuary configuration. Symbols 1–3 with filled dots indicate the sites of the fixed location moorings.

works for a quasi-linear system in which a steady state exists. Case B is not an appropriate test problem for comparing RRFK and EnKF because the system is fully nonlinear.

[37] The observations used in these experiments were constructed using the time series of surface elevation, currents, and salinity extracted from the FVCOM-computed fields at the three mooring sites shown in Figure 13. All these data were used in both cases A and B, with discussion focused on salinity and currents for case B. We also assumed that the observations were imperfect, and white noise was added using the same method described in sections 3 and 4.

[38] The initial perturbation fields were composed of 20 ensemble members of the hourly model-calculated field sampled randomly from day 5 to day 10. These ensemble members were directly used as the initial perturbation field for EnKF and EnSKF, while, to keep the same initial condition, the mean of ensemble members was used for RRFK. The assimilation started at day 10 and continued for an additional 10 days.

[39] Following the procedure used in the plume case, for the case B experiment, we also examined the relationship of assimilation methods to the convergence rate toward the true solution and the dependence of the convergence tendency on numbers and locations of the observations. The cases C1–C5 presented here have the same definition as those in section 4.

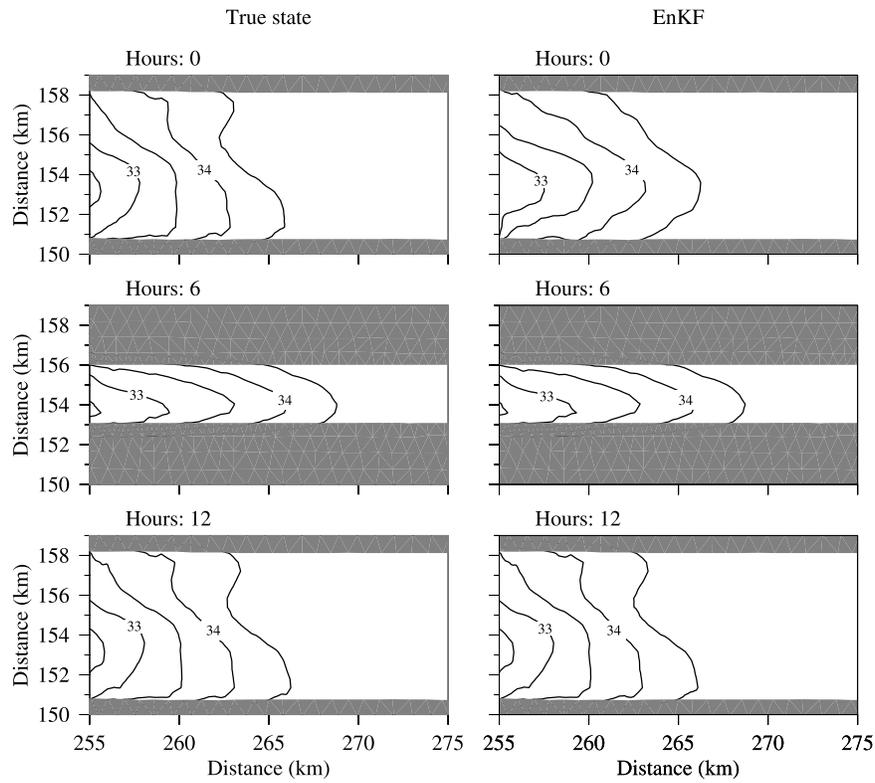
### 5.1. Case A

[40] RRFK, EnKF and EnSKF all show fast convergence of surface elevation and currents toward the true solution in the first 2 h after perturbation. Small differences appear in

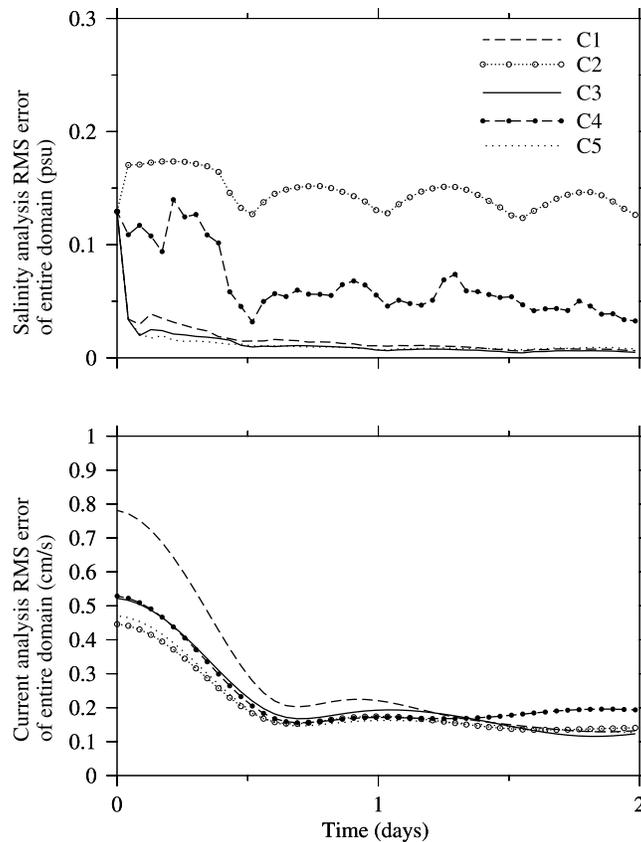
the domain analysis errors after the initial period of rapid decrease but all methods converge toward the same error level after 5 h (not shown). Without any assimilation, however, the model field in such a highly dissipated estuarine system also converged toward the true state in the same time period. In the context of KF testing, this represents a trivial case because the system is capable of restoring itself back to the true state without assimilation. On the other hand, it does imply that the tidal motion in a highly dissipated estuary is a self-restoring system in which the perturbation will be automatically filtered. This is not a surprising result, because the tidal dynamics in an estuary are driven by the open boundary. In an estuary with strong turbulent dissipation, realistic representation of local coastline and bathymetry, and prescribed tidal forcing at the open boundary, the model should be capable of accurately simulating the tidal motion in this estuary. As long as only the tide is considered, effort should be made to improve the accuracy of the local bathymetry and coastline geometry rather than data assimilation.

### 5.2. Case B

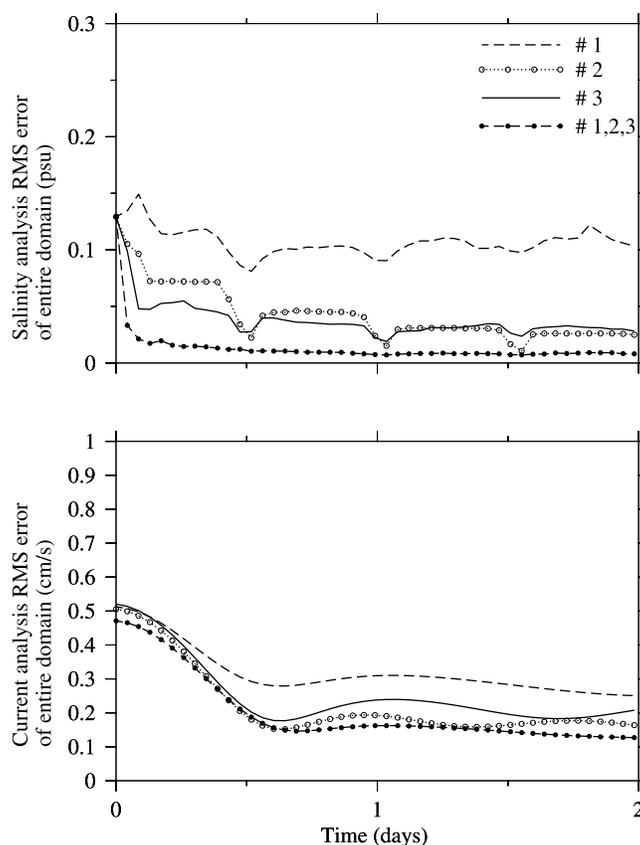
[41] The objective of this experiment is to examine if EnKF and EnSKF can influence the model fields of salinity and buoyancy-driven low-frequency currents to converge toward the true solution after a perturbation. The EnKF and EnSKF results show no significant difference, so the discussion given below is focused on EnKF. The model run without data assimilation clearly showed that this buoyancy-driven river system did not have a self-restoring nature. Once the system is perturbed, there is no mechanism to filter the error and return the model solution toward the true solution. EnKF works efficiently in this case. By



**Figure 14.** Distributions of the salinity at hours 0, 6 (low water), and 12 (high water) for the (left) true state and (right) EnKF analysis value.



**Figure 15.** Analysis RMS errors in (top) salinity and (bottom) velocity for the EnKF experiment for the five difference configurations (C1–C5).



**Figure 16.** Analysis RMS errors in (top) salinity and (bottom) velocity for the EnKF experiment for selection of the salinity and velocity measurements at sites 1, 2, and 3 and at all three sites.

including both salinity and velocity in the Kalman gain estimation, EnKF assimilation with salinity and velocity measurements at the three “mooring” sites (shown in Figure 13) are capable of filtering the perturbation error and restoring the model field to the true solution (Figure 14). This experiment clearly demonstrates that these two methods can be successfully applied to the flooding/drying process over the intertidal zone that is dry at low water and flooded at high water.

[42] Similar to the low-salinity plume case described in section 4, the success of EnKF in this estuary case also depends on univariate and multivariate configurations and what types of measurement are made (Figure 15). For the case in which the assimilation is conducted solely with the moored current data, the model was never able to restore the salinity field, even though the velocity error decreased rapidly after perturbation. The explanation given for the low-salinity plume case is valid for this case, too. Even when both velocity and salinity variables are used to calculate the Kalman gain, the assimilated salinity field still has a high error level when only the moored current data are used in the assimilation. Unlike the low-salinity plume over the continental shelf, in this highly dissipative estuary, both EnKF works well for the case in which only salinity is assimilated (a C1 experiment). The fastest convergence appears for the case in which EnKF is applied for both salinity and velocity with measurements of these two variables.

[43] In this idealized estuary case, the convergence rate in EnKF is related to the location of sampling in the estuary. When only salinity and velocity measurements at site 1 are selected (a C5 experiment), the convergence rate is significantly retarded in both salinity and velocity fields (Figure 16), since site 1 was out of the plume region in the perturbed initial salinity and velocity fields. Once sampling is made within the plume region (e.g., at site 2 or site 3), convergence rates toward the true solution increase, with the fastest convergence occurring in the case with sampling both inside and out of the plume.

## 6. Discussion

[44] The RRKF, EnKF, and EnSKF experiments for selected idealized cases show that the success of data assimilation in coastal ocean and estuarine model applications depends critically on assimilation methods, data sampling, and the dynamical nature of the system being studied. In all successful cases described in our experiments, the covariance of ensemble member deviations from ensemble mean is consistent with the difference between the ensemble mean and true state.

[45] Many shallow water regions (e.g., inner bays, inlets, lagoons, and estuaries) are highly turbulent and dissipative, with dynamics featuring multiscale processes in time and space. Determining an optimal assimilation for such regions should be considered on a case by case basis. The example given in section 5 demonstrates that an estuary is a self-

restoring system for tidal motion but not for the subtidal currents driven by the freshwater discharge at the head of the estuary. In the former example, better knowledge of the bathymetry, coastline, and intertidal zone geometry would lead to improved model simulation of the tidal motion, even without data assimilation. Inclusion of data assimilation with appropriately placed salinity and velocity measurement sites would on the other hand lead to improved simulation of the subtidal buoyancy-driven flow and plume structure.

[46] Our experiments on the low-salinity plume over the continental shelf (section 4) suggest that the univariate assimilation of salinity can produce strong near-inertial oscillations, even though the model does show an improvement in the salinity field. The fast convergence in the salinity field acts similarly to imposing external buoyancy forcing in the system, and numerical oscillations are generated during a rapid adjustment of the flow to the assimilated salinity. Note that we are not suggesting here that data assimilation should not be conducted separately for hydrographic and current fields because this approach works well for highly dissipative estuarine regions, but that attention should be paid to having independent comparisons with field measurements of the different variables to ensure that unrealistic secondary numerical responses (such as the generation of near-inertial oscillations) do not occur for the specific method and variable set being used. The twin experiments for our three selected idealized cases clearly show that the multivariate covariance approach works better than the univariate covariance approach. It probably is premature to make such a conclusion because the success of the multivariate covariance approach requires a reliable error covariance, which may be case specific.

[47] Government and public interest in establishing integrated coastal observing systems have been growing in recent years, with an objective to build an end-to-end predictive capability for the coastal ocean environment. Creating a scientifically based design of an in situ monitoring network is a significant intellectual challenge given the wide range of spatial and temporal scales of the dominant processes involved, the real limitations on observing assets, and the competing priorities for accurate predictions. Our experiments for tidal oscillations over the continental shelf show that the convergence rate of the model fields to the true solution depends on the sampling locations and differ between currents and elevations. Although the results from this highly idealized case have limited direct application because of the complexity of the real coastal ocean, the approach used here can be used to conduct model-based experiments needed for the design of optimal monitoring networks and adaptive sampling methods.

[48] It should be pointed out that the EnKF experiments made in this study followed the Monte Carlo sampling of ensemble members. Recently the SEEK and SEIK filters were developed to improve the computational efficiency by presenting the error covariance with lower ensemble numbers [Evensen, 2004; Nerger *et al.*, 2005]. Through comparison experiments with the EnKF, SEEK and SEIK filters, Nerger *et al.* [2005] pointed out that the selection of state ensemble members and assimilation schemes can influence the filter performance and the SEIK filter appears to be a more computationally efficient algorithm for nonlinear systems. A twin experiment was also made between EnKF

and EnKS by Ngodock *et al.* [2006], who showed that the success of the KF methods depends on the resolution of data sampling. We believe that those proof-of-concept twin experiment results are useful for improving the application of KF-based assimilation schemes to realistic coastal ocean models.

## 7. Conclusion

[49] The data assimilation twin experiments were conducted to validate RRKF, EnKF and EnSKF for three idealized coastal ocean and estuarine cases. Both EnKF and EnSKF are suitable to coastal ocean systems in which the dynamics vary significantly with time and in space. Experiments show that the convergence speed of the model field to the true solution depends on the sampling locations, assimilation methods, and the dynamical nature of the system. Because the coastal ocean features multiscale dynamics in time and space, a case-by-case approach should be used to determine the most effective and most reliable data assimilation method for different dynamical systems. Finally, we note that no matter how powerful an assimilation method is, it only works well when used with a model that is capable of resolving the underlying physical system. Assimilation should be used to improve the accuracy of the simulation for hindcast or forecast, but with limited sampling, it is impossible to create dynamical features that are not resolved in the forward prediction.

[50] **Acknowledgments.** The authors would like to thank the reviewers for their helpful suggestions that have done much to improve this manuscript. For this work, P. Malanotte-Rizzoli and J. Wei were supported by the Office of Naval Research (ONR grant N00014-06-1-0290); C. Chen and Q. Xu were supported by the U.S. GLOBEC/Georges Bank program (through NSF grants OCE-0234545, OCE-0227679, OCE-0606928, OCE-0712903, OCE-0726851, and OCE-0814505 and NOAA grant NA-16OP2323), the NSF Arctic research grants ARC0712903, ARC0732084, and ARC0804029, and URI Sea Grant R/P-061; P. Xue was supported through the MIT Sea Grant 2006-RC-103; Z. Lai, J. Qi, and G. Cowles were supported through the Massachusetts Marine Fisheries Institute (NOAA grants NA04NMF4720332 and NA05NMF4721131); and R. Beardsley was supported through U.S. GLOBEC/Georges Bank NSF grant OCE-02227679, MIT Sea Grant NA06OAR1700019, and the WHOI Smith Chair in Coastal Oceanography. The numerical experiments were conducted using the High Performance Computer Cluster of the Marine Ecosystem Dynamics Modeling Laboratory at the School of Marine Science and Technology, University of Massachusetts–Dartmouth, purchased through Massachusetts Marine Fisheries Institute NOAA grants listed above.

## References

- Bergamasco, A., P. Malanotte-Rizzoli, W. C. Thacker, and R. B. Long (1993), The seasonal steady circulation of the eastern Mediterranean determined with the adjoint method, *Deep Sea Res.*, *40*, 1269–1298, doi:10.1016/0967-0645(93)90070-4.
- Bishop, C. H., B. J. Etherton, and S. J. Majumdar (2001), Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects, *Mon. Weather Rev.*, *129*, 420–436, doi:10.1175/1520-0493(2001)129<0420:ASWTET>2.0.CO;2.
- Blanchet, I., C. Frankignoul, and M. A. Cane (1997), A comparison of adaptive Kalman filter for a tropical Pacific Ocean model, *Mon. Weather Rev.*, *125*, 40–58, doi:10.1175/1520-0493(1997)125<0040:ACOAKF>2.0.CO;2.
- Blumberg, A. F., and L. H. Kantha (1985), Open boundary conditions for circulation models, *J. Hydraul. Eng.*, *111*, 237–255, doi:10.1061/(ASCE)0733-9429(1985)111:2(237).
- Buehner, M., and P. Malanotte-Rizzoli (2003), Reduced-rank Kalman filters applied to an idealized model of the wind-driven ocean circulation, *J. Geophys. Res.*, *108*(C6), 3192, doi:10.1029/2001JC000873.
- Buehner, M., P. Malanotte-Rizzoli, A. J. Busalacchi, and T. Inui (2003), Estimation of the tropical Atlantic circulation from altimetry data using a

- reduced-rank stationary Kalman filter, in *Interhemispheric Water Exchanges in the Atlantic Ocean*, Elsevier Oceanogr. Ser., vol. 68, edited by G. Goni and P. Malanotte-Rizzoli, pp. 193–212, Elsevier, New York.
- Burchard, H. (2002), *Applied Turbulence Modeling in Marine Waters*, 215 pp., Springer, Berlin.
- Chen, C., H. Liu, and R. C. Beardsley (2003), An unstructured, finite-volume, three-dimensional, primitive equation ocean model: Application to coastal ocean and estuaries, *J. Atmos. Oceanic Technol.*, 20, 159–186, doi:10.1175/1520-0426(2003)020<0159:AUGFVT>2.0.CO;2.
- Chen, C., R. C. Beardsley, and G. Cowles (2006a), An unstructured grid, finite-volume coastal ocean model-FVCOM user manual, 2nd ed., *Tech. Rep. SMAST/UMASSD-06-0602*, 318 pp., Sch. for Mar. Sci. and Technol., Univ. of Mass.–Dartmouth, New Bedford.
- Chen, C., R. C. Beardsley, and G. Cowles (2006b), An unstructured grid, finite-volume coastal ocean model (FVCOM) system, *Oceanography*, 19(1), 78–89.
- Chen, C., H. Huang, R. C. Beardsley, H. Liu, Q. Xu, and G. Cowles (2007), A finite volume numerical approach for coastal ocean circulation studies: Comparisons with finite difference models, *J. Geophys. Res.*, 112, C03018, doi:10.1029/2006JC003485.
- Chen, C., Q. Xu, R. Houghton, and R. C. Beardsley (2008), A model-dye comparison experiment in the tidal mixing front zone on the southern flank of Georges Bank, *J. Geophys. Res.*, 113, C02005, doi:10.1029/2007JC004106.
- Cowles, G. (2008), Parallelization of the FVCOM coastal ocean model, *Int. J. High Perform. Comput. Appl.*, 22(2), 177–193, doi:10.1177/1094342007083804.
- Evensen, G. (1992), Using the extended Kalman filter with a multilayer quasi-geostrophic ocean model, *J. Geophys. Res.*, 97(C11), 17,905–19,924, doi:10.1029/92JC01972.
- Evensen, G. (1993), Open boundary condition for the extended Kalman filter with a quasi-geostrophic ocean model, *J. Geophys. Res.*, 98(C9), 16,529–16,546, doi:10.1029/93JC01365.
- Evensen, G. (1994), 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, doi:10.1029/94JC00572.
- Evensen, G. (2003), The ensemble Kalman filter: Theoretical formulation and practical implementation, *Ocean Dyn.*, 53(4), 343–367, doi:10.1007/s10236-003-0036-9.
- Evensen, G. (2004), Sampling strategies and square root analysis scheme for the EnKF, *Ocean Dyn.*, 54, 539–560, doi:10.1007/s10236-004-0099-2.
- Ghil, M. (1989), Meteorological data assimilation for oceanographers, I. description and theoretical framework, *Dyn. Atmos. Oceans*, 13(3–4), 171–218.
- Ghil, M., and P. Malanotte-Rizzoli (1991), Data assimilation in meteorology and oceanography, *Adv. Geophys.*, 33, 141–266.
- Houtekamer, P., and H. L. Mitchell (1998), Data assimilation using an ensemble Kalman filter technique, *Mon. Weather Rev.*, 126, 796–811, doi:10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2.
- Kalman, R. E. (1960), A new approach to linear filtering and prediction problems, *J. Basic Eng.*, 82D, 35–45.
- Le Dimet, F., and O. Talagrand (1986), Variational algorithm for analysis and assimilation of meteorological observations: Theoretical aspects, *Tellus, Ser. A*, 38, 97–110.
- Lorenz, A. C. (1981), A global three-dimensional multivariate statistical interpolation scheme, *Mon. Weather Rev.*, 109, 701–721, doi:10.1175/1520-0493(1981)109<0701:AGTDMS>2.0.CO;2.
- Mellor, G. L., and T. Yamada (1982), Development of a turbulence closure model for geophysical fluid problem, *Rev. Geophys.*, 20, 851–875, doi:10.1029/RG020i004p00851.
- Morrow, R. A., and P. De Mey (1995), Four-dimensional assimilation of altimetric and cruise data in the Azores current in 1992–93, *J. Geophys. Res.*, 100(C12), 25,007–25,025, doi:10.1029/95JC02315.
- Nerger, L., W. Hiller, and J. Schröter (2005), A comparison of error subspace Kalman filters, *Tellus, Ser. A*, 57, 715–735.
- Ngodock, H. E., G. A. Jacobs, and M. Chen (2006), The representer method, the ensemble Kalman filter and the ensemble Kalman smoother: A comparison study using a nonlinear reduced gravity ocean model, *Ocean Modell.*, 12, 378–400, doi:10.1016/j.ocemod.2005.08.001.
- Pham, D. T., J. Verron, and L. Gourdeau (1998a), Filters de Kalman singuliers évolutif pour l'assimilation de données en océanographie, *C. R. Acad. Sci., Ser. IIa Sci. Terre Planet.*, 326, 255–260.
- Pham, D. T., J. Verron, and M. C. Roubaud (1998b), Singular evolutive extended Kalman filter with EOF initialization for data assimilation in oceanography, *J. Mar. Syst.*, 16, 323–340, doi:10.1016/S0924-7963(97)00109-7.
- Pietrzak, J., J. B. Jakobson, H. Burchard, H. J. Vested, and O. Petersen (2002), A three-dimensional hydrostatic model for coastal and ocean modeling using a generalized topography following co-ordinate system, *Ocean Modell.*, 4, 173–205, doi:10.1016/S1463-5003(01)00016-6.
- Smagorinsky, J. (1963), General circulation experiments with the primitive equations: I. The basic experiment, *Mon. Weather Rev.*, 91, 99–164, doi:10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2.
- Thacker, W. C., and R. B. Long (1988), Fitting dynamics to data, *J. Geophys. Res.*, 93(C2), 1227–1240, doi:10.1029/JC093iC02p01227.
- Tziperman, E., and W. C. Thacker (1989), An optimal-control/adjoint-equations approach to studying the oceanic general circulation, *J. Phys. Oceanogr.*, 19, 1471–1485, doi:10.1175/1520-0485(1989)019<1471:AOCEAT>2.0.CO;2.
- Wang, X., and C. H. Bishop (2003), A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes, *J. Atmos. Sci.*, 60, 1140–1158, doi:10.1175/1520-0469(2003)060<1140:ACOBAE>2.0.CO;2.
- Whitaker, J. S., and T. M. Hamill (2002), Data assimilation without perturbed observations, *Mon. Weather Rev.*, 130, 1913–1924, doi:10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2.
- Zang, X., and P. Malanotte-Rizzoli (2003), A comparison of assimilation results from the ensemble Kalman filter and the reduced-rank extended Kalman filter, *Nonlinear Processes Geophys.*, 10, 6477–6491.

R. C. Beardsley, Department of Physical Oceanography, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA. (rbeardsley@whoi.edu)

C. Chen, G. W. Cowles, Z. Lai, J. Qi, Q. Xu, and P. Xue, School for Marine Science and Technology, University of Massachusetts–Dartmouth, New Bedford, MA 02744, USA. (c1chen@umassd.edu; gcowles@umassd.edu; g\_zlai@umassd.edu; jq1@umassd.edu; qxu@umassd.edu; pxue@umassd.edu)

S. Lyu, P. Malanotte-Rizzoli, and J. Wei, Department of Atmospheric, Earth and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02193, USA. (rizzoli@mit.edu; junwei@mit.edu)