Application of an inverse method to coastal modelling

Francis Auclair, Sophie Casitas, Patrick Marsaleix

To cite this version:

Francis Auclair, Sophie Casitas, Patrick Marsaleix. Application of an inverse method to coastal modelling. Journal of Atmospheric and Oceanic Technology, 2000, 17 (10), pp.1368-1391. 10.1175/1520-0426(2000)0172.0.CO;2 . hal-00160512

Application of an Inverse Method to Coastal Modeling

F. AUCLAIR, S. CASITAS, AND P. MARSALEIX

Laboratoire d'Aérologie, UMR CNRS/UPS 5560, Observatoire Midi-Pyrénées, Toulouse, France

(Manuscript received 21 October 1998, in final form 17 June 1999)

ABSTRACT

Free surface coastal models currently suffer from the difficulty of having to specify the global circulation during the initialization process and along the open boundaries. As an alternative to the long spinup periods, an original explicit approach based on inverse techniques has been developed. Data originating from in situ observations and/or ocean general circulation models are optimally interpolated over the small-scale grid in such a way that the tendency terms are reduced to physically consistent values. The errors on the ''true'' tendencies and the truncated nonlinear term are evaluated to compute the model covariance. The observation covariance matrix is divided into two parts: the homogeneous, isotropic matrix calculated with a global energy spectrum; and a parameterized nonhomogeneous, nonisotropic matrix. The inverse method is applied to the study of the interaction of a barotropic alongshore current over a narrow canyon. The transient processes following the initialization are drastically reduced and the analysis field can efficiently be used in a flow relaxation scheme along the open boundaries.

1. Introduction

Thanks to the numerous in situ observations now available and to the global synoptic data provided by satellites (Busalacchi 1997), oceanographic models can offer a time-dependent realistic map of large-scale currents, from the surface of the oceans, to their abyss. We have a fairly good idea of the main synoptic characteristics of their general circulation with a spatial resolution of a few minutes of degrees: description of the different water masses, position, and evolution of the main gyres (Bryan et al. 1995; Horton et al. 1997). However, computational costs remain prohibitive as soon as we are interested in meso- and small-scale dynamics. A resolution of only a few kilometers is, for instance, necessary to represent the first Rossby radii. One solution is to model only the most salient parts of the basin: the region of a jet where instabilities are most likely to develop, the interactions of the general circulation with small-scale topographic features. However, such a model is usually integrated over a small domain and, as a consequence, cannot generate the general circulation on its own. Different approaches have been proposed to overcome the problem. Spall and Holland (1991) developed a method based on a multidomain method. A general circulation model is dynamically coupled with a smaller-scale model. In this approach, both models are iterated in parallel and exchange data both ways (feedback). In particular the small-scale model is well forced by a time-dependent general circulation. However, the cost of these simulations is rather expensive, and the small-scale model cannot be used with different global models that would be well adapted to the ocean considered.

If the models are not dynamically linked together, it is necessary to specify the initial fields (velocity, temperature, salinity, surface elevation anomaly, . . .) and as importantly, the open boundary conditions. However, the initialization and boundary condition problems can be particularly difficult. A crude, unbalanced specification of the initial state can have numerous troublesome consequences, and the unrealistic motions exhibited by the numerical models during the initialization transient process were early recognized to be mostly inertial-gravity waves (Hinckelman 1951), and a long spinup would very often be required before some physically meaningful circulation be obtained. The treatment of this problem gave rise to a large literature in meteorology: see, for instance, Machenhauer (1977) for a nonlinear normal mode initialization, Bourke and McGregor (1983) for a physical-space initialization, or Courtier and Talagrand (1990) for a variational approach.

Due to the rigid-lid approximation, most large-scale oceanographic models filter out the quickly propagating inertia-gravity waves (Malanotte-Rizzoli et al. 1989), but coastal free-surface models do not and have to be

---

Corresponding author address: Dr. Francis Auclair, Laboratoire d'Aérologie/Observatoire Midi-Pyrénées, UMR CNRS/UPS 5560, 14, Av. E. Belin, 31400 Toulouse, France.
E-mail: aucf@aero.obs-mip.fr

carefully iterated over a period of adjustment (spinup) before some kind of well-behaved equilibrium state be reached for the general circulation. However, these fast waves are known to have undesirable consequences: partial destruction of the initial field, or divergence of the model. Indeed, the baroclinic part of the free-surface model is very sensitive to high-frequency motions (Woodgate and Killworth 1997), and in certain conditions, the consequences can be catastrophic. Coastal regions are particularly favorable to the generation of these transient problems. Indeed, the coastal steep slopes produce mesoscale noise when data are interpolated on the model grid. This noise generates unrealistic patterns that contaminate the whole domain. Woodgate and Killworth (1997) and Woodgate (1997) proposed a solution based on a ''nudging'' method. They showed that the addition of nudging relaxation terms in the primitive equations had important consequences on the period of the inertia gravity waves modelized by a free-surface model. The lowering of their frequency removed most of the initialization problems, the price to be paid being a loss of physical realism during the spinup period.

Another crucial problem to be addressed is the specification of the open boundary conditions. Open boundary conditions must fill a double objective: prevent the reflection of the various types of waves generated inside the domain, and give some hints about the large-scale circulation. Radiative boundary conditions, as prescribed by Orlanski (1976), offer an efficient answer to the spurious reflections of waves. Although Bennett and Kloeden (1978) showed that the open boundary problem was in fact ill-posed, Stevens (1990) proposed a complete set of boundary conditions for a rigid-lid model, in which a distinction is made between passive and active boundary conditions. Another approach was chosen by Davies (1976) or Engedahl (1995), namely, the Flow Relaxation Scheme (FRS). Small bands are added along the open boundary over which the interior flow is relaxed toward the exterior general circulation. However, all these methods have in common the necessity to have a good knowledge of the exterior general circulation, and to specify it in such a way that it agrees with the model dynamics. Gunson and Malanotte-Rizzoli (1996) lately showed that this problem could not be separated from the specification of the initial state.

It is therefore indubitable that before any type of observations be assimilated in a free-surface model, a trade must be reached between the realism of these observations and their physical consistency with the model. Different approaches have been developed in the last decade to clearly define this optimal state and help reconstruct data fields from a few observations. Following Sasaki's approach (Sasaki 1958), the smoothing spline methods all seek for such an optimal state (in a least squares sense). It must be close enough to the available data, and satisfies some physical constraints, the uniqueness of the solution being obtained with the addition of a smoothing operator. A clear mathematical formulation and some analytical solutions for this problem were provided by Wahba and Wendelberger (1980). Provost and Salmon (1986) introduced this spline-based method in oceanography using both the Sverdrup and geostrophic balance associated to a smoothing operator to constraint the interpolation of observations. Brasseur (1991) added an advection constraint to reconstruct the general circulation fields in the northern Bering Sea. On the other hand, the objective analysis method (Gandin 1965) uses the statistical variance and covariance of the observations to determine the relative weights to be given to the data and to the constraints. This approach is based on the Gauss–Markov theorem and leads to a minimum-variance or a best linear unbiased estimate (BLUE). Bretherton et al. (1976) give a method to evaluate the statistical mean, which provides the smallest minimum error variance in a stationary case, the nonstationary case being studied by Le Traon (1990). A scale-based approach of the objective analysis method was followed by Ooyama (1987). He showed that the geometrical distribution of the observing stations was determining the resolvable scales and that a precise knowledge of the second-moment statistics improved the analysis by dealiasing the amplitude of the resolvable scales but has no effect on the definition of the unresolvable scales. Based on observations on the continental shelf region off southwest Vancouver Island, Denman and Freeland (1985) provided analytical correlation functions and noise levels for future objective analysis work. McIntosh (1990) compared these two methods in the case of a data interpolation over the ocean. The objective analysis appeared to be much demanding: the mean fields are rather difficult to obtain and the method is quite sensitive to length scales. On the other hand, he showed that the reproducing kernel obtained with a spline method was necessary red, unlike the objective analysis, which allows statistical approaches. He eventually concluded that the data separation was more important than the method used.

These methods, which include both observations and theory, constitute what is called an inverse problem, a definition of which is provided by Wunsch (1996): ''*The Ocean Circulation Inverse Problem* is the problem of inferring the state of the ocean circulation, understanding it dynamically, and even perhaps forecasting it, through a quantitative combination of theory and observations.'' Wunsch (1977) uses, for instance, global constraints such as the conservation of tracers in an open domain in order to compute the reference velocity needed to integrate the thermal wind equations. A great amount of theoretical research has lately been dedicated to the optimal use of data in order to initialize and control the open boundaries of various types of models. It yielded to the emergence of the optimal control (Miller 1986) and adjoint equation approaches (Le Dimet and Talagrand 1986; Tacker and Long 1988; Bennett 1992), see also De Mey (1997) for a complete review. Seiler (1993) proposed an adjoint model approach to assimilate data into a quasigeostrophic model, while controlling the evolution of the boundary conditions. Using the same approach, Bennett and McIntosh (1982) and Bodgen et al. (1996) optimized the open boundaries for a shallow-water model, offering an opportunity to eliminate the spurious solutions due the illposeness of the problem. Gunson and Malanotte-Rizzoli (1996) went further and included both the initial and boundary conditions in their control variables for their barotropic vorticity equation model. To optimize the use of available data, Chu et al. (1997) developed lately a variational approach based on a multiperturbation method in order to use observations from inside the domain to determine the open boundary conditions. To our knowledge, Zhang and Marotzke (1998) are yet the only ones to apply an adjoint model to determine the initial state and the boundary conditions for a 3D primitive equation model, the boundary conditions remains time-independent and an equilibrium state is only sought for. However, even if their development remains time demanding and their computing cost remains important, the adjoint methods are comparatively very efficient. Indeed, if the problem studied is linear only two iterations are required to optimize a whole trajectory: one forward, and one adjoint backward. Yet, the control variables must be chosen close to their optimal values, and the use of the different types of observations to specify the initial and open boundary conditions still raises numerous problems.

In the present paper, our objective is thus to construct a state of the ocean that is consistent with the observations, the available statistics, and the physics of the free-surface model, at a cost that is affordable with respect of the actual computer resources. From the various types of observation data, we determine a ‘‘mean’’ field and its covariance and variance. The energy spectrum, which globally describes the circulation in the region, is used to reconstruct an isotropic covariance matrix, and some causes of nonhomogeneity and anisotropy are then identified and parameterized in the observation covariance matrix. The constraints chosen to reconstruct the optimal state are chosen in the tangent linear space of the coastal free-surface model. The Euler–Lagrange equations related to the optimization problem are explicitly derived and a solution is found by iterative methods. The initial field that has been reconstructed is also used as a boundary condition to specify the incoming and outgoing flows in the FRS. In this case, the open boundary conditions are therefore not only determined from the data found in the neighborhood of the domain, but also from interior observations. However, they are not allowed to vary continuously in time. They can either be kept constant till an equilibrium state is reached, or be varied linearly between two ‘‘analysis’’ fields.

The following section is dedicated to the presentation of the inverse method. Original methods are in particular described to construct the model and observation data covariance matrices. In section 3, we use a twin-experiment approach to show how the inverse approach can improve the initialization of a small-scale coastal simulation. The experiment chosen is the modeling of an upwelling over a canyon (Klinck 1989). Some sensitivity studies to the constraints and to the data covariance matrix are detailed in section 4. In a last section, the validity of the various assumptions in the light of the experiments is discussed.

2. Inverse method

a. Free-surface coastal model

The direct model used is the free surface coastal model SYMPHONIE (Estournel et al. 1997; Marsaleix et al. 1998). It is based on the primitive equations, using both the Boussinesq and hydrostatic approximations. The finite difference scheme makes use of a ‘‘staggered’’ C grid (Arakawa and Suarez 1983), with sigma coordinates. The velocity components u, v, and w are derived from the incompressibility relation, and the conservation of momentum:

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \tag{2.1}$$

$$\begin{aligned}\frac{\partial u}{\partial t} + \frac{\partial uu}{\partial x} + \frac{\partial vu}{\partial y} + \frac{\partial wu}{\partial z} - fv &= -\frac{g}{\rho_0}\frac{\partial}{\partial x}\int_{z'=z}^{\eta}(\rho - \rho_0)\,dz' - g\frac{\partial \eta}{\partial x}\\ &\quad + A\left[\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right] + \frac{\partial}{\partial z}\left(k\frac{\partial u}{\partial z}\right)\end{aligned} \tag{2.2}$$

$$\begin{aligned}\frac{\partial v}{\partial t} + \frac{\partial uv}{\partial x} + \frac{\partial vv}{\partial y} + \frac{\partial wv}{\partial z} + fu &= -\frac{g}{\rho_0}\frac{\partial}{\partial y}\int_{z'=z}^{\eta}(\rho - \rho_0)\,dz' - g\frac{\partial \eta}{\partial y}\\ &\quad + A\left[\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}\right] + \frac{\partial}{\partial z}\left(k\frac{\partial v}{\partial z}\right),\end{aligned} \tag{2.3}$$

where (x, y, z) is the Cartesian coordinate system; ρ the density, ρ_0 a reference density; f the Coriolis parameter; η the position of the sea surface or sea surface elevation; g the gravity acceleration; and k and A are, respectively, the vertical and horizontal diffusion coefficients and k is computed using the turbulent kinetic energy. For more details on the turbulent closure scheme refer to Estournel et al. (1997).

The fast surface waves are solved by integrating the set of depth-averaged equations with a smaller time step:

$$\frac{\partial \eta}{\partial t} + \frac{\partial}{\partial x}(h\overline{u}) + \frac{\partial}{\partial y}(h\overline{v}) = 0 \qquad (2.4)$$

$$\frac{\partial h\overline{u}}{\partial t} + \frac{\partial}{\partial x}(h\overline{uu}) + \frac{\partial}{\partial y}(h\overline{vu}) - hf\overline{v} + gh\frac{\partial \eta}{\partial x} = \Pi_x \qquad (2.5)$$

$$\frac{\partial h\overline{v}}{\partial t} + \frac{\partial}{\partial x}(h\overline{uv}) + \frac{\partial}{\partial y}(h\overline{vv}) + hf\overline{u} + gh\frac{\partial \eta}{\partial y} = \Pi_y, \qquad (2.6)$$

where h designates the total depth and the overbar a depth average. The right-hand side terms Π_x and Π_y correspond to the depth integration of the z-dependent terms in the right-hand side of (2.2) and (2.3) (see Estournel et al. 1997). This mode-splitting approach separating the calculation of $\overline{u}$, $\overline{v}$, and η from the internal mode variables, allows the modeling of rapidly propagating inertia–gravity waves at a reduced cost (Blumberg and Mellor 1987).

The density ρ is obtained from the equation of state:

$$\rho = \rho_0 f(T, S, P). \qquad (2.7)$$

The temperature T and salinity S are calculated using their respective conservation equations:

$$\frac{\partial(T, S)}{\partial t} + \frac{\partial u(T, S)}{\partial x} + \frac{\partial v(T, S)}{\partial y} + \frac{\partial w(T, S)}{\partial z} = A'\left[\frac{\partial^2}{\partial x^2}(T, S) + \frac{\partial^2}{\partial y^2}(T, S)\right] + \frac{\partial}{\partial z}\left(k'\frac{\partial(T, S)}{\partial z}\right). \qquad (2.8)$$

Different sets of boundary conditions are also available. For the experiments shown in this paper, the FRS (Engedahl 1995) is used. Along each open boundary of the domain, a band of points is added over which the temperature, salinity and transport fields are relaxed toward the exterior circulation. The relaxation coefficient goes from 0 (no relaxation) at the inner edge of the band to 1 (full relaxation) at the outer edge of the band and is a second-order polynomial of the distance to the edge of the domain. Such an approach is well adapted to relax a limited area model toward an exterior circulation, and is also efficient to avoid the reflection of waves generated inside the domain. Yet, the FRS does not necessary conserve mass and momentum (Davies 1976; Engedahl 1995), and one has to specify an external solution that is correct with regard to the model dynamics (Engedahl 1995).

b. *Derivation of the analysis field*

In the following we will use when it is possible the notations recommended by Ide et al. (1997). A classic procedure of initialization is to interpolate in situ observations or outputs of OGCM on the fine grid of the model in order to obtain a ‘‘first-guess’’ field $\mathbf{x}^b(t_0)$. Then, the nonlinear model M is iterated forward in time leading to the so-called forecast $\mathbf{x}^f(t_i)$:

$$\mathbf{x}^f(t_{i+1}) = M_i[\mathbf{x}^f(t_i)], \qquad (2.9)$$

with $\mathbf{x}^f(t_0) = \mathbf{x}^b(t_0)$.

During the initialization process, the tendency terms in Eqs. (2.2) to (2.8) can be divided into two parts:

$$\frac{\partial}{\partial t} = \left(\frac{\partial}{\partial t}\right)^t + \left(\frac{\partial}{\partial t}\right)^n, \qquad (2.10)$$

where the first term in the right-hand side stands for the physically meaningful evolution of the fields (the ‘‘true’’ tendency), and the second is the results of the adjustment of the model (the ‘‘numerical’’ tendency). For a perfectly adjusted initial field, this second term would vanish, but the specification of the true tendencies is a difficult problem and when the model is initialized without any particular care, spurious tendencies due to transient processes leads to

$$\left(\frac{\partial}{\partial t}\right)^n \gg \left(\frac{\partial}{\partial t}\right)^t, \qquad (2.11)$$

which can be written

$$\lim_{\Delta t \to 0^+}\left[\frac{\mathbf{x}^f(t_0 + \Delta t) - \mathbf{x}^f(t_0)}{\Delta t}\right] \gg \left.\frac{\partial \mathbf{x}^t}{\partial t}\right|_{t_0^+}. \qquad (2.12)$$

One way to correct the nonphysical transient terms could be to integrate the model till an observation is available, and to correct the forecast from time to time leading to a sequential assimilation [Kalman filter (Miller 1986)] or to propagate backward the updating in order to correct the first guess [adjoint methods (Bennett 1992)]. It is yet necessary to start from a first guess that is not too far from its true value, and a set of observations large and reliable enough to evaluate the correction that has to be made.

Based on this observation, we propose to use explicitly the tangent to the coastal model to reduce the spurious numerical tendencies in (2.10). We thus compute the initial field that is close to the data, and does not lead to meaningless tendencies at the first iterate of the free-surface model. If the true tendency term can be evaluated for each field, the tendency of the analysis field is made as close as possible to its value, otherwise, it is decreased to the same order of magnitude. Within the tangent linear approximation, we thus consider the tendency vector $\mathbf{y}_0^o(t_0)$ to be given by

$$\mathbf{y}_0^o = \mathsf{H}_0[\mathbf{x}^t(t_0)] - \mathbf{F}_0 + \boldsymbol{\varepsilon}_0, \qquad (2.13)$$

where $\mathbf{F}_0$ stands for the forcing terms (appendix A) and H_0 is the observation operator, given in the present case by the tangent linear to M_0 (detailed in appendix A):

$$\mathsf{H}_0 \equiv \mathsf{M}_0', \qquad (2.14)$$

and $\boldsymbol{\varepsilon}_0$ is the error due to the linearization process and to the difficulty to ‘‘observe’’ the true tendencies. It is assumed to have zero mean:

$$\langle \underline{\boldsymbol{\varepsilon}_0} \rangle = 0. \qquad (2.15)$$

$\langle\ \rangle$ stands for the expectation of the statistical variable

$\underline{\boldsymbol{\varepsilon}}_0$. The true field $\mathbf{x}^t(t_0)$ is thought as the ‘‘perfect’’ trajectory in the model phase space. Thus, $\mathbf{y}_0^o(t_0)$ is a ‘‘measure’’ of the tendencies of the field at the first iterate of the model. The innovation vector can then be written

$$\mathbf{d}_0 = \mathbf{y}_0^o - \mathsf{H}_0\,[\mathbf{x}^a(t_0)] + \mathbf{F}_0. \tag{2.16}$$

The solution of this optimization problem is obtained by minimization of the cost function **J**:

$$J = (\mathbf{x}_0^a - \mathbf{x}_0^b)^{\mathrm{T}}\mathsf{B}_0^{-1}(\mathbf{x}_0^a - \mathbf{x}_0^b) + \mathbf{d}_0^{\mathrm{T}}\mathsf{R}_0^{-1}\mathbf{d}_0, \tag{2.17}$$

and the analysis field is eventually given by

$$\mathbf{x}^a(t_0) = \mathbf{x}^b(t_0) + \mathsf{K}_0[\mathbf{y}_0^o - \mathsf{H}_0\mathbf{x}^b(t_0) - \mathbf{F}_0], \tag{2.18}$$

where the Kalman gain K_0 is provided by

$$\mathsf{K}_0 = \mathsf{B}_0\mathsf{H}_0^{\mathrm{T}}(\mathsf{H}_0\mathsf{B}_0\mathsf{H}_0^{\mathrm{T}} + \mathsf{R}_0)^{-1}, \tag{2.19}$$

and R_0 is the model covariance matrix and is given by

$$\mathsf{R}_0 = \langle \underline{\boldsymbol{\varepsilon}_0}\,\underline{\boldsymbol{\varepsilon}_0^{\mathrm{T}}} \rangle. \tag{2.20}$$

Here, B_0 is the approximate error covariance matrix for the interpolated field $\mathbf{x}^b(t_0)$. If we define $\delta\mathbf{x}(t_0)$ as the difference between the true field and the interpolated field, then

$$\delta\mathbf{x}(t_0) = \mathbf{x}^{\mathrm{t}}(\mathrm{t}_0) - \mathbf{x}^{\mathrm{b}}(t_0), \tag{2.21}$$

and the covariance matrix B_0 is given by

$$\mathsf{B}_0 = \langle \underline{\delta\mathbf{x}}(t_0)\underline{\delta\mathbf{x}}(t_0)^{\mathrm{T}} \rangle. \tag{2.22}$$

We make the classic assumption that the interpolated first-guess is statistically consistent. This holds as long as there is no important bias in the first-guess fields. It leads to

$$\langle \underline{\delta\mathbf{x}} \rangle = 0. \tag{2.23}$$

All the quantities defined refer to the initial state, so we will drop the indices ‘‘0’’ in the following.

c. Interpolation of the data

The choice of the interpolation scheme can be particularly critical to obtain $\mathbf{x}^b$. It is formally written

$$\mathbf{x}^b = \mathsf{Z}\mathbf{y}^{\mathrm{d}}, \tag{2.24}$$

where Z is the interpolation matrix, and $\mathbf{y}^d$ is observation data vector.

A good scheme must retain all the meaningful physical structures contained in the observation data, and filter out the noise. The interpolation scheme must be as transparent as possible and is only a way to construct a gridded set of observations from data with different origins, and particular specificities. The very first type of data one can think about are in situ observations. They include observations made during specific cruises, or satellite remote observations. In this case, the observation data fields being used to reconstruct an optimized field, are considered to be simultaneous. An instrumental error is also necessarily attached to the measurement procedure and must be evaluated and removed. The outputs from an OGCM can also be used to initialize and force the coastal model. In any events, the output fields are dealt with as if they were observation data, and the same types of problems have to be faced, and for convenience, the word ‘‘data’’ will be used in the following.

Each dataset is thus defined by a data vector ($\mathbf{y}^d$), and its variance and covariance matrix (B^{dd}). Depending on what type of data is available, the observation $\mathbf{y}^d$ can be multivariate and includes the temperature, salinity, velocity, transport, and surface elevation fields. The observations can obviously not be used to solve all the length scales present in the region, and we can define a ‘‘window of observation.’’ The lower bound (L) of this window is particularly important to interpolate the observation on the model small scale grid. When one is dealing with observations originating from an OGCM, L is equal to twice the model horizontal grid scale (the Nyquist critical length scale). With in situ data, the observation array is, most of the time, far from being regular, and we shall make the assumption that the horizontal scale L characterizing a minimum average distance between two neighboring observation points can be defined.

As a consequence, when we interpolate the data on the small-scale model grid, our objective is to retain all the physical structures related to scales larger than L. For the remaining observation scales, we need to filter out which should be considered as unresolved scales. Most of the classical interpolation schemes (distance weighted schemes, statistical objective analysis, . . .) can offer such a specificity (for details see Thiebaux and Pedder 1987). For the simple example given in the present study, a ‘‘radially symmetric Gaussian filter’’ is used to interpolate the simulated OGCM data (see Bretherton et al. 1976 for sensitivity tests and Thiebaux and Pedder 1987 for a description of the filter). A complete expression for the interpolation filter is also given in appendix B. This choice is also consistent with the a priori statistical machinery that is set up, in so far as no new length scale is brought in by the filter and the observed large scales are rather well reproduced. In any events, the ‘‘noise’’ produced by the interpolation scheme is taken into account in the data covariance matrix.

d. Covariance of the interpolated field

The data covariance matrix (2.22) depends on many different parameters: the quality of the observation data field; the horizontal and vertical interpolation procedures, which both generate errors; or the prior statistics of the fields, which are very often the only description of the ‘‘real field’’ we have. In oceanography and particularly along the coast, this matrix is neither homogeneous nor isotropic. A realistic and accurate treatment of the problem requires a large amount of observations and, as importantly, a great deal of computer resources. Consequently, one cannot afford to deal with a fully nonhomogeneous and anisotropic problem. As a first guess, we choose to compute the covariance matrices under homogeneity and isotropy assumptions. Then, we correct this first approximation using our knowledge of the bottom topography. This can be written

$$\mathbf{B} = \mathbf{B}_{\mathrm{iso}} + \mathbf{B}_{\mathrm{aniso}}. \tag{2.25}$$

Using (2.21), $\mathbf{B}_{\mathrm{iso}}$ is thus derived from the true and data prior covariance matrices,

$$\begin{aligned}
&\mathbf{B}_{\mathrm{iso}}(\alpha, \beta)\\
&\quad = \langle(\underline{\delta^t\mathbf{x}}(\alpha) - \langle\underline{\delta^t\mathbf{x}}\rangle(\alpha))(\underline{\delta^t\mathbf{x}}(\beta) - \langle\underline{\delta^t\mathbf{x}}\rangle(\beta))^{\mathrm{T}}\rangle\\
&\quad = \langle[(\underline{\mathbf{x}}^t(\alpha) - \langle\underline{\mathbf{x}}^t\rangle(\alpha)) - (\underline{\mathbf{x}}^b(\alpha) - \langle\underline{\mathbf{x}}^b\rangle(\alpha))]\\
&\qquad \times [(\underline{\mathbf{x}}^t(\beta) - \langle\underline{\mathbf{x}}^t\rangle(\beta)) - (\underline{\mathbf{x}}^b(\beta) - \langle\underline{\mathbf{x}}^b\rangle(\beta))]^{\mathrm{T}}\rangle\\
&\quad = \left\langle\left[(\underline{\mathbf{x}}^t(\alpha) - \langle\underline{\mathbf{x}}^t\rangle(\alpha)) - \sum_{\alpha'} \mathbf{Z}_{\alpha,\alpha'}(\underline{\mathbf{y}}^d(\alpha') - \langle\underline{\mathbf{y}}^d\rangle(\alpha'))\right]\right.\\
&\qquad \left.\times\left[(\underline{\mathbf{x}}^t(\beta) - \langle\underline{\mathbf{x}}^t\rangle(\beta)) - \sum_{\beta'} \mathbf{Z}_{\beta,\beta'}(\underline{\mathbf{y}}^d(\beta') - \langle\underline{\mathbf{y}}^d(\beta')\rangle)\right]^{\mathrm{T}}\right\rangle\\
&\quad = \mathbf{B}^{\mathrm{tt}}|_{\alpha,\beta} + \mathbf{B}^{\mathrm{td}}\mathbf{Z}^{\mathrm{T}}|_{\alpha,\beta} - \mathbf{Z}(\mathbf{B}^{\mathrm{td}})^{\mathrm{T}}|_{\alpha,\beta} + \mathbf{Z}\mathbf{B}^{\mathrm{dd}}\mathbf{Z}^{\mathrm{T}}|_{\alpha,\beta},
\end{aligned} \tag{2.26}$$

where $\mathbf{B}^{\mathrm{tt}}$ and $\mathbf{B}^{\mathrm{dd}}$ are the prior covariance matrices for the true and data fields, and $\mathbf{B}^{\mathrm{td}}$ is their joint covariance matrix. Here α and β stand for two grid points, and α' and β' two points of the observation array. Under the assumptions of homogeneity and isotropy, the computation of the three prior covariance matrices $\mathbf{B}^{\mathrm{tt}}$, $\mathbf{B}^{\mathrm{dd}}$, and $\mathbf{B}^{\mathrm{td}}$ only requires the knowledge of the spatial energy spectrum of the corresponding fields, the covariance matrices being computed using the inverse Fourier transforms of a global homogeneous spectrum (F). We define the normalized correlation and cross-correlation functions for $\mathbf{x}^t$ and $\mathbf{y}^d$:

$$\begin{aligned}
f^{\mathrm{tt}}(r) &= \int_{k=0}^{\infty} F^{\mathrm{tt}}(k)\exp[-(2i\pi)kr]\,dk\\
f^{\mathrm{td}}(r) &= \int_{k=0}^{\infty} F^{\mathrm{td}}(k)\exp[-(2i\pi)kr]\,dk\\
f^{\mathrm{dd}}(r) &= \int_{k=0}^{\infty} F^{\mathrm{dd}}(k)\exp[-(2i\pi)kr]\,dk,
\end{aligned} \tag{2.27}$$

with $f^{\mathrm{tt}}(0) = f^{\mathrm{td}}(0) = f^{\mathrm{dd}}(0) = 1$.

Hence,

$$\begin{aligned}
\mathbf{B}^{\mathrm{tt}}(\alpha, \beta) &= \langle(\underline{\mathbf{x}}^t(\alpha) - \langle\underline{\mathbf{x}}^t\rangle(\alpha))(\underline{\mathbf{x}}^t(\beta) - \langle\underline{\mathbf{x}}^t\rangle(\beta))^{\mathrm{T}}\rangle\\
&= \sigma_{\mathrm{tt}}^2 f^{\mathrm{tt}}(r)\\
\mathbf{B}^{\mathrm{td}}(\alpha, \beta) &= \langle(\underline{\mathbf{x}}^t(\alpha) - \langle\underline{\mathbf{x}}^t\rangle(\alpha))(\underline{\mathbf{x}}^d(\beta) - \langle\underline{\mathbf{x}}^d\rangle(\beta))^{\mathrm{T}}\rangle\\
&= \sigma_{\mathrm{td}}^2 f^{\mathrm{td}}(r)\\
\mathbf{B}^{\mathrm{dd}}(\alpha, \beta) &= \langle(\underline{\mathbf{x}}^d(\alpha) - \langle\underline{\mathbf{x}}^d\rangle(\alpha))(\underline{\mathbf{x}}^d(\beta) - \langle\underline{\mathbf{x}}^d\rangle(\beta))^{\mathrm{T}}\rangle\\
&= \sigma_{\mathrm{dd}}^2 f^{\mathrm{dd}}(r) + \langle\underline{\boldsymbol{\varepsilon}}^d\,\underline{\boldsymbol{\varepsilon}}^{d\mathrm{T}}\rangle I,
\end{aligned} \tag{2.28}$$

where r is the distance between the points α and $\boldsymbol{\varepsilon}^d$ the observation noise. The spectra can be computed based on previous simulations on the same domain, or on dedicated observations (see Denman and Freeland 1985). It can also be constructed from a prior knowledge of the persistent structure of the general circulation (see section 4). The covariances between different variables ($u - \rho$, $\overline{u} - \eta$, . . .) are computed following Bretherton et al. (1976).

The standard deviations σ_{tt}, σ_{td} and σ_{dd} are very dependent on the physics of the domain and cannot be determined a priori. We thus evaluate them using the observation data:

$$\sigma_{\mathrm{dd}}^2 = \sigma_{\mathrm{td}}^2 = \sigma_{\mathrm{tt}}^2 = (\mathbf{y}^d - \tilde{\mathbf{y}}^d)(\mathbf{y}^d - \tilde{\mathbf{y}}^d)^t, \tag{2.29}$$

where the statistical means have been replaced by the spatial mean $\tilde{\mathbf{y}}^d$ over the domain. This is a rather strong approximation, all the more as we saw that the observation data field $\mathbf{y}^d$ described the circulation only for scale larger or equal than L. This approach usually gives us useful hints about the magnitude of the variables and can be refined with the results of posterior runs on the same domain.

The nonhomogeneity and the anisotropy of the fields is of particular importance in coastal zones (Denman and Freeland 1985). Many different sources of nonhomogeneity and anisotropy can indeed be found. At turbulent scales, the direction of the current or the shear stress are for instance a possible cause of the anisotropic distribution of energy (Turner 1973). The local stratification, or the well-determined directions of propagation of internal waves are responsible for larger scale anisotropic motions. The closeness to the coast and more generally the bottom topography can also be two relevant geographical indicators of the nonhomogeneity and anisotropy of the currents (see for instance Brink 1982 or Huthnance 1995). Purely barotropic geostrophic currents are indeed constraint to follow the line of constant topography in order to conserve vorticity (Roisin 1994). On the other hand, the density flows follow the lines of steepest topography gradients. Thus, the bottom topography brings in the flow specific length scales and as importantly specific directions of currents. These effects are not very often taken into account in the statistical description of the fields though they are may be the most obvious sources of nonhomogeneity and anisotropy in the coastal circulation. To parameterize these effects, the nonhomogeneous, anisotropic part of the covariance matrix is itself divided into two parts:

$$\mathbf{B}_{\text{aniso}}(\alpha, \beta) = \xi_1\mathbf{B}_{\text{topo}}(\alpha, \beta) + \xi_2\mathbf{B}_{\text{delta}}(\alpha, \beta), \quad (2.30)$$

where ξ_1, and ξ_2 are two constants chosen so that the elements of the anisotropic matrices be less than 10% of the corresponding elements of the isotropic matrix, and

$$\mathbf{B}_{\text{topo}}(\alpha, \beta) = \exp\left(-\frac{d^2(\alpha, \beta)}{d_{\text{ref}}^2}\right)\exp\left(-\frac{[h_x(\alpha) - h_x(\beta)]^2}{h_{\text{ref}}^2}\right) \times [1 - \delta(\alpha, \beta)] \quad (2.31)$$

$$\mathbf{B}_{\text{delta}}(\alpha, \beta) = \mathbf{B}_{\text{iso}}(\alpha, \beta)\frac{[h_x(\alpha) - h_y(\beta)]^2}{h_x(\alpha)^2}\delta(\alpha, \beta), \quad (2.32)$$

where h_x and h_y are the bottom topography fields on, respectively, the small-scale grid and the observation array. If the data y^d is based on observations, then h_y is the smoothed topography of the area, the length scales smaller than L being filtered out. If, on the other hand, the data y^d have been computed with an OGCM, the topography field h_y should be the rather smoothed field used in this model. Here, δ is the Kronecker delta function, $d(\ )$ is the distance between points α and β, and d_{ref} and h_{ref} are two parameters. Also, d_{ref} is chosen as the smallest solution of

$$\frac{f_{\text{tt}}(d_{\text{ref}})}{f_{\text{tt}}(0)} = \frac{1}{e}, \quad (2.33)$$

and is therefore a radius of influence along the lines of constant topography, whereas h_{ref} has the same meaning but in the orthogonal direction. The importance of the currents in this direction, or the magnitude of the topography gradients can give indications on its value.

Thus, the first part of the anisotropic covariance matrix, $\mathbf{B}_{\text{topo}}$ specifies in the covariance matrix the preferred directions of flow. Such an expression proved to be well adapted to cases where the geostrophic currents are dominant. Two Gaussian-shaped functions have been chosen in (2.31). One is related to the distance between the points α and β (the smaller the distance the larger the covariance), and the other concerns the difference between the bottom topography at these points (the smaller the difference the larger the covariance).

On the other hand, $\mathbf{B}_{\text{delta}}$ is the normalized difference between the bottom topographies h_{x} and h_{y} and describes how different the small-scale bottom topographies are in the fields x^{t} and y^d. It is added as an ''error'' term to the diagonal of the covariance matrix.

Other parameterizations could have been chosen, but (2.31) and (2.32) are simple enough to be physically understandable (see section 4b) and the normal and square laws used are smoothed enough to avoid undesirable sharp gradients in the variance covariance matrices.

e. Model covariance matrix

The set of linearized constraints being different from the coastal model equations, the magnitude ($\boldsymbol{\varepsilon}_0$) of the nonlinear contribution to the primitive equations must be evaluated. The errors on the evaluation of the physical tendency terms must also be known. Based on (2.20), this is achieved by computing the second order moment of the nonlinear contribution to the primitive equations (appendix A) and, if necessary, of the tendency error terms.

A preliminary experiment is thus conducted over the same domain. During this experiment, the amplitude of the nonlinear terms is carefully recorded and their variance and covariance are evaluated.

It is shown in a sensitivity study (section 4d) that the part of the matrix related to the nonlinear truncations can be computed with a very good accuracy by such a direct evaluation.

f. Mode splitting approach and general algorithm

Based on the same mode splitting approach as the direct model, the analysis field is computed separately and iteratively for the external and internal mode variables. We thus introduce the indices E and I referring to the external or internal variables. Using these notations, $\mathbf{x}_E^b$ and $\mathbf{x}_I^b$ are respectively referring to

$$\mathbf{x}_E^b = \begin{bmatrix} (\overline{u}_{i,j})_{(i,j)\in V} \\ (\overline{v}_{i,j})_{(i,j)\in V} \\ (\eta_{i,j})_{(i,j)\in V} \\ (\overline{\rho}_{i,j})_{(i,j)\in V} \end{bmatrix}, \qquad \mathbf{x}_I^b = \begin{bmatrix} (u_{i,j,k})_{(i,j,k)\in V'} \\ (v_{i,j,k})_{(i,j,k)\in V'} \\ (\rho_{i,j,k})_{(i,j,k)\in V'} \end{bmatrix}, \quad (2.34)$$

where V and V' are the sets of grid points in two and three dimensions.

The conservation of mass (2.4) must be treated with a lot of care. Indeed, this equation is crucial in the initialization process for two reasons. A nonzero transport divergence in the initial field [right-hand side of (2.4)] leads to transient processes without any physical meanings. On the other hand, if the initial field is used to specify the incoming and outgoing transport through the open boundaries, it must be globally balanced. If not, even a small misfit can rapidly alter the mean surface elevation over the domain. We also mentioned that the use of FRS could produce artificial divergence (Engedahl 1995), and, as a consequence, a ''correct'' external solution had to be specified. Equation (2.4) is also a linear equation and it is equal to its tangent equation: no approximation has to be made in the tangent linear and the equation can be rigorously satisfied. However, the use of a strong constraint such as (2.4) rises numerous numerical problems. Indeed, in an explicit formulation the matrix of a system including both strong and weak constraints, is neither symmetric nor positive definite, unlike the one for a system made of only weak constraints. The minimization algorithm are therefore much more expensive in both CPU time and memory storage (for an evaluation of the computer resources needed see the following section). As a first approximation, we thus consider (2.4) to be a weak constraint, then we constrain the analysis field just found to satisfy (2.4) exactly. With such a procedure, the system giving the Lagrange multiplier is much smaller and can be solved numerically (see appendix C). When necessary, the whole procedure is iterated until it converges. However, the Lagrange multipliers are rather small, and the associated modification of the transport field does not perturb drastically the previous equilibrium.

Following this iterative procedure, a minimum is first reached for the external mode problem. The problem can sometimes be simplified by interpolating the barotropic streamfunction instead of the transport components. With a simple change of variables in the model tangent equations, the transport can be replaced by the derivative of the streamfunction, avoiding the use of the mass conservation equation and therefore reducing the number of equations in the tangent linear M'_{E}.

The internal mode variables are then corrected so that their depth average be equal to the corresponding optimized vertically integrated variable. A minimum is eventually obtained for the internal mode problem. The contribution of x_I^b to the vertical advection and diffusion terms is neglected and these terms do not appear explicitly in the tangent linear equations. Consequently, the different internal mode equations are not coupled with each other and the associated Euler–Lagrange equations for each sigma surface can be solved independently. The internal or external mode optimization problems are thus very similar. This algorithm can be iterated until the process converges. Practically, this can be reached in only one or two iterations if the nonlinear contributions are small enough.

For small systems (less than 3000 × 3000), the global optimization matrices are inverted using a preconditioned conjugate gradient algorithm. When large memory storage is required, limited-memory quasi-Newton methods are now being tested. They indeed requires low and adjustable storage requirements (see for instance Zou et al. 1993 for their comparative study of such methods for oceanographic problems). To avoid the problems of ill-conditioning, a scaling of the optimization parameters is necessary (Gunson and Malanotte-Rizzoli 1996).

Very often the temperature and salinity fields behave in similar ways, that is, the same structures can be found on their respective maps and the set of internal mode constraints can be simplified. This is obviously far from being true at molecular scales, where the physical processes controlling these quantities are very different—compare, for instance, their respective molecular diffusions. However, at larger scales, their diffusion coefficients are very similar. The main difference concerns in fact the forcings. Thus, if one is working in an estuary or if strong surface forcings are applied the temperature and salinity can exhibit very different structures. In such a case, the temperature and salinity equations have to be used as constraints in the internal mode optimization procedure. Yet, if such is not the case, one can use a single equation for density, reducing significantly both the number of unknowns and variables: the size of the system of Euler–Lagrange equations is reduced by about 25%. Thus, new density and velocity fields are obtained and the resulting modifications on the temperature and salinity fields are computed using a short optimization algorithm. This leads to a system of one equation (the equation of state) for two unknowns (the temperature and salinity modifications): the problem is underdetermined. One solution is to consider a system of three unknowns—the density (ρ_f), the temperature (T_f), and the salinity (S_f), submitted to four constraints, the first one being applied as a strong constraint:

$$\rho^{pa} = \rho^{a} \qquad T^{pa} = T^{a} \qquad S^{pa} = S^{a}$$
$$\rho^{pa} = \rho_0 f(T^{pa}, S^{pa}, P^{a}). \qquad (2.35)$$

This system is now overdetermined and an optimal solution in a least squares sense is found via the minimization of the cost function J_{TS}:

$$J_{TS}(i, j, k) = 2\mu_\rho[\rho_0 f(T^{pa}, S^{pa}, P^{a}) - \rho^{pa}] + \frac{(T^{pa} - T^{a})^2}{\sigma_T^2} + \frac{(S^{pa} - S^{a})^2}{\sigma_S^2}, \qquad (2.36)$$

where μ_ρ is the Lagrange multiplier associated to the density. The solution for this last optimization problem is found analytically, and its numerical computation is rather cheap in so far as J_{TS} can be minimized independently at each grid point. The density being exactly conserved in the process, the equilibrium state previously reached is conserved. So the initial field is not altered by the final adjustment of T and S. However, the change in temperature and salinity depends on the values of σ_T and σ_S, their respective prior standard deviation.

3. Twin experiments

We chose to illustrate the usefulness of the inverse model on a well-known coastal problem: the interaction of a large-scale current with a canyon on a coastal shelf. A twin-experiment approach was followed. To begin with, the coastal model is integrated on a small-scale grid (5 km), over a large domain including the continental shelf and its canyon (Fig. 1). In the following, this very first experiment is referred to as the ‘‘reference experiment.’’ This is in fact the best representation the free surface model could give of the flow over the canyon. A barotropic current is forced parallel to the coast and stands for the general circulation over the canyon. Such a situation was studied by Klinck (1989) as the ‘‘weakly stratified upwelling case.’’

To simulate the results that a large-scale model could

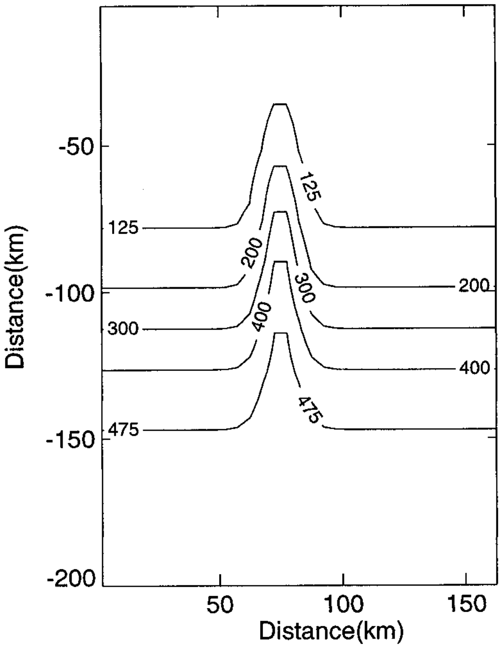

FIG. 1. Bottom topography (m) of the large reference domain including the continental shelf and a canyon. Dash line: small domain for the twin experiments.

give on such a problem, the same coastal model is integrated over the same domain, but using a large-scale grid (25 km). Such a model cannot represent at all the canyon whose transverse horizontal scale is smaller than 25 km (around 20 km). The results of this very coarse resolution model are used to initialize two twin experiments over a smaller domain above the canyon, using a small-scale grid (5 km). A first experiment is achieved over the small domain without any prior optimization. The observation data field is interpolated on the 5 km × 5 km grid using the Gaussian filter described in the previous chapter. In parallel to this initialization, an optimized simulation is presented. To begin with, the initialization fields are compared with each other, then the rms errors are studied after a 100-h run.

a. Reference and large-scale experiments

1) REFERENCE EXPERIMENT

A map of the bottom topography is shown on Fig. 1. To generate this topography the same analytical functions as Klinck (1996) were used but the canyon is a little larger:

$$h(x, y) = h_m - \frac{h_s}{2}\left[1 - th\left(\frac{y - Y_0(x)}{a}\right)\right]$$

$$Y_0(x) = Y_n + Y_b\left[1 - \exp\left(-\frac{x - X_0}{2b^2}\right)\right], \quad (3.1)$$

where h is the depth, $h_m = 500$ m, $h_s = 400$ m, $a = 25$ km, $b = 5.2$ km, $y_n = 60$ km, and $Y_b = 50$ km. The canyon has an horizontal width between the fine grid (5 km) and the large-scale grid (25 km).

The reference and large-scale experiments are both initialized with a surface elevation gradient orthogonal to the coast. This gradient creates an easterly barotropic current of 10cm s^{-1}. The vertical density profile corresponds to a weakly stratified flow with a first baroclinic Rossby radius around 22 km, well represented by the 5 km × 5 km grid of the model. It is initially given by

$$\rho(z) = \rho_0(1 + 1.928 \times 10^{-4} z), \quad (3.2)$$

which roughly corresponds to a decrease of 4°C every 500 m. The salinity field is kept constant, and a single prognostic equation for temperature is used.

This experiment has been integrated till an equilibrium was reached. Along the eastern and western open boundaries, 10-point bands are added to relax the flow toward the easterly barotropic current over which the FRS is used. It is integrated long enough so that all spurious inertia-gravity waves be damped (a 5-day spinup was necessary). The use of the FRS drives quickly the solution toward an equilibrium with alongshore barotropic current along the eastern and western boundaries. The good behavior of the FRS concerning the reflection of gravity waves is studied in the following section.

Klinck (1989) showed that this configuration yields to an upwelling at the canyon head. The mechanism can be explained writing the conservation of potential vorticity. The water inside the canyon is initially at rest, but is subjected to the surface elevation gradient. The lines of vorticity are thus stretched at the canyon head and compressed at the other end. An upwelling is therefore created at the head. The dense water initially in the canyon is advected upward, and once the geostrophic adjustment is reached, a baroclinic pressure counter gradient is created, inducing an anticyclonic surface current. Figures 2–4 illustrate these processes. At the canyon head, the mean current separates into two parts (Fig. 2). One of them has enough anticyclonic vorticity to follow the bottom topography. The second part flows in the alongshore direction and creates a coastal jet. The surface elevation anomaly is shown on Fig. 3. The gradient of the surface elevation, initially orthogonal to the coast, is smoothed above the canyon, inducing locally a positive anomaly. A horizontal section of the density at 90-m depth (Fig. 4 for the density anomaly with respect to $\rho_0 = 1024.8$ kg m^{-3}) shows that dense water is pushed upward in the upwelling, giving a negative temperature anomaly at the canyon head. Part of the water is advected downstream with the alongshore and cross-shore branches of the downstream current.

2) SENSITIVITY TO THE FRS

In order to study the generation and propagation of the gravity waves, the surface elevation and the vertical

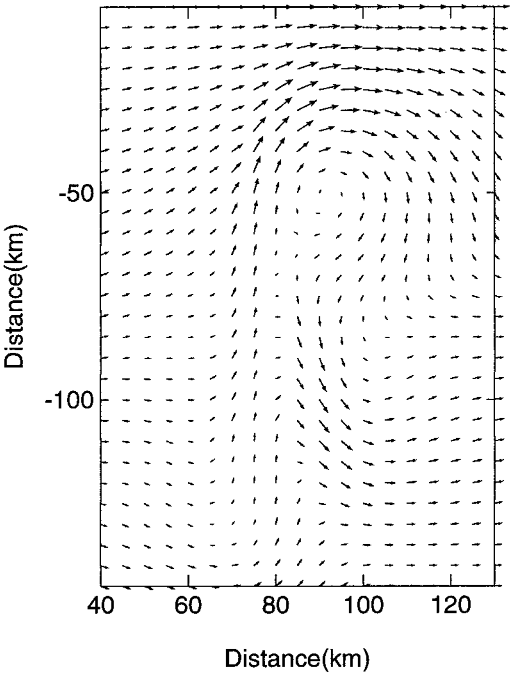

Fig. 2. Horizontal mean velocity: reference simulation over the canyon.

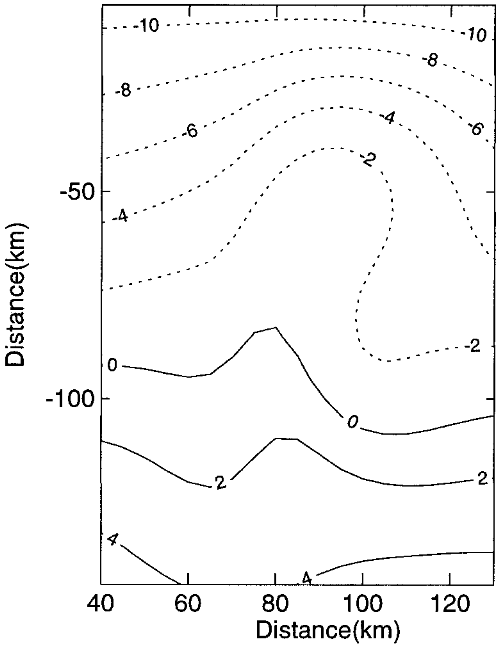

Fig. 3. Surface elevation anomaly (cm): reference simulation over the canyon.

velocity (Ω, in σ coordinates) are recorded during the first hours of simulation, Ω is given by

$$\Omega = (H + \eta)\left(\frac{\partial}{\partial t} + u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y} + w\frac{\partial}{\partial z}\right)\left[\frac{z + h}{\eta + h}\right]. \tag{3.3}$$

Deleersnijdner (1989) shows that Ω gives a better representation of the true upwelling processes than the physical vertical velocity w. Figures 5a and 5b show the time evolution of the surface elevation (η) and the vertical velocity (Ω) of the fields at respectively 75 and 110 km off the coast. On both plots, the abscissa is the distance in the y direction from the western open boundary. Figures 5a and 5b show waves propagating eastward and westward from the center of the canyon (distance = 90 km) with an amplitude decreasing with space and time. The phase speed for these waves can be computed using the lines of constant phase on the space–time plot. During the very first hours of simulation, this measurement gives, on Fig. 5a, a speed of about 32 ± 5 m s^{-1} for the waves propagating downstream and 30 ± 5 m s^{-1} for the others, which is very close to the theoretical phase speed for external waves: $c = \sqrt{gh} = 34.8$ m s^{-1}. The evaluation of the phase speed for the waves shown on Fig. 5b is a little more difficult. Indeed these waves do not propagate horizontally and we can only see them leaving the canyon. However, the slopes indicate a speed of about 0.95 ± 0.20 m s^{-1}. Using the procedure detailed by Tintore et al. (1995), the computation of the internal wave phase speed gives a value of the same order of magnitude. These results indicate the generation of both external and internal gravity waves inside the canyon right after the initialization. Indeed, the canyon, like any other topographic small-scale variations, induces a perturbation of the initial fields, with rather large values of the vertical velocity. These perturbations propagate at first all over the

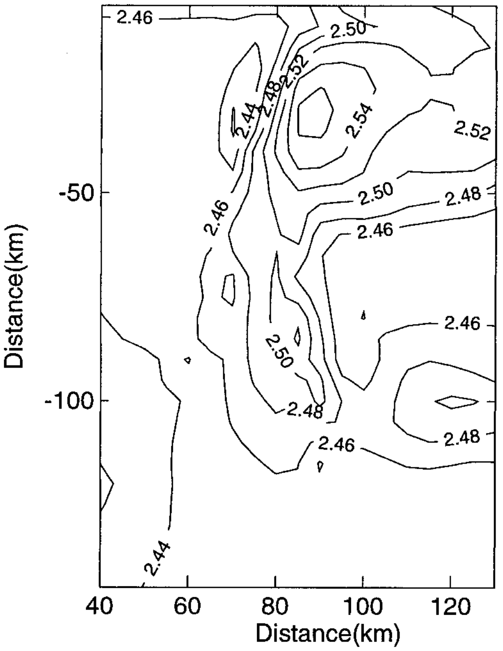

Fig. 4. Density anomaly (kg m^{-3}) at 80-m depth: reference simulation over the canyon.

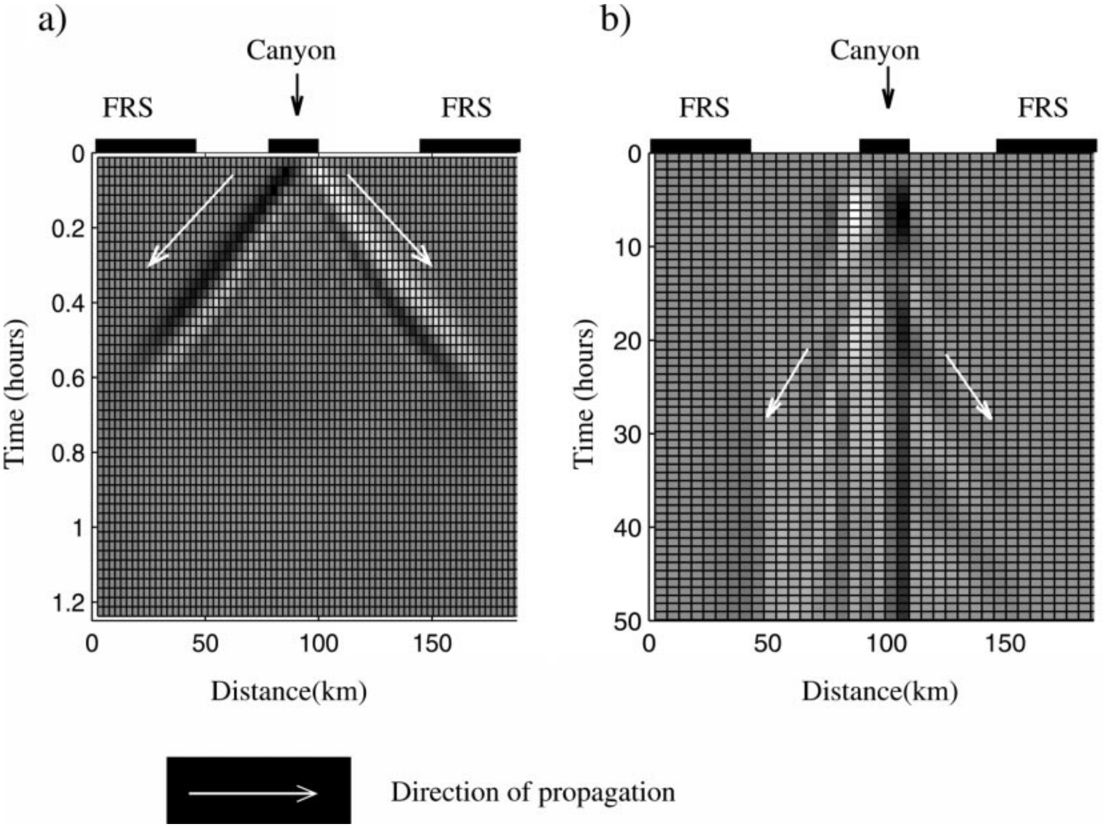

FIG. 5. (a) Time evolution of the surface elevation at 75 km off the coast (color map), from −29 cm (black) to 22 cm (white); (b) time evolution of the vertical velocity (Ω) at a depth of $H/2$ and at 110 km off the coast (color map), from -0.047 m h^{-1} (black) to 0.047 m h^{-1} (white).

domain via gravity waves. This is consistent with the results presented by Lacarra and Talagrand (Lacarra and Talagrand 1988) who show that inertia–gravity waves are generated during the first hours of simulation, that their amplitude is decreasing exponentially with time before slowly propagating Rossby waves became eventually dominant and energy was transferred toward the large-scale geostrophic motion.

The usefulness of the FRS is thus particularly evident in Fig. 5a where the external gravity wave amplitude is drastically damped when entering the FRS (after only 0.6 h). There is also no evidence of large wave reflections. Yet, a more careful examination of the results shows that small amplitude reflected waves can be found, but they are several orders of magnitude smaller. This rather good results are a consequence of the large number of relaxation points along the boundary and the simplicity of the barotropic general circulation. We can thus conclude that no spurious reflections contaminated the reference solution, and that the long spinup dedicated to the adjustment leads to a physically consistent circulation.

3) LARGE-SCALE EXPERIMENT

The 25-km grid does not allow the representation of the canyon bottom topography. This type of simulation stands for what a general circulation model would give in the same coastal region. For this simulation the same FRS boundary conditions as for the reference simulation were used. The velocity, temperature and salinity fields are extracted once the model has reached a geostrophic equilibrium.

b. Forced simulations over a small domain

1) TWIN EXPERIMENTS

The general circulation given by the large-scale experiment is interpolated on the small-scale grid in order to initialize two twin experiments: one is started with a simple interpolation, although the other is ‘‘optimized’’ using the inverse method.

We make the assumption that no statistics is available about the ‘‘real’’ circulation, and we choose a standard normal spectrum S:

$$F_{tt}(k) = S(k) = \exp\left(\frac{-a^2k^2}{4}\right), \tag{3.4}$$

with $a = 12.5$ km. Here $F_{dd}(k)$ and $F_{td}(k)$ are equal to $F_{tt}(k)$ but are truncated for $k = 1/25$ km. Indeed, no structures with a length scale smaller than the grid scale

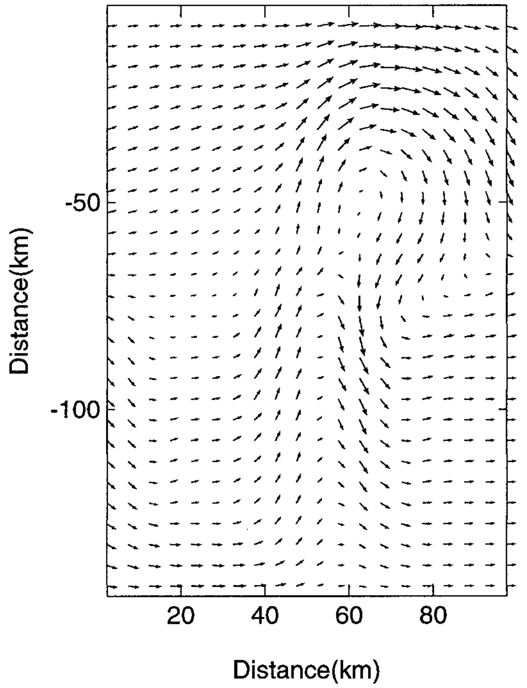

FIG. 6. Horizontal mean velocity optimally interpolated on the small-scale grid.

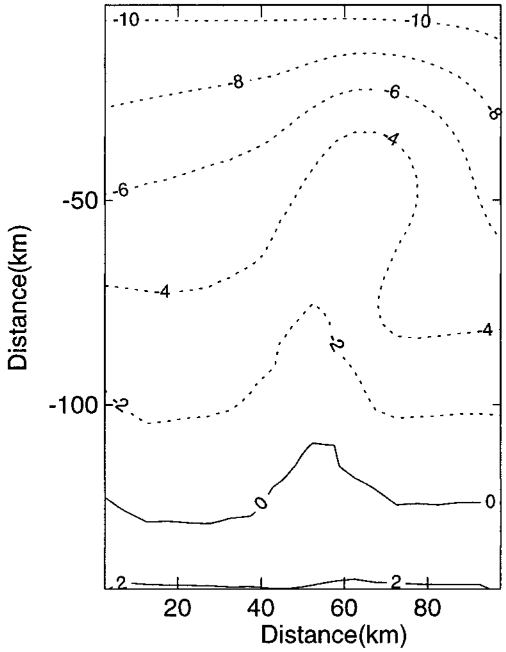

FIG. 7. Surface elevation anomaly (cm) optimally interpolated on the small-scale.

(25 km) can be found in the large-scale experiment. At the shelf break, the difference of topography between two points separated by 25 km is roughly equal to 50 m. As a consequence, we chose for the anisotropic covariance parameter: $h_{\mathrm{ref}} = 50$ m, which is a correlation length orthogonal to the lines of constant topography.

The measurement noise $\boldsymbol{\varepsilon}^d$ was considered to vanish. The addition of a measurement noise $\boldsymbol{\varepsilon}$ would have very straightforward consequences: less trust is put into the observations, and the influence of the physical constraints is thus larger. The model covariance matrix is computed by evaluating the truncated nonlinear terms following the algorithm given in section (2.5). As an equilibrium field is only searched for, the "true" tendencies are expected to vanish.

Both the external and internal parts of the inverse model were used for the optimization. However, the salinity field being kept constant, the equation for the density field is used. In the present experiment, the nonlinear effects are also rather small. Indeed, the magnitude of the currents remains small (10 cm s^{-1}) and no surface forcing is applied, which limits the advective and turbulent adjustments. As a consequence, the results shown were obtained after only one iteration of the algorithm presented in section 2f.

The optimized fields are presented in Figs. 6–8. As in the reference circulation (Fig. 2), the coastal jet is about 45 km wide at the canyon head, and the second offshore branch along the downstream edge of the canyon shows a meander of the same length. The consequences of the upwelling can also be seen on the pressure fields. Indeed, it induces a positive surface elevation anomaly above the canyon (see Fig. 7), its cross-shore gradient has been smoothed over the canyon. The density anomaly shown on Fig. 8 is higher at the canyon head and downstream, cold salty water being brought from the bottom of the canyon.

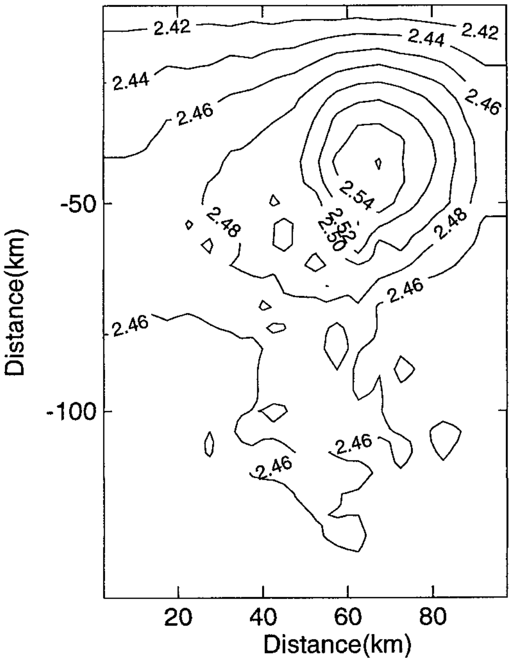

FIG. 8. Density anomaly (kg m^{-3}) at 80-m depth optimally interpolated on the small-scale grid.

The complete inversion was realized on an HP server ESS-800 with 256 Mbytes memory. Most of the matrices are stored and handled with Matlab 5.0 sparse matrix package (Gilbert et al. 1992). The optimization of the external mode for the 30 × 20 point grid requires the inversion of a 2350-point square matrix containing only 35% of nonzero elements. For larger systems, the nonzero elements can reach less than 8% of the total number of elements. The matrix being in most cases definite positive, Cholesky preconditioned conjugate gradients can be used. In so far as a single equation is used for density, the internal mode matrices have very similar characteristics. The CPU time necessary for a complete iteration is very dependent on the algorithm used for the inversion of the optimization matrix, which requires about 80% of the total CPU time.

2) RMS ERRORS WITH RESPECT TO THE REFERENCE EXPERIMENT

The previous fields, constructed respectively with and without optimization, are used to initialize and force the direct model along its open boundaries during 100 hours. The evolution of the rms errors is plotted in Figs. 9a–d. For a given variable x, the rms error is given by

$$\mathrm{rms}(x_E) = \sqrt{\frac{\sum_{i=m}^{(M-m)} \sum_{j=n}^{(N-n)} [x_{i,j} - (x_{i,j})^{\mathrm{ref}}]^2}{(M - 2m + 1)(N - 2n + 1)}} \qquad (3.5)$$

and

$$\mathrm{rms}(x_I) = \sqrt{\frac{\sum_{i=m}^{(M-m)} \sum_{j=n}^{(N-n)} \sum_{k=1}^{L} [x_{i,j,k} - (x_{i,j,k})^{\mathrm{ref}}]^2}{(M - 2m + 1)(N - 2n + 1)L}}, \qquad (3.6)$$

where M, N, and L stand for the number of grid points in, respectively, the x, y, and z directions and x^{ref} is the reference simulation field. The rms error is recorded inside the domain, but not along the relaxation bands (their width being given by m and n). The flow is indeed submitted to different types of boundary conditions along the frontiers and its comparison with the reference simulation does not give any useful hints about the dynamical behavior of the model.

Figures 9a to 9d show the evolution of the global rms errors for the surface elevation anomaly, the cross and alongshore mean velocity and the density. For each variable, the rms error has been taped during the first 100 h of the experiment for an optimized (‘‘Optimization I’’ curves) and a nonoptimized run (‘‘Interpolation’’ curves). The generation and propagation of gravity waves, can be observed on the plot of the nonoptimized surface elevation and mean velocities rms errors (Figs. 9a–c). We saw in a previous section (Fig. 5) that these waves were generated in the canyon. They last for a few hours and are followed after one and a half inertial period by a linear increase of the surface elevation rms error. Indeed, the divergence of the crudely interpolated transport toward which the simulation is relaxed along its opened boundaries does not globally vanish. Thus, as the numerical schemes of the direct model conserve both mass and energy, these boundary conditions lead to an increase of the surface elevation, more water being forced inside the domain than outside.

Both problems, the initial generation of gravity waves and the increasing of the surface elevation anomaly have been solved by the inverse model. Indeed, the plot of the rms in the optimized case shows a quasi-constant rms error during the whole simulation.

4. Sensitivity studies

a. Sensitivity to the physical constraints

The previous optimized fields have been computed using all the available constraints: conservation of mass, of momentum (external and internal modes), and conservation of density. Various initializations have also been performed using only some of the constraints.

The equations for the conservation of momentum or density could have been inverted without the linearized advective or the bottom friction terms. Figure 10 shows the mean velocity field that would be obtained if these terms were neglected. Two major differences can be noticed. At the canyon head, the coastal jet is much larger when the bottom friction terms are used, which is consistent with the reference experiment results given on Fig. 2. This can be explained by the strong bottom current on the shallow continental shelf, which induces large bottom friction terms in the conservation equations. The second major difference between Figs. 6 and 10 is the creation of a nonsymmetrical current above the canyon when advective terms are used and linearized in the constraint equations. The downstream offshore current meanders, which had been reconstructed by the full optimization (Fig. 6) has disappeared on Fig. 10. The rms error of the second optimized initialization (without advection nor bottom friction), plotted on Fig. 9, Optimization II, is larger and adjusts slower than in the first optimized initialization. The CPU time necessary to invert the various matrices is only slightly decreased by the removal of the few advective and bottom friction terms (the matrices are only a little sparser). This test also insures that both the geostrophic and ageostrophic parts of the flow are corrected by the inverse method.

One initialization field was also obtained using only the external-mode Eqs. (2.4)–(2.6) as physical constraints. The rms errors are shown on Fig. 9, Optimization III. As can be expected when considering the barotropic structure of the general circulation, the rms curves for the surface elevation and mean velocity are only slightly influenced by the absence of adjustment of the internal equations. The few small amplitude inertial oscillations present in the full optimization have

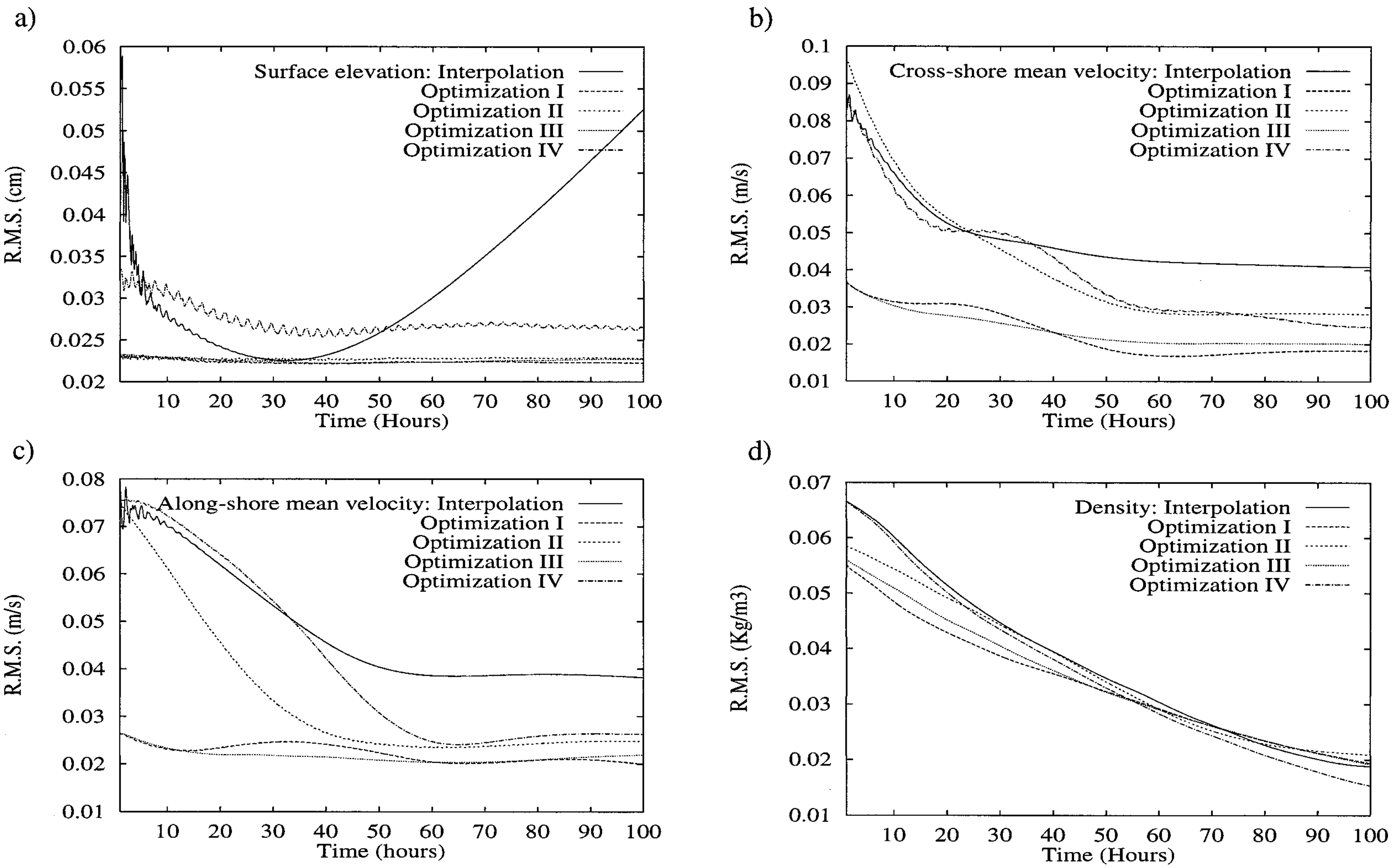

FIG. 9. Evolution during the first 100 h of simulation of the global rms error of the (a) surface elevation anomaly, (b) the cross-shore, and (c) alongshore (d) mean velocity and the density. Solid line: interpolated field; long-dashed line: optimized field I (complete optimization); short dashed line: optimized field II (no advection or bottom friction); dotted line: optimized field III (external-mode equations); dash–dotted line optimized field IV (conservation of mass).

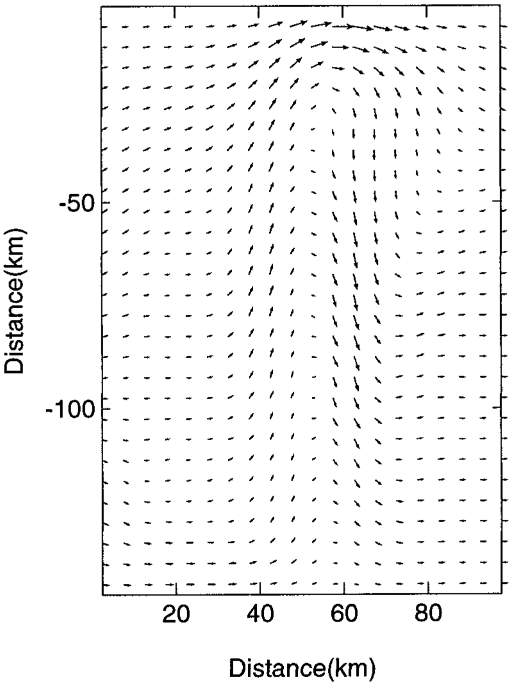

FIG. 10. Horizontal mean velocity optimized without the advective and bottom friction terms.

disappeared. They were indeed mainly due to the small misfit introduced in the external mode equations by the separate optimization of the internal mode constraints. They could easily have been reduced by one additional iteration of the algorithm. The influence of the internal mode optimization is, however, more important on the density field. The rms error is indeed increased if they are neglected (especially for the deeper sigma levels, not plotted here). Yet, the slope of the curve remains the same. It is in fact closely linked to the adjustment of the turbulent diffusion coefficients which are only crudely taken into account by the inverse model. These coefficients are considered to be constant ($A = 10^{-4}$ m^2 s^{-1}). For the present experiment, only one inversion was required: the 2350 × 2350 external mode matrix.

A crude optimization using only one constraint [the conservation of mass, given by (2.4)] has also been performed. This type of initialization has been used successfully in meteorology for more than two decades (see, for instance, Washington and Baumhefner 1974). In this case, no geostrophic adjustment is initially realized and the local upwelling is not reconstructed above the canyon. The rms error is given on Fig. 9, Optimization IV. We shall first compare this optimization with the simple interpolation case. The surface elevation rms does not increase any more with time. Indeed, the divergence of the transport, which is relaxed at the opened boundaries vanishes. Oscillations of the rms errors of the various parameters have been reduced but can still be observed and last longer, showing the adjustment of the geostrophic circulation. It takes a few hours to the different parameters to adjust. The flow is also of very poor quality along the boundaries (not shown). It is indeed relaxed toward the initial field which is not dynamically adjusted. Such a configuration is, however, very attractive since it is much less expensive and only requires the inversion of 580 × 580 sparse matrix. This optimization is yet insufficient if only a few observations are known. In this case, the boundary conditions cannot be efficiently reconstructed and one cannot assert that the general circulation is optimally represented. A barotropic current being used in the experiments shown, we can infer that the situation is going to be even worth if the conservation of mass is used alone in a strongly baroclinic and nonlinear case.

It is therefore clear that a trade-off has to be reached between the number of physical constraints which are optimized and the computer resources devoted to the optimization. This obviously depends strongly on the physics involved in the circulation studied, and if one has any prior knowledge of the main phenomena involved, we saw that a well-behaved solution could be approached with only a very few constraints. In more elaborate cases, one can infer from the present results that the prior adjustment of the dominant physical phenomenon leads to a drastic diminution of the transient processes after the initialization. The sensitivity study also confirms the importance of the balance of mass to avoid the generation of inertia-gravity waves and if the field is to be used in the boundary conditions. The choice of the constraint does not have to be restricted to the set of tangent equations and one can think about combinations of these constraints in order to specify for instance the conservation of energy.

b. Data covariance matrix

Based on (2.25), the data covariance matrix is divided into two parts. Figure 11 gives the diagonal and off-diagonal terms corresponding to the density points in matrix $\mathbf{B}_E$. The spectrum S in (3.4) is still used, and the parameters of influence of the anisotropic matrices are the same as for the experiment presented in the previous section. Figure 11a shows the diagonal elements of the isotropic matrix. Local minima can be observed in the vicinity of the forcing points, that is, every 25 km. Then, the further the forcing points, the larger the variance. The shape of the hills and holes depends on the type of interpolation used. In the present case, a Gaussian interpolation with a small radius $R = 10$ km (appendix B) gives rather sharp slopes in the vicinity of the minima and maxima. In Fig. 11b, the diagonal terms of the delta covariance matrix are maximum where the discrepancies between the bottom topography on 5 km and 25 km grids are maximum, that is, above the canyon and along the slope. Indeed, the canyon is too narrow to be described by the large-scale grid, and the sharp slope cannot be accurately represented either.

Figures 11c and 11d give the off-diagonal terms cor-

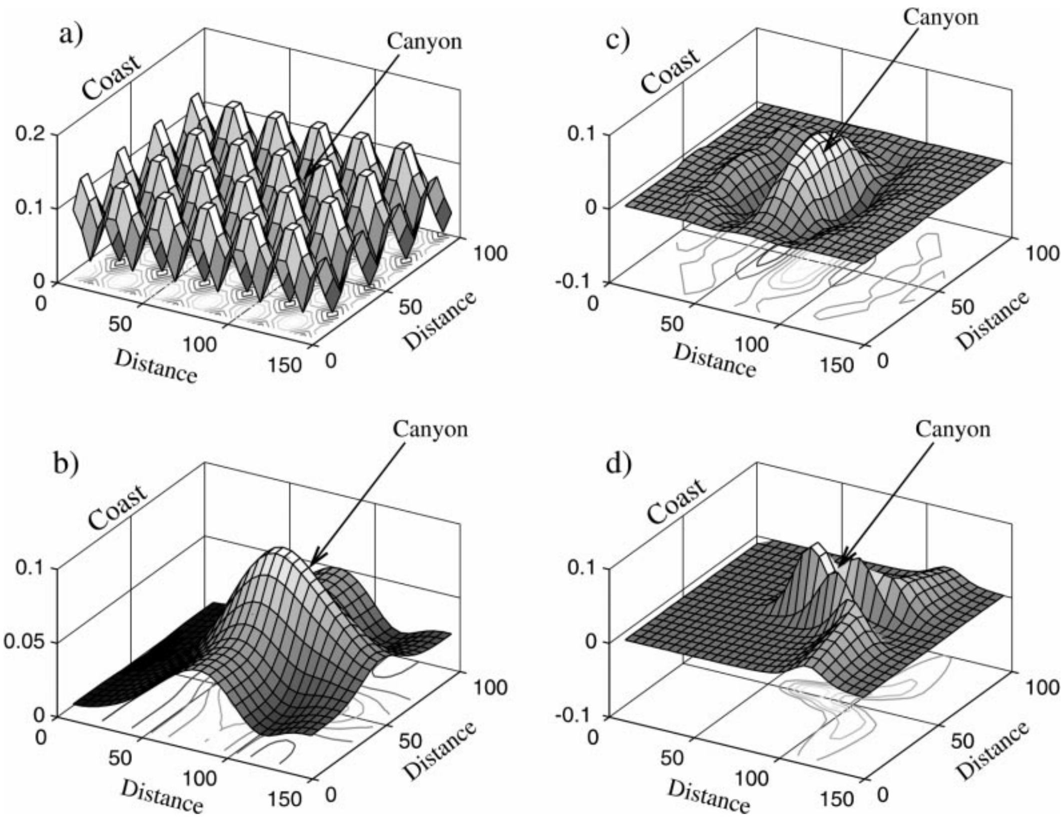

FIG. 11. Data covariance matrices: (a) diagonal terms of the isotropic matrix (2.26), (b) diagonal terms of the anisotropic delta matrix (2.32), (c) off-diagonal terms of the isotropic matrix (2.26), and (d) off-diagonal terms of the anisotropic matrix (2.31). Distances are given in kilometers.

responding to the density at a point located in the middle of the canyon. Figure 11c shows a bell-shaped curve, which is a consequence of the use of the Gaussian global energy spectrum *S.* A small anisotropy can be observed, and the alongshore covariance terms are larger than the cross-shore ones. This is a consequence of the greatest number of ‘‘data’’ points in the along shore direction. Even if the global energy spectrum is isotropic, the interpolation matrix $\mathbf{Z}$ is not necessarily, and, due to (2.26), $\mathbf{B}_{\mathrm{iso},E}$ is not exactly isotropic either. Figure 11d shows for the same point the greatest correlations along the lines of constant topography and perpendicular to them. The chosen point is right in the middle of the canyon, and we can easily recognize the shape of the canyon in the contour plot. The influence of the parameters $\mathrm{d}_{\mathrm{ref}}$ and $\mathrm{h}_{\mathrm{ref}}$ can easily be studied on such a plot in so far as they are responsible for the thickness of the ‘‘hill’’ along the lines of constant bottom topography and perpendicular to it.

Fig. 11 gives a detailed presentation of the numerous phenomena that are taken into account in the data covariance matrix such as the type of interpolation, the spatial distribution of the observation points, the accuracy of the bottom topographies, the isotropic energy spectrum in the region, and the local shape of the bottom topography. These are combined to describe the multivariate field (x_E^s) statistically. The same procedure is followed and the same type of conclusions can be reached for the data covariance matrix $\mathbf{B}_{\mathrm{iso},I}$. As a comparison, Fig. 12 shows the same type of results but for the model error covariance matrix ($\mathbf{R}$)(see section 4d).

For the twin experiments, the optimized initial field was obtained with a standard normal spectrum. The specific influence of the off-diagonal elements of the data covariance matrix on the analysis field is quite difficult to separate from the remaining of the solution. The only direct and trivial influence is linked to their magnitude: the larger it is, the larger the modification of the initial field (see, for instance, the discussion about the influence of the decay length scale in Denman and Freeland 1985). However, the influence of the shape of the spectra is not straightforward. The use of various types of energy spectra for the previous experiment shows that the main characteristics of the circulation (the upwelling) was strongly persistent. The only significant change that could be obtained by specifying a particular covariance matrix is the creation of small amplitude meanders on the upstream side of the canyon (not shown) when the spectrum contains energy around the first Rossby radius (22 km). This observation leads to the conclusion that the specification of a covariance matrix has a strong influence on the analysis fields only when it induces processes that agree with the physics of the coastal model. However, further tests based on more realistic experiments are still necessary. In fact, this implies that the inverse method does not ‘‘create’’ structures, but rather ‘‘corrects’’ the interpolated fields where they contradict the most the physics of the model on the small-scale grid (topographic constraints, geostrophic balance, . . .). One can also infer that even if no a priori statistics about the global energy spectra is available, this one can be approximated with some prior knowledge of the main physical structures of the circulation of the region, and in cases where no statistics at all can be provided, a Gaussian spectra can be adjusted to obtain a smoothed solution, the nonhomogeneous anisotropic part of the matrix being used to ‘‘adapt’’ this crude isotropic covariance matrix to the topography of the region.

c. Isotropic spectrum and nonisotropic parameterization in a simple case

For the twin experiments, the optimized initial field was obtained with a standard normal spectrum (S). We investigate here the sensitivity of the optimized solution to the type of global energy spectrum and to the anisotropic part of the covariance matrix. The influence of the extradiagonal terms of the covariance matrix is investigated in term of the length scales of the circulation.

The specific influence of the off-diagonal elements of the data covariance matrix are quite difficult to separate from the remaining of the solution. The only direct and trivial influence is linked to their magnitude: the larger it is, the larger the modification of the initial field (see, for instance, the discussion about the influence of the decay length scale in Denman and Freeland 1985). However, no clear conclusion can be drawn concerning the influence of the shape of the spectra.

To go a little further, we propose a very simple experiment. We consider the same easterly current as for the twin experiments, but the only constraint that is kept in the inverse method is the conservation of mass (2.4) (as a weak constraint). In this section, $\mathbf{R}_E$ is not computed following the algorithm given in section 2f. It is instead chosen diagonal with rather large misfit [10^{-5} $(\mathrm{m\ s^{-1}})^2$]. In such a configuration, no strong physical length scales are introduced in the flow by the constraints and the influence of the off-diagonal elements of the covariance matrix becomes crucial.

Figure 13 shows the results obtained with the isotropic spectra given in Table 1, and a fourth mean velocity field obtained when the anisotropic part of the data covariance matrix is added. The mean velocity fields obtained with the pick spectra S_2 and S_3 (Figs. 13a and 13b) clearly show the appearance of meanders and eddies of 10 and 20 km, respectively. A comparison of Figs. 13b and 13c shows that with the parameterized anisotropy (2.32) the strongest meanders and eddies are trapped by the bottom topography on the shelf or inside the canyon. We have shown in the previous section that a larger variance was specified right above the canyon and along the slope (Fig. 13b). In this region, the data covariances are larger, and the constraints are well adjusted. The situation is now different if we compare Figs. 13b and 13d. The latest was obtained using only the anisotropic matrix given by Eq. (2.31). We observe that unlike in Fig. 13c, some specific directions of flow have appeared. Indeed, we can notice 20-km meanders along the coast in the northern part of the domain, and a meandering current going around the canyon. We specified that the line of constant bottom topography were preferred directions of flow. As a consequence, the inverse method adjusts the conservation of mass along these directions.

The results of this latest sensitivity test have to be considered with a lot of care. It would be a mistake to conclude that one can ‘‘artificially’’ create any types of flow specifying the corresponding covariance matrix. Indeed, it must be reminded that the previous results were obtained with only one (very) weak constraint. This obviously does not ensure the stability of the solution, and the ‘‘transient’’ tendency terms are very large in (2.9). Moreover, the use of various type of energy spectra in the previous experiment have shown that the main characteristics of the circulation (the upwelling) was strongly persistent. The only significant change that could be obtained by specifying a particular covariance matrix is the creation of very small amplitude meanders on the upstream side of the canyon when the spectrum contains energy around the first Rossby radius. This very first observation, which still has to proven on a more elaborated case, would lead to the conclusion that the specification of a physically inconsistent covariance matrix does not lead to spurious meandering or eddy generation.

d. Model covariance matrix with randomly varying initial field

The same domain, grid, and boundary conditions as for the reference simulation are eventually used to perform a set of statistical reference simulations in order to evaluate the magnitude and spatial correlations of the nonlinear terms in the direct model. With such a procedure we expect to show that the direct evaluation of the nonlinear components of the model covariance matrix computed following the algorithm given in (2.5) is statistically realistic.

Every simulation is integrated till an equilibrium state is reached (5 days). However, the runs differ from each other by the addition of spatially correlated noise on the initial mean velocity (Evensen 1994) and FRS relaxation fields. The main correlation length is chosen to be five times the grid scale, that is 25 km. Figure 12a shows the evolution of the 2-norm of the covariance matrix $\mathbf{R}_E$ as a function of the number of statistical runs. As can be seen, the procedure converges for about 100 realizations. The interesting point is the L_2-norm converges

TABLE 1. Classic spectra used in the sensitivity study (J_0 is the Bessel function).

S	Normal Gaussian spectrum (3.4)	$\exp(-a^2k^2/4)$	$a = 12.5$ km
S_2	Pick spectrum at 10 km (Denman and Freeland 1985)	$\int_0^\infty \exp(-r^2/a^2)J_0(r/b)\exp[-(2i\pi)kr]\,dr$	$a = 1$ km $b = 1.6$ km
S_3	Pick spectrum at 20 km	$\int_0^\infty \exp(-r^2/a^2)J_0(r/b)\exp[-(2i\pi)kr]\,dr$	$a = 2$ km $b = 3.2$ km

to a value that is very close to the direct evaluation which was used in the previous section (dash line in Fig. 12a). This implies that, at least in simple cases, such an approximate matrix can be used with confidence. Tests have been conducted in more realistic cases using different correlation scales, and they also show that the direct evaluation of the model covariance matrix gave at least the correct order of magnitude.

Figures 12b and 12c show the diagonal elements of the covariance matrix. The regions of larger currents are also the regions of larger variance. On both plots, the variance is maximum at the bottom of the canyon, where the circulation follows the bottom topography and changes direction. A second, smaller maximum can be observed at the canyon head. This is consistent with the large bottom currents and, as a consequence, with the large bottom friction observed in the region (not shown).

5. Discussion

We proposed an inverse method that is quite affordable in terms of computing resources and which allows the initialization of a coastal free-surface model with various types of data. It is also well adapted to specify the incoming flow along the open boundaries. If different OGCM output fields are known at various dates, a linear interpolation of the reconstructed fields can be used at the open boundary. Methodological, physical, and dynamical assumptions have been made to reach such a result. The ''methodological'' assumptions are related to the least squares optimization, and by ''physical'' assumptions we mean the approximations or parametrizations that were used to deal with the observations themselves, lastly the ''dynamical'' assumptions concern the use of the tangent linear.

The most fundamental type of assumption concerns in fact the least squares optimization. It can be shown that whatever the a priori statistics in the observation and model spaces are, the least squares estimator minimizes the variance (Tarantola 1987). However, the minimization of the variance is not such a good criterion most of the time, and mathematically speaking, this can be a serious restriction to the use of least squares optimization. It must therefore be kept in mind that the covariance matrix must always remain close to Gaussian approximations. If such is not the case, the best validation one can think about, is the good behavior of the reconstructed fields: weak transient processes during the initialization procedure, physically realistic behavior of the circulation around the FRS and close resemblance with the observations. Anyway, one must always remember that only ''quasi-optimal'' solutions are dealt with, and that posterior validations are always necessary.

Two physical assumptions emerge. First, concerning the interpolation scheme: one needs to grid the data. We thus focused on the lower length scale (L) that could be extracted from the observations in order to choose an interpolation method that does not bring in too much small-scale noise and as importantly that does not smooth in excess the large-scale structures. For the canyon experiment, the large-scale circulation was very simple, and the interpolation scheme was not crucial. However, the diagonal elements of the data covariance matrix (Fig. 11a) take into account the position of the ''observations'' and the interpolation scheme [via (2.26)]. To end up with, we considered that the data did not include any bias (2.23). This assumption is related to the difficulty to evaluate the variable mean in optimal analysis (''kriging''; Le Traon 1990). This rather classical problem is a consequence of the fact that, very often, the observation noise can be characterized adequately, but no data about the signal itself is available. Moreover, only very poor tests of the validity of the assumption can be made.

The second physical assumption is related to the choice of the data covariance matrix. We showed that, in coastal oceanography, an anisotropic correction was to be added to the conventional homogeneous, isotropic covariance matrix. The various parameters involved in this specification have been determined based on physical considerations such as the bottom topography, and the a priori strength of the geostrophic current. In cases where the isotropic energy spectrum is not available, an analytical spectrum can be derived based on a priori knowledge of the general circulation or on smoothing requirements. The study of the structure of the data covariance matrix in section 4b has shown that the various assumptions made to derive the matrix lead to a result which was physically acceptable. Another point concerns the computation of the standard deviations of the different fields which has to be considered with care. Indeed, the underestimation of the standard deviation of one of the fields can lead to small modifications of this field by the inverse method, which in turn leads to larger modifications of all the other fields to reach the objec-

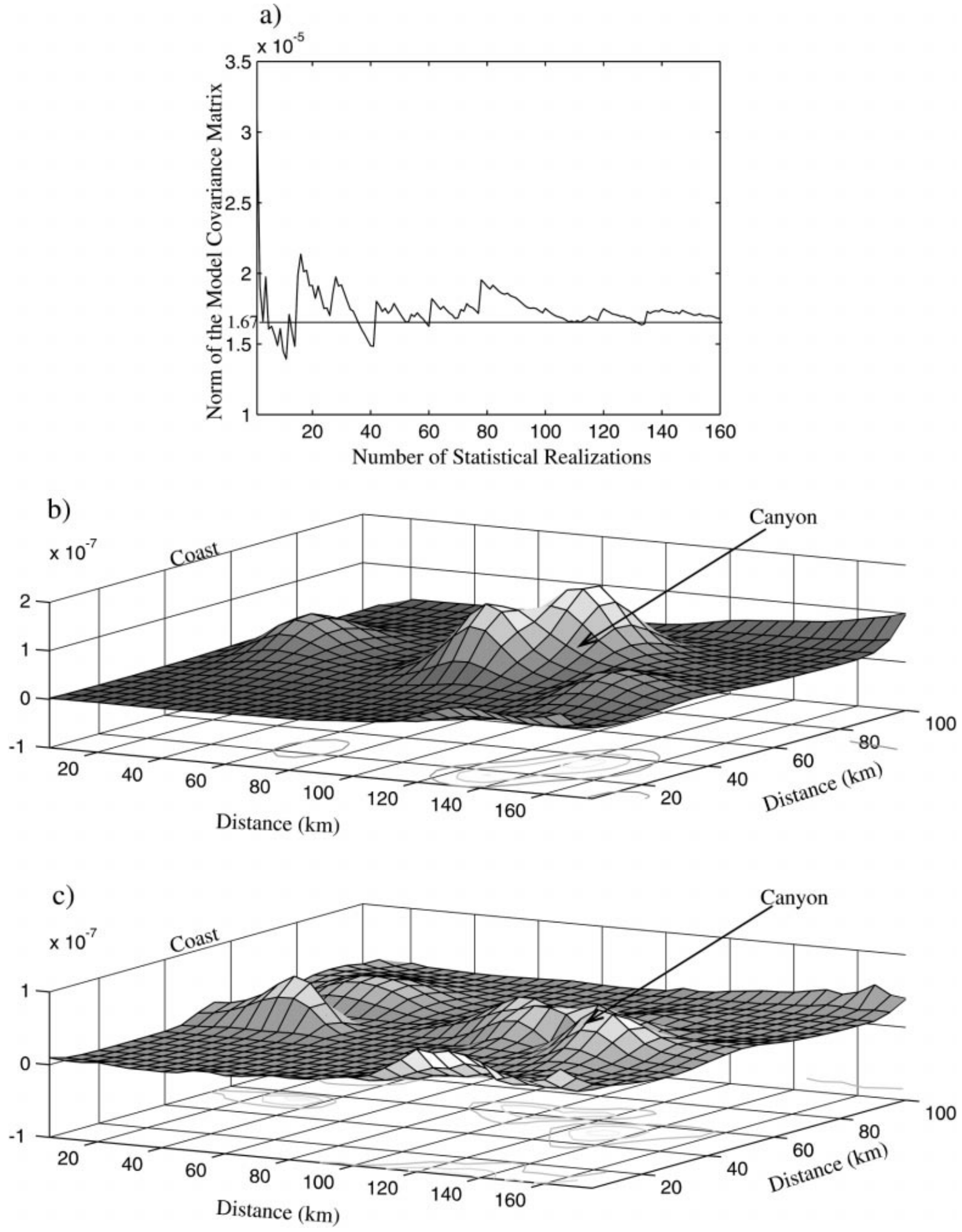

FIG. 12. (a) Solid line: norm of the model covariance matrix as a function of the number of statistical realizations; dashed line: norm of the model covariance matrix with no additional noise; (b) diagonal elements of the model covariance matrix: momentum equation in the *x* direction; (c) diagonal elements of the model covariance matrix: momentum equation in the *y* direction.

tive. This can be important to weight the relative effect of the free-surface elevation and of the integrated density, yet, only the order of magnitude proved to be necessary. In the experiment proposed the adjustment is easily obtained using a spatial average. In more elaborate cases, some physical insights might be necessary to adjust the variance, if not the method has to be iterated.

The last type of assumptions concerns the treatment of the free-surface model dynamics. The use of the tangent linear equations can be a crude approximation when strongly nonlinear flows are treated. In such cases, the

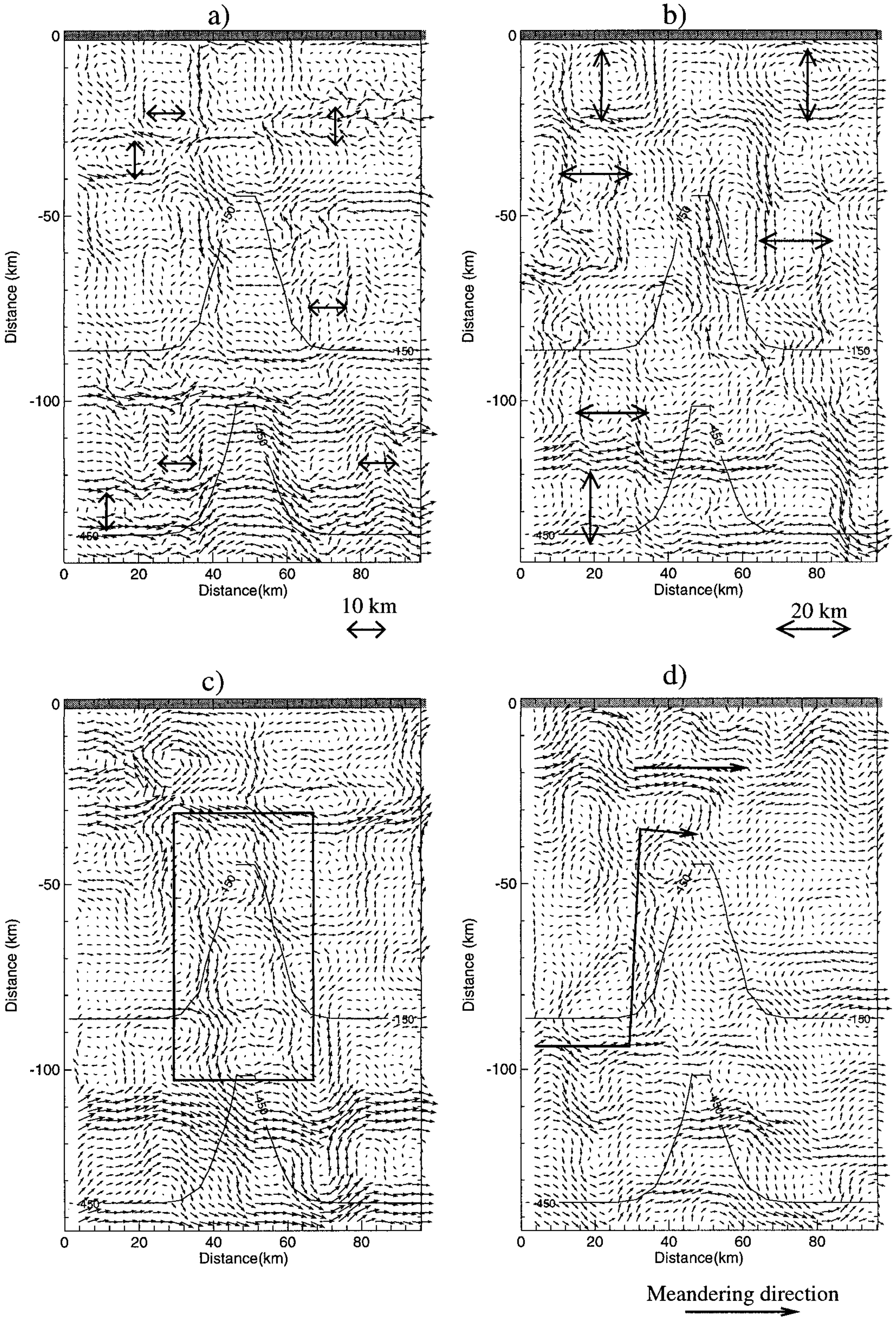

FIG. 13. Sensitivity of the mean velocity to the data covariance matrix. Solid lines: bottom topography contours. (a) Isotropic matrix with spectrum S_2 (10-km pick); (b) Isotropic matrix with spectrum S_3 (20-km pick); (c) Anisotropic matrix: spectrum S_3 (20-km pick), $\mathbf{B}_{\mathrm{delta}}$; (d) Anisotropic matrix: spectrum S_3 (20-km pick), $\mathbf{B}_{\mathrm{topo}}$.

method hopefully converges after a few iterations. We also showed that the variables and constraints could be chosen depending on the physics of the circulation. In certain cases, a single internal constraint for density can in particular be used. In this case, the computation of the new temperature and salinity fields based on the optimized density did not change the dynamical behavior of the solution (the density is not modified), but reduces significantly the number of variables. In fact, this assumption can be met more often that one can first think of. Indeed, the inverse method does not reconstruct the whole fields, but only makes little modifications of the various variables so that the constraint misfits be reduced. As a first approximation, one can consider that the adjustments of the temperature and salinity fields are similar.

The computation of the model covariance matrix using a Monte Carlo method shows that the direct computation of the truncated nonlinear terms leads to a result that was statistically acceptable. In more elaborated cases where the ‘‘true’’ tendencies do not vanish, a second error term has to be added to the matrix to take into account the errors on the evaluation of the tendencies.

6. Conclusions

Based on optimal interpolation techniques, the inverse method we have developed offers an original approach to the problem of the realistic initialization of a free-surface model. We have shown that the optimized run exhibits a realistic circulation after only a few iterations of the direct model. This is particularly interesting in coastal oceanography where short runs are very often designed to study particular feature of the circulation. The initial field can also be used to specify the large-scale circulation at the open boundary in a way that is consistent with the dynamical model and thus enables the use of an FRS. The physical constraints used for the optimization can be adapted to the physics of the region considered. This flexible procedure allows the adaptation of the optimization to the computing resources.

An original approach is also detailed to construct the data covariance matrix when very few statistical realizations are available. The covariance matrices are indeed divided into two parts: the isotropic matrix, computed from the global energy spectrum; and the nonhomogeneous, anisotropic matrix, parametrized using the bottom topography. It appeared that an approximation of the global energy spectra can be found based on the knowledge of the circulation of the region. In the worst case, a smoothing analytical Gaussian spectrum can be used. The second covariance matrix, namely, the model covariance matrix, contains statistics about the nonlinear terms approximated in the tangent equations. It is computed by evaluating the errors on the tendencies and the nonlinear terms truncated in the tangent linear model.

Acknowledgments. This study was supported by the EPSHOM under the contract EPSHOM/CNRS 27/95, the Programme National d’Océanologie Côtière (PNOC) and was undertaken in the framework of MAST-III-ELOISE European Union METRO-MED Project (Dynamics of Matter Transfer and Biochemical Cycles: Their Modeling in Coastal Systems of the Mediterranean Sea). We acknowledge the support of the European Commission’s Marine Science and Technology (MAST) Program under Contract CT 960049. We thank S. Prieur for his technical support. Finally, we gratefully thank Pierre DeMey for his constructive comments.

APPENDIX A

Tangent Equations

The tangent equations are a linearized version of the primitive Eqs. (2.1)–(2.6) in the vicinity of the ‘‘first-guess’’ solution (Courtier and Talagrand 1990). The validity of a similar linear tangent assumption has been studied carefully by Lacarra and Talagrand (1988) based on perturbations of an atmospheric field. A rather interesting characteristic of the tangent equations is that, following the free-surface coastal model approach, they are based on a mode-splitting approach. In order to simplify these equations, we first define the ‘‘advection operators’’ $ADV_2(b;\, a_1,\, a_2)$ and $ADV_3(b;\, a_1,\, a_2,\, a_3)$, which give the advection of the quantity b by the velocity components (a_1, a_2, a_3),

$$ADV_2(b;\, a_1,\, a_2) = \frac{\partial a_1 b}{\partial x} + \frac{\partial a_2 b}{\partial y}$$

$$ADV_3(b;\, a_1,\, a_2,\, a_3) = \frac{\partial a_1 b}{\partial x} + \frac{\partial a_2 b}{\partial y} + \frac{\partial a_3 b}{\partial z}.$$

Substituting $(\overline{u}, \overline{v})$ by $(\overline{u}^b + \delta\overline{u},\, \overline{v}^b + \delta\overline{v})$, the conservation of mass (2.4) can be written

$$Q_1(\delta\overline{u},\, \delta\overline{v}) = -Q_1(\overline{u}^b,\, \overline{v}^b) - \left.\frac{\partial \eta}{\partial t}\right|^t, \qquad \text{(A1)}$$

where

$$Q_1(\overline{u},\, \overline{v}) = \frac{\partial}{\partial x}(h\overline{u}) + \frac{\partial}{\partial y}(h\overline{v}).$$

When the second-order term in $\delta\overline{u}$ and $\delta\overline{v}$ are neglected, the conservation of momentum for the external mode [(2.5) and (2.6)] can be written

$$\begin{aligned} &Q_2(\delta\overline{u},\, \delta\overline{v},\, \delta\overline{\rho},\, \delta\eta) + ADV_2(\delta\overline{u};\, H\overline{u}^b,\, H\overline{v}^b) \\ &\quad + ADV_2(\overline{u}^b;\, H\delta\overline{u},\, H\delta\overline{v}) + \Gamma'^x_B \\ &\qquad = -\left.\frac{\partial H\overline{u}}{\partial t}\right|^t - Q_2(\overline{u}^b,\, \overline{v}^b,\, \overline{\rho}^b,\, \eta^b) \\ &\qquad\quad - \frac{g}{\rho_0}\int_{z=0}^{\eta} \frac{\partial}{\partial x}\int_{z'=z}^{\eta} (\rho^b - \overline{\rho}^b + \rho_0)\, dz'\, dz \\ &\qquad\quad - ADV_2(\overline{u}^b;\, H\overline{v}^b,\, H\overline{u}^b) + \Gamma^x_W - \Gamma^x_B \qquad \text{(A2)} \end{aligned}$$

$$
\begin{aligned}
&Q_3(\delta\overline{u},\ \delta\overline{v},\ \delta\overline{\rho},\ \delta\eta) + ADV_2(\delta\overline{v};\ H\overline{u}^b,\ H\overline{v}^b)\\
&\quad + ADV_2(\overline{v}^b,\ H\delta\overline{u},\ H\delta\overline{v}) + \Gamma_B^{y'}\\
&\quad = -\frac{\partial H\overline{v}}{\partial t}\bigg|^t - Q_3(\overline{u}^b,\ \overline{v}^b,\ \overline{\rho}^b,\ \eta^b)\\
&\qquad - \frac{g}{\rho_0}\int_{z=0}^{\eta}\frac{\partial}{\partial y}\int_{z'=z}^{\eta}(\rho^b - \overline{\rho}^b + \rho_0)\,dz'\,dz\\
&\qquad - ADV_2(\overline{u}^b,\ H\overline{u}^b,\ H\overline{v}^b) + \Gamma_W^y - \Gamma_B^y, \qquad \text{(A3)}
\end{aligned}
$$

where

$$
\begin{aligned}
Q_2(\overline{u},\ \overline{v},\ \rho,\ \eta) &= -Hf\overline{v} + \frac{g}{\rho_0}\int_{z=0}^{\eta}\frac{\partial}{\partial x}\int_{z'=z}^{\eta}\rho\,dz'\,dz\\
&\quad - A\left[\frac{\partial^2\overline{u}}{\partial x^2} + \frac{\partial^2\overline{u}}{\partial y^2}\right] + gH\frac{\partial\eta}{\partial x},\\
Q_3(\overline{u},\ \overline{v},\ \rho,\ \eta) &= Hf\overline{u} + \frac{g}{\rho_0}\int_{z=0}^{\eta}\frac{\partial}{\partial y}\int_{z'=z}^{\eta}\rho\,dz'\,dz\\
&\quad - A\left[\frac{\partial^2\overline{v}}{\partial x^2} + \frac{\partial^2\overline{v}}{\partial y^2}\right] + gH\frac{\partial\eta}{\partial y}.
\end{aligned}
$$

The bottom friction is given by

$$\Gamma_B^x = C_D\|\overline{V}^b\|u_B^b \quad \text{and} \quad \Gamma_B^y = C_D\|\overline{V}^b\|v_B^b,$$

and, for its linearized version,

$$\Gamma'^x_B = \frac{1}{2}C_D\frac{u_B^b}{\|\overline{V}^b\|}(2\overline{v}^b\delta\overline{v} + 2\overline{u}^b\delta\overline{u}) + C_D\|\overline{V}^b\|\delta\overline{u} \quad \text{and}$$

$$\Gamma'^y_B = \frac{1}{2}C_D\frac{v_B^b}{\|\overline{V}^b\|}(2\overline{v}^b\delta\overline{v} + 2\overline{u}^b\delta\overline{u}) + C_D\|\overline{V}^b\|\delta\overline{v}$$

where $\|\overline{V}^b\| = (\overline{u}^{b^2} + \overline{v}^{b^2})^{1/2}$ is the norm of the mean velocity, Γ_W^x and Γ_W^y are the wind stress components, C_D is the drag coefficient, and $(u_B^b,\ v_B^b)$ is the bottom velocity.

For the internal mode, the linearization with respect to δu, δv and $\delta\rho$ of the momentum Eqs. (2.2) and (2.3) now read

$$
\begin{aligned}
&Q_4(\delta u, \delta v, \delta\rho) + ADV_3(\delta u; u^b, v^b, w^b) + ADV_2(u^b; \delta u, \delta v)\\
&\quad = -\frac{\partial u}{\partial t}\bigg|^t - Q_4(u^b, v^b, \rho^b - \rho_0) - ADV_3(u^b; u^b, v^b, w^b)\\
&\qquad - g\frac{\partial(\eta^b + \delta\eta)}{\partial x} + \frac{\partial}{\partial z}\left(K\frac{\partial u^b}{\partial z}\right), \qquad \text{(A4)}
\end{aligned}
$$

$$
\begin{aligned}
&Q_5(\delta u, \delta v, \delta\rho) + ADV_3(\delta v; u^b, v^b, w^b) + ADV_2(v^b; \delta u, \delta v)\\
&\quad = -\frac{\partial v}{\partial t}\bigg|^t - Q_5(u^b, v^b, \rho^b - \rho_0) - ADV_3(v^b; u^b, v^b, w^b)\\
&\qquad - g\frac{\partial(\eta^b + \delta\eta)}{\partial y} + \frac{\partial}{\partial z}\left(K\frac{\partial v^b}{\partial z}\right), \qquad \text{(A5)}
\end{aligned}
$$

where

$$
\begin{aligned}
Q_4(u,\ v,\ \rho - \rho_0) &= -fv + \frac{g}{\rho_0}\frac{\partial}{\partial x}\int_{z'=z}^{\eta}(\rho - \rho_0)\,dz'\\
&\quad - A\left[\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right],\\
Q_5(u,\ v,\ \rho - \rho_0) &= fu + \frac{g}{\rho_0}\frac{\partial}{\partial y}\int_{z'=z}^{\eta}(\rho - \rho_0)\,dz'\\
&\quad - A\left[\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y}\right].
\end{aligned}
$$

To decrease the number of variables, the equations for T and S can be replaced by one single equation for the density ρ, and the last constraint is thus written for $\psi \in \{T,\ S,\ \rho\}$:

$$
\begin{aligned}
&Q_6(\delta\psi) + ADV_3(\delta\psi;\ u^b,\ v^b,\ w^b) + ADV_2(\psi^b;\ \delta u,\ \delta v)\\
&\quad = -\frac{\partial\psi}{\partial t}\bigg|^t - Q_6(\psi^b) - ADV_3(\psi^b;\ u^b,\ v^b,\ w^b)\\
&\qquad + \frac{\partial}{\partial z}\left(K\frac{\partial\psi^b}{\partial z}\right), \qquad \text{(A6)}
\end{aligned}
$$

where

$$Q_6(\psi) = -A[(\partial^2\psi/\partial x^2) + (\partial^2\psi/\partial y^2)].$$

In the case where $\psi = \rho$, (A6) is the linearized version of the conservation of density:

$$
\begin{aligned}
&\frac{\partial\rho}{\partial t}\bigg|_{\text{physical}} + \frac{\partial u\rho}{\partial x} + \frac{\partial v\rho}{\partial y} + \frac{\partial w\rho}{\partial z}\\
&\quad = A\left[\frac{\partial^2\rho}{\partial x^2} + \frac{\partial^2\rho}{\partial y^2}\right] + \frac{\partial}{\partial z}\left(K\frac{\partial\rho}{\partial z}\right). \qquad \text{(A7)}
\end{aligned}
$$

Last, the vertical velocity (w^b) is computed from the diagnostic equation (2.1)

$$\frac{\partial w^b}{\partial z} = -\left(\frac{\partial u^b}{\partial x} + \frac{\partial v^b}{\partial y}\right). \qquad \text{(A8)}$$

The tendency terms (written $\partial/\partial t|^t$) are rather difficult to estimate but can be computed when OGCM outputs are used. However, even in that case, the present method does not allow a control of the posterior trajectory of the model, and the specification of the first derivative can be pointless if the circulation is complex. However, the magnitude of the error due this shortage is taken into account in the model covariance matrices ($\mathbf{R}_E$ or $\mathbf{R}_I$). As far as the present study is concerned, an equilibrium field is seek for and the tendency terms all vanish.

In matrix notations, $\mathbf{M}'_E$ is defined from Eqs. (A1)–(A3) and is written

$$M'_E\delta x = y^o - M'_E x^b + F_E, \qquad \text{(A9)}$$

and M'_I from Eqs. (A4)–(A6):

$$M_I'\delta x = y^o - M_I' x^b + F_I. \qquad \text{(A10)}$$

APPENDIX B

Gaussian Interpolation

The external and internal mode variables x^b are interpolated using a radially symmetric Gaussian filter:

$$x^b_{i,j} = \frac{\displaystyle\sum_{(\alpha,\beta)} \exp\left(-\frac{r^2}{R_L^2}\right) y^d_{\alpha,\beta}}{\displaystyle\sum_{(\alpha,\beta)} \exp\left(-\frac{r^2}{R_L^2}\right)} \qquad \text{(B1)}$$

$$x^b_{i,j,k} = \frac{\displaystyle\sum_{(\alpha,\beta)} \exp\left(-\frac{r^2}{R_L^2}\right)[y^d_{\alpha,\beta,\gamma_1} d + y^d_{\alpha,\beta,\gamma_2}(1 - d)]}{\displaystyle\sum_{(\alpha,\beta)} \exp\left(-\frac{r^2}{R_L^2}\right)}, \qquad \text{(B2)}$$

where (i, j) is a point of the small-scale grid and (α, β) a point of the observation field. The distance between the points (i, j) and (α, β) is given by r and R_L is a constant which has the same order of magnitude as L. The vertical distance between points (i, j, k) and (α, β, γ) is given by d and the vertical interpolation is linear, γ_1 and γ_2 satisfying $z(\gamma_2) \le z(k) \le z(\gamma_1)$. Using matrix notations, (B1) and (B2) can be written

$$x^b = Zy^d. \qquad \text{(B3)}$$

APPENDIX C

Euler–Lagrange Equations for Mass Conservation

Starting from the analysis field, we define a new cost function J_μ in order to correct the remaining misfits in the mass conservation equation

$$J_\mu = 2\mu_{\mathrm{mc}}[\mathbf{M}_{\mathrm{mc}}\delta\mathbf{x} - (\mathbf{y}^o - \mathbf{M}_{\mathrm{mc}}\mathbf{x}^a)], \qquad \text{(C1)}$$

where $\mu_{\mu\chi}$ is a Lagrange multiplier, and $\mathbf{M}_{\mathrm{mc}}$ is defined by (A1), written as

$$\mathbf{M}_{\mathrm{mc}}\delta\mathbf{x} = \mathbf{y}^o - \mathbf{M}_{\mathrm{mc}}\mathbf{x}^a. \qquad \text{(C2)}$$

The new analysis field is eventually given by

$$\mathbf{x}^A = \mathbf{x}^a + \mathbf{M}_{\mathrm{mc}}^{-1}(\mathbf{y}^o - \mathbf{M}_{\mathrm{mc}}\mathbf{x}^a). \qquad \text{(C3)}$$

REFERENCES

Arakawa, A., and M. J. Suarez, 1983: Vertical differencing of the primitive equations in sigma coordinates. *Mon. Wea. Rev.,* **111,** 34–45.

Bennett, A. F., 1992: *Inverse Methods in Physical Oceanography.* Cambridge University Press, 346 pp.

——, and P. E. Kloeden, 1978: Boundary conditions for limited area forecasts. *J. Atmos. Sci.,* **35,** 990–996.

——, and P. C. McIntosh, 1982: Open ocean modeling as an inverse problem: Tidal theory. *J. Phys. Oceanogr.,* **12,** 1004–1018.

Blumberg, A. F., and G. L. Mellor, 1987: A description of a three-dimensional coastal circulation model. *Three Dimensional Coastal Ocean Models,* N. S. Heaps, Ed., Coastal and Estuarine Sciences Series, Vol. 4, Amer. Geophys. Union, 1–16.

Bodgen, P. S., P. Malanotte-Rizzoli, and R. Signell, 1996: Open-ocean boundary conditions from interior data: Local and remote forcing of Massachusetts Bay. *J. Geophys. Res.,* **101** (C3), 6487–6500.

Bourke, W., and J. L. Mc Gregor, 1983: A nonlinear vertical mode initialization scheme for a limited area prediction model. *Mon. Wea. Rev.,* **111,** 2285–2297.

Brasseur, P., 1991: A variational inverse methods for the reconstruction of general circulation fields in the northern Bering Sea. *J. Geophys. Res.,* **96** (C3), 4891–4907.

Bretherton, F. P., R. E. Davis, and C. B. Fandry, 1976: A technique for objective analysis and design of oceanographic experiments applied to MODE-73. *Deep-Sea Res.,* **23,** 559–582.

Brink, K. H., 1982: The effect of bottom friction on low-frequency coastal trapped. *J. Phys. Oceanogr.,* **12,** 127–133.

Bryan, F. O., C. W. Boning, and W. R. Holland, 1995: On the mid-latitude circulation in a high-resolution model of the North Atlantic. *J. Phys. Oceanogr.,* **25,** 289–305.

Busalacchi, A. J., 1997: Oceanic observations. *J. Meteor. Soc. Japan,* **75,** 131–154.

Chu, P. C., F. Chenwu, and L. L. Ehret, 1997: Determination of open boundary conditions with an optimization method. *J. Atmos. Oceanic Technol.,* **14,** 723–734.

Courtier, P., and O. Talagrand, 1990: Variational assimilation of the meteorological observations with the direct and adjoint shallow-water equations. *Tellus,* **42A,** 531–549.

Davies, H. C., 1976: A lateral boundary formulation for multi-level prediction models. *Quart. J. Roy. Meteor. Soc.,* **102,** 405–418.

Deleersnijder, E., 1989: Upwelling and upsloping in three-dimensional marine models. *Appl. Math. Modelling,* **13,** 462–467.

De Mey, P., 1997: Data assimilation at the oceanic mesoscale: A review. *J. Meteor. Soc. Japan,* **75,** 415–427.

Denman, K. L., and H. J. Freeland, 1985: Correlation scales, objective mapping and a statistical test of geostrophy over the continental shelf. *J. Mar. Res.,* **43,** 517–539.

Engedahl, H., 1995: Use of the flow relaxation scheme in a three-dimensional baroclinic ocean model with realistic topography. *Tellus,* **47A,** 365–382.

Estournel, C., V. Kondrachoff, P. Marsaleix, and R. Vehil, 1997: The plume of the Rhone: Numerical simulation and remote sensing. *Contin. Shelf Res.,* **17,** 899–924.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.,* **99** (C5), 10 143–10 162.

Gandin, L. S., 1965: *Objective Analysis of Meteorological Fields.* Israel Program for Scientific Translations, 242 pp.

Gilbert, J. R., C. Moler, and R. Schreiber, 1992: Sparse matrices in MATLAB: Design and implementation. *SIAM J. Matrix Anal. Appl.,* **13,** 333–356.

Gunson, J. R., and P. Malanotte-Rizzoli, 1996: Assimilation studies of open-ocean flows, 1: Estimation of initial and boundary conditions. *J. Geophys. Res.,* **101** (C12), 28 457–28 472.

Hinkelmann, K., 1951: Der Mechanimus des meteorologischen Lämes. *Tellus,* **3,** 285–296.

Horton, C., M. Clifford, J. Schmitz, and L. H. Kantha, 1997: A real-time oceanographic nowcast/forecast system for the Mediterranean Sea. *J. Geophys. Res.,* **102** (C11), 25 123–25 156.

Huthnance, J. M., 1995: Circulation, exchange and water masses at the ocean margin: The role for physical processes at the shelf edge. *Progress in Oceanography,* Vol. 35, Pergamon, 353–431.

Ide, K., P. Courtier, M. Ghil, and A. C. Lorenc, 1997: Unified notation for data assimilation: Operational, sequential and variational. *J. Meteor. Soc. Japan,* **75,** 181–189.

Klinck, J. M., 1989: Geostrophic adjustment over submarine canyons. *J. Geophys. Res.,* **94** (C5), 6133–6144.

——, 1996: Circulation near submarine canyons: A modeling study. *J. Geophys. Res.,* **101** (C1), 1211–1223.

Lacarra, J. F., and O. Talagrand, 1988: Short-range evolution of small perturbations in a barotropic model. *Tellus,* **40A,** 85–95.

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

Le Traon, P. Y., 1990: A method for optimal analysis of fields with spatially variable mean. *J. Geophys. Res.,* **95** (C8), 13 543–13 547.

Machenhauer, B., 1977: On the dynamics of gravity oscillations in a shallow water, with applications to normal mode initialization. *Beitr. Phys. Atmos.,* **50,** 253–271.

Malanotte-Rizzoli, P., R. E. Young, and D. B. Haidvogel, 1989: Initialization and data assimilation experiments with a primitive equation model. *Dyn. Atmos. Oceans,* **13,** 349–378.

Marsaleix, P., C. Estournel, V. Kondrachoff, and R. Vehil, 1998: A numerical study of the formation of the Rhône River plume. *J. Mar. Res.,* **14,** 99–115.

McIntosh, P. C., 1990: Oceanographic data interpolation: Objective analysis and splines. *J. Geophys. Res.,* **95** (C8), 13 529–13 541.

Miller, R. N., 1986: Towards the application of the Kalman filter to regional open ocean modeling. *J. Phys. Oceanogr.,* **16,** 72–86.

Ooyama, K. V., 1987: Scale-controlled objective analysis. *Mon. Wea. Rev.,* **115,** 2479–2506.

Orlanski, I., 1976: A simple boundary condition for unbounded hyperbolic flows. *J. Comput. Phys.,* **21,** 255–261.

Provost, C., and R. Salmon, 1986: A variational method for inverting hydrographic data. *J. Mar. Res.,* **44,** 1–34.

Roisin, B. C., 1994: *Introduction to Geophysical Fluid Dynamics.* Prentice Hall, 320 pp.

Sasaki, Y., 1958: An objective analysis band on the variational method. *J. Meteor. Soc. Japan,* **36,** 77–88.

Seiler, U., 1993: Estimation of open boundary conditions with the adjoint method. *J. Geophys. Res.,* **98,** 22 855–22 870.

Spall, M. A., and W. R. Holland, 1991: A nested primitive equation model for oceanic applications. *J. Phys. Oceanogr.,* **21,** 205–220.

Stevens, D. P., 1990: On open boundary conditions for three-dimensional primitive equation ocean circulation models. *J. Geophys. Astrophys. Fluid Dyn.,* **51,** 103–133.

Tacker, W. C., and R. B. Long, 1988: Fitting dynamics to data. *J. Geophys. Res.,* **93** (C2), 1227–1240.

Tarantola, A., 1988: *Inverse Problem Theory.* Elsevier, 613 pp.

Thiebaux, H. J., and M. A. Pedder, 1987: *Spatial Objective Analysis with Applications in Atmospheric Science.* Academic Press, 299 pp.

Tintore, J., D.-P. Wang, E. Garcia, and A. Viudez, 1995: Near-inertial motions in the coastal ocean. *J. Mar. Syst.,* **6,** 301–312.

Turner, J. S., 1973: Buoyancy effects in fluids. *Mechanics and Applied Mathematics,* Cambridge University Press, 367 pp.

Wahba, G., and J. Wendelberger, 1980: Some new mathematical methods for variational objective analysis using splines and cross validation. *Mon. Wea. Rev.,* **108,** 1122–1143.

Washington, W. M., and D. P. Baumhefner, 1974: A method of removing Lamb waves from initial data for primitive equation models. *J. Appl. Meteor.,* **14,** 114–119.

Woodgate R. A., 1997: The effects of assimilation on the physics of an ocean model. Part II: Baroclinic identical-twin experiments. *J. Atmos. Oceanic Technol.,* **14,** 910–924.

——, and P. D. Killworth, 1997: The effects of assimilation on the physics of an ocean model. Part I: Theoretical model and barotropic results. *J. Atmos. Oceanic Technol.,* **14,** 897–909.

Wunsch, C., 1977: Determining the general circulation of the oceans: A preliminary discussion. *Science,* **196,** 871–875.

——, 1996: *The Ocean Circulation Problem.* Cambridge University Press, 442 pp.

Zhang, Q., and J. Marotzke, 1998: The importance of open-boundary estimation for an Indian Ocean GCM—Data synthesis. Center for Global Change Science Rep. No. 56, 57 pp. [Available from Woods Hole Oceanographic Institution, Woods Hole, MA 02548.]

Zou, X., I. M. Navon, M. Berger, K. H. Phua, T. Schlick, and F. X. LeDimet, 1993: Numerical experiences with Limited-Memory Quasi-Newton and Truncated Newton Methods. *SIAM J. Optimi.,* **3,** 582–608.