

A variational approach for the reconstruction of regional scale Eulerian velocity fields from Lagrangian data

Vincent Taillandier ^{a,*}, Annalisa Griffa ^{a,b}, Anne Molcard ^c

^a *CNR-ISMAR, Forte Santa Teresa, 19036 Lerici, La Spezia, Italy*

^b *RSMAS, Miami, FL, United States*

^c *CNR-ISAC, Torino, Italy*

Received 23 June 2005; received in revised form 23 September 2005; accepted 27 September 2005
Available online 21 October 2005

Abstract

A method for the reconstruction of the mesoscale Eulerian velocity field based on Lagrangian data at given sampling period is presented. A variational approach is used, where information on the float positions are combined with a simple model constraint describing the motion of particles advected in a velocity field. The velocity field of a priori specified space scale is estimated minimizing a cost function which measures the distance between the observed float positions and the model predicted ones for each sampling period. The method is first implemented considering the time-independent approximation of the velocity correction during a time interval shorter than the typical time scale of the mesoscale field. In a second step, this approximation is relaxed to consider inertial oscillations superimposed on the mesoscale field. The method is tested on a numerical regional circulation on the Northwestern Mediterranean Sea, characterized by an Eulerian time scale significantly longer than the inertial period. The observational coverage of the region varies between 80% and 16% and the corresponding error in the velocity reconstruction varies from less than 10% to approximately 50%, with a sampling period of the day. For smaller sampling intervals, the reconstruction is further improved in the time-dependent approximation. Finally, the method is used to combine information from Lagrangian data and General Circulation Models. The results suggest the useful application of the method for assimilation studies.

© 2005 Elsevier Ltd. All rights reserved.

1. Introduction

Lagrangian instruments are floating buoys moving in good approximation with the ocean current (Davis, 1991), and providing information at intervals of time Δt about their position \mathbf{r} and about other environmental parameters, depending on the specific applications. The position data \mathbf{r} provide direct information on the velocity field and on the associated transport.

Lagrangian instruments are an integral part of the Ocean Observing system, and they have generated in the last decades extensive data sets both at the surface and in the interior of the ocean. Statistical analysis of

* Corresponding author. Tel.: +39 0187 978305; fax: +39 0187 970585.
E-mail address: vincent.taillandier@sp.ismar.cnr.it (V. Taillandier).

historical data sets, especially in terms of positions \mathbf{r} , have provided a great amount of information on the world ocean in terms of mean flow and variability (e.g. Davis, 1991; Owens, 1991; Swenson and Niiler, 1996; Lavander et al., 2000; Fratantoni, 2001; Richardson, 1993; Zhang et al., 2001; Bauer et al., 2002; Zhou et al., 2002; Reverdin et al., 2003; Veneziani et al., 2004), greatly contributing to our knowledge and understanding of the ocean circulation. Various methodologies have been developed and used to estimate the mean Eulerian velocity \mathbf{u} from Lagrangian data \mathbf{r} (e.g. Davis, 1991; Poulain, 2001; Bauer et al., 1998; Lumpkin, 2003). Typically, the “pseudo-Lagrangian” approach is used, where velocities are computed along trajectories from successive positions, $\Delta\mathbf{r}/\Delta t$, and they are interpreted as proxy for the Eulerian velocity \mathbf{u} .

Recently, a growing attention has been devoted to the use of Lagrangian data not only to provide statistical information on the velocity field, but also to provide Near Real Time (NRT) information. Such information are of primary importance in many relevant applications such as search and rescue, forecasting and containment of oil spill or other pollutants, and warfare problems such as mine tracking. Lagrangian data are especially well suited for all these problems, since they can be easily and operationally deployed and can be directly used in Rapid Environmental Assessment (REA) of the velocity and of the transport. NRT data are provided via satellite by surface drifters and by ARGO floats, which drift at a given reference depth and resurface to communicate (e.g. Davis et al., 2001). While in applications such as search and rescue and mine detection, the main interest lies on the ocean surface velocity; for pollutant spreading also the information on the sub-surface velocity in the water column can be of great interest. The aim of this paper is to investigate a methodology for the reconstruction of the Eulerian velocity field from Lagrangian data that can be of interest for NRT applications.

A number of different methods have been proposed in the literature aimed at using NRT Lagrangian data to provide a rapid assessment of Eulerian velocity and transport. Özgökmen et al. (2000) developed a technique aimed at reconstructing the trajectories of nonobserved particles using information from nearby observed floats (Castellari et al., 2001; Özgökmen et al., 2001), rather than Eulerian velocity. The problem of velocity reconstruction has been addressed by Toner et al. (2001b,a) considering pseudo-Lagrangian velocities computed along trajectories differencing the position data. These velocity estimates, assumed to be representative of the Eulerian velocity, are projected on a set of global basis functions (such as Geometrical Orthogonal Functions) and the time-dependent coefficients are evaluated by least square minimization. Other authors have considered a different approach, where information from Lagrangian data and models are combined via assimilation (e.g. Kamachi and O'Brien, 1995; Molcard et al., 2003, 2005; Özgökmen et al., 2003; Kuznetsov et al., 2003). In many of these works, the specific Lagrangian nature of the position data is taken into account, and the Eulerian velocity field \mathbf{u} is modified requiring that the distance between observed and model forecasted particle positions is minimized. The Lagrangian data assimilation approach has been thoroughly tested with positive results using idealized point vortex models (Kuznetsov et al., 2003) and GCM with simplified stratification (1.5 layer) (e.g. Kamachi and O'Brien, 1995; Molcard et al., 2003; Özgökmen et al., 2003). Further developments involve method testing with multi-level or multi-layer GCMs (Molcard et al., 2005).

All together, the results of the various methods for assessing \mathbf{u} from Lagrangian data appear very promising. On the other hand, there are still some open questions, especially regarding rapid assessment of the velocity field in practical applications such as search and rescue, pollutant spreading or mine detection. In these cases, a fast response is required in areas which are not a priori known in details, and the methodology has to be easy to apply and generalize. Methods based on global functions (Toner et al., 2001b) can be very powerful, but they require a significant a priori knowledge of the area. Basis functions might not be easy to identify especially in regions with complex bathymetry or open boundaries. Regarding Lagrangian assimilation for practical applications, realistic models with high resolution (in the horizontal and in the vertical) and complex physics are required. For these models, correcting the velocity field at the level of the floats is likely not to be sufficient, and techniques for mass field correction (still under development, e.g. Molcard et al., 2005) have to be used. In particular, for coastal applications a geostrophic based correction might not be appropriate. Finally, both the reconstruction method of Toner et al. (2001b) and the various assimilation methods (e.g. Kamachi and O'Brien, 1995; Molcard et al., 2003, 2005; Kuznetsov et al., 2003) deal only with the large scale or mesoscale component of the velocity field, neglecting the effects of high frequency motion such as tidal or inertial frequency oscillations. In regional and coastal applications, the high frequency component of the

velocity field is likely to play an important role and should be taken into account. All these points motivate the present work.

In this paper we present a method for the reconstruction of the Eulerian velocity field based on Lagrangian data, which is portable in the sense that can be easily applied to a given region provided that only a few general statistical information are known, and it is flexible in the sense that it can take into account the presence of both the mesoscale velocity field and the high frequency fluctuations. The method takes into account the nature of the Lagrangian data as position data. Lagrangian approaches are expected to have some advantages with respect to the pseudo-Lagrangian ones. First of all they are more general, given that the pseudo-Lagrangian approximation is valid only for small Δt to ensure that Eulerian and Lagrangian velocities coincide. Also, Lagrangian methods optimize \mathbf{u} using a cost function based on particle transport, which is often one of the most important quantities to be evaluated in practical application. The proposed method can be used directly for the \mathbf{u} reconstruction based on the Lagrangian data, or it can be used to combine the data with other a priori information, such as for instance results from GCMs. As such it can be considered as a step to improve the estimate of \mathbf{u} to be used in the framework of assimilation developments.

The methodological approach presented here consists in combining in an optimal way the Lagrangian data with a model constraint simply given by the numerical description of the float advection in the velocity field, to produce an estimate of the Eulerian velocity field. This inverse problem is considered in the variational approach (LeDimet and Talagrand, 1986; Talagrand and Courtier, 1987), in which the adjoint of the float advection equation is used to control the velocity field. The estimate is then chosen in the optimal sense as the Eulerian velocity field which minimizes the distance between the model predicted positions and their respective observation.

The method is tested in the framework of results from the circulation model OPA configured on the Mediterranean Sea, with a regional scale application in the Northwestern Mediterranean. An approach conceptually similar to the “twin experiment” approach is used. A main control run, regarded as the “truth” is considered, where numerical drifting floats are released and position data are taken at intervals Δt . The Eulerian field \mathbf{u} is then reconstructed on the basis of the position data and the method skills are evaluated through a quantitative comparison with the control velocity field. The sensitivity of the method results to varying the values of Δt and the number of floats is considered.

The main features of the Mediterranean circulation and associated transports appear satisfactory represented by the OPA model (Béranger et al., 2005). Also, the circulation model is characterized by the presence of strong inertial oscillations, in agreement with previous experimental and theoretical results (e.g. Millot and Crépon, 1981; Flexas et al., 2002; Van Haren and Millot, 2003). Regarding the mesoscale field, since the horizontal resolution is relatively low (1/8 degree), the circulation model results are expected to overestimate the space and time scales, especially at the surface where the variability is the highest (e.g. Garraffo et al., 2001). For the subsurface, in situ data from RAFOS floats (Testor and Gascard, 2003) and ARGO floats (see e.g. <http://www.bo.ingv.it/mfstep>) in the considered region of the Northwestern Mediterranean suggest the presence of scales of motion which appear in qualitative agreement with those of the circulation model. Given that the mesoscale field is expected to be more realistic in the subsurface, the experiments presented in this paper are obtained releasing floats at 350 m. The results are scaled with respect to the main flow parameters, so that they can be considered as indicative of the general methodology, rather than restricted to the specific flow under consideration.

The paper is organized as follows. In Section 2 the methodology is presented, while in Section 3 details on the OPA model and on the region of interest are provided. The experimental setup is given in Section 4, and the results of the experiments are presented in Section 5. A summary and a discussion are given in Section 6.

2. Methodology

2.1. Statement of the problem and definition of the main parameters

We consider an Eulerian velocity field \mathbf{u} characterized by mesoscale structures with typical time scale T_E and space scale R_E . From the Lagrangian point of view, the mesoscale flow is characterized by a time scale T_L , typical of the autocorrelation of Lagrangian particles released in the flow. Inertial frequency oscillations

with time scale T_I are superimposed to the mesoscale field, and they are assumed to be characterized by $T_I \ll T_E$. This assumption is realistic for regional and open ocean flows in mid-latitudes (such as the application considered here), since T_E is typically of the order of one week or more, while $T_I \leq 1d$.

We consider P Lagrangian floats released in the flow at the same time $t = 0$ and at different locations. They provide information on their position at the same discrete times $m\Delta t$, $m = 0, M$, over a period $T_M = M\Delta t$. These observations are denoted by $\mathbf{r}^{\text{obs}}(m\Delta t)$, where the $2P$ -dimensional vector \mathbf{r}^{obs} is composed of the plane coordinates of the P float positions.

Our objective is to provide an estimate \mathbf{u}^{est} of the Eulerian velocity field, given the Lagrangian data \mathbf{r}^{obs} during the period T_M . $T_M \ll T_E$ is assumed, so that \mathbf{u}^{est} can be considered representative of the mesoscale field. This estimation problem is performed in the variational approach (LeDimet and Talagrand, 1986; Talagrand and Courtier, 1987), by using the complete data set from the P floats over the time T_M and a model constraint which numerically describes the float advection in the velocity field.

The reconstruction of \mathbf{u}^{est} is carried out in the following way. The M sequences of length Δt are consecutively considered. For each sequence $((m-1)\Delta t, m\Delta t)$, a first guessed velocity field \mathbf{u}^{bck} and its associated prior trajectories \mathbf{r}^{bck} are considered and, by using the optimization procedure detailed in Sections 2.2 and 2.3, corrected to fit the model predicted positions $\mathbf{r}(m\Delta t)$ onto their respective observation $\mathbf{r}^{\text{obs}}(m\Delta t)$. The corresponding velocity increment $\delta\mathbf{u}$, which corrects \mathbf{u}^{bck} on each sequence, is assumed to be indicative of the flow anomalies characterized by structures with typical space scale R . So $\delta\mathbf{u}$ is approximated as a time-independent Eulerian velocity field, since $T_M \ll T_E$. Notice that in the first sequence ($m = 1$), the first guessed velocity \mathbf{u}^{bck} is null, while in the other cases it is given by the reconstructed field obtained in the previous sequences. In consequence, the obtained estimate \mathbf{u}^{est} is time-independent as indicative of the mesoscale mean flow during the period T_M . In cases when the Lagrangian data are combined with GCM results, \mathbf{u}^{bck} is provided for each sequence by the GCM velocity field. Finally, given a reconstructed Eulerian velocity field in the time-independent approximation and a data set sampling the period T_I , velocity signals at the inertial frequency can be reconstructed in a second step using the same optimization procedure, as detailed in Section 2.4.

2.2. Estimation of velocity increments in the time-independent approximation

The float advection in the Eulerian velocity field can be described as the solution of the nonlinear ordinary differential equation $d\mathbf{r}/dt = \mathbf{u}(\mathbf{r})$ with initial condition $\mathbf{r}(t_0)$. For a sequence of length Δt , this can be formulated as $\mathbf{r}(m\Delta t) = \mathbf{H}_{\text{NL}}(\mathbf{u})$, where the operator \mathbf{H}_{NL} relates the initial float position $\mathbf{r}(t_0)$ taken at the beginning of the sequence $t_0 = (m-1)\Delta t$, to the position $\mathbf{r}(m\Delta t)$ at the end of the sequence. So \mathbf{H}_{NL} describes the float advection during the sequence in within the flow \mathbf{u} .

In numerical modeling, the float advection equation is discretized over an horizontal mesh grid of spatial resolution Δs , and with a time stepping δt assumed for simplicity to be a fraction of Δt , as

$$\mathbf{r}(t_n) = \mathbf{r}(t_{n-1}) + \delta t \cdot \mathbf{L}_{\text{NL}}(\mathbf{r}(t_{n-1})) \cdot \mathbf{u}(t_{n-1}) \quad n = 1, \Delta t/\delta t \quad (1)$$

where \mathbf{L}_{NL} is the bilinear Lagrange interpolator used for the velocity projection. So the predicted float positions $\mathbf{r}(m\Delta t)$ are computed by the recurrence given in Eq. (1) initialized with the observed positions $\mathbf{r}(t_0) = \mathbf{r}^{\text{obs}}((m-1)\Delta t)$.

Time-independent perturbations $\delta\mathbf{u}$ on the first guessed velocity field \mathbf{u}^{bck} are now considered. The corresponding variation on the prior float positions \mathbf{r}^{bck} is obtained at each time step (t_{n-1}, t_n) by the linearized perturbation equation relative to Eq. (1)

$$\delta\mathbf{r}(t_n) = \delta\mathbf{r}(t_{n-1}) + \delta t \cdot \mathbf{L}_{\text{NL}}(\mathbf{r}^{\text{bck}}(t_{n-1})) \cdot \delta\mathbf{u} + \delta t \cdot \mathbf{L} \cdot \delta\mathbf{r}(t_{n-1}) \cdot \mathbf{u}^{\text{bck}}(t_{n-1}) \quad n = 1, \Delta t/\delta t \quad (2)$$

where the linear operator \mathbf{L} is obtained by deriving \mathbf{L}_{NL} with respect to \mathbf{r} around the prior float positions \mathbf{r}^{bck} . So the variation $\delta\mathbf{r}(m\Delta t)$ of the float positions around their prior estimate at time $m\Delta t$ are computed by the recurrence given in Eq. (2), assuming unperturbed floats positions at time $(m-1)\Delta t$ (i.e. $\delta\mathbf{r}(t_0) = \mathbf{0}$). It can be formulated as $\delta\mathbf{r}(m\Delta t) = \mathbf{H} \cdot \delta\mathbf{u}$, where the operator \mathbf{H} relates the linearized equation of float advection around their prior trajectories. When considering a one-step integration, Eq. (2) is then reduced to

$$\delta\mathbf{r}(m\Delta t) = \Delta t \cdot \mathbf{L}_{\text{NL}}(\mathbf{r}^{\text{obs}}((m-1)\Delta t)) \cdot \delta\mathbf{u} \quad (3)$$

The variation of float positions at time $m\Delta t$ is identified with the Lagrangian velocity increment at the location $\mathbf{r}^{\text{obs}}((m-1)\Delta t)$. In this case, the operator \mathbf{H} simply attributes weights at each neighboring grid-point value of the Eulerian velocity increment $\delta\mathbf{u}$. This repartition can either be treated by a Gaussian distribution (as done by Molcard et al., 2003), or directly by the spatial interpolation coefficients defined from each cell containing $\mathbf{r}^{\text{obs}}((m-1)\Delta t)$. When using more than one step for the trajectory computation given in Eq. (2), the velocity increment $\delta\mathbf{u}$ is projected along the first guessed trajectories, i.e. at each $\mathbf{r}^{\text{bck}}(t_n)$, $n = 1, \Delta t/\delta t$. As a matter of fact, the trajectory sampling δt defines the resolution in space of this projection; it can be related to the accuracy with which \mathbf{H} provides the predicted variations $\delta\mathbf{r}(m\Delta t)$ from velocity increments $\delta\mathbf{u}$ taken at the resolution Δs . For relatively coarse spatial resolutions Δs compared to float displacements during Δt , the formulation given by Eq. (3) is sufficient. Instead, when going to regional scales with higher resolution velocity fields, the formulation given by Eq. (2) is necessary to cover the whole structure described by the float displacements during Δt . As a consequence, the linearized equation of float advection is augmented with a term correcting the spatial interpolation of the first guessed velocity according to the position variations, which increases the accuracy on the prediction of $\delta\mathbf{r}(m\Delta t)$.

The observed float positions at time $m\Delta t$ are now considered to provide an estimation of the velocity increment $\delta\mathbf{u}$ according to the first guess velocity field. So the distance to minimize is defined by the corresponding variation of the float positions $\delta\mathbf{r}(m\Delta t) = \mathbf{H} \cdot \delta\mathbf{u}$, as expressed by the cost function

$$J(\delta\mathbf{u}) = \frac{1}{2} (\mathbf{H}_{\text{NL}}(\mathbf{u}^{\text{bck}}) + \mathbf{H} \cdot \delta\mathbf{u} - \mathbf{r}^{\text{obs}}(m\Delta t))^{\text{T}} \cdot (\mathbf{H}_{\text{NL}}(\mathbf{u}^{\text{bck}}) + \mathbf{H} \cdot \delta\mathbf{u} - \mathbf{r}^{\text{obs}}(m\Delta t)) \quad (4)$$

where $^{\text{T}}$ denotes the vector transpose. Notice that the components of this model-data misfit are supposed independent, and associated to Gaussian and homogeneous errors, to properly define this measure in the least square sense. Notice also that background error information are specified as exact optimization constraints (see Section 2.3), which avoids the introduction of a regularization term in the cost function. Regarding to Eq. (4), the cost function J depends on the velocity increment. So its minimization requires the knowledge of its gradient ∇J according to $\delta\mathbf{u}$. It is formally expressed from Eq. (4) as

$$\nabla J = \mathbf{H}^{\text{T}} \cdot (\mathbf{H}_{\text{NL}}(\mathbf{u}^{\text{bck}}) + \mathbf{H} \cdot \delta\mathbf{u} - \mathbf{r}^{\text{obs}}(m\Delta t)) \quad (5)$$

It is the exact adjoint \mathbf{H}^{T} of the operator \mathbf{H} which is implicated in the expression of the gradient. Notice that the practical implementation of these two linear operators is done using differentiation hand coding techniques (Talagrand, 1991). Finally ∇J is numerically used to perform the steepest descent step toward the optimal value for $\delta\mathbf{u}$ in the iterative procedure described hereafter.

2.3. Method implementation and background error correlation modeling

For each sequence $((m-1)\Delta t, m\Delta t)$, the first guessed Eulerian velocity field \mathbf{u}^{bck} is iteratively corrected by velocity increments $\delta\mathbf{u}$ following the incremental formulation of variational problems (Courtier et al., 1994). In order to reconstruct the P float trajectories starting from the observed positions $\mathbf{r}^{\text{obs}}((m-1)\Delta t)$, updates of prior trajectories \mathbf{r}^{bck} are achieved by the nonlinear operator \mathbf{H}_{NL} imposed as a time-dependent model constraint. In between each prior trajectory update, an optimal velocity increment $\delta\mathbf{u}$ is determined in the time-independent approximation as the model constraint is degraded to the linear operator \mathbf{H} previously described.

A priori information on the length scale R of $\delta\mathbf{u}$ can be inserted through the characteristic error correlations for \mathbf{u}^{bck} . These are classically specified in the definition of the background error covariance matrix \mathbf{B} . Following e.g. Derber and Bouttier (1999), the optimization problem is rewritten in terms of a new variable $\mathbf{v} = \mathbf{B}^{-\text{T}/2} \cdot \delta\mathbf{u}$, deduced from the factorization of the background error covariance matrix as $\mathbf{B} = \mathbf{B}^{\text{T}/2} \cdot \mathbf{B}^{1/2}$. This allows to circumvent the explicit specification of the inverse matrix \mathbf{B}^{-1} , while providing an efficient preconditioner for minimization (Lorenc, 1988). Considering this variable change, the gradient of the cost function J with respect to \mathbf{v} can be deduced from Eq. (5), $\nabla_{\mathbf{v}} J = \mathbf{B}^{1/2} \nabla J$. The minimization of J is performed by computing the steepest descent step from $\mathbf{v} = \mathbf{0}$ in the direction of $\nabla_{\mathbf{v}} J$. In addition to the forward integration of \mathbf{H} to compute J (Eq. (4)) and the backward integration of \mathbf{H}^{T} to compute ∇J (Eq. (5)), the transformation from \mathbf{v} to $\delta\mathbf{u}$ involving $\mathbf{B}^{\text{T}/2}$ and the adjoint transformation from ∇J to $\nabla_{\mathbf{v}} J$ involving $\mathbf{B}^{1/2}$ are actually required for the

minimization. The steepest descent step is achieved at a reasonable computational cost using the limited memory quasi-Newton algorithm MIQN3 (Gilbert and Lemaréchal, 1989).

For our specific purpose, the background error covariance matrix can be neither fully estimated nor stored explicitly. It is therefore modeled as a composition of operators from which the \mathbf{B} factorization can be deduced. The covariances can be split into two contributions which define two distinct operators (as described in Derber and Bouttier, 1999). The bivariate component represents the cross-covariances between grid points associated to the zonal and the meridional velocities; it is assumed negligible in our case where these two variables are supposed unbalanced. The univariate component represents the auto-covariances between grid points corresponding to the zonal or the meridian velocity. It can be modeled by the solution of a diffusion equation which is interpreted as a covariance operator (Derber and Rosati, 1989). This Laplacian grid-point smoother has been generalized to take into account complex boundaries imposed by coastlines and also for anisotropic correlations (Weaver and Courtier, 2001); both are necessary when considering background errors in regional or coastal basins. Our practical covariance modeling considers an homogeneous diffusion coefficient defined with respect to the grid resolution Δs . If the Laplacian smoother is applied K times on \mathbf{v} and ∇J , the background error covariance matrix is expressed in the horizontal plane as

$$\mathbf{B} = \mathbf{W}^{-1} \cdot (\mathbf{I} + 1/4\Delta s^2 \cdot \mathbf{D})^{2K} = (\mathbf{I} + 1/4\Delta s^2 \cdot \mathbf{D})^K \cdot \mathbf{W}^{-T/2} \cdot \mathbf{W}^{-1/2} \cdot ((\mathbf{I} + 1/4\Delta s^2 \cdot \mathbf{D})^T)^K \quad (6)$$

The identity matrix \mathbf{I} and the matrix \mathbf{D} , which denotes the discrete Laplacian operator, have both the self-adjointness property to get a suitable expression of $\mathbf{B}^{T/2}$ and $\mathbf{B}^{1/2}$. \mathbf{W} is a diagonal normalization matrix determined a priori from the $2K$ -iteration of the Laplacian smoother to a single unit value placed at each grid point. Notice that the obtained numerical solutions can be approximated by a Gaussian distribution with characteristic length scale expressed by $R = \Delta s \cdot K^{1/2}$. So an approximate value for the correlation scale of background errors, which is by definition the length scale R of the velocity increment $\delta \mathbf{u}$, can be specified and parameterized a priori with the value of K . Notice that the choice of the parameter R depends on the specific application considered. For the time-independent approximation, where the reconstructed field is the Eulerian mesoscale velocity, R is chosen as $R = R_E$. This is not the case for the time-dependent reconstruction, as discussed in Section 2.4.

2.4. Refinement of the method taking into account the inertial signal

The method developed in the previous section provides a reconstruction of the time-independent velocity field from the observation of float trajectories at the resolution Δt . So the obtained velocity field \mathbf{u}^{est} is representative of the mesoscale mean circulation during T_M . The assumption of time-independence can be relaxed considering \mathbf{u}^{est} as a reference state and computing for each sequence Δt the discrepancy $\hat{\mathbf{u}}_m$, $m = 1, M$, with respect to \mathbf{u}^{est} . This set of perturbed velocity fields $\hat{\mathbf{u}}_m$ can be used to estimate time-varying patterns of the reconstructed velocity field, with a particular attention for inertial currents. In the idealized case, the flow can be assumed as the superposition of the mesoscale circulation \mathbf{u}^{est} and of an inertial signal with zonal and meridional components given by

$$\hat{\mathbf{u}}(t) = [\mathbf{a}_x \cdot \cos(T_1^{-1}t) - \mathbf{b}_x \cdot \sin(T_1^{-1}t), \mathbf{a}_y \cdot \sin(T_1^{-1}t) + \mathbf{b}_y \cdot \cos(T_1^{-1}t)] \quad (7)$$

where $\mathbf{a}_x, \mathbf{b}_x, \mathbf{a}_y, \mathbf{b}_y$ are time-independent parameters. When the observation time scale Δt is able to sample the inertial period T_1 , these parameters can be estimated from $\hat{\mathbf{u}}_m$ by minimizing

$$Q = \sum_{m=1, M} (\hat{\mathbf{u}}(m\Delta t) - \hat{\mathbf{u}}_m)^T \cdot (\hat{\mathbf{u}}(m\Delta t) - \hat{\mathbf{u}}_m) \quad (8)$$

Analytical solutions are given by e.g. Priestley (1981), and they have been numerically implemented for each grid point.

In the practical implementation, for each sequence $((m-1)\Delta t, m\Delta t)$, a perturbed field $\hat{\mathbf{u}}_m$ is estimated using the same technique as for the time-independent field (Sections 2.2 and 2.3). Each perturbed field have a regular spatial pattern due to the background error correlation modeling. The spatial scale R chosen for the reconstruction is usually larger than the mesoscale R_E . This can be justified at least in open ocean or regional flows

since the inertial space scales are expected to be associated with wind patterns, and it allows for a smooth field given a reduced number of data. The time-dependent perturbed fields $\hat{\mathbf{u}}_m$ are first averaged over the time $T_M = M\Delta t$. This average is then removed from the time series $\hat{\mathbf{u}}_m$ in order to have a zero-mean fluctuation signal, and it is used to correct \mathbf{u}^{est} . The cost function Q (Eq. (8)) is successively minimized over a moving inertial period to get $(\mathbf{a}_{xm}, \mathbf{b}_{xm}, \mathbf{a}_{ym}, \mathbf{b}_{ym})$. Then the signal reconstruction (lead by Eq. (7)) is processed considering the moving average of $(\mathbf{a}_{xm}, \mathbf{b}_{xm}, \mathbf{a}_{ym}, \mathbf{b}_{ym})$, $m = 1, M$, on the inertial period T_I .

In summary, the time-dependent refinement of the method provides on the one side a correction of the mesoscale mean field \mathbf{u}^{est} , and on the other side an estimate of the time-dependent velocity in terms of time varying parameters. Notice that, while the method is expected to capture the main aspects of the time dependence, the details of the inertial oscillations might be only partially estimated. This is due to several reasons. First of all, the accuracy of the estimation depends on the mean velocity reconstruction. Second, the method is based on a clear decomposition into a stationary part and an oscillating part; so the presence of signals at other frequencies is ignored. Finally, inertial signals are mostly related to the dynamics of inertial oscillations which specifies a quadrature of phase between the two velocity components; such a bivariate constraint is not taken into account in the background error construction and estimated velocity components are treated independently.

3. Numerical circulation model and region of application

The circulation model is an extended version of the primitive equation numerical model OPA with a rigid lid (Madec et al., 1998), configured in the Mediterranean Sea. Model solutions have been previously analyzed by Béranger et al. (2005), focusing on the characteristics of the general circulation and of the transport in the straits of Gibraltar, Sicily and Sardinia. The results show a good agreement with recent estimates from in situ observations. The circulation model version used here has horizontal resolution of $1/8$ degree and 43 geopotential levels in the vertical. Its time step is equal to 20 min. The model has been initialized from rest and from climatological hydrological conditions from MEDATLAS; it is forced by daily winds and heat fluxes from the ECMWF atmospheric fields during the period August 1998–July 2002. A 12 year spin-up has been performed, repeating three times the 4 year forcing cycle, and the convergence of the solution has been assessed in terms of transport in the straits. Results presented here are initialized after the spin-up.

We focus on a region in the Northwestern Mediterranean, situated northeast of the Balearic Islands (Fig. 1a). An example of the circulation at 350 m averaged over a 5 days period in March 1999 of the simulation is shown in Fig. 1b. Even though this is just one specific configuration, it appears to be quite representative of the flow generated by the model in this region during the winter period. A number of mesoscale structures can be observed in Fig. 1b, including an energetic cyclonic eddy in the southern part (region A), an eastward flow with associated recirculation cells in the central area (region B), and a westward flow in the northern part (region C). The space scale R_E of the structures can be roughly evaluated by visual inspection and it appears to be of the order 20–30 km. The velocity reaches values of approximately 20 cm s^{-1} in the energetic cyclonic gyre, while it is of the order of few cm s^{-1} in the northern area. The Eulerian velocity auto-covariance (computed from velocity time series of 30 days from all the 250 grid points of the region) is shown in Fig. 2a. As it can be seen the flow is characterized by two distinct time scales. The mesoscale time scale T_E , roughly evaluated as the e-folding scale of the auto-covariance, is of the order 20–30 days. Superimposed to this, there is a clear high frequency inertial signal with $T_I \approx 19 \text{ h}$.

The Lagrangian structure of the flow is depicted in Fig. 1c, where 250 floats, launched at each grid point of the control area (union of the regions A, B, C) and advected for 5 days, are shown. The numerical advection is carried out in the time-dependent field using a fourth order Runge–Kutta scheme in time and a bilinear interpolation in space. The same recirculation structures as in the Eulerian field are indicated by the trajectories. The Lagrangian auto-covariance (computed from 30 day velocity time series along the 250 trajectories) is shown in Fig. 2b, and it shows the same two-scale superposition as for the Eulerian velocity auto-covariance. The Lagrangian mesoscale time T_L , though, is significantly shorter than T_E , as $T_L \approx 3 \text{ d}$.

The presence of a complex mesoscale flow is confirmed by in situ float observations in the area. Trajectories of MEDARGO floats at 350 m launched in the framework of the MFSTEP project (<http://www.bo.ingv.it/>)

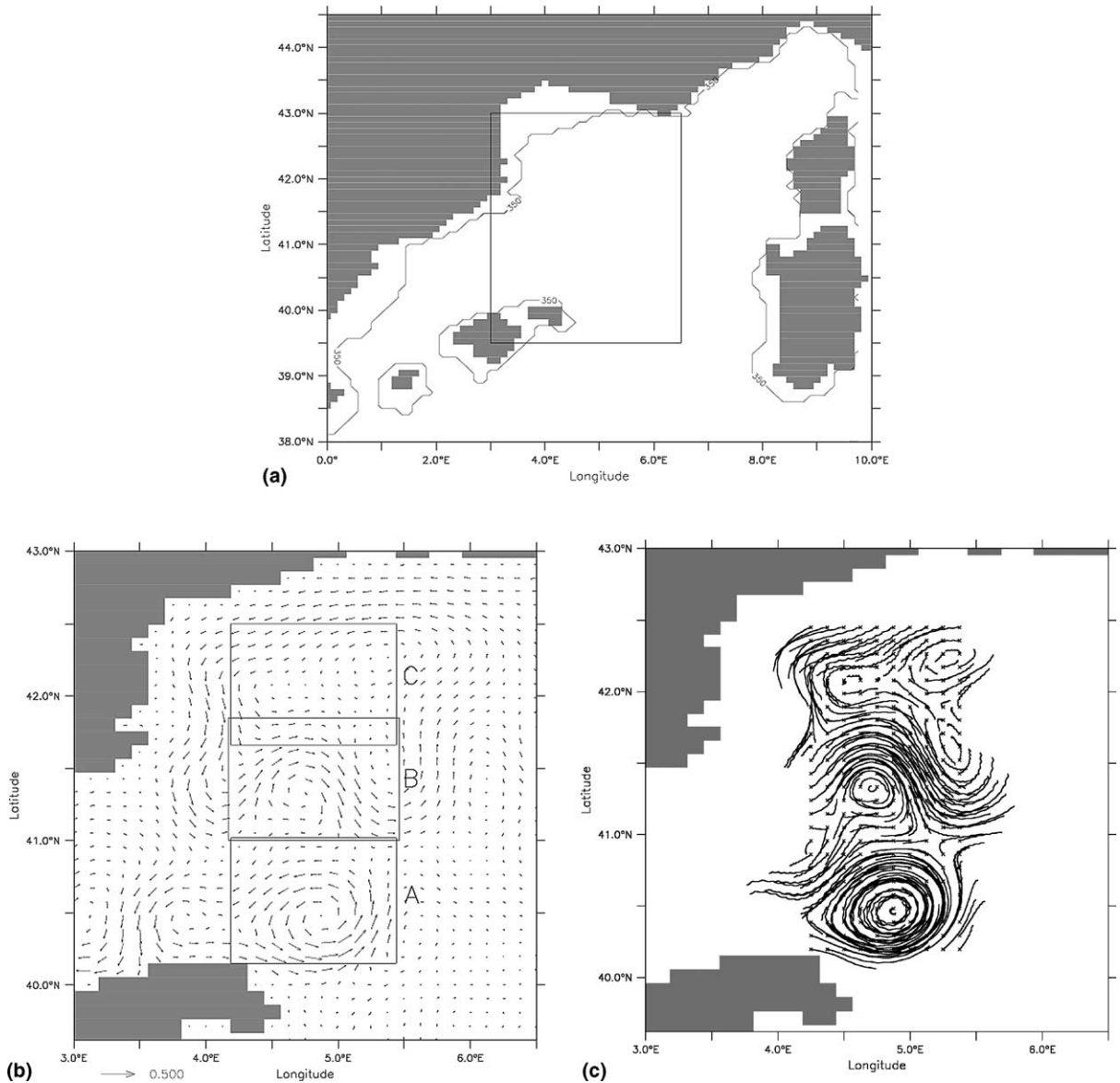


Fig. 1. (a) Location of the computational domain in the Northwestern Mediterranean Sea, including the 350 m isobath. (b) Control velocity field (in m s^{-1}) averaged over $T_M = 5$ d at 350 m. The three sub-regions A, B, C where the diagnostics are performed, are indicated. (c) Lagrangian structure of the control flow from 250 floats advected for 5 days.

mfstep; http://doga.ogs.trieste.it/WP4/real_time_west.html) and RAFOS floats (Testor and Gascard, 2003) at various depths from 350 m to 1450 m show significant mesoscale structures with space scales comparable with the ones in Fig. 1. MEDARGO trajectories also show a significant variability from the point of view of energetics. Some of them are characterized by average velocity of a few cm s^{-1} , while others reach values of 20 cm s^{-1} , corresponding to displacements of roughly 100 km in 5 days, similarly to what shown in Fig. 1c. Notice that, in addition to the mesoscale structures, the RAFOS floats also indicate the presence of energetic submesoscale vortices (Testor and Gascard, 2003), with significant scales less than 10 km, probably connected to the processes of deep water formation and propagation. We do not expect to find these in the model solutions, given the relatively coarse resolution.

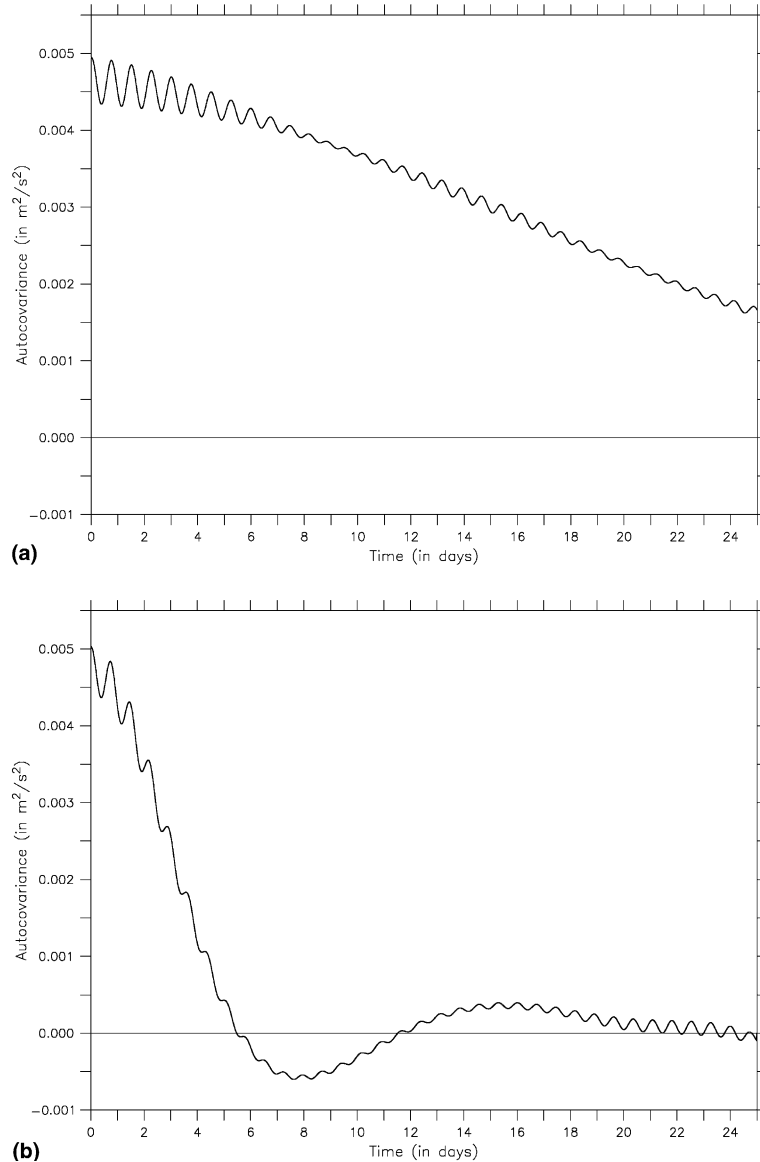


Fig. 2. Representation of the auto-covariances (a) of the zonal Eulerian velocity and (b) of the zonal Lagrangian velocity.

4. Experimental setup

The flow configuration in Fig. 1 is used to test the methodology introduced in Section 2. The presence of various types of structures, including the highly energetic cyclone as well as the eastern and western flows and recirculation cells, provides a very good test ground for the method. We recall that the main flow parameters as shown in Section 3 are $R_E \approx 20\text{--}30$ km, $T_E \approx 20$ d, $T_L \approx 3$ d, $T_I \approx 19$ h. We are therefore in the regime $T_E > T_L$, $T_I \ll T_E$, $T_I < T_L$, which is quite typical of upper ocean flows (Lumpkin et al., 2002). The averaging time is chosen to be $T_M = 5$ d, so $T_M \ll T_E$.

Our goal is to reconstruct the mesoscale Eulerian velocity \mathbf{u}^{est} averaged over T_M , given $M + 1$ consecutive pseudo-observed positions \mathbf{r}^{obs} taken at intervals Δt during $T_M = M\Delta t$ from P floats. The float positions are extracted from the control trajectories (Fig. 1c) with sampling interval Δt . Float positions are taken without

noise. We explore the sensitivity of the method to the number of floats P and to the sampling period Δt . We also consider different implementations of the method, for instance comparing results obtained with and without the time-dependent correction, and considering the combination of data with GCM results. A number of preliminary experiments (not shown) have been performed, varying the scale parameter R . For the time-independent reconstruction of the mesoscale field \mathbf{u}^{est} , the mesoscale parameter R_E has been used, varying its value in the range between 10 km and 30 km. No significant differences have been found, so that all the experiments presented in the following are obtained with $R = R_E = 20$ km. For the time-dependent perturbation field related to the inertial oscillation, a greater value of R has been used, $R = 100$ km. Also, preliminary runs have been done to test the convergence of the model for high P values, considering $P = 250$, i.e. floats launched at each grid point. The results confirm that the reconstruction converges toward the control mean field of Fig. 1b.

A total of 34 experiments are presented and discussed in Section 5. Seven main values of Δt are considered, $\Delta t = 60$ h (i.e. 2.5 d), 24 h (i.e. 1 d), 12 h, 10 h, 6 h, 4 h, 1 h. They are within the range $\Delta t < T_L$, in agreement with previous works suggesting that significant velocity estimates can be obtained for Δt up to T_L (e.g. Molcard et al., 2003). For $\Delta t > T_L$ information on the velocity are highly degraded, as the auto-covariance tends to zero. Only one case has been considered here for $\Delta t > T_L$, $\Delta t = T_M = 5$ d, aimed to test the limit of applicability of the method. We notice that in realistic applications for ARGO floats at 350 m, values of $\Delta t < 1$ d are not quite realistic since the floats needs few hours to ascend (and descend) to the surface. The results then have to be considered as methodological, aimed at exploring the impact of resolving inertial oscillations, rather than directly related to ARGO application at 350 m. Regarding the P values, we consider four main values, $P = 60, 24, 12, 6$. The highest value, $P = 60$, is considered as a reference upper limit, not necessarily realistic for a region of the size we consider. The lower values of P , on the other hand, are in a realistic range for practical applications.

The results are first qualitatively assessed by visual comparison between the control field in Fig. 1 and the reconstructed field. From the quantitative point of view, two metrics are introduced and used to assess the method performance. The first metric E provides an estimate of the reconstruction error for the Eulerian velocity field. It is given by the difference between the control \mathbf{u}^{tru} (Fig. 1b) and the reconstruction \mathbf{u}^{est} both averaged during T_M , normalized by the norm of the mean control:

$$E(\Omega) = \frac{\iint_{\Omega} \|\mathbf{u}^{\text{est}} - \mathbf{u}^{\text{tru}}\|^2 d\Omega}{\iint_{\Omega} \|\mathbf{u}^{\text{tru}}\|^2 d\Omega} \quad (9)$$

where $\|\mathbf{u}\|^2 = \mathbf{u}^T \cdot \mathbf{u}$ and Ω is the region of interest. In the experiments of Section 5, different definitions of Ω have been considered, depending on the applications. In order to check more closely the impact of the various structures in the field, Ω is either considered as three separate regions (indicated as A, B, C in Fig. 1b), or as the union of them. Region A is characterized by the energetic cyclonic gyre, region B by an eastward flow with recirculation cells, while C is primarily dominated by a westward flow. Normalization of metric E is necessary especially for comparison between the three regions, since they are characterized by very different regimes (in terms of structures and energy). The second metric L is based on the reconstruction of the Lagrangian trajectories and it is given by the rms distance between control trajectories \mathbf{r}^{tru} (Fig. 1c) and trajectories computed in the reconstructed field \mathbf{u}^{est} starting from the same initial conditions,

$$L(\tau) = 1/\tau \int_{t \in [0, \tau]} 1/P \cdot \|\mathbf{r}(t) - \mathbf{r}^{\text{tru}}(t)\| dt \quad (10)$$

where $\tau \leq T_M$. Notice that L is computed from numerical trajectories inside the reconstructed field for the period τ . It cannot therefore be directly linked to the cost function J given in Eq. (4), which is given from reconstructed trajectories at successive sequences of duration Δt .

5. Results

The experiment results are now presented and discussed. In Section 5.1, the sensitivity to the number of floats P is studied, considering the time-independent reconstruction method introduced in Section 2. In Section 5.2 the effects of varying Δt are studied, considering also the time-dependent correction. Finally, in Section 5.3, we consider an application where Lagrangian observations are combined with GCM results.

5.1. Sensitivity to observation coverage

The dependence of the method on spatial coverage is discussed considering the case $\Delta t = 24$ h and varying P . The results can be considered representative also of different values of Δt , as shown in Section 5.2. Examples of trajectory ensembles with different values of P are shown in Fig. 3 for homogeneous releases in the control area. As it can be seen for $P = 60$, the flow appears well sampled in the three sub-regions A, B, C. The length of the trajectories indicates that the flow is significantly more energetic in A, in presence of the strong cyclone, than in B and C. The presence of inertial oscillations is indicated by the wiggles in the trajectories, especially evident in the least energetic areas. The coverage of the three sub-regions is maintained for $P = 24$ and, at least at some extent also for $P = 12$, even though the recirculation cells do not appear well sampled. For $P = 6$, minimal information are retained for the cyclone and for the large scale flow.

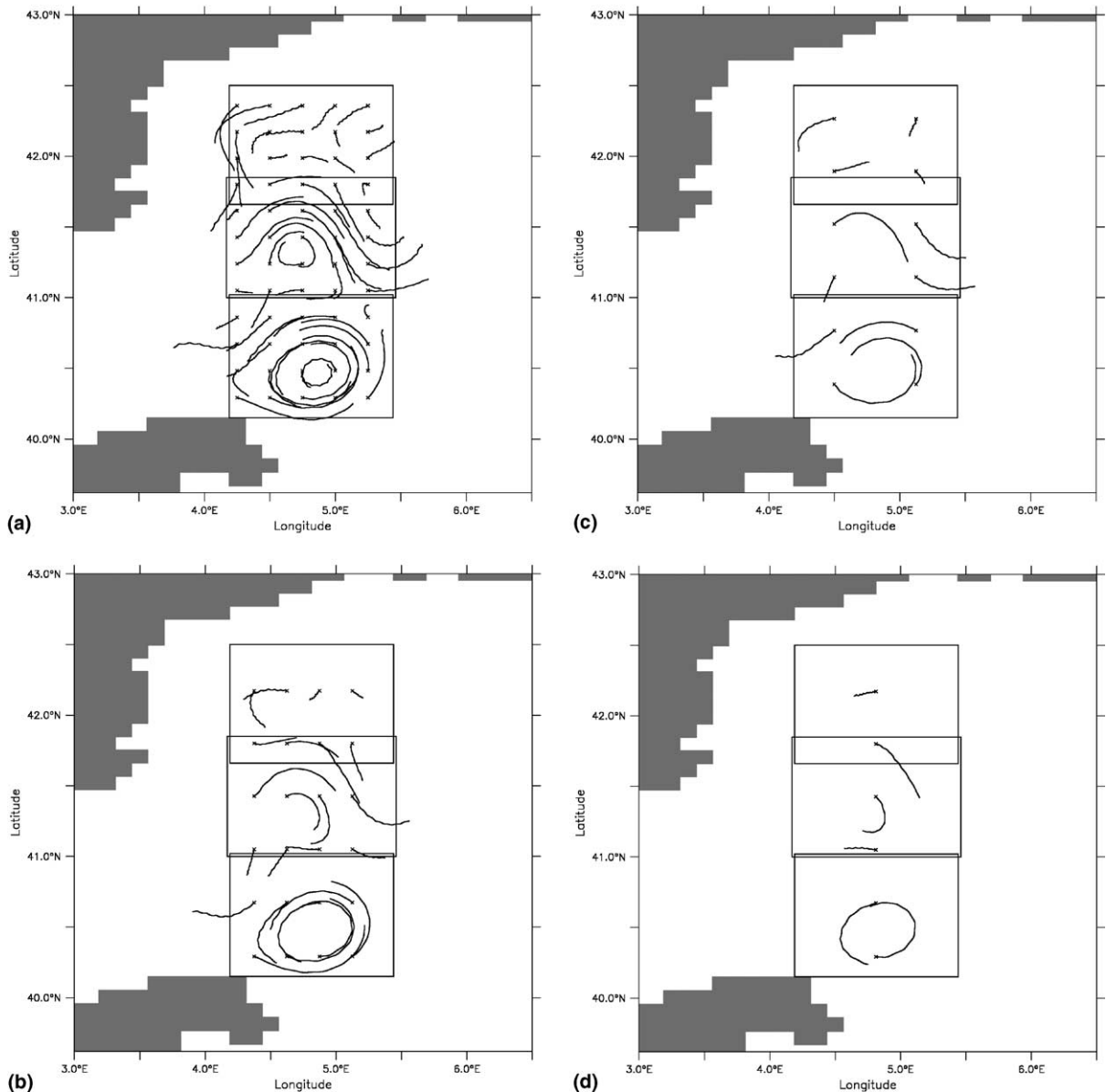


Fig. 3. Control trajectories over T_M for (a) $P = 60$, (b) $P = 24$, (c) $P = 12$ and (d) $P = 6$. Their associated initial coordinates (indicated by crosses) homogeneously cover the regions A, B, C.

The reconstruction results, obtained using the time-independent method are shown in Fig. 4, in terms of estimated velocity field \mathbf{u}^{est} and of trajectories integrated within the field. A qualitative comparison can be done considering the control velocity field \mathbf{u}^{tru} in Fig. 1b and the control trajectories in Fig. 3. For $P = 60$ (Fig. 4a), the patterns of the circulation appear well reproduced. The position and the extension of the cyclonic gyre in sub-region A are well captured, as well as the eastern and western flows and the recirculation cells in B and C. Concerning the reconstructed trajectories in this field, their path remains close to the observed one. Notice that the wiggles characterizing the inertial oscillations have disappeared. For $P = 24$ (Fig. 4b), the reconstruction appears still satisfactory, while it starts to deteriorate for $P = 12$ and $P = 6$ (Fig. 4b and c). The cyclonic gyre and the main eastern and western flows are still present, but the details of the recirculation cells are lost, since they are not sufficiently sampled. On the other hand, in the areas where observations are available, both the velocity field and the trajectories appear satisfactorily reproduced.

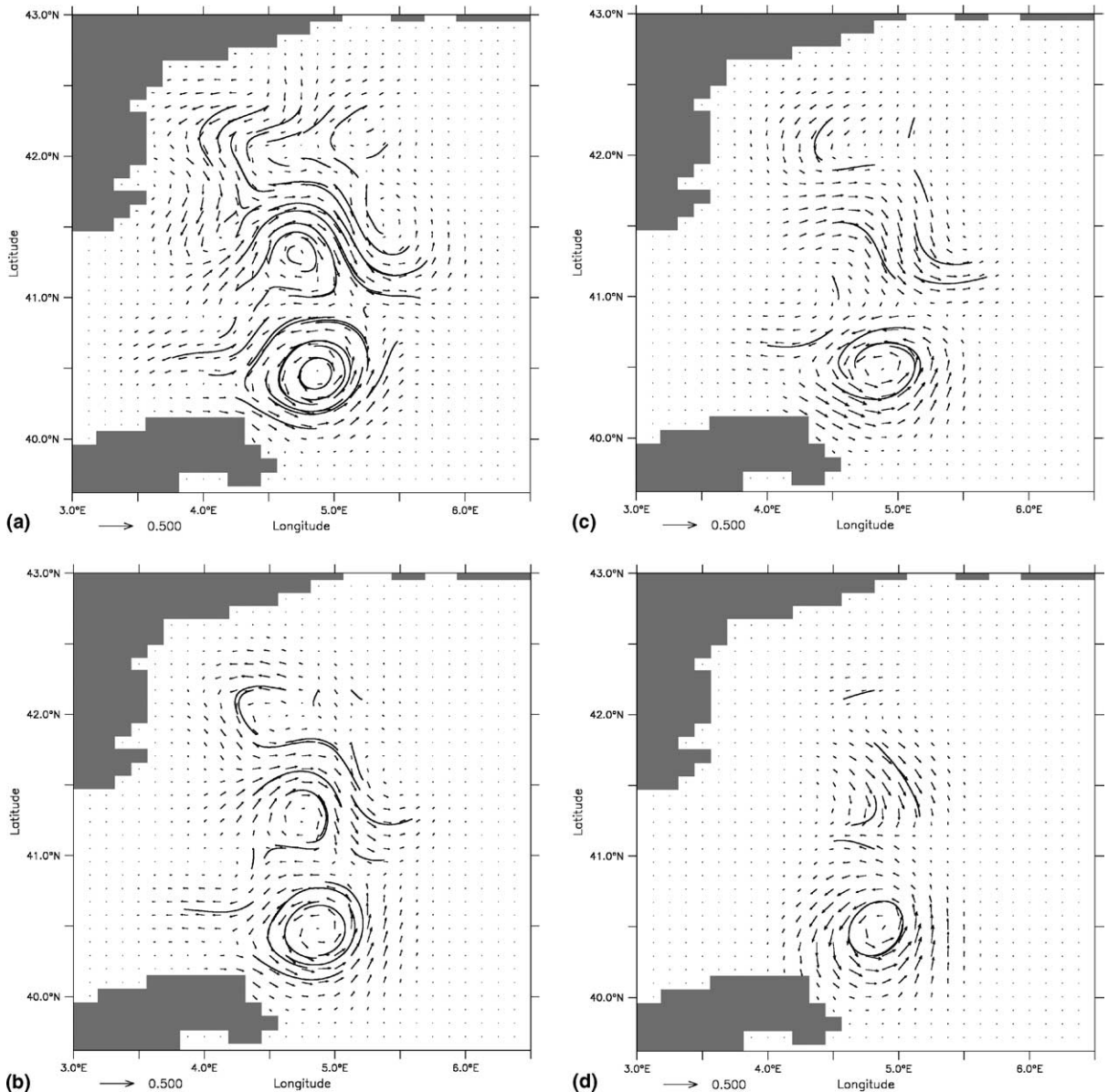


Fig. 4. Reconstruction of the time-independent Eulerian velocity field (in m s^{-1}) and corresponding trajectories, from the float positions which sample every day the control trajectories (Fig. 3) for (a) $P = 60$, (b) $P = 24$, (c) $P = 12$ and (d) $P = 6$.

A quantitative analysis of the results is performed using the two metrics given in Eqs. (9) and (10). First L is plotted in Fig. 5. It is computed during T_M for the four sets of trajectories in the sub-region A (Fig. 5a) and in the sub-regions B + C (Fig. 5b) respectively. As it can be seen, in all cases L increases with time and the evolution remains monotonic, indicating that the reconstructed trajectories progressively diverge from the observed ones. This divergence is accentuated when the number of observed floats P decreases, and for the more energetic region A with respect to B + C. These results can be conceptually expected, given that particle trajectories are known to be highly sensitive to the details of the underlying Eulerian field (e.g. Aref, 1984; Samelson, 1996). Even small changes in the estimation of the ocean currents can drastically change particle trajectories. In particular, Griffa et al. (2004) have studied the effects of smoothing the velocity field on particle trajectories, and they have shown that L is expected to grow initially exponentially, reaching then an inertial asymptotic regime in $t^{1/2}$. Our reconstruction problem is under many respects similar to smoothing, especially for small values of P . For $P = 6$, the initial exponential regime is evident in the most energetic region A, while in B + C the asymptotic regime seems to be reached more quickly. It is also important to compare the L values with dispersion values of the control trajectories, i.e. with the rms distance covered by the trajectories. In A, the dispersion after 5 days is ≈ 25 km, so that L reaches

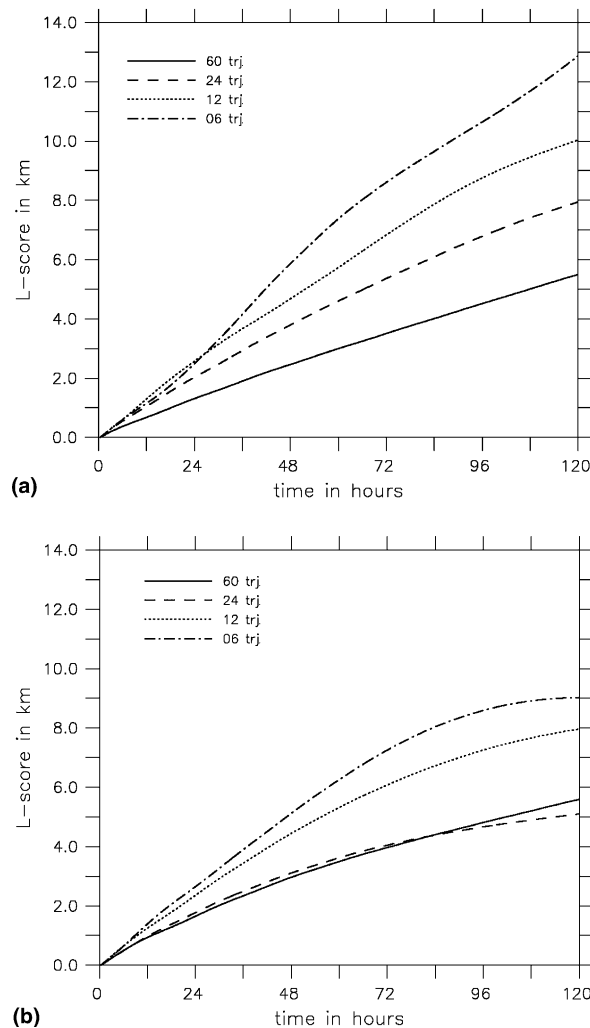


Fig. 5. Evolution of the Lagrangian metric L during the period of reconstruction, for $P = 60$ (plain lines), $P = 24$ (dashed lines), $P = 12$ (dotted lines), and $P = 6$ (dot-dashed lines). Trajectories used to compute the metric start (a) inside the region A or (b) inside the region B + C.

20–50% of this dispersion value. In B + C on the other hand, the dispersion after 5 days is ≈ 15 km, so that L reaches values of 35–60%. This indicates that region A, characterized by a more tight and smooth structure, is relatively better predicted in Lagrangian terms than region B + C, even though the absolute values of L are larger in region A.

The E metric integrated over the whole region is shown in Fig. 6a. The results are presented considering the E values versus the “coverage” corresponding to the number of floats P . The coverage is estimated (Toner et al., 2001b) as the ratio between the number of grid cells crossed by at least one observed trajectory with respect to the total number of grid cells (see Table 1). Results in terms of coverage are expected to be more easily generalized than results in terms of float number (Toner et al., 2001b). As it can be seen, the total coverage ranges between approximately 80% for $P = 60$ to 16% for $P = 6$. Notice that the coverage values, especially for small values of P , are expected to be dependent on the specific float realization considered. Here, in order to maintain consistency at varying P , we have considered realizations corresponding to homogeneous releases. The dependence on different sampling strategies is studied in details e.g. in Toner et al. (2001b) and Molcard et al. (in press). In Fig. 6a, E obviously increases with decreasing coverage, going from less than 10% for the maximum coverage to approximately 50% for the minimum one.

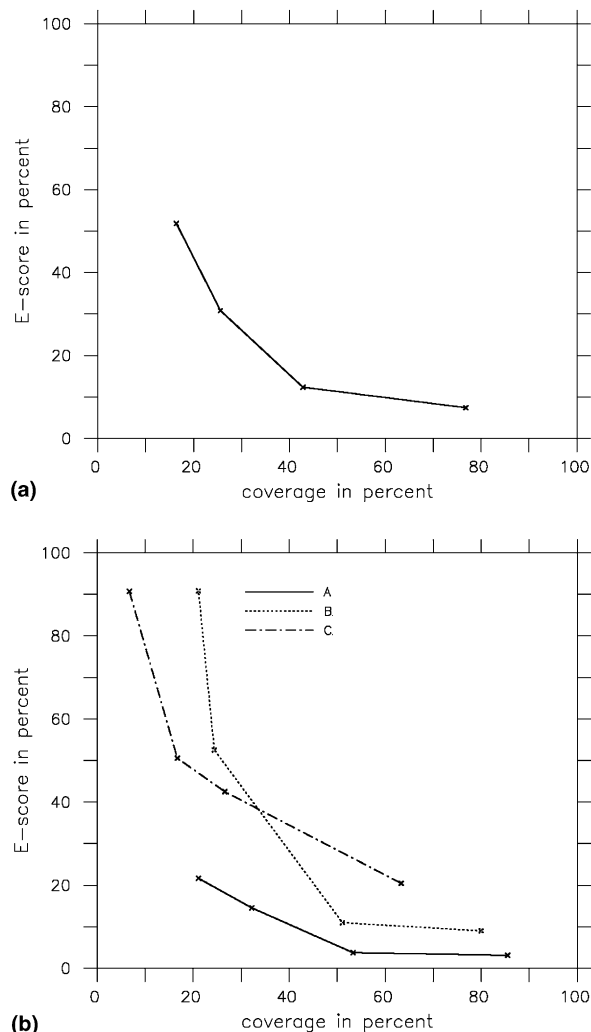


Fig. 6. Representation of the Eulerian metric E in function of the coverage rate as reported in Table 1. (a) $E(A + B + C)$. (b) $E(A)$ in plain line, $E(B)$ in dashed line, and $E(C)$ in dash-dotted line.

Table 1

Area (in number of grid cells) and observation coverage (i.e. the number of grid cells crossed at least by one control trajectory during T_M) for each region of interest

Region	Area	$P = 60$	$P = 24$	$P = 12$	$P = 6$
A	90	77	48	29	19
B	90	72	46	22	19
C	90	57	24	15	6
A + B + C	250	192	107	64	41

Notice that one grid cell represents $\sim 10 \text{ km} \times 10 \text{ km}$.

In order to better understand the dependency of the results on the type of flow to be reconstructed, we consider E versus coverage for the three separate sub-regions A, B, C (Fig. 6b). The values of coverage at given P are different in the three sub-regions (see Table 1), with A having the highest values and C the lowest. This is due to the different energetics of the sub-regions. In A, where the energy level is the highest, float trajectories are significantly longer and cover more grid cells than in region C, where the energy is the lowest. Minimum coverage (corresponding to $P = 6$) is approximately 20% for A and 7% for C. Also, the metric E appears to increase significantly more slowly for A than for B and C, even considering the same range of coverage. For 20% coverage, for instance, the metric E is approximately 20% for A, 45% for C and reaches 80% for B. This can be understood considering the characteristics of the reconstruction in Fig. 4, where the major features of the flow are retained even at small P , while the details of the recirculation cells are lost, starting from $P = 12$. The recirculation cells are mostly located in B, while A is dominated by the strong cyclonic structure and C by a broad flow. The E results then confirm the L results in Fig. 5 indicating that sub-region A, dominated by an energetic and coherent flow, is more easily reconstructed.

5.2. Sensitivity to observation sampling intervals

The sensitivity of the method to the sampling interval Δt is explored by varying Δt in the range between 60 h and 1 h for the four considered values of P . Values of the E metric versus Δt are shown in Fig. 7a, computed over the whole region. First we consider results obtained using the time-independent approximation (blue lines in Fig. 7). As it can be seen at given P the error does not change significantly in the range $T_1 < \Delta t < T_L$, i.e. between $\Delta t = 60 \text{ h}$ and $\Delta t = 24 \text{ h}$. This is in qualitative agreement with previous results (in the framework of assimilation, e.g. Molcard et al., 2003) showing that Lagrangian methods stay robust and provide reliable results in the range $\Delta T < T_L$. More in details, we notice that as Δt decreases within the range $T_1 < \Delta t < T_L$, we can expect two competing effects to occur. The flow in fact is influenced by the two different components, the mesoscale and the inertial oscillations. The mesoscale component is better resolved as Δt decreases, and the curvature of the trajectories over T_L is better reproduced. The inertial oscillations, on the other hand, are sub-sampled in all cases and can lead to aliasing in the trajectories. Possible biases in the average field are expected to increase at decreasing Δt as the ratio signal to noise increases. The balance between these two competing effects is difficult to predict. The results presented in Fig. 7a suggest that for high P the second effect tends to prevail and the error increases slightly with decreasing Δt , while the opposite happens for small P . This different behavior might be due to the fact that at high P the flow curvature is well represented even at relatively high values of Δt due to the presence of many trajectories. Results in the range $\Delta t < T_1$ (Fig. 7a) show a consistent increase of error at decreasing Δt for all P . This clearly shows that the time-independent approximation is not adequate when the inertial oscillations are even partially resolved. The method tends to reproduce the “wiggles” of the trajectories and the associated velocity fluctuations occurring at each sequence Δt . Because of these fluctuations, the reconstruction of the average field therefore becomes increasingly noisier. The results therefore suggest that, if the simple time-independent approximation is considered, the data should be first sub-sampled at $\Delta t > T_1$.

We now explore the effects of using the time-dependent method of Section 2.4 in the range $\Delta t < T_1$. The method is implemented in the following way. A reference field reconstruction is first performed, using the time-independent method and considering sub-sampled data at $\Delta t = 24 \text{ h}$. Instantaneous fields of velocity

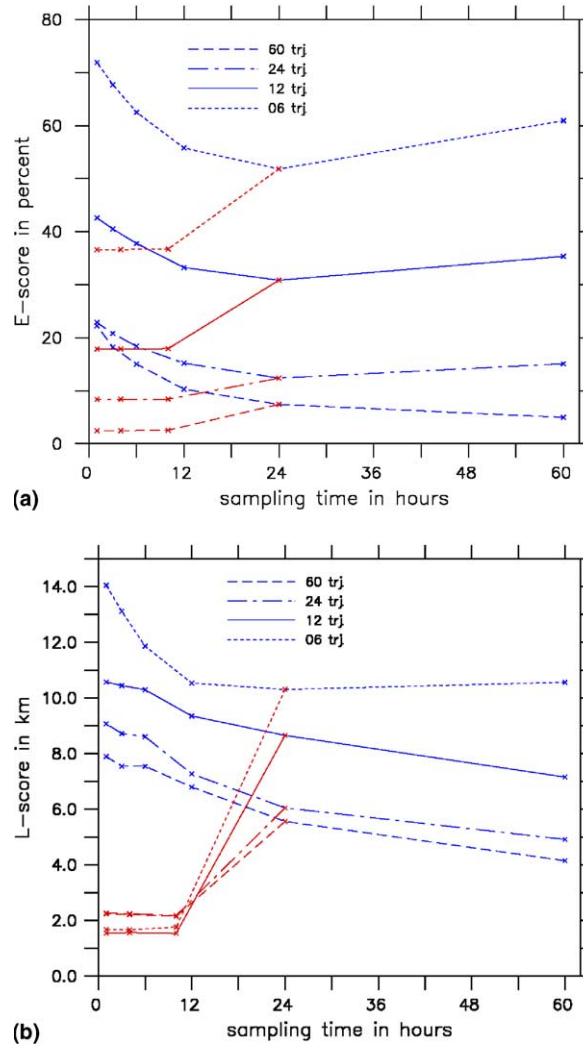


Fig. 7. Representation of (a) $E(A + B + C)$ and (b) $L(T_M)$ in function of the observation sampling time. Estimation performed in the time-independent approximation (time-dependent approximation) are reported in blue lines (red lines).

increments with respect to the reference field are then computed from the high frequency data as a function of time. From these fields, an average in time is computed and added to the reference field to provide a correction to the time-independent flow. The results in Fig. 7a (red lines) show the E metric computed using this corrected average flow. The time-dependent residuals are used to evaluate amplitude and phase of the inertial oscillations using the cost function given in Eq. (8). Examples of the time-dependent reconstructed field are shown in Fig. 8, in terms of velocity through a cross gyre section in the sub-region A. Results in Fig. 7a show that the average field reconstruction significantly improves using the time-dependent method. The E values significantly decrease with respect to those obtained with $\Delta t = 24$ h, dropping to less than half for $P = 60$ and decreasing from more than 50% to 35% for $P = 6$. Conceptually, the time-dependent method tends to correct biases introduced in the reconstruction by the sub-sampling of the inertial frequency. The method can be considered as an alternative to the common practice methods of pre-filtering the trajectory data or the pseudo-Lagrangian velocity time series and then using them to compute the average field (e.g. Toner et al., 2001b). These methods are generally not optimized for Lagrangian data and tend to convolute the space and time dependence of the signal. Here instead the Lagrangian nature of the data is used to reconstruct the space and time-varying field, over which the average flow is evaluated.

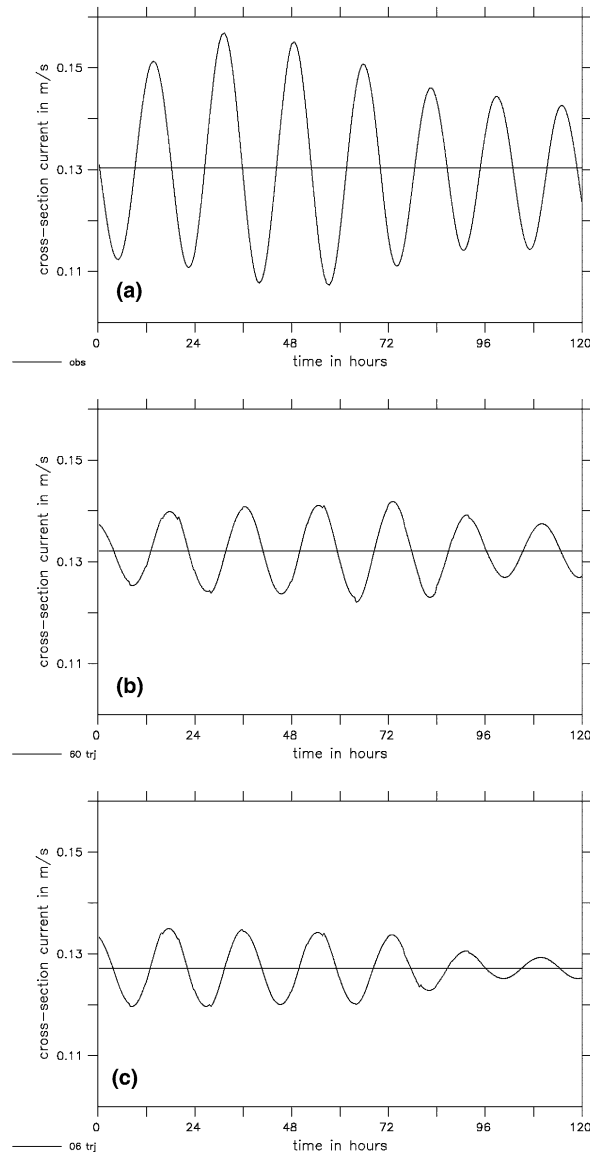


Fig. 8. Evolution of the meridional velocity averaged along the zonal section [4.8E, 5.4E] located at latitude 40.45N. (a) the control case, estimation with $\Delta t = 4$ h and (b) $P = 60$ or (c) $P = 6$. The straight line indicates the time-averaged value.

While the evaluation of the average field appears satisfactory, the reconstruction of the detailed inertial oscillations is only partially achieved. This is shown in the example of Fig. 8, where the reconstruction of the time-dependent gyre flux for $\Delta t = 4$ h and $P = 60$ and $P = 6$ is shown, together with the control time series. As it can be seen, the average flux is well reproduced, but the amplitude and the phase of the oscillations are only partially reproduced. As discussed in Section 2.4, this can be due to a number of reasons, including spatial inhomogeneity of the inertial oscillations and other simplifications used in the method such as independence of the velocity component.

Results in terms of the L metric computed over 5 days are shown in Fig. 7b, in the same range of Δt and P as for the metric E . The behavior of L versus Δt for the time-independent reconstruction (blue lines) is similar to the behavior of E , except that the error appears to increase slightly for $\Delta t = 24$ h with respect to $\Delta t = 60$ h for almost all P ($P = 24, 12, 6$). This suggests that the L metric is more sensitive to the presence of noise

introduced by inertial oscillations in the reconstruction. The time-dependent reconstruction reduces L even more drastically than in the E case. We notice that the reconstructed trajectories are obtained by integration in the time-dependent reconstructed field, rather than in the corrected average field considered in Fig. 7a. Examples of reconstructed trajectories for $\Delta t = 24$ h and $\Delta t = 4$ h are shown in Fig. 9, together with the corresponding control trajectories. The 4 h time-dependent reconstruction provides trajectories much closer to the control than the 24 h time-independent one. Nevertheless, when a zoom is performed on the trajectories and the details of the inertial oscillations are considered (Fig. 9b), amplitude and phase of the reconstructed trajectories appear only marginally reproduced, as already discussed for the Eulerian statistics (Fig. 8). In summary, we can say that the time-dependent method improves drastically the reconstruction of the mean velocity field and of the trajectories, as the exhaustive data set involved is able to efficiently constrain their estimation. On the other hand, the detailed dynamics of inertial currents are not completely identified, in terms of amplitude and phase, which would need further explorations (refer Park et al., 2004) that are out of the scope of the present paper.

We conclude by briefly discussing the problem of reconstruction in the case of $\Delta t = T_M$, i.e. when the meso-scale field is sub-sampled. For this value of Δt , the field reconstruction is necessarily performed using the time-independent approximation. The results are found to satisfactorily converge toward the control mean field (Fig. 1b) in the limit of very high float number ($P = 250$), but they quickly deteriorate at decreasing P . For $P = 24$, for instance, E is already of the order of 100%. In order to understand the reasons for the deterioration, we consider the structure of the reconstructed field \mathbf{u}^{est} in Fig. 10. The reconstruction appears reasonably similar to Fig. 1b in regions B and C, while it is clearly different in region A, where an unrealistic convergence region appears instead of the cyclonic gyre. The failure of the method can be understood considering the control trajectories in Fig. 3b, and taking into account the fact that for $\Delta t = T_M$ only the extrema, i.e. the initial and final points, of the trajectories are taken as data (i.e. $\mathbf{r}^{\text{obs}}(0)$ and $\mathbf{r}^{\text{obs}}(T_M)$). In the field reconstruction procedure, during the first iteration the velocity field is determined by the reconstructed trajectories connecting two consecutive positions along a straight line. In the case of Fig. 3b, two of the trajectories in the cyclonic gyre have extrema such that the reconstructed straight line trajectories cross, so that the velocities along them produce a convergence region. Since there are no other trajectories providing additional information on the velocity structure in the gyre interior, the convergence cannot be corrected by the method and it leads to the catastrophic error shown in Fig. 10. The problem is inherent in the nature of the data, which do not resolve the curvature of the trajectories. A similar convergence region in the reconstruction is expected to occur also

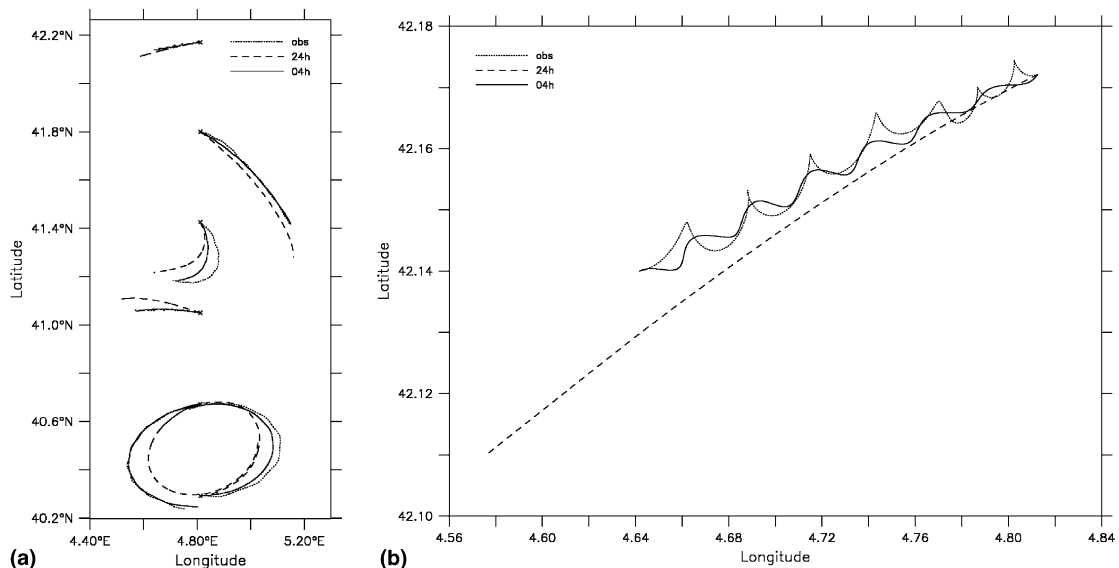


Fig. 9. (a) Trajectory reconstruction for $P = 6$; control trajectories (dotted lines), reconstructed trajectories with $\Delta t = 4$ h (plain line) and $\Delta t = 24$ h (dashed line). (b) Zoom on the northern reconstructed trajectory path.

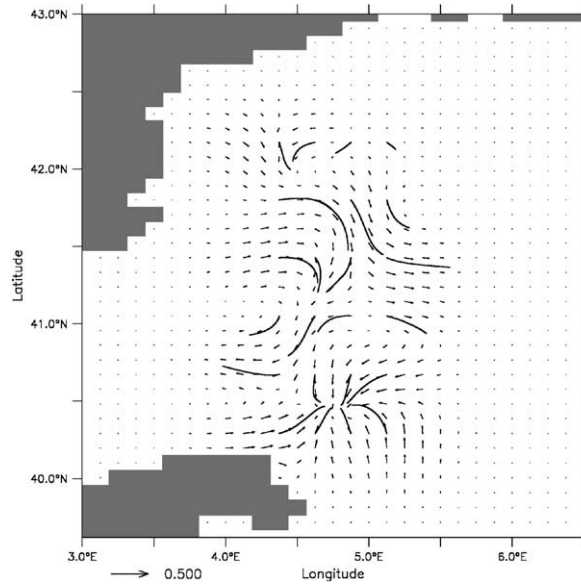


Fig. 10. Reconstruction of the Eulerian velocity field (in m s^{-1}) and corresponding trajectories, from the float positions with $\Delta t = 5$ d and $P = 24$.

for different methods, both Lagrangian and pseudo-Lagrangian, since it is directly linked to the “crossing” of the trajectory data. These results then point out to a potential problem occurring when using Lagrangian data with $\Delta t > T_L$. Since the typical mesoscale curvature is not resolved, it is possible, especially inside a coherent structure, that they lead to an unrealistic convergence region in the reconstruction. Only if other data or other a priori information (for instance from GCM results) are available, the information on the curvature can be recovered and a correct estimate can be obtained.

5.3. Combining observations and GCM results

As mentioned in Section 2.1, the method can be used also to combine information from data with information from GCM results. Technically, in this case the first guess \mathbf{u}^{bck} is provided by the GCM results, instead of being assumed null as in the reconstruction based on the data only. Here we show an example that illustrate the method performance in this modality. An example of a different realization of the OPA model velocity field in the same computational domain (Fig. 1) is shown in Fig. 11a. The velocity field (averaged over T_M) is shown, with superimposed six trajectories launched from the same initial conditions as the control trajectories of Fig. 3d. As it can be seen the velocity field is relatively similar to the control in terms of large scale flow, showing a main westward current along the coast, but the mesoscale flow is significantly different especially in the areas of interest A, B, C. The cyclonic gyre in A is not present, and the eastward return current in B has a different structure and different recirculation cells. Also the paths of the trajectories appear significantly different, especially in A and B. The time-independent method is used to reconstruct the control field using the data from $P = 6$ with $\Delta t = 24$ h, and combing them with the information of the field in Fig. 11a, used as first guessed \mathbf{u}^{bck} . The results, shown in Fig. 11b, indicate that the correction with respect to the mean \mathbf{u}^{bck} is significant in the region of interest, despite the small number of floats considered. The cyclonic vortex is present in the reconstructed field \mathbf{u}^{est} (Fig. 11b), and the eastward flow in B is at least partially corrected where data are available. Also the trajectories are strongly modified with respect to the background trajectories of Fig. 11a, and are quite similar to the control ones.

These results suggest that the method in this modality can be useful to blend information from data and GCM results, and, as such, can be used in the framework of assimilation studies for correcting GCM initial conditions. In the assimilation studies of Molcard et al. (2003, 2005) and Özgökmen et al. (2003), the velocity

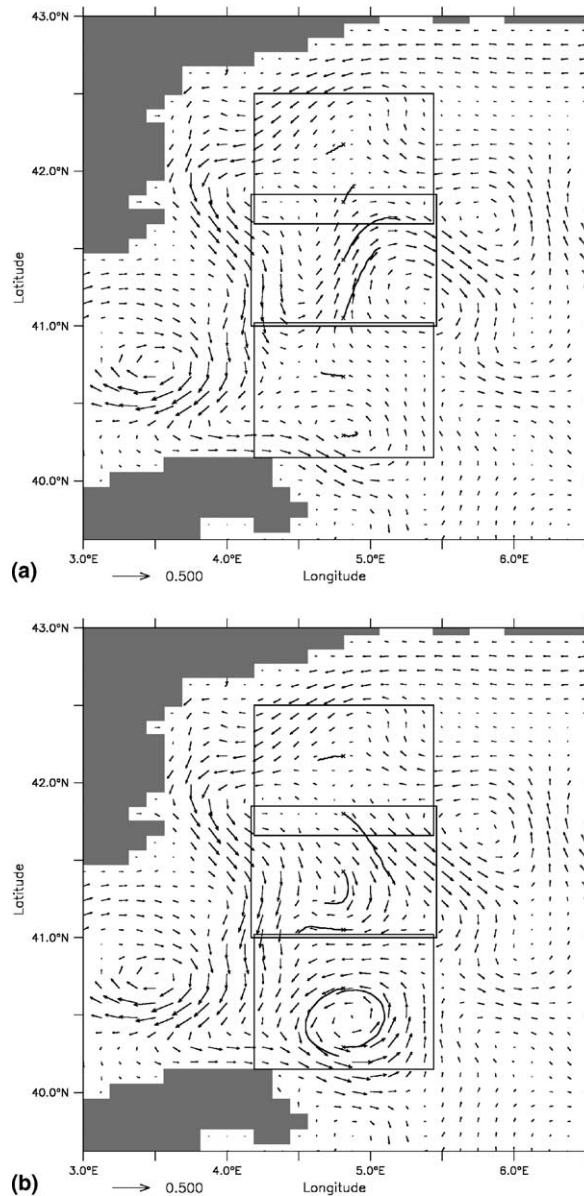


Fig. 11. Representation of the Eulerian velocity field (in m s^{-1}) averaged over T_M and associated trajectories. (a) first guess from GCM, (b) estimation obtained with $\Delta t = 1$ d and $P = 6$.

reconstruction from Lagrangian data is performed considering a velocity correction only in the neighborhood of the initial point of each sequence (at $\mathbf{r}^{\text{obs}}((m-1)\Delta t)$ for the m th sequence), and using a diagonal covariance matrix \mathbf{B} . In the present method, the correction is performed along the estimated trajectories and the covariance matrix \mathbf{B} is appropriately modeled. It can be expected that these improvements lead to significant advantages especially in the case of high resolution GCMs with complex mesoscale structures. For large scale, low resolution models (as the ones considered in Molcard et al., 2003, 2005), the correction in the neighborhood of the position is likely to be appropriate since the trajectory is not likely to cross many grid points during Δt and the field is sufficiently smooth. For high resolution models, on the other hand, expanding the correction along the trajectories and through the covariance matrix \mathbf{B} can be relevant, helping to capture the more complex mesoscale structures.

6. Summary and concluding remarks

In this paper, a method for the reconstruction of the mesoscale Eulerian velocity field based on Lagrangian observations is presented. A variational approach is used, where information from the observations, i.e. the positions of P floats, are combined with a simple model constraint, given by the advection equation describing the motion of particles in a velocity field. The velocity field is estimated minimizing a cost function which measures the distance between the observed float positions and the positions predicted by the advection model during the sampling period Δt . The method is first implemented considering the time-independent approximation of the velocity correction during a time interval T_M , which is shorter than the typical time scale of the mesoscale field T_E , $T_M \ll T_E$. The method uses all the information from the P floats during T_M to provide an estimate of the mesoscale velocity \mathbf{u}^{est} during that time. In a second step, the time-independent approximation is relaxed, considering the case of high frequency inertial oscillations superimposed to the mesoscale field. The deviation of the velocity field from \mathbf{u}^{est} at each sampling interval is computed and it is used on the one side to correct \mathbf{u}^{est} itself, and on the other side to estimate the time-dependent parameters of the inertial oscillations, such as amplitude and phase. Notice that the method directly depends only on one flow parameter, i.e. the space scale R . For the reconstruction of the time-independent mesoscale field, the Eulerian space scale $R_E = 20$ km has been chosen, and preliminary tests have shown that the results are robust in the range $R_E = 10\text{--}30$ km. For the reconstruction of the time-dependent inertial field, a larger R has been used, $R = 100$ km. The Eulerian and Lagrangian time scales of the field, T_E and T_L do not enter directly in the reconstruction procedure. Nevertheless, their range of values has to be known at least approximately since the methodological setup depends on them. T_M in fact has to be $T_M \ll T_E$ in order to provide a sensible estimate of the mesoscale field, and appropriate choices of Δt depends on the value of T_L .

The method is tested in the framework of results from the OPA circulation model in the Mediterranean Sea, considering a regional application. An area of approximately 100×250 km² is considered in the northwestern Mediterranean Sea, located north of the Balearic Islands and characterized by a vigorous mesoscale field superimposed with a strong inertial signal. The field is characterized by an Eulerian time scale T_E significantly longer than the Lagrangian time scale T_L , while both scales exceed the inertial period: $T_I < T_L < T_E$. This condition is typical of mid-latitude flows in the open ocean or in regional seas, in the upper-middle column (e.g. Lumpkin et al., 2002). The specific realization of the mesoscale velocity chosen as a control field for testing the reconstruction is characterized by a number of different structures, therefore providing an interesting test ground for the method. A strong and coherent cyclonic vortex is present in the southern part of the region, while an eastern flow with recirculation cells is found in the central part.

The method is tested varying the number of floats P and the value of the sampling time Δt . The float number is varied between $P = 60$ and $P = 6$ and the results are presented in terms of coverage, to allow for a more general interpretation (Toner et al., 2001b). The coverage varies between approximately 80% for $P = 60$ and 16% for $P = 6$, and the corresponding error in the field reconstruction varies from less than 10% to approximately 50% for $\Delta t = 24$ h and considering the time-independent approximation. For smaller sampling intervals and using the time-dependent reconstruction the results can be further improved. From a qualitative point of view, the main features of the control field appear to be captured by the reconstruction \mathbf{u}^{est} even for the lowest coverage, especially in the cyclonic gyre. The reconstruction of the recirculation cells is more problematic, since their space scales are smaller and they are only partially sampled for low values of P .

The method sensitivity to the sampling interval is mostly tested considering the range $\Delta t < T_L$, for which the position data are expected to retain significant information on the velocity (e.g. Molcard et al., 2003). A first set of tests is performed considering the time-independent reconstruction. They indicate that the reconstruction is weakly sensitive to Δt in the range $T_I < \Delta t < T_L$, while it deteriorates significantly in the range $\Delta t < T_I$ for decreasing Δt , signaling the limit of validity of the time-independent assumption. The time-dependent correction is then tested for $\Delta t < T_I$, showing significantly improved results. The mesoscale reconstruction \mathbf{u}^{est} is corrected, taking into account the biases introduced by the sub-sampling of the inertial oscillations. The procedure can be considered as a valid and advantageous alternative to the commonly used method of pre-filtering the trajectories or the associated (pseudo-Lagrangian) velocity time-series. Also, the method provides an estimate of the time-dependent field during T_M . The corresponding trajectories approximate quite closely the control trajectories, at least in terms of main paths. The details of the inertial oscillations in terms of amplitude

and phase, on the other hand, are only partially reproduced. This can be due to a number of effects that are not included in the present methodology, such as bivariate constraints in the error covariance, high spatial inhomogeneity or presence of other high frequency signals. The detailed description of the inertial signal is out of the scope of the present paper, which is centered on the mesoscale reconstruction, and it will be considered in future work. A more restricted set of tests have been performed considering a case with $\Delta t > T_L$, and varying P . The results point out to a potential problem, which can occur especially for floats in coherent structures. Since the data do not resolve the mesoscale flow curvature, it can happen that data sets induce trajectories that cross each other, therefore producing a convergence in the reconstructed velocity field. This “crossing” problem is inherent in the data, and it would affect also different reconstruction methods. In the majority of cases, on the other hand, when this problem does not occur in the data, the reconstruction appears qualitatively satisfactory even though the field is necessarily smoothed. Finally, an example where the method is used to combine information from the data with GCM information has been considered. The results are very encouraging and they suggest that the method can be useful in the framework of assimilation studies. With respect to previous assimilation works by Molcard et al. (2003, 2005), the present method has the advantage of expanding the velocity correction along the estimated trajectories and using an appropriate matrix covariance. These aspects are expected to be particularly advantageous in the case of high resolution GCMs describing complex mesoscale fields.

The present results suggest various directions of future work. They include the already mentioned further refinements in the estimation of inertial parameters, as well as the practical applications in assimilation problems especially for correcting GCMs initial conditions. Also, the present results are obtained using the assumption of “perfect” data. The impact of data error on the reconstruction is an important topic to be explored, especially for subsurface ARGO floats, since shear and surface drifts can significantly affect the float positions. Finally, in this work we have assumed a separation of scales between the high frequency inertial signal and the mesoscale flow. Even though this assumption is reasonable for regional flows in mid-latitude, the separation might not be so clear for instance for applications in coastal regions where T_L can be of the order of a day (e.g. Maurizi et al., 2004), or for low latitude applications where T_L can be of the order of several days. In these cases a further relaxation of the time-independent approximation might be necessary.

Acknowledgements

V. Taillandier and A. Griffa greatly appreciate the support of the ECC Mediterranean Forecasting Project MFSTEP and the Office of Naval Research (grant N00014-97-1-0620). A. Molcard was supported by the EC project “Stirring and Mixing in Turbulence”. All authors acknowledge support from the Consiglio Nazionale Ricerche (CNR, Italy). Numerical calculations were performed on the IDRIS computing center.

References

- Aref, H., 1984. Stirring by chaotic advection. *J. Fluid Mech.* 143, 1–21.
- Bauer, S., Swenson, M.S., Griffa, A., 2002. Eddy-mean flow decomposition and eddy-diffusivity estimates in the tropical Pacific Ocean. 2. Results. *J. Geophys. Res.* 107, 3154–3171.
- Bauer, S., Swenson, M.S., Griffa, A., Mariano, A.J., Owens, K., 1998. Eddy-mean flow decomposition and eddy-diffusivity estimates in the tropical Pacific Ocean. *J. Geophys. Res.* 103, 30855–30871.
- Béranger, K., Mortier, L., Crépon, M., 2005. Seasonal variability of transports through the Gibraltar, Sicily and Corsica straits from a high resolution Mediterranean model. *Prog. Oceanogr.* 66 (2–4), 341–364.
- Castellari, S., Griffa, A., Özgökmen, T.M., Poulain, P.M., 2001. Prediction of particle trajectories in the Adriatic Sea using Lagrangian data assimilation. *J. Mar. Syst.* 29, 33–50.
- Courtier, P., Thépaut, J.-N., Hollingsworth, A., 1994. A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.* 120, 1367–1388.
- Davis, R.E., 1991. Observing the general circulation with floats. *Deep-Sea Res.* 38, 5531–5571.
- Davis, R.E., Sherman, J.T., Dufour, J., 2001. Profiling ALACEs and other advances in autonomous subsurface floats. *J. Atmos. Oceanic Tech.* 18, 982–993.
- Derber, J., Bouttier, F., 1999. A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus* 51A, 195–221.
- Derber, J., Rosati, A., 1989. A global oceanic data assimilation system. *J. Phys. Oceanogr.* 19, 1333–1347.

- Flexas, M., Durrieu de Madron, X., Garcia, M.A., Canals, M., Arnau, P., 2002. Flow variability in the Gulf of Lions during the MATER HFF experiment (March–May 1997). *J. Mar. Syst.* 33–34, 197–214.
- Fratantoni, D.M., 2001. North Atlantic surface circulation during the 1990s observed with satellite-tracked drifters. *J. Geophys. Res.* 106, 22067–22093.
- Garraffo, Z.D., Mariano, A.J., Griffa, A., Veneziani, C., Chassignet, E.P., 2001. Lagrangian data in a high-resolution numerical simulation of the North Atlantic. 1. Comparison with in situ drifter data. *J. Mar. Syst.* 29, 157–176.
- Gilbert, J.-C., Lemaréchal, C., 1989. Some numerical experiments with variable-storage quasi-Newton algorithms. *Math. Programm.* 45, 407–435.
- Griffa, A., Piterbarg, L.I., Özgökmen, T.M., 2004. Predictability of lagrangian particle trajectories: effects of the underlying eulerian flow. *J. Mar. Res.* 35, 1–35.
- Kamachi, M., O'Brien, J.J., 1995. Continuous assimilation of drifting buoy trajectories into an equatorial Pacific Ocean model. *J. Mar. Syst.* 6, 159–178.
- Kuznetsov, L., Ide, K., Jones, C.K.R.T., 2003. A method for assimilation of lagrangian data. *Mon. Wea. Rev.* 131, 2247–2260.
- Lavander, K.L., Davis, R.E., Owens, W.B., 2000. Direct velocity measurements in the Labrador and Irminger Seas describe pathways of Labrador Sea Water. *Nature* 407, 66–69.
- LeDimet, F.X., Talagrand, O., 1986. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus* 38A, 97–110.
- Lorenc, A.C., 1988. Optimal nonlinear objective analysis. *Quart. J. Roy. Meteor. Soc.* 114, 205–240.
- Lumpkin, R., 2003. Decomposition of surface drifter observations in the Atlantic Ocean. *Geophys. Res. Lett.* 30, 1753.
- Lumpkin, R., Treguier, A., Speer, K., 2002. Lagrangian eddy scales in the Northern Atlantic. *J. Phys. Oceanogr.* 32, 2440–2452.
- Madec, G., Delecluse, P., Imbard, M., Levy, C., 1998. OPA8.1 ocean general circulation model reference manual. Technical Report LODYC/IPSL. Av[<http://www.lodyc.jussieu.fr/opa>].
- Maurizi, A., Griffa, A., Poulain, P.M., Tampieri, F., 2004. Lagrangian turbulence in the Adriatic Sea as computed from drifter data: effects of inhomogeneity and nonstationarity. *J. Geophys. Res.* 109, C04010.
- Millot, C., Crépon, M., 1981. Inertial oscillations on the continental shelf of the Gulf of Lions. Observations and Theory. *J. Phys. Oceanogr.* 11, 639–657.
- Molcard, A., Griffa, A., Özgökmen, T.M., 2005. Lagrangian data assimilation in multi-layer primitive equation ocean models. *J. Atmos. Oceanic Tech.* 22 (1), 70–83.
- Molcard, A., Piterbarg, L.I., Griffa, A., Özgökmen, T.M., Mariano, A.J., 2003. Assimilation of drifter positions for the reconstruction of the eulerian circulation field. *J. Geophys. Res.* 108 (C3), 3056.
- Molcard, A., Poje, A.J., Özgökmen, T.M., in press. Directed drifter launch strategies for Lagrangian data assimilation using hyperbolic trajectories. *Ocean Modell.*
- Owens, W.B., 1991. A statistical description of the mean circulation and eddy variability in the northwestern Atlantic using SOFAR floats. *Prog. Oceanogr.* 28, 257–303.
- Özgökmen, T.M., Griffa, A., Piterbarg, L.I., Mariano, A.J., 2000. On the predictability of the Lagrangian trajectories in the ocean. *J. Atmos. Oceanic Tech.* 17, 366–383.
- Özgökmen, T.M., Molcard, A., Chin, T.M., Piterbarg, L.I., Griffa, A., 2003. Assimilation of drifter positions in primitive equation models of midlatitude ocean circulation. *J. Geophys. Res.* 108 (C7), 3238.
- Özgökmen, T.M., Piterbarg, L.I., Mariano, A.J., Ryan, E.H., 2001. Predictability of drifter trajectories in the tropical Pacific Ocean. *J. Phys. Oceanogr.* 31, 2691–2720.
- Park, J.J., Kim, K., Crawford, W.R., 2004. Inertial currents estimated from surface trajectories of ARGO floats. *Geophys. Res. Lett.* 31, L13307.
- Poulain, P.M., 2001. Adriatic Sea surface circulation as derived from the drifter data between 1990 and 1999. *J. Mar. Syst.* 29, 3–32.
- Priestley, M.B., 1981. Spectral analysis and time series. Academic Press.
- Reverdin, G., Niiler, P.P., Valdimarsson, H., 2003. North Atlantic Ocean surface currents. *J. Geophys. Res.* 108 (C1), 3002.
- Richardson, P.L., 1993. A census of eddies observed in North Atlantic SOFAR float data. *Prog. Oceanogr.* 31, 1–50.
- Samelson, R.M., 1996. Chaotic transport by mesoscale motions. In: Adler, R.J., Muller, P., Rozovskii, B.L. (Eds.), *Stochastic Modeling in Physical Oceanography*. Birkhauser, pp. 423–438.
- Swenson, M.S., Niiler, P.P., 1996. Statistical analysis of the surface circulation of the California Current. *J. Geophys. Res.* 101, 22631–22645.
- Talagrand, O., 1991. The use of adjoint equations in numerical modeling of the atmospheric circulation. In: Proc. Workshop on Automatic Differentiation of Algorithms, Theory Implementation and Application, ECMWF.
- Talagrand, O., Courtier, P., 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation. *Quart. J. Roy. Meteor. Soc.* 113, 1311–1328.
- Testor, P., Gascard, J.C., 2003. Large-scale spreading of deep waters in the Western Mediterranean Sea by submesoscale coherent eddies. *J. Phys. Oceanogr.* 33, 75–87.
- Toner, M., Kirwan, A.D., Kantha, L.H., Choi, J.K., 2001a. Can general circulation models be assessed and their output enhanced with drifter data? *J. Geophys. Res.* 106, 19563–19579.
- Toner, M., Poje, A.C., Kirwan, A.D., Jones, C.K.R.T., Lipphardt, B.L., Grosch, C.E., 2001b. Reconstructing basin-scale Eulerian velocity fields from simulated drifter data. *J. Phys. Oceanogr.* 31, 1361–1376.
- Van Haren, H., Millot, C., 2003. Seasonality of internal gravity waves kinetic energy spectra in the Ligurian Sea. *Oceanol. Acta* 26, 635–664.

- Veneziani, M., Griffa, A., Reynolds, A.M., Mariano, A.J., 2004. Oceanic turbulence and stochastic models from subsurface Lagrangian data for the North-West Atlantic Ocean. *J. Phys. Oceanogr.* 34, 1884–1906.
- Weaver, A.T., Courtier, P., 2001. Correlation modeling on the sphere using a generalized diffusion equation. *Quart. J. Roy. Meteor. Soc.* 127, 1815–1846.
- Zhang, H.M., Prater, M.D., Rossby, T., 2001. Isopycnal Lagrangian statistics from the North Atlantic Current RAFOS floats observations. *J. Geophys. Res.* 106, 13817–13836.
- Zhou, M., Niiler, P.P., Hu, J., 2002. Surface currents in the Bransfield and Gerlache Straits measured by surface Lagrangian drifters. *Deep-Sea Res.* 49, 267–280.