6-1-2011

A Hybrid Background Error Covariance Model for Assimilating Glider Data into a Coastal Ocean Model

Max Yaremchuk
Stennis Space Center

Dmitri A. Nechaev
University of Southern Mississippi, dmitri.nechaev@usm.edu

Chudong Pan
University of Southern Mississippi, chudong.pan@usm.edu

A Hybrid Background Error Covariance Model for Assimilating Glider Data into a Coastal Ocean Model

MAX YAREMCHUK

Naval Research Laboratory, Stennis Space Center, Mississippi

DMITRI NECHAEV AND CHUDONG PAN

Department of Marine Science, University of Southern Mississippi, Stennis Space Center, Mississippi

(Manuscript received 24 May 2010, in final form 24 January 2011)

ABSTRACT

A hybrid background error covariance (BEC) model for three-dimensional variational data assimilation of glider data into the Navy Coastal Ocean Model (NCOM) is introduced. Similar to existing atmospheric hybrid BEC models, the proposed model combines low-rank ensemble covariances $\mathbf{B}_m$ with the heuristic Gaussian-shaped covariances $\mathbf{B}_0$ to estimate forecast error statistics. The distinctive features of the proposed BEC model are the following: (i) formulation in terms of inverse error covariances, (ii) adaptive determination of the rank m of $\mathbf{B}_m$ with information criterion based on the innovation error statistics, (iii) restriction of the heuristic covariance operator $\mathbf{B}_0$ to the null space of $\mathbf{B}_m$, and (iv) definition of the BEC magnitudes through separate analyses of the innovation error statistics in the state space and the null space of $\mathbf{B}_0$.

The BEC model is validated by assimilation experiments with simulated and real data obtained during a glider survey of the Monterey Bay in August 2003. It is shown that the proposed hybrid scheme substantially improves the forecast skill of the heuristic covariance model.

1. Introduction

In recent years, development of hybrid background error covariance (BEC) models has been an area of active research in atmospheric data assimilation (Hamill and Snyder 2000; Etherton and Bishop 2004; Wang et al. 2007). It has been shown in particular that hybrid models tend to be more robust than conventional ensemble-based data assimilation schemes, especially when the model errors are larger than observational ones (Wang et al. 2007, 2008, 2009). This feature is attractive for the regional assimilation problems in oceanography, where information on the background state is often scant and incomplete.

Sequential data assimilation schemes developed so far for regional oceanographic studies can be classified in two categories. The first one is the Kalman filter (KF)-type algorithms with low-rank BEC matrices $\mathbf{B}_m$ derived from ensemble statistics. These applications encompass many flavors of the reduced-order KF techniques (e.g., Evensen 2003; Tippett et al. 2003; Brasseur and Verron 2006). They proved to be especially useful for monitoring comparatively large domains continuously covered by sea surface height/sea surface temperature (SSH/SST) observations at the surface with sporadic vertical temperature/salinity (T/S) soundings by Argo drifters and ships. The second type of assimilation algorithms employ steady-state covariances $\mathbf{B}_0$ derived from long-term model integrations (Yin et al. 2011) or heuristic Gaussian-shaped covariance operators with simple dynamical constraints (Weaver and Courtier 2001; Pannekoucke and Massart 2008). The latter type of the BEC models has recently gained considerable attention because of its flexibility and convenience in introducing prior information into the covariance model in cases when the background model solutions are biased and/or contain large errors.

A typical oceanographic setting of such kind is a near-coastal survey by autonomous gliders, which have recently become a fast-developing operational technology in oceanography (Rudnick et al. 2004). Gliders are

Corresponding author address: Max Yaremchuk, Naval Research Laboratory, Code 7321, Bldg. 1009, Stennis Space Center, MS 39529.
E-mail: max.yaremchuk@nrlssc.navy.mil

DOI: 10.1175/2011MWR3510.1

capable of making remotely controllable surveys of limited areas at high spatiotemporal resolution. Such a dense 4D coverage is usually accompanied by a relatively poor knowledge of the background ocean state: near-coastal regions are often affected by poorly known peculiarities of the bottom topography and the associated tidal/inertial motions that cannot be resolved by global OGCMs. Considerable model error covariances also persist at scales comparable with the size of the domain due to inconsistencies in the boundary conditions and/or local atmospheric forcing.

Because of the relative novelty of glider technology, examples of glider assimilation are rare in literature (Heaney et al. 2007; Shulman et al. 2009). Recently, Dobricic et al. (2010) have shown that three-dimensional variational data assimilation (3DVar) assimilation of glider data significantly improves the forecast skill of a regional model. Most importantly, glider data were able to capture basin-scale BE correlations, which improved the model's forecast skill several weeks after termination of glider observations. Dobricic et al. utilized the second category 3DVar algorithm based on stationary Gaussian-shaped BECs in the horizontal combined with EOF decomposition in the vertical (Dobricic and Pinardi 2008) and did not explicitly include adaptive error covariances inferred from model statistics.

In this study we propose a hybrid 3DVar assimilation system specifically targeted on preserving survey-scale correlations that could be resolved by gliders in coastal areas. Similar to the existing atmospheric hybrid models, the "flow dependent" part of the covariance $\mathbf{B}_m$ is defined as a low-rank matrix derived from ensemble statistics. The heuristic part of the covariance $\mathbf{B}_0$ is represented by the propagator of the diffusion equation for temperature and salinity. To gain extra computational efficiency, the action of the propagator is modeled by a semi-implicit scheme (Weaver and Ricchi 2004; Yaremchuk et al. 2011, manuscript submitted to *Ocean Modell.*). For that reason the proposed BEC model is formulated in terms of the inverse covariances and the assimilation problem is solved in the state space $\mathcal{R}^M$.

Another distinctive feature of the BEC model is an explicit separation of the covariance components in $\mathcal{R}^M$: the action of $\mathbf{B}_0$ is restricted to the null space of $\mathbf{B}_m$. This is done to better preserve the above-mentioned regional-scale error correlations. Since low-rank approximations of large covariance matrices tend to be more uncertain at larger distances (Hamill et al. 2001), we paid special attention to the determination of the statistically reliable number of modes m and the magnitudes (scaling coefficients) for both $\mathbf{B}_m$ and $\mathbf{B}_0$. The respective algorithms are based upon the Bayesian information criterion and analyses of the innovation statistics.

The rest of the paper is organized as follows. We start with the description of the hybrid BEC model (section 2), then briefly review the Navy Coastal Ocean Model (NCOM) forecast model and the experimental design for the Monterey Bay area (section 3). We continue with an examination of the forecast skills of the assimilation system for the twin-data experimental setting and subsequent real-data experiment (section 4). Section 5 concludes the paper.

2. A hybrid 3DVar assimilation scheme

a. The BEC model

The analysis increment $\delta\mathbf{x}$ of the sequential data assimilation scheme considered here is obtained by minimizing the cost function:

$$J(\delta\mathbf{x}) = \frac{1}{2}[\delta\mathbf{x}^{\mathrm{T}}\mathbf{B}^{-1}\delta\mathbf{x} + (\mathbf{H}\delta\mathbf{x} - \delta\mathbf{y})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x} - \delta\mathbf{y})] \to \min_{\delta\mathbf{x}}, \tag{1}$$

where $\mathbf{B}$ is the BEC matrix, $\mathbf{R}$ is the $K \times K$ observation error covariance matrix, T denotes transposition, and $\mathbf{H}$ is the linear operator projecting model state $\mathbf{x} \in \mathcal{R}^M$ on the innovation vector $\delta\mathbf{y} \in \mathcal{R}^K$, whose K components are the model-data misfits of the background solution.

To define linear operations with multivariate vectors $\delta\mathbf{x}$, we introduce a diagonal matrix $\mathbf{G}$ approximating the background error variance. Elements of $\mathbf{G}$ depend on the physical nature of the fields contributing to $\delta\mathbf{x}$ and spatial coordinates. Farther below we will assume that all the quantities in (1) are normalized by the respective error variances and introduce new variables:

$$\delta\mathbf{x}_* = \mathbf{G}^{-1/2}\delta\mathbf{x}, \quad \delta\mathbf{y}_* = \mathbf{R}^{-1/2}\delta\mathbf{y}.$$

The matrices $\mathbf{B}$, $\mathbf{H}$ are appropriately transformed to keep J invariant:

$$\mathbf{B}_*^{-1} = \mathbf{G}^{1/2}\mathbf{B}^{-1}\mathbf{G}^{1/2}, \quad \mathbf{H}_* = \mathbf{R}^{-1/2}\mathbf{H}\mathbf{G}^{1/2}.$$

Dropping the asterisks for convenience of further treatment, the cost function (1) and the normal equation $\partial J/\partial\delta\mathbf{x} = 0$ now take the following form:

$$J(\delta\mathbf{x}) = \frac{1}{2}[\delta\mathbf{x}^{\mathrm{T}}\mathbf{B}^{-1}\delta\mathbf{x} + (\mathbf{H}\delta\mathbf{x} - \delta\mathbf{y})^{\mathrm{T}}(\mathbf{H}\delta\mathbf{x} - \delta\mathbf{y})], \tag{2}$$

$$[\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{H}]\delta\mathbf{x} = \mathbf{H}^{\mathrm{T}}\delta\mathbf{y}. \tag{3}$$

Farther below we assume $\mathbf{R}$ to be known and focus on the structure of the BEC matrix $\mathbf{B}$.

The hybrid covariance models developed so far (Hamill and Snyder 2000; Etherton and Bishop 2004; Wang et al.

2007) utilize linear combinations of the heuristic or "static" covariance operator $\mathbf{B}_0$ and the flow-dependent operator $\mathbf{B}_m$ derived from the statistics of an ensemble of analyses/forecasts:

$$\mathbf{B} = \alpha^{-1}\mathbf{B}_m + \beta^{-1}\mathbf{B}_0. \tag{4}$$

Here α^{-1} and β^{-1} are empirically defined positive scalar parameters often constrained by the requirement $\alpha^{-1} + \beta^{-1} = 1$ (e.g., Wang et al. 2008). We adopt the traditional representation of $\mathbf{B}_m$ in the form of a matrix defined on a subspace $\mathcal{R}^m \in \mathcal{R}^M$ spanned by an orthogonal basis $\{\mathbf{e}_k\}$ derived from the eigenvector analysis of the ensemble covariance:

$$\mathbf{B}_m = \mathbf{P}\mathbf{\Lambda}_m\mathbf{P}^{\mathrm{T}}. \tag{5}$$

Here $\mathbf{P}$ is a rectangular $m \times M$ matrix with the columns $\mathbf{e}_k$, $k = 1, \ldots, m$ and $\mathbf{\Lambda}_m$ is the diagonal $m \times m$ matrix whose nonzero elements represent the variances of $\mathbf{e}_k$.

In the absence of the additional prior information, $\mathbf{B}_0$ is often represented by the propagator of the diffusion equation (e.g., Weaver and Courtier 2001; Pannekoucke and Massart 2008):

$$\mathbf{B}_0 = \exp(\tau\mathbf{D}); \quad \mathbf{D} = -\nabla^{\mathrm{T}}\nu\nabla, \tag{6}$$

where the diffusion tensor ν depends on spatial coordinates to simulate inhomogeneity and anisotropy of the background flow and τ is the scalar parameter, specifying in 3D the local correlation radii ρ^i via the eigenvalues λ^i_ν of ν: $\rho^i \sim \sqrt{2\lambda^i_\nu\tau}$, $i = 1, 2, 3$. The parameter τ can be also interpreted as "integration time" of the corresponding finite-difference diffusion equation.

In the present study, we adopt the diffusion model (6) and define the inverse of the BEC operator as

$$\mathbf{B}^{-1} = \alpha\mathbf{P}\mathbf{\Lambda}_m^{-1}\mathbf{P}^{\mathrm{T}} + \beta\mathbf{P}_\perp\mathbf{B}_0^{-1}\mathbf{P}_\perp^{\mathrm{T}}, \tag{7}$$

where $\mathbf{P}_\perp = \mathbf{I}_M - \mathbf{P}\mathbf{P}^{\mathrm{T}}$ is the projector on the orthogonal supplement of $\mathcal{R}^m$ and $\mathbf{I}_M$ is the identity operator in $\mathcal{R}^M$. This definition statistically separates the ensemble-generated components of the increment $\mathbf{P}\mathbf{P}^{\mathrm{T}}\delta\mathbf{x}$ from those described by the heuristic BEC model $\mathbf{B}_0$. Another reason for formulating the BEC model (7) in terms of the inverse covariances are computational advantages of the numerical approximation of (6) and solving the normal equation (3) in state space (Yaremchuk et al. 2011, manuscript submitted to *Ocean Modell.*).

Since $\mathbf{B}^{-1}$ has a two-cell structure in an orthogonal basis containing $\mathbf{e}_k$, the respective background error covariance matrix can be readily written as

$$\mathbf{B} = \frac{1}{\alpha}\mathbf{P}\mathbf{\Lambda}_m\mathbf{P}^{\mathrm{T}} + \frac{1}{\beta}[\mathbf{P}_\perp \exp(-\tau\mathbf{D})\mathbf{P}_\perp^{\mathrm{T}}]^{-1}, \tag{8}$$

with inversion in the second term standing for the generalized (Moore–Penrose) inverse. The normal equation (3) takes the following form:

$$[\alpha\mathbf{P}\mathbf{\Lambda}_m^{-1}\mathbf{P}^{\mathrm{T}} + \beta\mathbf{P}_\perp \exp(-\tau\mathbf{D})\mathbf{P}_\perp^{\mathrm{T}} + \mathbf{H}^{\mathrm{T}}\mathbf{H}]\delta\mathbf{x} = \mathbf{H}^{\mathrm{T}}\delta\mathbf{y}. \tag{9}$$

Among other parameters, the BEC model (8) depends on the inverse magnitudes α, β of its components and the number of eigenvectors m spanning the $\mathcal{R}^m$. In the proposed algorithm α, β, and m are determined from the model states and the data.

b. Definition of m and α

Accurate determination of the first term in the BEC model in (8) is important because this term is responsible for capturing error correlations on scales comparable with the size of the domain. In oceanographic applications these errors are generated by poorly known open boundary conditions and errors in atmospheric forcing, which tend to have larger scales than those of the internal oceanic variability. In addition, $\mathbf{B}_m$ may contain valuable information on the dynamical structure of the model error field because it is derived from the prior statistics of the forecast errors.

In many applications, the domain surveyed by gliders is rarely well-observed beforehand and the first-guess estimate of the background state may be far from reality. So the leading eigenvectors $\mathbf{e}_k$ of the first-guess BEC estimate provide poor approximation to the true eigenvectors of the background error covariance. To assess reliability of $\mathbf{e}_k$ we employ the Bayesian information criterion (Schwarz 1978) and define the optimal number of "trusted" eigenvectors as the minimum of

$$C(m) = m + \frac{N}{\ln N}\ln\sigma_m^2 \to \min_m, \tag{10}$$

where

$$\sigma_m^2 = \overline{\sum_{k=1}^{K_n}\left[\sum_{i=1}^{m} f_i^n\mathbf{e}_i(\mathbf{r}_k^n) - \delta\mathbf{x}(\mathbf{r}_k^n)\right]^2} \Big/ \overline{\sum_{k=1}^{K_n}\delta\mathbf{x}^2(\mathbf{r}_k^n)} \tag{11}$$

is the relative residual approximation error of N data samples by m modes. Here $\mathbf{r}_k^n$ denotes the kth observation location at the nth analysis time, the coefficients f_i^n are obtained by minimizing the numerator in (11) at a time layer n for a given number of modes $m \ll K_n$, and the overbar denotes averaging over N time layers.

The relationship in (10) gives an asymptotic ($N \gg 1$) approximation to the Bayesian posterior probability for a model with m parameters (linear regression on m eigenfunctions) given N observations (T/S fields sampled by gliders at the analysis times) under the assumption that model-data misfits are normally distributed. A similar, but less restrictive m criteria could be also used (Akaike 1974; Hannah and Quinn 1979).

The magnitude of $\mathbf{B}_m$ is determined by considering the optimization problem (1) in the m-dimensional subspace $\mathcal{R}^m$ spanned by $\{\mathbf{e}_k\}$. Because the Gaussian part of $\mathbf{B}$ is defined in the orthogonal supplement of $\mathcal{R}^m$, an approximate formula for the covariance matrix between the projections of $\delta\mathbf{x}$ on $\mathbf{e}_k$ can be obtained (see the appendix):

$$\langle \delta\mathbf{e}\delta\mathbf{e}^{\mathrm{T}} \rangle = [\alpha\mathbf{\Lambda}_m^{-1} + \mathbf{Q}]^{-1}\left[\frac{1}{\alpha}\mathbf{Q}\mathbf{\Lambda}_m\mathbf{Q}^{\mathrm{T}} + \mathbf{Q}\right][\alpha\mathbf{\Lambda}_m^{-1} + \mathbf{Q}]^{-1}. \tag{12}$$

Here $\mathbf{Q} = \mathbf{P}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{P}$ and $\delta\mathbf{e}$ is the m-dimensional vector of the expansion coefficients such that $\delta\mathbf{x} = \mathbf{P}\delta\mathbf{e}$.

Equation (12) can be used to compute α in several ways. The matrices $\mathbf{Q}$ and $\mathbf{\Lambda}$ are known and the matrix on the left-hand side can be estimated by approximating $\delta\mathbf{y}$ by the linear combinations of $\mathbf{e}_k$ at N analysis times and computing the time-averaged covariances between the vectors $\delta\mathbf{e}$ of the optimal fit coefficients. Optimal α can be then computed by minimizing a norm of the difference between the left- and right-hand sides of (12). Since all the matrices in (12) are positive definite, a convenient option is to set the difference between their traces to zero. In the application considered below, the background model errors are much larger than observational errors ($|\mathbf{Q}| \gg \alpha|\mathbf{\Lambda}_m{}^{-1}|$), and we use the simplified relationship

$$\langle \delta\mathbf{e}\delta\mathbf{e}^{\mathrm{T}} \rangle \simeq \frac{1}{\alpha}\mathbf{\Lambda}_m \tag{13}$$

to estimate α: its value is found by minimizing the mean squared difference between the diagonal elements of $\langle \delta\mathbf{e}\delta\mathbf{e}^{\mathrm{T}} \rangle$ and $\mathbf{\Lambda}_m/\alpha$.

In principle, one can generalize the covariance model in $\mathcal{R}^m$ and exactly fit the observed variances, $\mathrm{diag}\langle \delta\mathbf{e}\delta\mathbf{e}^{\mathrm{T}} \rangle$, by adjusting the diagonal elements of $\mathbf{\Lambda}_m$. In oceanographic applications, however, there is no reason to refine the covariance model by finetuning the eigenvalues because even the leading eigenvectors of $\mathbf{B}_m$ are known very poorly. Besides, the minimization problem is nonlinear and computationally expensive. We therefore choose a simpler model (13) with a single scaling factor α.

c. Definition of β

Having established the structure of the dynamical part of the covariance model (7) we can now determine the magnitude β of the Gaussian part by equating the trace of the sample forecast error covariance $\mathrm{Tr}\langle \delta\mathbf{y}\delta\mathbf{y}^{\mathrm{T}} \rangle$ derived form the innovation statistics to the trace of $\mathbf{HBH}^{\mathrm{T}} + \mathbf{I}_K$, a technique routinely used in computation of the inflation factor in the Kalman filtering schemes (e.g., Wang et al. 2007). Substituting $\mathbf{B}$ from (8) into the expression $\mathbf{HBH}^{\mathrm{T}}$, we obtain

$$\langle \delta\mathbf{y}^{\mathrm{T}}\delta\mathbf{y} \rangle = \mathrm{Tr}\left\{\frac{1}{\alpha}\mathbf{HP}\mathbf{\Lambda}_m\mathbf{P}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}} + \frac{1}{\beta}\mathbf{H}[\mathbf{P}_\perp \exp(-\tau\mathbf{D})\mathbf{P}_\perp^{\mathrm{T}}]^{-1}\mathbf{H}^{\mathrm{T}}\right\} + K, \tag{14}$$

so that

$$\beta = \frac{\mathrm{Tr}\left\{\mathbf{H}[\mathbf{P}_\perp \exp(-\tau\mathbf{D})\mathbf{P}_\perp^{\mathrm{T}}]^{-1}\mathbf{H}^{\mathrm{T}}\right\}}{\langle \delta\mathbf{y}^{\mathrm{T}}\delta\mathbf{y} \rangle - K - \mathrm{Tr}[\mathbf{HB}_m\mathbf{H}^{\mathrm{T}}]/\alpha}. \tag{15}$$

The numerator of this expression can be computed by the Monte Carlo technique (Bai and Golub 1997) at the expense of several iterative solutions of the $M \times M$ system of equations with random right-hand sides.

d. Numerical implementation

In the present study we used a simple diagonal model of the diffusion tensor $\nu = \mathrm{diag}\ \nu^i$, assuming that local decorrelation radii ρ^i are directly proportional to the model grid steps Δx^i, $i = 1, 2, 3$ spatially varying in 3D:

$$\sqrt{\nu^i} = \Delta x^i/\sqrt{2}. \tag{16}$$

Since gliders directly measure only the temperature and salinity fields, the operator $\mathbf{B}_0$ was only applied to the T/S components of the state vector under the prior assumption of zero correlations between them. The temperature/salinity background error correlations were taken into the account by the $\mathbf{B}_m$, which was also operating in the reduced space (i.e., eigenvectors $\mathbf{e}_k$ were only estimated for the temperature and salinity fields).

The BEC operator $\mathbf{B}_0$ in (6) was approximated by an implicit "time integration scheme" (see, e.g., Yaremchuk et al. 2011, manuscript submitted to *Ocean Modell.*):

$$\exp(\tau\mathbf{D}) \simeq \left[\mathbf{I}_M - \frac{\tau\mathbf{D}}{n}\right]^{-n}, \tag{17}$$

where τ/n is the length of the implicit "time step" and n is the number of explicit time steps. With the definition (16), the square root of the integration time τ has the

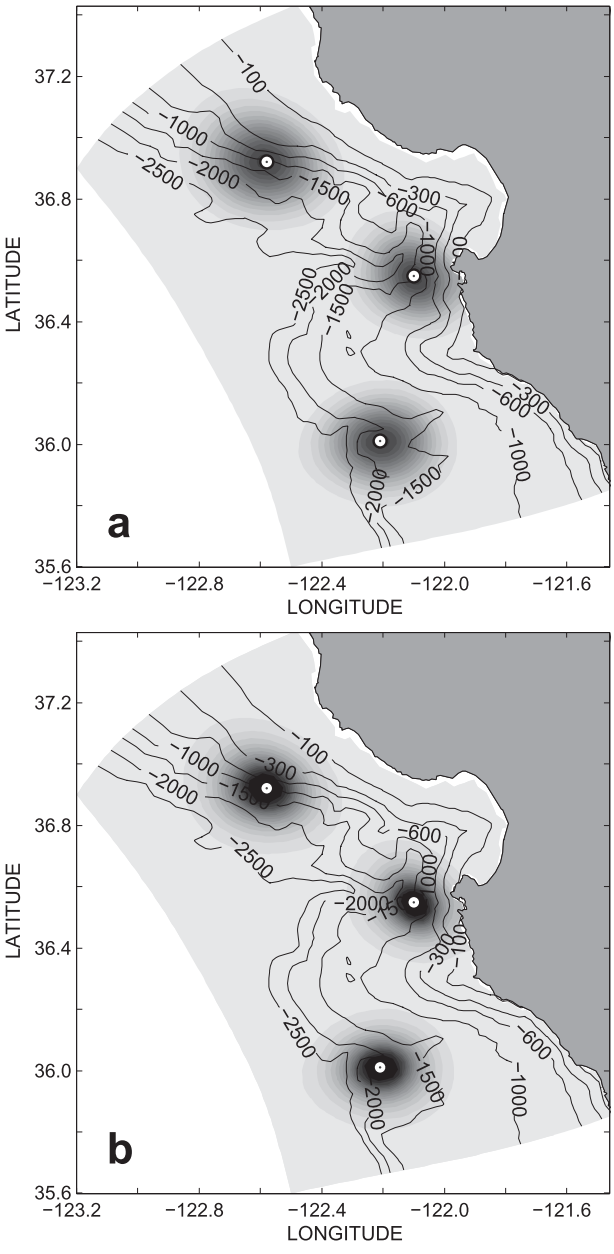

FIG. 1. 2D slices of the rows of the numerical approximations of $\exp(\mathbf{D}\tau)$. The patterns were obtained by applying (a) the high-order explicit approximation $(\mathbf{I} + \mathbf{D}t/100)^{100}$ and (b) the implicit approximation in (16) with $n = 2$ to the δ-shaped disturbances of the temperature field at three points shown by white circles. The bathymetry contours are in m.

meaning of the proportionality coefficient between the local decorrelation scales ρ^i and the grid steps in the corresponding directions: $\rho^i = \sqrt{2\nu_i\tau} = \Delta x^i\sqrt{\tau}$. The value of $\tau = 20$ was determined through preliminary experiments described in section 3. The second parameter n in (17) was chosen as a compromise between the numerical complexity and accuracy in approximation of $\exp(\tau\mathbf{D})$. For $n = 2$ the approximation error is close to 15% (Fig. 1), which is quite reasonable given the overall uncertainty in the definition of the heuristic covariance operator $\mathbf{B}_0$. At the boundaries, the operator $\mathbf{D}$ was specified by prescribing zero normal derivatives.

The analysis increment $\delta\mathbf{x}$ was obtained by solving (9) with a generalized minimum residual solver (Saad 2003). In correspondence with the approximation (17), the inverse of $\exp(\tau\mathbf{D})$ was represented by

$$\exp(-\tau\mathbf{D}) \simeq \left[\mathbf{I}_M - \frac{\tau\mathbf{D}}{n}\right]^n. \tag{18}$$

Depending on the number of analyzed observations K, the solution of (9) required, as a rule, 150–300 iterations, keeping the computational cost of the analysis well below the cost of a 12-h model run between the assimilations. In the reported experiments, the state space dimension M was 515 102, which is the number of observations K varied between 1500 and 3000, and m never exceeded 2. Within these ranges of K and m the CPU time required for the estimation of α and m was negligible compared to the time τ_{cpu} of solving the normal equation. Conversely, estimation of β required several (usually 4–7) iterative inversions of $\mathbf{P}_\perp \exp(-\tau\mathbf{D})\mathbf{P}_\perp^{\mathrm{T}}$ at the expense of $3–5\tau_{\text{cpu}}$, and was the most expensive part of the analysis.

The only type of data used in the present study were temperature and salinity profiles from gliders. Therefore, balance constraints were introduced by applying the linearized equation of state and the geostrophic–hydrostatic relationships directly to the temperature and salinity increments (e.g., Li et al. 2008) obtained from minimization of the cost function (1).

3. Experiment design

The BEC model was verified by 3DVar assimilation experiments with the Navy Coastal Ocean Model (NCOM) configured in the Monterey Bay (Fig. 2) for processing of the data acquired during the Autonomous Ocean Sampling Network (AOSN II) experiment (Ramp et al. 2008). The experiment was conducted in the summer of 2003 with the ultimate goal of developing an adaptive sampling technique that combines numerical forecasts with the data flows from controllable observation platforms. Observations were performed by several types of autonomous underwater vehicles (AUVs) including gliders, high-frequency radars, two moorings, bottom-mounted ADCPs, surface drifters, and CTD casts. In the present study, we focus the analysis on the temperature/salinity data from gliders only: space–time coordinates of the gliders are used to define observation operators $\mathbf{H}(t)$ in both twin- and real-data assimilation experiments with the hybrid 3DVar scheme.

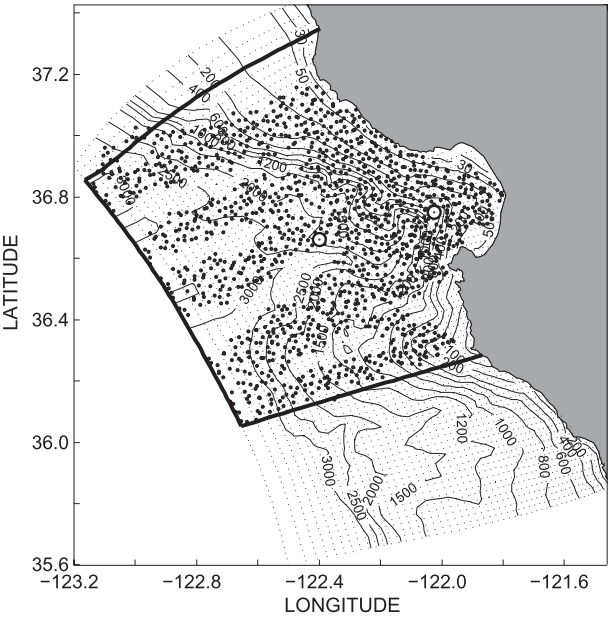

FIG. 2. Locations of glider profiles during the experiment (solid dots) and model grid (smaller dots). The bathymetry contours are in m. The domain used for estimation of the distances r^s_ξ between the model states in twin-data experiments is shown by the solid black line. Circles denote locations of the two moorings used for validation of the real-data experiments.

a. *Numerical model, observations, and validation technique*

To simulate oceanic variability during the experiment we used a version of NCOM forced by the Coupled Ocean–Atmosphere Mesoscale Prediction System (COAMPS; Hodur et al. 2002) winds in the time period between 1 August ($t = 0$) and 27 August ($t = 27$) of 2003. The model was configured on a curvilinear orthogonal grid (Fig. 2) with horizontal resolution ranging from 1 to 4 km, and a hybrid σ/z vertical coordinate system with 9 σ levels in the upper ocean and 32 z levels below. At the open boundaries, the model was one-way coupled to the global NCOM model (Shulman et al. 2009).

Glider observations during the experiment covered the central part of the model domain (Fig. 2). With a typical dive cycle of about 1 h, a glider would travel approximately 0.5 km between surfacings, which is well below the grid resolution. For that reason we prescribed observational operators $\mathbf{H}$ to measure instantaneous vertical profiles of temperature and salinity at the model grid point closest to the average of the surface locations of a glider before and after a dive. In the assimilation experiments we used a 12-h analysis cycle, so only those glider profiles occurring within 1-h window around 0000 and 1200 PST were assimilated. On average, the model domain was covered by 20–40 profiles every 12 h.

To measure distances between the model states, a diagonal metric $\mathbf{g}$ was used. The diagonal elements of $\mathbf{g}$ (g_T, g_S, g_u, g_v, and g_ζ) were depth dependent and were obtained as horizontally averaged time variances of temperature T, salinity S, horizontal velocity $\{u, v\}$, and SSH ζ, respectively, at a grid point $\mathbf{r}$:

$$g_\xi(z) = \langle \overline{[\xi(\mathbf{r}) - \overline{\xi(\mathbf{r})}]^2}^{1/2} \rangle_z.$$

In the above equation, ξ stands for either T, S, u, v, or ζ and angular brackets denote the horizontal average at level z.

Distances r^s and r^g between the model states were computed in both observational and state spaces:

$$r^s_\xi(\mathbf{x}_1, \mathbf{x}_2) = \langle (\xi_1 - \xi_2)^2 g_\xi^{-2} \rangle^{1/2};$$
$$r^g_\xi(\mathbf{x}_1, \mathbf{x}_2) = \langle (\xi_1 - \xi_2)^2 R_\xi^{-1} \rangle_g^{1/2}. \quad (19)$$

Here the angular brackets denote averaging over the 3D model domain covered by gliders (Fig. 2) and over the glider locations $\mathbf{r}^g_k$, respectively.

In the twin-data experiments, glider "observations" of temperature $\mathbf{y}_T$ and salinity $\mathbf{y}_S$ were extracted from the "true" fields T^t, S^t (Fig. 3, left panel) at glider locations $\mathbf{r}^g_k$ every 12 h and contaminated by white noise ε with zero mean and 0.1 rms variation:

$$\mathbf{y}_T = \mathbf{H}(T^t + \varepsilon g_T); \quad \mathbf{y}_S = \mathbf{H}(S^t + \varepsilon g_S).$$

To simulate model errors and assess the impact of assimilation in the twin-data experiments, the "first guess" model solution $\mathbf{x}^{\mathrm{fg}}(t)$ was generated by integrating the model for 27 days starting from the initial condition specified by $\mathbf{x}^t(t = 8.5)$ (Fig. 3, right panel).

Using $\mathbf{x}^{\mathrm{fg}}(0)$ as the background state at $t = 0$, a series of 3DVar assimilation experiments were performed: on every 12-h assimilation cycle, model forecasts $\mathbf{x}^f$ were updated with the analyses increment $\mathbf{x}^a = \mathbf{x}^f + \delta\mathbf{x}$ and the next 12-h integration was started from $\mathbf{x}^a$. The skill of assimilation $q(t)$ was assessed in both observational and state spaces by calculating the normalized distances between the 12-h model forecasts and the true states:

$$q^{g,s}_\xi(t) = \frac{r^{g,s}_\xi(\mathbf{x}^t, \mathbf{x}^f)|_t}{r^{g,s}_\xi(\mathbf{x}^t, \mathbf{x}^{fg})|_0}. \quad (20)$$

Experiments with real data were conducted and the results were validated in a similar manner, except that $\mathbf{x}^t$ was taken as the first guess and q^s_ξ values were not computed because the true state was unknown. Instead of q^s_ξ, we assessed the forecast skill of the model using independent temperature, salinity, and velocity observations at

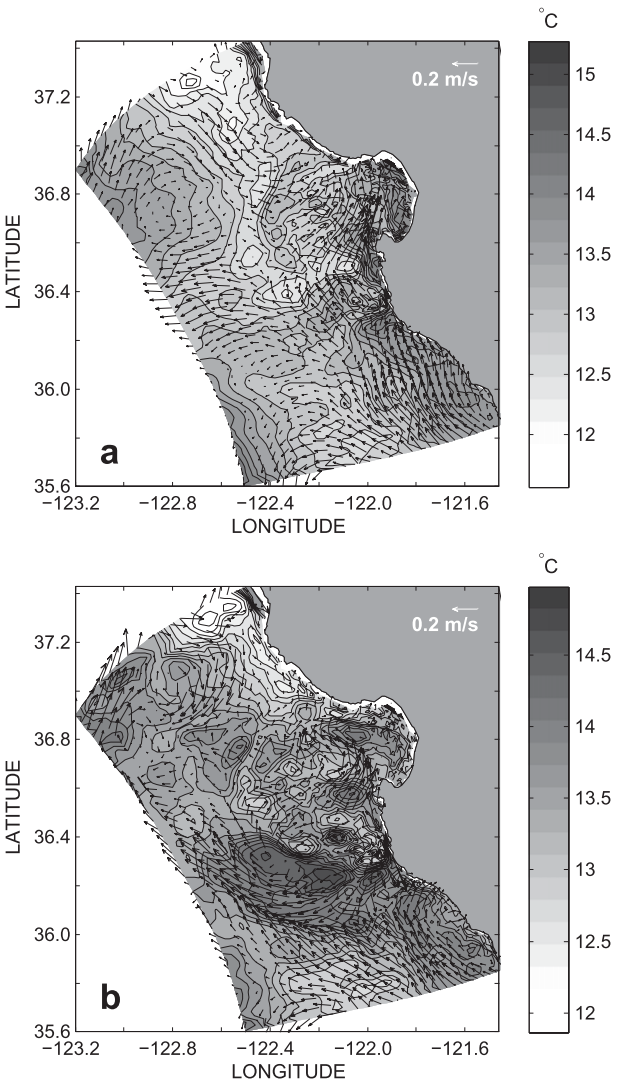

FIG. 3. An example of temperature and velocity fields for the (a) true and (b) first-guess solutions ($z = 28$ m, $t = 1.5$ days).

two moorings shown in Fig. 2. The respective distances r_ξ^m and skills q_ξ^m were computed similarly to (19) and (20), but the spatial average was taken over two moorings and the differences between variables in (19) were normalized by the respective rms temporal variations observed at the moorings.

b. Parameters of the hybrid covariance model

In contrast to atmospheric applications, regional oceanographic problems have more difficulties with the BEC estimation from ensembles. The reason is that realistic ensembles simulating ocean variability on regional scales are rarely available. In the present study, the first-guess background error statistics was obtained from the ensemble of the differences $\delta\mathbf{x}_l^0$ between the first-guess states $\mathbf{x}_l^{\mathrm{fg}}$ at times enumerated by l and the respective 12-h forecasts (background states) $\mathbf{x}_l^f$ derived from the assimilation run with $\alpha = 0$. The latter were treated as a rough approximation to the ensemble of the true ocean states. Since the number of ensemble members ($l = 1, \ldots, 55$) was limited by the duration of the glider survey, the expected number of statistically sensible eigenvectors of the ensemble covariance matrix was rather small and never exceeded 2 (see section 4).

In the hybrid assimilation runs (with $\alpha \neq 0$) this first-guess ensemble $\{\delta\mathbf{x}_l^0\}$ of the background errors was continuously updated: its members on the time layers l preceding the current analysis time were replaced by the members derived from the forecasts already made with the hybrid scheme. The quality of updated ensembles was monitored by the number of eigenvectors accepted by the information criterion (10) and by the percentage of $\delta\mathbf{y}$ variance explained by these eigenvectors.

To increase the robustness in estimating α and β, we utilized the method of Wang et al. (2007) and performed additional time averaging while computing the sample variances in (13) and (15). This averaging was done over the ensemble of 30 states (15 days) preceding the analysis time. In the initial 15 days of the assimilation run, the missing background states were taken from the respective forecasts $\mathbf{x}^{\mathrm{fg}}(t)$ generated by the first-guess solution. Similar averaging over $N = 30$ samples following the analysis time was done when estimating σ_m in (11).

There are two parameters in the definition of $\mathbf{B}_0$ that may affect the performance of the assimilation scheme. One is the "time of integration" τ controlling decorrelation length scales, and the other is the order n of approximation of the exponent in (17). These parameters were tuned by numerical experimentation.

Comparison of the model solutions (Fig. 3) with the grid (Fig. 2) gives an indication that horizontal correlations are likely to decay at 3–6 grid steps. We checked this hypothesis by twin experiments with $\alpha = 0$ and computed the forecast skill of the assimilated solutions with various values of $\sqrt{\tau}$. In these experiments, n was also varied in the range between 1 and 4. The best overall result was obtained with $\tau = 20$ ($\rho_i \sim 4.5\Delta x_i$) and $n = 2$. Although assimilation quality (with $\tau = 15$, $n = 3$) was similar and in some periods slightly better, the computational cost appeared to be much larger. We therefore selected $\tau = 20$, and $n = 2$ as basic parameters for the assimilation experiments.

4. Results

a. Twin-data experiments

Figure 4 compares the skill of 3DVar assimilation runs performed with the Gaussian and hybrid BEC models. During the first 8 days of assimilation, the hybrid scheme

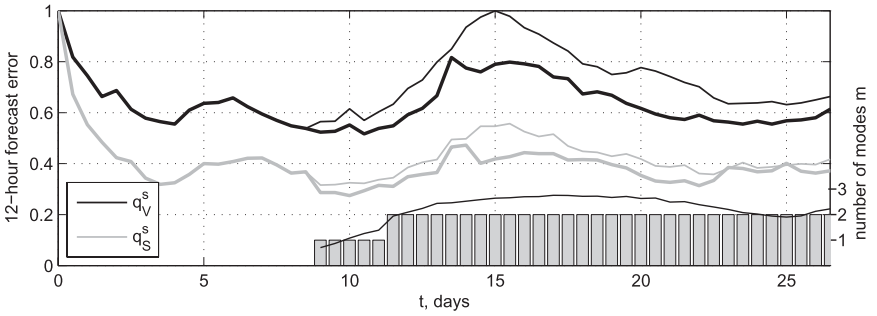

FIG. 4. Normalized distances q^s_S and q^s_v between the true solution and the 12-h forecasts of assimilated solutions with $\alpha = 0$ (thin lines) and $\alpha \neq 0$ (hybrid model, thick lines). The plot below shows the number of detected eigenvectors (bars, right axis) and the portion of the model/data misfit variance explained by those modes (thin line above the bars, left axis).

was unable to detect any statistically reliable modes. Between days 8 and 11 the first mode was detected, accounting for 8% of the forecast error on day 8, 14% on day 11 (Fig. 4), and 17% on day 17. On day 12 the second mode was detected, accounting for 4% of the forecast error variance. Contribution of the second mode increased to almost 10% on day 18. Later, the modes appear to lose their predictive skill with the contributions dropping to 12% and 7%, respectively, on day 25.

The 12-h forecast errors measured in terms of the normalized distances q^s_ξ from the true state are found to be approximately 15% smaller than for the assimilation run with $\alpha = 0$ (thin lines in Fig. 4). A similar level of error reduction was observed by Wang et al. (2008) in twin-data experiments with a hybrid assimilation into the Weather Research and Forecasting (WRF) model. In terms of the 12-h forecasts of temperature and salinity in glider observation points the error reduction is somewhat smaller (11% for temperature and 13% for salinity), but can still be considered as a satisfactory improvement (Fig. 5).

Assimilation experiments with different noise in observations have shown that the patterns in Figs. 4–5 are robust up to the noise levels of 0.5. At higher noise levels, the approximation (13) becomes less accurate and it is necessary to use the relationship (12) for estimating α. Larger errors in estimating α result in the loss of accuracy in estimating the number of modes m and the magnitude β of the Gaussian part of the covariance. We therefore assume that the proposed algorithm is valid when observation errors are considerably smaller than the background errors. This is not a severe restriction for regional assimilation problems in oceanography where the first-guess/background model solutions are rarely preconditioned by data and often appear to be rather far from reality.

b. Real-data experiments

Figure 6 shows a typical situation we encountered in the experiments with real data in the Monterey Bay: The first-guess model solution does not have much in common with the mooring record at 40 m (left panel). Moreover,

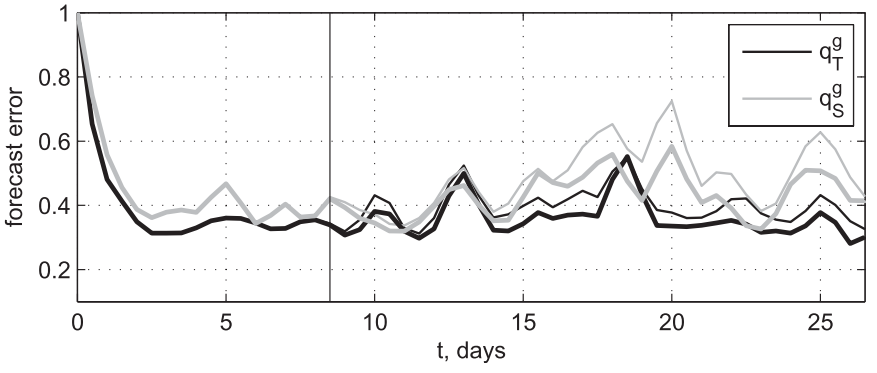

FIG. 5. Normalized distances q^g_T and q^g_S between the true solution and the 12-h forecasts of assimilated solutions with $\alpha = 0$ (thin lines) and $\alpha \neq 0$ (hybrid model, thick lines). Distances in observational space are normalized by the value at $t = 0$. The thin vertical line marks the detection time of the first mode (see Fig. 4).

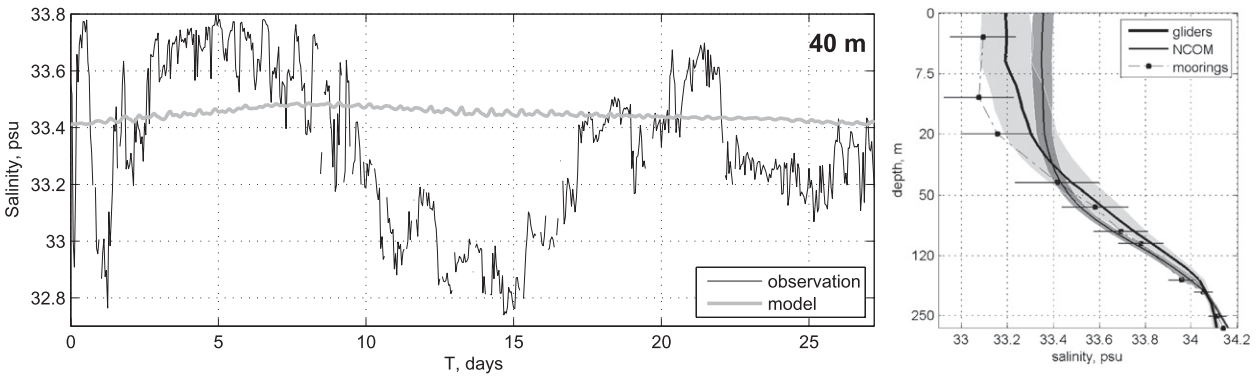

FIG. 6. (left) Salinity recorded by the offshore mooring in Fig. 2 (black line) and the corresponding salinity of the first-guess NCOM solution (gray line). (right) Profiles of the average salinity measured by gliders (solid bold line), moorings (dashed line), and extracted from the first-guess model solution (solid thin line). Shading and horizontal bars show rms variability.

the mean profiles in the right panel demonstrate considerable salinity biases above 30 m and in the depth range between 50 and 200 m. The rms variations of salinity measured by gliders and moorings are generally consistent with each other in magnitude (cf. horizontal bars and the width of light shading around the thick profile in the right panel). A noticeable bias between the mean salinity measured by moorings (solid dots) and gliders (thick line) could be attributed to differences in averaging: the glider profile is obtained by averaging over all the glider positions (Fig. 2), whereas the mooring profile (solid dots) is obtained as the mean of only two moorings. Similar biases between the first-guess solution and observations were obtained for the temperature field (not shown).

To estimate observation errors, we compiled the glider *T*/*S* records at times when gliders passed closer than 200 m of either of the moorings and compared these data with the corresponding observations at moorings. In total, 168 of such ''pass-by events'' were found. Comparison of these observations has shown that the rms discrepancies in temperature and salinity were fairly stable with depth and varied within 0.26–0.35 after normalization by the rms variances $\sigma_m(z)$ recorded at the moorings (horizontal bars in the right panel of Fig. 6). Based on these computations, the observation error variances were estimated as $R^{1/2}(z) = 0.3\sigma_m(z)$ and assumed not to vary in the horizontal.

In all other aspects (the first guess, the background error variance $\mathbf{G}$, etc.), assimilation experiments with real data were configured in the same way as the twin-data experiments. Because the ''true ocean state'' in the real-data experiments was unavailable, we introduced an additional parameter r_ξ^m to gauge the algorithm's performance. Similar to r_ξ^g r_ξ^m was computed as the normalized distance between the model forecast field ξ^f and temperature, salinity, or velocity ξ^m measured at the points of moored observations:

$$r_\xi^m = \langle (\xi_f - \xi^m)^2 \sigma_m^{-2} \rangle^{1/2}.$$

Angular brackets denote averaging in the vertical (11 levels of *T*/*S* observations or 18 levels of ADCP data) and over two moorings shown in Fig. 2.

Figure 7 demonstrates the differences δq between the salinity forecast skills $q_S^{m,g}$ of the assimilation run with $\alpha = 0$ and similar forecast skills obtained with the hybrid BEC model. Although the skill improvement does not look as good as in twin-data experiments (Fig. 4), it appears to be robust: the differences in skill δq_S^m and δq_S^g remain positive for most of the time after detection of the first mode on day 3. The time mean values for δq_T^m, δq_v^m, and δq_T^g were found to be 1.3%, 2.8%, and 2.0%, respectively.

Compared with the time-averaged assimilation skill in twin-data experiments (e.g., $\overline{q}_S^s \sim 0.4$ for salinity in Fig. 4), the values of $\overline{q}_\xi$ in real-data experiments were much higher (0.6–0.7 for temperature/salinity and 0.9 for velocity). This difference is due to larger observation noise, its more complex structure, and a considerable bias (left panel in Fig. 6) inconsistent with the prior statistical assumptions. The Bayesian algorithm (10) indicated an occasional presence of only one informative mode: detection events disappeared on day 14 and reemerged only at the end of the assimilation period (Fig. 7). Such behavior could be attributed to the poor quality of the first-guess solution and insufficient statistics of the 30-member ensemble in use. Experiments with changing the ensemble size n_e (20–55 members) have shown that with $n_e = 20$, the number of detection events dropped to 3, whereas with $n_e = 50$, it increased from 30 (Fig. 7) to 35 without any substantial improvement of the forecast skill. Further

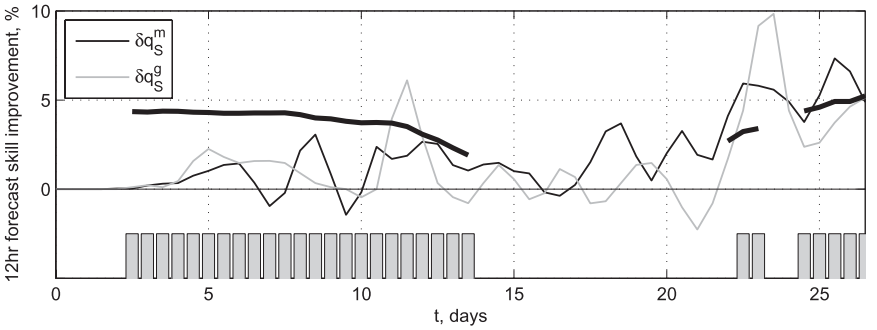

FIG. 7. Improvement of the 12-h salinity forecasts at glider observation points (gray line) and at the moorings (black line). Positive values correspond to smaller forecast errors for the hybrid scheme. Vertical bars indicate occasional detection of only one mode and the thick black line shows the percentage of the error variance that the mode explains.

increase of the ensemble size was limited by the duration of the assimilation experiments.

In principle, the ensemble could be expanded, for example, by the breeding technique, but the problem with the poor quality of the first-guess solution (Fig. 6) may still persist, because the bred vectors would still show unstable modes of the background state that is rather far from reality. In the present study, we used a simple approximation to the error fields by considering an ensemble of differences between a free model run and an assimilation run with the Gaussian covariance model. This ensemble was able to generate just a few members in the 27-day period. One may hope, however, that for longer observation periods the BEC model will gain enough skill to show better performance.

Figure 8 shows the time evolution of the ratio between the weighting parameters α and β in the twin- and real-data experiments. By an order of magnitude, the ratio γ is consistent with the results of Wang et al. (2008) who set $\gamma = \beta/\alpha = \text{const}$ in time and found the optimal γ to vary between 1 and 4 in a series of twin-data experiments with the WRF model. In our case, the relative weight β of the Gaussian term in the cost function appeared to be approximately 2 times smaller in the twin-data experiment (thin curve in Fig. 7). This is consistent with a better skill in explaining model-data misfits by the modes retrieved in the twin-data experiment (cf. Figs. 4 and 7). Larger relative values of α on days 10–13 (before the mode rejection) can be explained by the tendency of the algorithm to keep the deteriorating mode ''alive.''

We also investigated the impact of the algorithms for definition of m and α on the forecast skill. In the twin-data experiments with fixed m the 27-day-averaged skill was always worse than that in Fig. 4 for 3 tested values of $\gamma =$ 0.5, 1, and 2. When m was computed through (10) and γ was kept constant at 0.95, the forecast skill was virtually the same as in Fig. 4, but somewhat below using other values of γ. Similar results were obtained with real data: keeping $m = \text{const}$ degraded the forecast skill, often below the one obtained with Gaussian BEC model. Several runs with an adjustable m and $\gamma = \text{const}$ were difficult to interpret as the skill improvements were small, highly variable, and did not show any deterministic dependence on the value of $\gamma \in [0.5, 2.5]$.

5. Summary and discussion

In this study we proposed a hybrid BEC model specifically designed for 3DVar analysis of regional circulations supported by glider surveys. The model is supplied by an algorithm for weighting the ensemble-generated error covariance $\mathbf{B}_m$ against the heuristic covariance $\mathbf{B}_0$ represented by the propagator of the diffusion equation. Another distinctive feature of the algorithm is the detection of the statistically confident eigenvectors of $\mathbf{B}_m$ by means of the Bayesian information criterion (Schwarz 1978). The method is based on the assessment of the modes' skill in approximation of the forecast error fields accumulated in the course of the assimilation run.

The proposed BEC model is formulated in terms of the inverse covariances with the restriction of $\mathbf{B}_0$ to the null space of $\mathbf{B}_m$. This is done to better preserve the covariances detected by the information criterion and captured by $\mathbf{B}_m$. Formulation of the minimization problem in the state space allows us to gain extra computational efficiency by approximating the action of $\mathbf{B}_0$ via a semi-implicit scheme.

The hybrid BEC model was validated by numerical experiments with simulated and real data. In the twin-data setting, the hybrid formulation was capable of improving the model's forecast skill by 15%–20%, which is comparable with the improvement reported by Wang et al. (2008) for a hybrid scheme with the atmospheric WRF model. Results of the experiments with real data

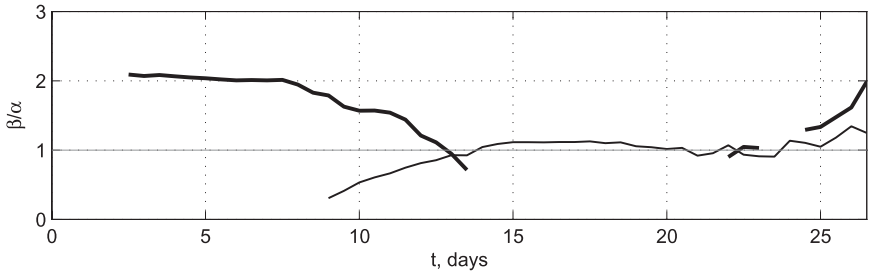

FIG. 8. Time variation of the ratio β/α in the twin-data (thin line) and real-data experiments.

showed a few percent improvement with sporadic detection of only one mode. We attribute this to a poor quality of the background solution, which was heavily biased and demonstrated considerably lower time variation in the temperature and salinity fields (left panel in Fig. 6). Thus, finding a better background solution appears to be the first priority in upcoming studies of the algorithm.

Other developments may include elaboration of the structure of the diffusion tensor in $\mathbf{B}_0$ and improvement of the ensemble generation technique. In particular, diffusion could be enhanced along the f/H contours and/or isopycnals of the geostrophically balanced modes if the latter are detected. The ensemble could be enriched by the vectors bred from the eigenvectors of $\mathbf{B}_m$ or just using a standard breeding technique, if the background state acquires a reasonable forecast skill in the course of assimilation. Finally, detected eigenvectors can be propagated by the model with the methods used in Kalman filtering schemes. This approach will naturally combine the advantages of statistical and dynamical methods and increase versatility of the hybrid algorithm.

One of the drawbacks of the proposed model is the computational cost of estimating the weight β of the static covariance [(15)]. Our experience shows, however, that the numerator in (15) weakly depends on the structure of $\mathbf{H}$ for a given number of eigenvectors and can be efficiently parameterized by a linear function of the number of observations. In fact, the predictive skill of the system did not change when such a linear parameterization was used. A similar kind of parameterization could also be employed to estimate α and β when the background errors are comparable with the observation errors.

The benefit of the proposed hybrid model may also be diminished for global assimilation problems where some sort of localization is needed and the impact of the ensemble-generated covariances may be smaller with higher observation density (e.g., Whitaker et al. 2008). Nevertheless, we assume that the proposed approach may have a prospect for further development for regional data assimilation problems with poorly known background states.

Acknowledgments. This study was supported by the Office of Naval Research (Program Element 0602435N) and NSF Grant 0629400. Helpful discussions with Dr. Shulman are acknowledged.

APPENDIX

Derivation of Equation (12)

Consider the optimization problem (1) in $\mathcal{R}^m \equiv \mathrm{span}\{\mathbf{e}_k\}$ by introducing a new variable $\delta\mathbf{e}$ such that $\delta\mathbf{x} = \mathbf{P}\delta\mathbf{e}$:

$$J = \frac{1}{2}[\delta\mathbf{e}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}}\mathbf{B}^{-1}\mathbf{P}\delta\mathbf{e} + (\mathbf{HP}\delta\mathbf{e} - \delta\mathbf{y})^{\mathrm{T}} \times (\mathbf{HP}\delta\mathbf{e} - \delta\mathbf{y})] \rightarrow \min_{\delta\mathbf{e}\in\mathcal{R}^m}. \quad \text{(A1)}$$

The normal equation (3) is now reduced to

$$[\mathbf{P}^{\mathrm{T}}\mathbf{B}^{-1}\mathbf{P} + \mathbf{Q}]\delta\mathbf{e} = \mathbf{E}\delta\mathbf{y}, \quad \text{(A2)}$$

where the operator $\mathbf{E} \equiv \mathbf{P}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}$ projects the data on $\mathcal{R}^m$ and $\mathbf{Q} = \mathbf{E}\mathbf{E}^{\mathrm{T}}$. Substituting the adopted inverse covariance model (7) into (A2) and taking into account the identities $\mathbf{P}^{\mathrm{T}}\mathbf{P} = \mathbf{I}_m$, $\mathbf{P}_{\perp}\mathbf{P} = 0$, the normal equation is further simplified to

$$[\alpha\mathbf{\Lambda}_m^{-1} + \mathbf{Q}]\delta\mathbf{e} = \mathbf{E}\delta\mathbf{y}. \quad \text{(A3)}$$

Introducing the notation $\mathbf{Y} = \langle\delta\mathbf{y}\delta\mathbf{y}^{\mathrm{T}}\rangle$ for $\delta\mathbf{y}$ covariance, covariances of $\delta\mathbf{e}$ could be written as follows:

$$\langle\delta\mathbf{e}\delta\mathbf{e}^{\mathrm{T}}\rangle = (\alpha\mathbf{\Lambda}_m^{-1} + \mathbf{Q})^{-1}\mathbf{E}\mathbf{Y}\mathbf{E}^{\mathrm{T}}(\alpha\mathbf{\Lambda}_m^{-1} + \mathbf{Q})^{-1}. \quad \text{(A4)}$$

On the other hand, in accordance with the observation model, misfits $\delta\mathbf{y}$ between the background state and the data have the following covariance:

$$\mathbf{Y} = \mathbf{HBH}^{\mathrm{T}} + \mathbf{I}_K, \quad \text{(A5)}$$

which, after projecting on the eigenvectors $\mathbf{e}_k$, is

$$\mathbf{EYE}^{\mathrm{T}} = \mathbf{EHBH}^{\mathrm{T}} + \mathbf{EE}^{\mathrm{T}} = \frac{1}{\alpha}\mathbf{Q}\boldsymbol{\Lambda}_m\mathbf{Q}^{\mathrm{T}} + \frac{1}{\beta}\mathbf{EH}[\mathbf{P}_{\perp}\mathbf{B}_0^{-1}\mathbf{P}_{\perp}]^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{E}^{\mathrm{T}} + \mathbf{Q}. \tag{A6}$$

Consider now the middle term in the rhs of (A6):

$$\mathbf{EH}[\mathbf{P}_{\perp}\mathbf{B}_0^{-1}\mathbf{P}_{\perp}]^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{E}^{\mathrm{T}} = \mathbf{P}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{H}[\mathbf{P}_{\perp}\mathbf{B}_0^{-1}\mathbf{P}_{\perp}]^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{HP} \equiv \tilde{\mathbf{Q}}.$$

For ''pointwise'' observations (local observational operators) the matrix $\mathbf{H}^{\mathrm{T}}\mathbf{H}$ is equal to the identity matrix $\mathbf{I}_M$ with diagonal elements masked by zeroes in the points without observations. Therefore, in the limit of a perfectly observed state ($\mathbf{H}^{\mathrm{T}}\mathbf{H} = \mathbf{I}_M$), this term vanishes. For glider observations, which densely populate the domain during the survey, one may assume that $\langle \mathbf{H}^{\mathrm{T}}\mathbf{H} \rangle \sim \mathbf{I}_M$ and neglect this term in the time average. We checked the validity of this assumption by estimating the ratio $|\tilde{\mathbf{Q}}|/|\mathbf{Q}|$ for all the values of $\mathbf{H}(t)$ and several $\mathbf{P}$s containing the first 10 eigenmodes. The ratio was found to be on the order of 10^{-2}, allowing the relationship in (A6) to be reduced to

$$\mathbf{EYE}^{\mathrm{T}} = \frac{1}{\alpha}\mathbf{Q}\boldsymbol{\Lambda}_m\mathbf{Q}^{\mathrm{T}} + \mathbf{Q}. \tag{A7}$$

Substitution of (A7) into (A4) yields (12).

REFERENCES

Akaike, H., 1974: A new look at the statistical model identification. *IEEE Trans. Automat. Contrib.,* **19** (6), 716–723.

Bai, Z., and B. H. Golub, 1997: Bounds for the trace of the inverse and the determinant of symmetric positive definite matrices. *Ann. Numer. Math.,* **4,** 29–38.

Brasseur, P., and J. Verron, 2006: The SEEK filter method for data assimilation in oceanography: A synthesis. *Ocean Dyn.,* **56,** 650–661.

Dobricic, S., and N. Pinardi, 2008: An oceanographic three-dimensional variational data assimilation scheme. *Ocean Modell.,* **22,** 89–105.

——, ——, P. Testor, and U. Send, 2010: Impact of data assimilation of glider observations in the Ionian Sea (Eastern Mediterranean). *Dyn. Atmos. Oceans,* **50,** 78–92, doi:10.1016/j.dynatmoce.2010.01.001.

Etherton, B. J., and C. H. Bishop, 2004: Resilience of hybrid ensemble/3DVAR analysis schemes to model error and ensemble covariance error. *Mon. Wea. Rev.,* **132,** 1065–1080.

Evensen, G., 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.,* **53,** 343–357, doi:10.1007/s10236-003-0036-9.

Hamill, T. M., and C. Snyder, 2000: A hybrid ensemble Kalman filter–3D variational analysis scheme. *Mon. Wea. Rev.,* **128,** 2905–2919.

——, J. S. Whitaker, and C. Snyder, 2001: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.,* **129,** 2776–2790.

Hannah, E. J., and B. G. Quinn, 1979: The determination of the order of an autoregression. *J. Roy. Stat. Soc.,* **41B,** 190–195.

Heaney, K. D., G. Gawarkiewicz, T. F. Duda, and P. F. J. Lermusiaux, 2007: Non-linear optimization of autonomous undersea vehicle sampling strategies for oceanographic data-assimilation. *J. Field Robot.,* **24** (6), 437–448.

Hodur, R. M., J. Pullen, J. Cummings, X. Hong, J. D. Doyle, P. J. Martin, and M. A. Rennick, 2002: The coupled ocean/atmospheric mesoscale prediction system (COAMPS). *Oceanography,* **15,** 88–98.

Li, Z., Y. Chao, J. C. McWilliams, and K. Ide, 2008: A three-dimensional variational data assimilation scheme for the regional ocean modeling system. *J. Atmos. Oceanic Technol.,* **25,** 2074–2090.

Pannekoucke, O., and S. Massart, 2008: Estimation of the local diffusion tensor and normalization for heterogeneous correlation modelling using a diffusion equation. *Quart. J. Roy. Meteor. Soc.,* **134,** 1425–1438.

Ramp, S. R., and Coauthors, 2008: Preparing to predict: The second Autonomous Ocean Sampling Network (AOSN-II) experiment in the Monterey Bay. *Deep-Sea Res. II,* **56,** 68–86, doi:10.1016/j.dsr2.2008.08.013.

Rudnick, D. L., R. E. Davis, C. C. Eriksen, D. M. Fratantoni, and M. J. Perry, 2004: Underwater gliders for Ocean Research. *J. Mar. Technol. Soc.,* **38,** 73–84.

Saad, Y., 2003: *Iterative Methods for Sparse Linear Systems*. 2nd ed. Society for Industrial and Applied Mathematics, 528 pp.

Schwarz, G. E., 1978: Estimating the dimension of a model. *Ann. Stat.,* **6** (2), 461–464.

Shulman, I., and Coauthors, 2009: Impact of glider data assimilation on the Monterey Bay model. *Deep-Sea Res. II,* **56,** 188–198, doi:10.1016/j.dsr2.2008.08.003.

Tippett, M., J. L. Anderson, C. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble square root filters. *Mon. Wea. Rev.,* **131,** 1487–1490.

Wang, X., T. M. Hamill, J. S. Whitaker, and C. Bishop, 2007: A comparison of hybrid ensemble transform Kalman filter-OI and ensemble square-root filter analysis systems. *Mon. Wea. Rev.,* **135,** 1055–1076.

——, D. M. Barker, C. Snyder, and T. M. Hamill, 2008: A hybrid ETKF-3DVAR data assimilation scheme for the WRF model. Part I: Observing System Simulation Experiment. *Mon. Wea. Rev.,* **136,** 5116–5131.

——, T. M. Hamill, J. S. Whitaker, and C. H. Bishop, 2009: A comparison of the hybrid and EnSRF analysis schemes in the presence of model error due to unresolved scales. *Mon. Wea. Rev.,* **137,** 3219–3232.

Weaver, A., and P. Courtier, 2001: Correlation modeling on a sphere using a generalized diffusion equation. *Quart. J. Roy. Meteor. Soc.,* **127,** 1815–1846.

——, and S. Ricchi, 2004: Constructing a background-error correlation model using generalized diffusion operators. *Proc. ECMWF Seminar Series on Recent Advances in Atmospheric and Ocean Data Assimilation,* Reading, United Kingdom, ECMWF, 327–340.

Whitaker, J. S., T. M. Hamill, X. Wei, Y. Song, and Z. Toth, 2008: Ensemble data assimilation with the NCEP Global Forecast System. *Mon. Wea. Rev.,* **136,** 463–482.

Yin, Y., O. Alves, and P. Oke, 2011: An ensemble ocean data assimilation system for seasonal prediction. *Mon. Wea. Rev.,* **139,** 786–808.