Journal of Marine Research, 62, 1–35, 2004

Journal of
MARINE RESEARCH

Volume 62, Number 1

Predictability of Lagrangian particle trajectories: Effects of smoothing of the underlying Eulerian flow

by Annalisa Griffa[1,2], Leonid I. Piterbarg[3] and Tamay Özgökmen[1]

ABSTRACT

The increasing realism of ocean circulation models is leading to an increasing use of Eulerian models as a basis to compute transport properties and to predict the fate of Lagrangian quantities. There exists, however, a significant gap between the spatial scales of model resolution and that of forces acting on Lagrangian particles. These scales may contain high vorticity coherent structures that are not resolved due to computational issues and/or missing dynamics and are typically suppressed by smoothing operators.

In this study, the impact of smoothing of the Eulerian fields on the predictability of Lagrangian particles is first investigated by conducting twin experiments that involve release of clusters of synthetic Lagrangian particles into "true" (unmodified) and "model" (smoothed) Eulerian fields, which are generated by a QG model with a flow field consisting of many turbulent coherent structures. The Lagrangian errors induced by Eulerian smoothing errors are quantified by using two metrics, the difference between the centers of mass (CM) of particle clusters, ρ, and the difference between scattering of particles around the center of mass, *s*. The results show that the smoothing has a strong effect on the CM behavior, while the scatter around it is only partially affected.

The QG results are then compared to results obtained from a multi-particle Lagrangian Stochastic Model (LSM) which parameterizes turbulent flow using main flow characteristics such as mean flow, velocity variance and Lagrangian time scale. In addition to numerical results, theoretical results based on the LSM are also considered, providing asymptotics of ρ, *s* and predictability time. It is shown that both numerical and theoretical LSM results for the center of mass error (ρ) provide a good qualitative description, and a quantitatively satisfactory estimate of results from QG experiments. The scatter error (*s*) results, on the other hand, are only qualitatively reproduced by the LSM.

1. RSMAS/MPO, University of Miami, Miami, Florida, 33149, U.S.A. ***email: agriffa@rsmas.miami.edu***
2. Consiglio Nazionale delle Ricerche, ISMAR, LaSpezia, Italy.
3. Center for Applied Mathematical Sciences, University of Southern California, Los Angeles, California, 90089, U.S.A.

1. Introduction

In the last decade, a great effort has been devoted to improving our capability of predicting ocean currents through the combined use of high-resolution circulation models (e.g., Stammer and Chassignet, 2000; Smith *et al.,* 2000) and data assimilation (e.g., Ghil and Malanotte-Rizzoli, 1991; Lorenc, 1986). Many important applications of ocean current prediction are related to transport problems (Coulliette and Wiggins, 2000; Toner *et al.,* 2001; Mariano *et al.,* 2002; Özgökmen *et al.,* 2003). They include biological and ecological problems, such as assessment of larvae paths, search and rescue problems, and prediction of pollution spreading from industrial plants as well as oil spills due to accidents at sea. Motivated by the interest in transport problems, a number of different methods have been developed to accurately compute transport in a given velocity field in the ocean. In general, two categories of techniques have been developed to estimate the Lagrangian trajectories. The first approach relies on existing correlated Lagrangian particle data to predict trajectories of other particles (e.g., Özgökmen *et al.,* 2000, 2001; Castellari *et al.,* 2001; Piterbarg, 2001a; Piterbarg and Özgökmen, 2002). The second approach is based on the knowledge of the underlying Eulerian field, either from high-resolution surface velocity data (e.g., Shay *et al.,* 2002) and/or ocean general circulation models (OGCMs). Given the Eulerian velocity field, Lagrangian particles can be advected with the flow (e.g. Blanke *et al.,* 1999; Dutkiewicz *et al.,* 1993) or methods based on dynamical system theory can be applied to describe the general Lagrangian flow structure (e.g., Wiggins, 1992; Poje and Haller, 1999; Couliette and Wiggins, 2000; Kuznetsov *et al.,* 2002).

Despite the advances in Lagrangian techniques, transport prediction in the ocean is still problematic. The well-known conceptual reason is that particle trajectories are extremely sensitive to the details of the underlying Eulerian flow (e.g., Aref, 1984; Samelson, 1996) and are highly chaotic. Even small errors in the estimation of ocean currents can drastically change particle trajectories and greatly reduce Lagrangian predictability. Sources of errors for the Eulerian fields produced by prediction models include uncertainties in the forcing functions, bathymetry, coastline, internal stratification and details of model numerics. At a more fundamental level, ocean general circulation models have a finite horizontal space resolution, typically 5–20 km, and subgrid effects are usually modeled by using simple closure schemes such as Laplacian or biharmonic operators (Chassignet and Garraffo, 2001). However, these closures are simply smoothing operators and fail to capture many complex oceanic processes at the submeso-scale and small-scale range. An example of such missing turbulent processes is strong eddies, with vorticity of $\sim 7 \times 10^{-4}\,\mathrm{s}^{-1}$ (or 10 times the Coriolis frequency at midlatitudes) and diameter of 1–3 km, observed using high-resolution Doppler radar off the Florida coast (Shay *et al.,* 2002; Peters *et al.,* 2002). Also, most ocean models are subject to the hydrostatic approximation, which fails at space resolutions of less than several km even if the computations are feasible. In contrast, oceanic Lagrangian particles presumably feel forces acting at scales of the O (m). Therefore, truncation of turbulent scales below the typical grid spacing of 5–20 km in OGCMs is highly likely to impact the accuracy of simulated Lagrangian trajectories. Note

that similar errors may also arise from other gridded data, such as satellite data or radar data used for studying transport in coastal environments.

While the inherent difficulty in Lagrangian prediction is well known, a quantitative assessment of Lagrangian errors due to uncertainties in the underlying Eulerian velocity field is still lacking. The question of how to quantify Lagrangian errors versus Eulerian model errors is conceptually relevant as well as important in many practical applications, and it provides the main motivation for the present work. Here, we focus on a single error source, i.e. on the ubiquitous smoothing error due to removal of subgrid-scale processes in an Eulerian field. In other words, we assume that the Eulerian model is perfect except for finite resolution and smoothing, and we investigate quantitatively the impact of Eulerian smoothing on the prediction of particle trajectories. Of course, in real applications many other errors will come into play but at least the present results will provide a lower bound to Lagrangian uncertainty estimates and it will take into account the effects of subgrid-scale suppression.

The work is performed following two convergent lines of investigations. A numerical study is carried out using a reduced-gravity quasi-geostrophic (QG) model. Even though this model does not contain small-scale motion and nonhydrostatic effects, the flow field is rich with meso- and sub-mesoscale coherent structures that are driven by the stochastic western boundary conditions (Mariano *et al.,* 2003). A smoothing Kernel is then applied to this flow field as an approximation for the loss of small-scale information. Thus, this is a first approach in the context of a simple ocean model and simplified removal of velocity information, in order to explore the impact on Lagrangian predictability. The fundamental problem we pose can be further pursued using more comprehensive models and schemes, in light of the results presented in this study.

The original flow field is considered here as our "true" flow. Smooth versions of the truth are also computed and regarded as "model" results. Particle clusters are launched in the "truth" and in the "model" flows. The impact of Eulerian smoothing on particle trajectories is quantified considering primarily the difference between the center of mass positions, ρ, of clusters launched in the same positions in the "truth" and in the "model" flows. This metric characterizes the average path of the cluster and its sensitivity to smoothing. A second metric is also introduced to characterize the sensitivity of the spreading within each cluster, and it is given by the difference between particle scattering, s, in the "true" and "model" flows. The two metric results, obtained in two different energetic regimes, provide an assessment of the impact of smoothing on Lagrangian predictability.

In parallel, we consider recent theoretical results by Piterbarg (2004) obtained in the framework of an idealized Lagrangian Stochastic Model (LSM) (Piterbarg, 2001a,b). The LSM allows parameterization of the dynamics of multiple particles in a turbulent flow knowing the Eulerian mean flow $\mathbf{U}(\mathbf{x})$ and a restricted number of parameters characterizing the turbulent fluctuation, i.e. the Eulerian space scale R, the velocity variance σ_u^2 and the Lagrangian time scale τ. In the case of single particles, the LSM reduces to the well-known Langevin equation (e.g. Risken, 1989), which is widely used to represent particle motion in

the atmosphere and ocean (e.g. Thomson, 1986; Dutkiewicz *et al.,* 1993; Falco *et al.,* 2000). The multi-particle behavior is introduced through an appropriate space-correlated noise (Piterbarg, 2001a,b). This feature is essential in prediction studies like the one presented here. In Piterbarg (2004) the LSM is used to evaluate the smoothing impact in terms of ρ and s as function of the fluctuation parameters. Simple asymptotic formulas are derived, the validity of which are well supported by the stochastic simulations.

The present study has two objectives. The first aim is to get a quantitative measure of the sensitivity of Lagrangian predictability as a function of the degree of smoothing of Eulerian fields simulated by the QG model. The second objective is to compare and validate the LSM results of Piterbarg (2004) using results from QG experiments. While single-particle models have been tested in many situations, including model results (Garraffo *et al.,* 2001) and Lagrangian ocean data (Griffa, 1996; Falco *et al.,* 2000; Bauer *et al.,* 1998, 2002), multi-particle LSMs (e.g. Thomson, 1990; Sawford, 1993; Piterbarg, 2001a,b) are still relatively new and have been only partially tested in an oceanographical context (LaCasce and Bower, 2000). Validation of the multi-particle LSM (Piterbarg, 2004) with Lagrangian particle results from the Eulerian fields generated by the QG model would indicate that the LSM results can be used as a guideline to give estimates of (smoothing) Lagrangian errors also in the real ocean, thus providing significant advantages by allowing a direct and simple error estimation in terms of few basic flow parameters (e.g., R, σ_u, τ).

The paper is organized as follows. In Section 2, the general methodology to compare "true" and "model" results is outlined, the QG model is described, the background on the LSM (Piterbarg, 2001a,b) and the theoretical results of Piterbarg (2004) are provided. The results from the QG experiments are presented in Section 3 and those from LSM are shown in Section 4. A comparison between LSM and QG results is given in Section 5. A summary is provided in Section 6 together with remarks on future studies and applications. The Appendix documents the sensitivity of the metrics to sample size in QG and LSM experiments.

2. Methodology

a. Notation and metrics

Consider a particle trajectory $\mathbf{r}(t)$ as the solution of

$$\dot{\mathbf{r}} = \mathbf{u}(t, \mathbf{r}), \quad \mathbf{r}(0) = \mathbf{r}_0,$$

where $\mathbf{u}(t, \mathbf{r})$ is the true Eulerian velocity field and $\mathbf{r}_0$ is the trajectory initial condition.

We assume that $\mathbf{u}$ is modeled using a "perfect model" except for finite resolution. The model result can then be considered as a smoothed version of $\mathbf{u}$ in space and time. Here we concentrate on the issue of smoothing in space, since space resolution is often the critical issue in practical applications. The model result is then represented simply as the true field $\mathbf{u}(t, \mathbf{r})$ smoothed in space with scale h and is indicated as $\mathbf{u}_h(t, \mathbf{r})$,

$$\mathbf{u}_h = K_h * \mathbf{u}, \tag{1}$$

where K_h is a smoothing scalar kernel and the star means convolution.

Note that the smoothing procedure (1) is not equivalent to a velocity field obtained by lowering the resolution of a numerical dynamical model. When resolution is lowered in a dynamical model, the Reynolds number of the solution, and, therefore, its degree of nonlinearity, changes and other effects rather than simple smoothing can occur (e.g. Siegel *et al.,* 2001). Since our interest is in taking a first step in determining the effects of uncertainty in the Eulerian fields, we focus here on the smoothing effects, which are more straightforward to control as function of the parameter h, rather than considering the more complex problem of running a model at different resolutions. On the other hand, the procedure (1) exactly corresponds to considering a "statistical model" (rather than a dynamical model), i.e. a smoother and interpolator applied to a set of discrete data point. This is a practical and important problem for transport studies based on velocity fields obtained from HF radar measurements or coastal survey data (e.g. Paris *et al.,* 2002).

The trajectories advected by the smoothed model field are indicated as $\mathbf{r}_h$,

$$\dot{\mathbf{r}}_h = \mathbf{u}_h(t, \mathbf{r}_h).$$

In this study, the following kernel is used:

$$K_h(\mathbf{r}) = \frac{1}{2\pi h^2} \exp(-\mathbf{r}^2/2h^2).$$

We consider clusters consisting of M particles released randomly within a circle of small diameter ($\ll R$) centered at random locations. The particles are then advected starting from these initial positions with two flow fields, $\mathbf{u}$ and $\mathbf{u}_h$. We address the question of how different the particle trajectories become under these two flow fields at a later time as a function of the smoothing parameter h. We consider two different metrics of cluster differences. The first one is the mean square distance, ρ, between their centers of mass (CM)

$$\rho(t) = (E\{(\bar{\mathbf{r}}(t) - \bar{\mathbf{r}}_h(t))^2\})^{1/2}, \tag{2}$$

where $E\{\ \}$ is the expectation,

$$\bar{\mathbf{r}}(t) = \frac{1}{M}\sum_1^M \mathbf{r}_m(t), \quad \bar{\mathbf{r}}_h(t) = \frac{1}{M}\sum_1^M \mathbf{r}_{h,m}(t),$$

the sub m indicates the particle number in a cluster. The second metric is the difference between the scattering radia around the CMs

$$s(t) = S(t) - S_h(t), \tag{3}$$

where the scattering for the true flow is defined by

$$S(t) = \left(\frac{1}{M} \sum_{1}^{M} E\{|\mathbf{r}_m(t) - \bar{\mathbf{r}}(t)|^2\} \right)^{1/2}$$

and for the model flow in the same way.

For a flow **u** characterized by a typical space scale R, we introduce the predictability time T_{pred} defined by

$$\rho(T_{pred}) = R. \tag{4}$$

T_{pred} characterizes the typical time after which particles in the truth and in the model can be considered independent, i.e. separated by the correlation scale R.

b. Quasi-geostrophic model

A reduced-gravity, quasi-geostrophic model is used to generate the Eulerian fields. Using nondimensionalization

$$\psi = (T_0 H^{-1})\psi^*, \; (x, y) = R_d(x^*, y^*), \; t = f_0^{-1} t^*,$$

where ψ is the streamfunction in the horizontal (x, y) plane, T_0 the net northward transport across the domain, H the active layer depth, $R_d = \sqrt{g'H}/f_0$ the radius of deformation, and g' the reduced gravity. Dropping (*), the vorticity transport equations are:

$$\frac{\partial q}{\partial t} + aJ(\psi, \nabla^2\psi) + b\frac{\partial \psi}{\partial x} = d\nabla^4\psi, \tag{5}$$

$$\nabla^2\psi - \psi = q, \tag{6}$$

where the nondimensional parameters are defined as $a \equiv T_0/(HR_d^2 f_0)$, $b \equiv \beta R_d/f_0$ and $d \equiv \nu/(R_d^2 f_0)$, β is the meridional gradient of the Coriolis frequency, and ν is the horizontal viscosity. The prognostic equation (5) is advanced in time using a predictor-corrector type leapfrog method (Gazdag, 1976). The Jacobian operator is computed using the formulation proposed by Arakawa (1966) that conserves kinetic energy and enstrophy, while accurately maintaining the property $J(a, b) = -J(b, a)$. All other differential operators are approximated by central differences. The diagnostic equation (6) is inverted using a Fast Fourier Transform solver (Swarztrauber, 1977).

The model is configured in a regular rectangular domain. The meridional length is 3000 km, which is adequate to contain several wavelengths of characteristic instabilities of the western boundary current. The zonal length is 1000 km such that the domain is much wider than the mesoscale eddies. The domain is centered around a reference latitude of 26°. The equilibrium layer thickness ($H = 1000$ m) and the stratification ($g' = 0.01$ m s^{-1}) are such that the Rossby radius of deformation is approximately 52 km, typical of midlatitude circulation. The viscosity is taken either as $\nu = 100$ m^2 s^{-1} (in the higher energy experiment) or $\nu = 200$ m^2 s^{-1} (in the lower energy experiment), which are high

enough to ensure numerical stability at the selected grid scale, but also small enough to allow existence of turbulent eddy fields in the simulation ($a \gg d$). The horizontal grid size is 10 km, which resolves both the mesoscale eddy scale of $R_d = 52$ km, and the viscous western boundary current scale of 17 km $\leq (\nu/\beta)^{1/3} \leq 22$ km. The model time step of 900 s is much smaller than the time it takes for the fastest current to cross the grid ($\sim$9000 s). The model is forced by specifying a net transport between the eastern and western boundaries. The transport is $T_0 = 30$ Sv, which is the average transport of the Florida Current (Leaman *et al.,* 1989). Periodic boundary conditions are used in the northern and southern boundaries, and free-slip boundary conditions are applied along the eastern boundary.

The primary novel feature that differentiates this model configuration from more conventional applications of similar models (e.g., Özgökmen and Chassignet, 1998; Berloff and McWilliams, 1999; Özgökmen *et al.,* 2001; Lee, 2001) is the use of stochastic boundary conditions along the western boundary (Mariano *et al.,* 2003). Unlike a no-slip boundary condition which is based on laboratory scale fluid dynamics (Batchelor, 1967) and a free-slip condition which is based on the ad-hoc argument that unknown small-scale processes would allow western boundary currents to slip smoothly along boundaries, a stochastic boundary condition is an empirical representation of boundary vorticity based on Ocean Surface Current Radar (OSCR) measurements along the southeastern Florida coastline with spatial and temporal resolutions of $\sim$250 m and $\sim$ 15 mins, respectively (Shay *et al.,* 2002; Peters *et al.,* 2002).

More specifically, the boundary vorticity $\hat{\zeta}$ is modeled as

$$\hat{\zeta}(y, t) = g(y - ct), \tag{7}$$

where the generating function g is evaluated as the sum of the regressive oscillator and a white noise process γ which represents (simplistically) the ambient, wide-band spectral energy,

$$g'_n = a_1 g'_{n-1} + a_2 g'_{n-2} + w_n$$

$$g_n = g'_n + \gamma_n$$

where the variances of the driving noise w_n and ambient noise γ_n are both obtained empirically as $1.64 \times 10^{-10}\,\mathrm{s}^{-2}$, or $(0.2f_0)^2$. A sampling interval of $c\Delta t$ (=234 m) is used for discretization based on the northward advection speed of $c = 0.26\,\mathrm{m\,s}^{-1}$ estimated from the radar observations. For further details of the boundary condition, the reader is referred to Mariano *et al.* (2003).

It was found in Mariano *et al.* (2003) that the interior model fields with stochastic boundary conditions display episodic events characterized by energetic meanders, dipoles and eddies at a broad-band of scales, whereas conventional no-slip and free-slip boundary conditions yield quasi-periodic finite-amplitude modes with little variability. Underestimation of variability with respect to reality is usually associated with the lack of knowledge in the forcing functions, and it is a chronic problem even in realistic, state-of-the-art ocean

models (e.g., Garraffo *et al.,* 2001; McClean *et al.,* 2002). As stochastic boundary conditions impose high variability and greater range of scales on the interior Eulerian field, this QG setup is the choice to conduct our investigation for the objectives of this study.

Using an approach which is conceptually similar to the classical "twin experiments" in data assimilation studies, we consider the eddy-rich solutions of the QG model as "true" solutions, $\mathbf{u}$, while the "model" solutions, $\mathbf{u}_h$, are computed by smoothing the original QG solutions using (1) at each time step, as an approximation of the effect of removing subgrid-scale processes in models with coarse resolution and/or missing physics. Lagrangian particles are then released in both true and smoothed solutions and advected using a fourth order Runga-Kutta method and a time step equal to the model time step to minimize discretization errors due to advection technique. A bilinear interpolation is used to approximate the velocity field at the location of each Lagrangian particle. The real ocean certainly has a much higher degree of complexity than the QG solutions, which is itself limited by its finite resolution and many approximations, but on the other hand the approach has the great advantage that the truth is exactly known so that the uncertainty can be quantified using the ρ and s metrics.

c. Lagrangian stochastic model

In this section, a background on a multi-particle LSM (Piterbarg, 2001a,b) is provided. The LSM has been used in Piterbarg (2004) to obtain theoretical asymptotic results on Lagrangian predictability, which are briefly summarized in the following.

The multi-particle model (Piterbarg, 2001a,b) is based on two postulates:

(1) One particle motion is covered by the well-known Langevin equation (also denoted as "random flight" model; Thomson, 1987) characterized by two parameters for each velocity component: Lagrangian correlation time, τ, and velocity variance σ_u^2.

(2) Velocities of any two particles are correlated and this correlation is characterized by a space correlation radius, R.

By assuming homogeneity of velocity fluctuations, one immediately arrives at the stochastic differential equations describing motion of any M particles

$$d\mathbf{r}_m = (\mathbf{U}(\mathbf{r}_m) + \mathbf{v}_m)dt, \quad d\mathbf{v}_m = -(\mathbf{v}_m/\tau)dt + \sigma_{mj}d\mathbf{w}_j, \quad m = 1, \ldots, M, \tag{8}$$

where $\mathbf{U}(\mathbf{r})$ is the mean flow, $\mathbf{w}_j$, $j = 1, \ldots, M$ are independent two dimensional standard Brownian motions modeling stochastic forcing, the summation over j is meant, and for 2×2 matrices σ's satisfy

$$\sigma_{kj}\sigma_{lj}^T = \mathbf{B}((\mathbf{r}_k - \mathbf{r}_l)/R),$$

where $\mathbf{B}(\mathbf{r})$ is the space covariance tensor. In other words, for any M particles their positions and velocities form a classical multi diffusion process in $4M$ dimensions with drift explicitly expressed in terms of the mean flow and τ, and a diffusion matrix completely determined by the covariance tensor $\mathbf{B}$. Thus, particle positions and velocities in (8) are correlated through the noise terms. In addition we assume that the forcing is

isotropic and its divergence is zero, and thereby **B** is completely defined by a scalar covariance $B(r/R)$ (covariance of a streamfunction determining the forcing), where $r = |\mathbf{r}|$.

The LSM (8) has been used in Piterbarg (2004) to study the behavior of $\rho(t)$ (2) and $s(t)$ (3), as metrics of the discrepancy between two stochastic flows, one of which is the smoothed version of the other. More exactly, for a fixed value of the smoothing parameter, h, define a new forcing, $\mathbf{w}_h = K_h * \mathbf{w}$, and let $\mathbf{r}_{h,m}$ and $\mathbf{v}_{h,m}$ be the particles' positions and velocities obtained from (8) when **w** is replaced by $\mathbf{w}_h$. In this context, we call the positions $\mathbf{r}_m$ and $\mathbf{r}_{h,m}$ 'true' and 'model,' respectively, and h is the 'model' resolution limit (loosely speaking, grid size) while there is no limit for the 'true' motion resolution. Thus, $\rho(t)$ is the mean square distance between the centers of mass for the 'real' and 'model' clusters. Some of the main theoretical results obtained in Piterbarg (2004) are summarized in the following.

If $h \ll R$, then during the initial stage $\rho(t)$ grows exponentially

$$\rho(t) \approx \frac{C_0 h^2}{R} (\exp(\Lambda t) - P(t)), \tag{9}$$

where $\Lambda > 0$ is the so-called second Lyapunov moment, $P(t)$ is a second order polynomial with $P(0) = 1$, and $C_0 = C_0(\alpha)$ is a dimensionless constant depending on the nonlinearity parameter

$$\alpha = \frac{\tau \sigma_u}{R}, \tag{10}$$

shape of the spatial covariance and type of smoothing kernel. Exact expressions for Λ in terms of σ_u, τ, and R are given in Piterbarg (2001a). For large t, the following inertial regime holds true

$$\rho(t) \approx 2\sqrt{2}\sigma_u (\tau t/M)^{1/2}, \tag{11}$$

regardless of whether $h \ll R$ or not.

Asymptotic (9), which describes the case $\rho < R$, can be used to determine the predictability time T_{pred}, under the same condition $h \ll R$. By neglecting the power terms in t, one gets

$$T_{pred} \approx 2\Lambda^{-1} \ln \frac{R}{h}. \tag{12}$$

A scale for the scattering difference $s(t)$ can be estimated as

$$s(t) \approx \frac{h^2 \sigma_u}{R^2} \sqrt{\frac{(M-1)\tau t}{M}}. \tag{13}$$

Table 1. Parameters of QG and LSM experiments.

Exp	ν (m^2 s^{-1})	h (km)	M	L
QG100	100	20, 30, 50	10	500
QG200	200	20, 30, 50	10	500

Exp	σ_u (cm s^{-1})	τ (day)	R (km)	h (km)	M	L
LSM100	35	3.5	75–100	20, 30, 50	1, 10, 50	500
LSM200	21	4.5	50–75	20, 30, 50	1, 10, 50	500

Notice that, while ρ is mostly characterized by an exponential growth in time (9), s is characterized by a $t^{1/2}$ growth (13). The reason for this difference can be understood in the following way. By definition (3), s is the difference between the truth and model scatterings. It can be shown that the truth scattering also evolves in two stages, exponential

$$S(t) \sim d_0 \exp(\Lambda t), \tag{14}$$

where d_0 is the initial cluster radius, and inertial

$$S(t) \approx 2\sigma_u \sqrt{\tau t (M-1)/M}. \tag{15}$$

Of course, the same asymptotics hold true for the model scattering with replacing Λ and σ_u by the corresponding values for the model flow.

The crucial point is that the two-stage evolution (14, 15) of S and two-stage evolution (9, 11) of ρ are quite different. The exponential growth stage of the latter is much longer than that of the former, because $\rho(t)$ starts from zero while S starts from d_0. Thus, the exponential stage makes the main contribution in the overall behavior for $\rho(t)$ while evolution of S is more influenced by the inertial stage. As a conclusion, the scattering difference should be estimated primarily by using the inertial asymptotic (15) for the truth scattering and the corresponding inertial asymptotic for the model scattering. That results in the estimate (13) for s.

As a matter of curiosity, note that the truth scattering S goes to the single particle dispersion (e.g. Zambianchi and Griffa, 1994)

$$D(t) = E\{\mathbf{r}(t, \mathbf{a})^2\}^{1/2} \sim 2\sigma_u \sqrt{\tau t} \tag{16}$$

as $M \to \infty$, while for the two-particle ($M = 2$) relation, (15) turns out to be the inertial asymptotic for the relative diffusion with a factor $\frac{1}{2}$. Asymptotics (9, 11–16) were carefully validated by simulations in Piterbarg (2004).

3. Results from QG experiments

Two main experiments are performed using the QG model setup (Table 1). The experiments differ in the value of viscosity ν, while all other parameters are kept identical. Varying model viscosity is a reasonable approach to change the mean energy levels, size

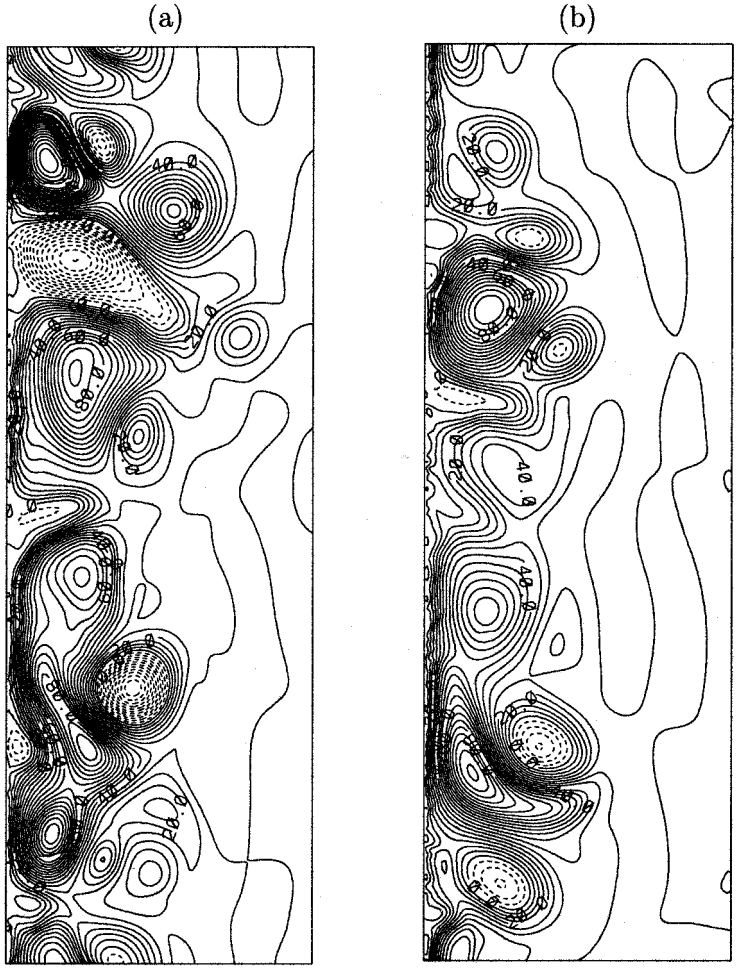

Figure 1. Snapshots of streamfunction ψ fields for (a) high energy QG100, and (b) low energy QG200 cases (ci: 5 Sv).

and strength of the coherent turbulent structures in the flow field. Both experiments reach a statistical steady state after approximately two years starting from rest, but they are integrated for five years before the release of synthetic drifters. The first experiment, denoted as QG100, has $\nu = 100\ \mathrm{m^2\ s^{-1}}$ and is characterized by a basin-averaged kinetic energy level of $EKE \approx 1000\ \mathrm{cm^2\ s^{-2}}$, whereas the second experiment, QG200, has $\nu = 200\ \mathrm{m^2\ s^{-1}}$ and has relatively lower $EKE \approx 400\ \mathrm{cm^2\ s^{-2}}$.

Snapshots of the QG100 and QG200 streamfunction fields are shown in Figure 1. Both fields are characterized by intense vortices forming along the western boundary and propagating northward. The spatial scales of the vortices are of the order of $\approx$100 km, but they differ slightly between the experiments, those in QG100 being $\approx$ 25% larger than

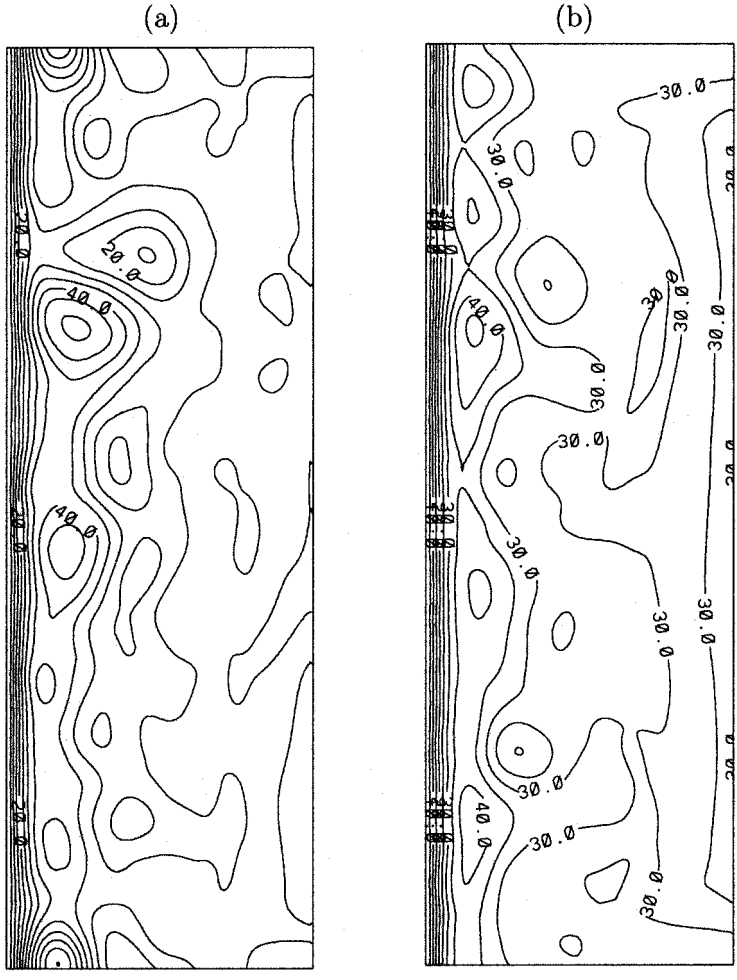

Figure 2. Time-averaged mean transport streamfunction fields $\bar{\psi}$ for (a) high energy QG100, and (b) low energy QG200 cases (ci: 5 Sv).

those in QG200. The time-averaged mean flows over the last three years of the integrations show a localized region of strong shear close to the western boundary, which is due to western intensification of the total net induced transport, and it is similar in both experiments (Fig. 2).

The flow fields shown in Figure 1 are considered as the true ocean, **u**. The Kernel K_h is applied via (1) at each time step in order to obtain a time series of the model smoothed fields $\mathbf{u}_h$. Typical results from the application of the smoothing operator to the QG100 flow field are illustrated in Figure 3 for smoothing parameters 20 km $\leq h \leq$ 600 km. As it can be seen, the smoothing operator is very effective in removing spatial structures. The effects

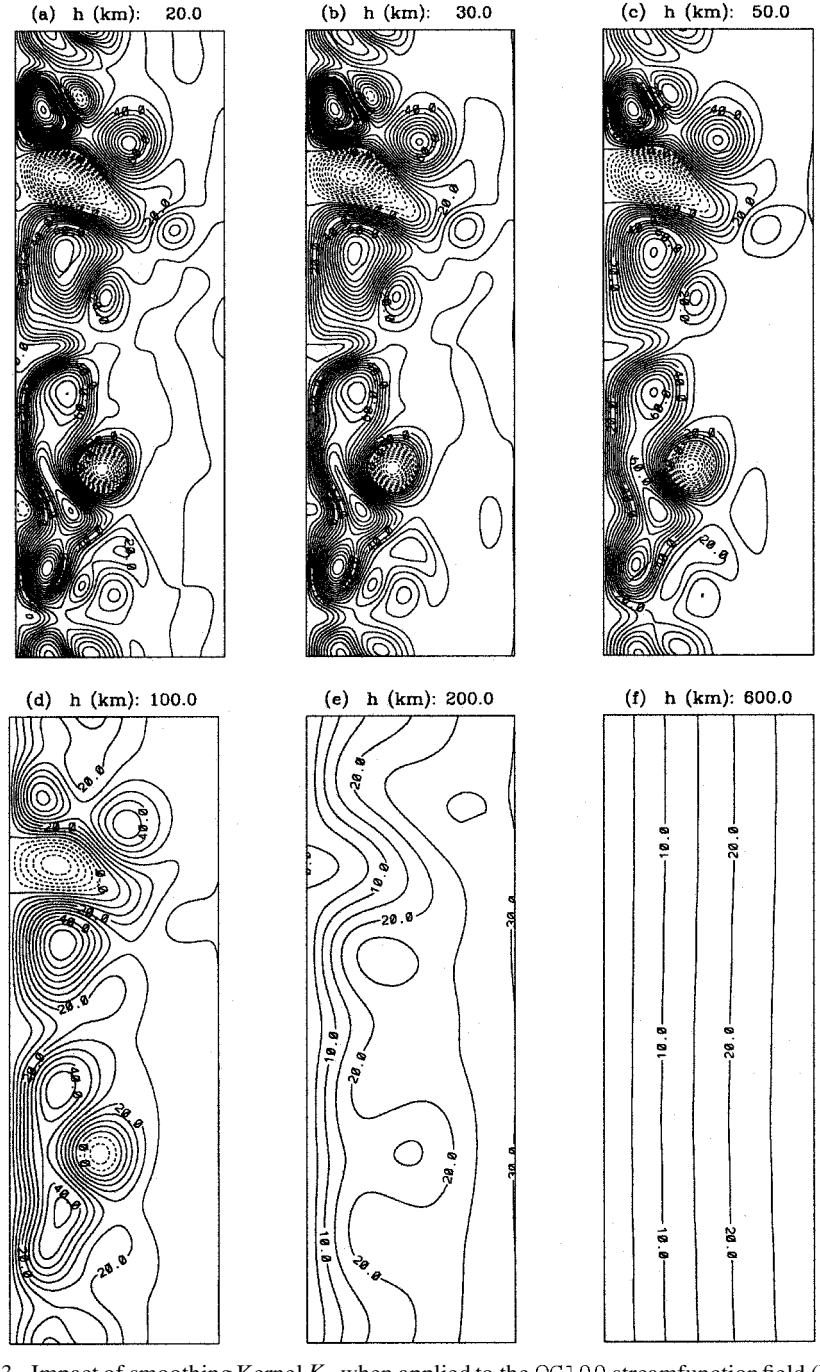

Figure 3. Impact of smoothing Kernel K_h when applied to the QG100 streamfunction field (Fig. 1a) for 20 km $\leq h \leq$ 600 km (ci: 5 Sv).

are already visually evident starting from $h = 20$–30 km, they lead to a suppression of the structures for $h > 100$ km and to a virtually flat field for $h = 600$ km. The energy progressively decreases with increasing h.

The QG experiments (Table 1) are performed considering clusters of $M = 10$ particles (the impact of varying M is considered in Section 4, in the framework of LSM). Three values of the smoothing parameter h are considered, $h = 20, 30, 50$ km. These values are higher than the "true" model resolution (10 km), and they cover a realistic range going from values significantly smaller than the eddy size ($h = 20$ km) to comparable values ($h = 50$ km). A total of $L = 500$ different cluster realizations (corresponding to a total of 5000 particles) are considered for each experiment and for each value of h, and they are used to estimate the two metrics ρ (2) and s (3). The cluster realizations are performed considering five different initial conditions for the flow field and releasing 100 clusters in each of them. The clusters are deployed in the western half of the basin (as the flow field is weak on the eastern side) but they have a random distribution otherwise. The initial radius of the clusters is limited to be smaller than the radius of deformation; the maximum radius is $R_d/6 \approx 8$ km $\ll R$. The particles are launched in the same positions and at the same time in the original $\mathbf{u}$ and in the smoothed $\mathbf{u}_h$. They are advected by the flow field generating trajectories $\mathbf{r}$ and $\mathbf{r}_h$, some examples of which are shown in Figure 4 for the original `QG100` and for $h = 50$ km. As it can be seen, the trajectories tend to be shorter as h increases, due to decreasing energy.

A quantitative analysis of the trajectories is performed by computing their autocovariance function $\mathbf{R}_\mathbf{u}$ ($R_{uu}(t_{lag}) = \langle u(t)u(t + t_{lag})\rangle$, $R_{vv}(t_{lag}) = \langle v(t)v(t + t_{lag})\rangle$), which is used also in Section 4 to set the parameters of the LSM. The effect of smoothing on the autocovariances is shown in Figure 5 for $h = 0, 20, 50$ km in the case of `QG100`. The energy decreases strongly with increasing h, dropping by approximately 50% for $h = 50$ km. The time scale τ shows a less pronounced change. A rough estimate of τ as e-folding time indicates $\tau \approx 3.5$ days for $h = 0, 20$ km, and $\tau \approx 4.5$ days for $h = 50$ km. The asymmetry in the two velocity components is mostly due to the mean flow. This has been checked by subtracting the mean flow along trajectories and computing the autocovariance of the residual fluctuation $\mathbf{u}'$, $\mathbf{R}_{\mathbf{u}'}$ (Fig. 6a). The fluctuation appears isotropic with the same time scales τ as in Figure 5a. The results for `QG200` are qualitatively similar, while the fluctuations are less energetic and have longer time scales. $\mathbf{R}_{\mathbf{u}'}$ for `QG200` is shown in Figure 6b, indicating also a more "wavy" behavior, with relatively more pronounced negative lobes with respect to that from `QG100`.

The effect of smoothing on the predictability of particle trajectories is quantified by calculating the two error metrics introduced in Section 2, i.e. the error between centers of clusters ρ(t) (2), and the difference in particle scatter around the center of mass, $s(t)$ (3). An investigation, reported in some details in the Appendix, has been carried out to check the robustness of the results in terms of convergence at increasing sampling size, L. Results obtained by varying L between 200 and 500 show that the ρ statistics are robust, with a variability of the order of 10–15% at varying L (Fig. 13a, c). The s statistics, on the other

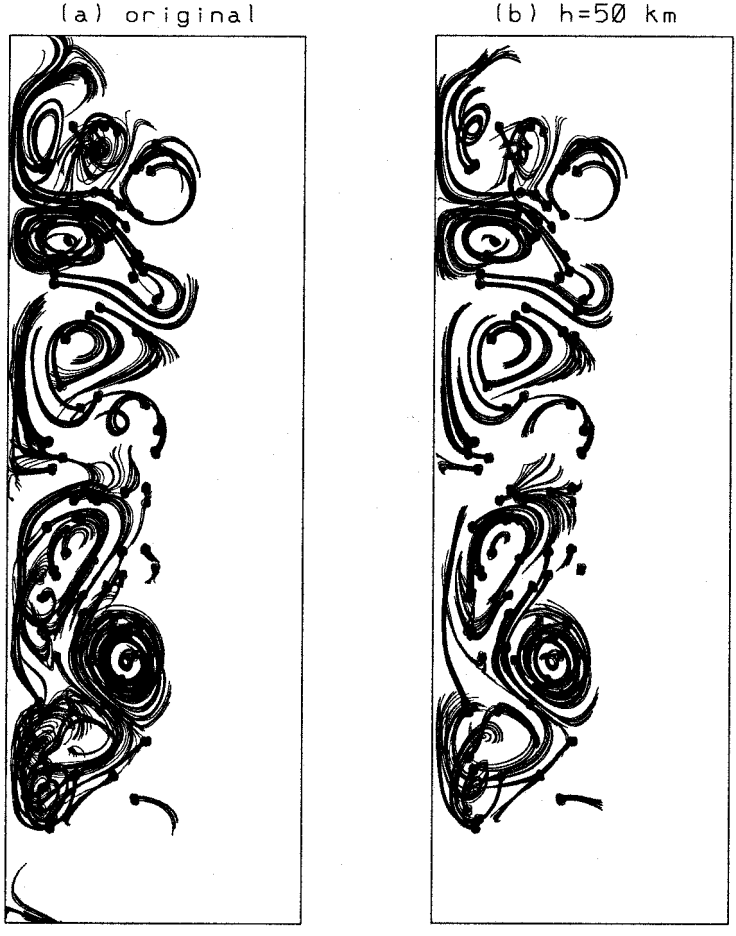

Figure 4. Samples of 15 day-long trajectories from clusters released in (a) the original flow field QG100, and (b) the field smoothed with $h = 50$ km.

hand, appear more sensitive to the sampling size and to the specific realization considered, showing a variability up to 100% at varying L (Fig. 13b, d). For this reason, the following quantitative analysis and comparison with LSM results are performed primarily in terms of $\rho(t)$, while the $s(t)$ comparison is performed in more qualitative terms.

The $\rho(t)$ results for QG100 and QG200 and for $h = 20$, 30, 50 km are shown in Figure 7 and summarized in Table 2. The results are qualitatively similar in both cases even though $\rho(t)$ reaches slightly higher values in the higher energy case QG100. The change of $\rho(t)$ as a function of h illustrates the impact of smoothing on the prediction of CM. Even for $h = 20$ km, when the impact of smoothing on the Eulerian field is minute (see Figs. 1–3), the CM in the smoothed fields significantly diverges from the true CM after only a few days. As h increases, the CM error becomes stronger and it starts occurring at

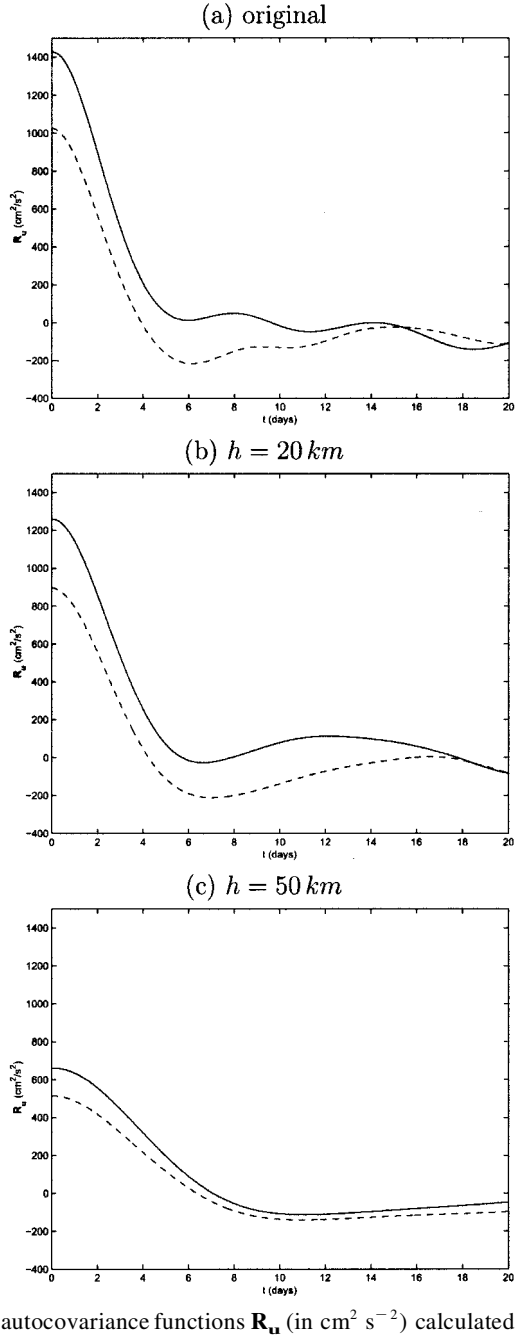

Figure 5. Lagrangian autocovariance functions $\mathbf{R_u}$ (in cm^2 s^{-2}) calculated without the removal of the mean flow as a function of t_{lag} (in days) from (a) original trajectories from QG100, (b) trajectories from smoothed field with $h = 20$ km, and (c) trajectories from smoothed field with $h = 50$ km. Solid (dashed) lines show meridional (zonal) velocity component.

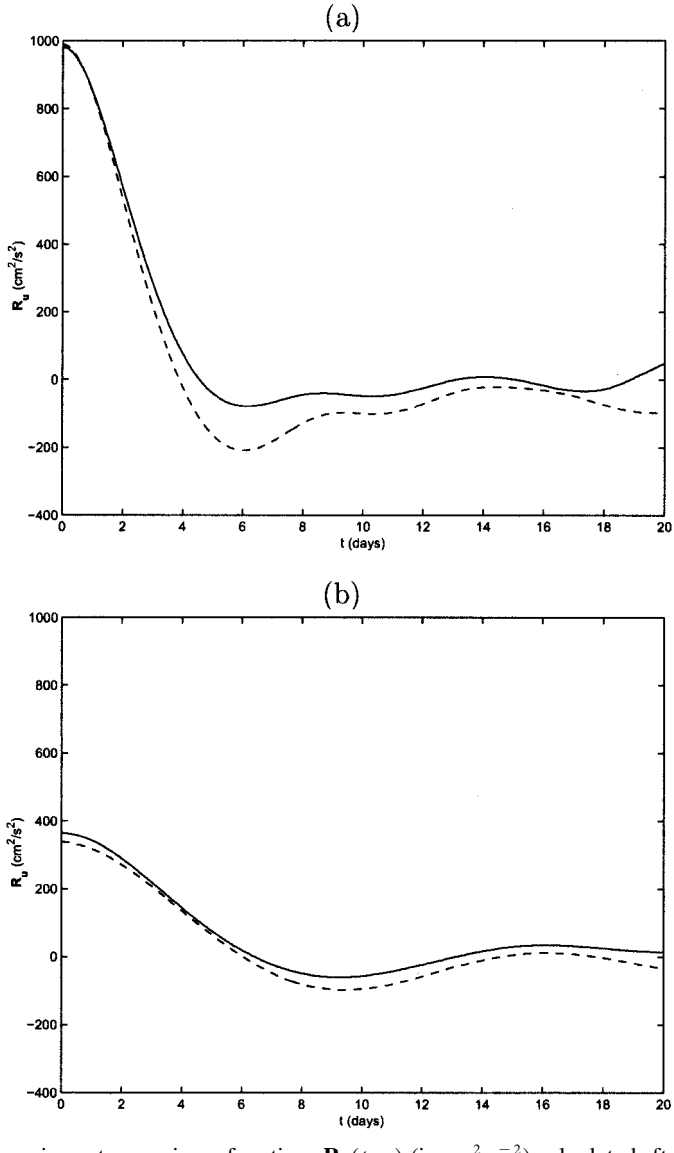

Figure 6. Lagrangian autocovariance functions $\mathbf{R_u}(t_{lag})$ (in cm^2 s^{-2}) calculated after removing the mean flow for (a) QG100, (b) QG200. Solid (dashed) lines show meridional (zonal) velocity component.

earlier times, with $\rho(t)$ increasing quickly in time. The $\rho(t)$ values at 30 days (Table 2) indicate that the error is almost doubled for $h = 50$ km with respect to $h = 20$ km. The values of T_{pred} (4), which are roughly evaluated by assuming $R \approx 75$ km for QG100 and $R \approx 50$ km for QG200, are summarized in Table 3. They indicate $T_{pred} \sim 15$ and 5 days

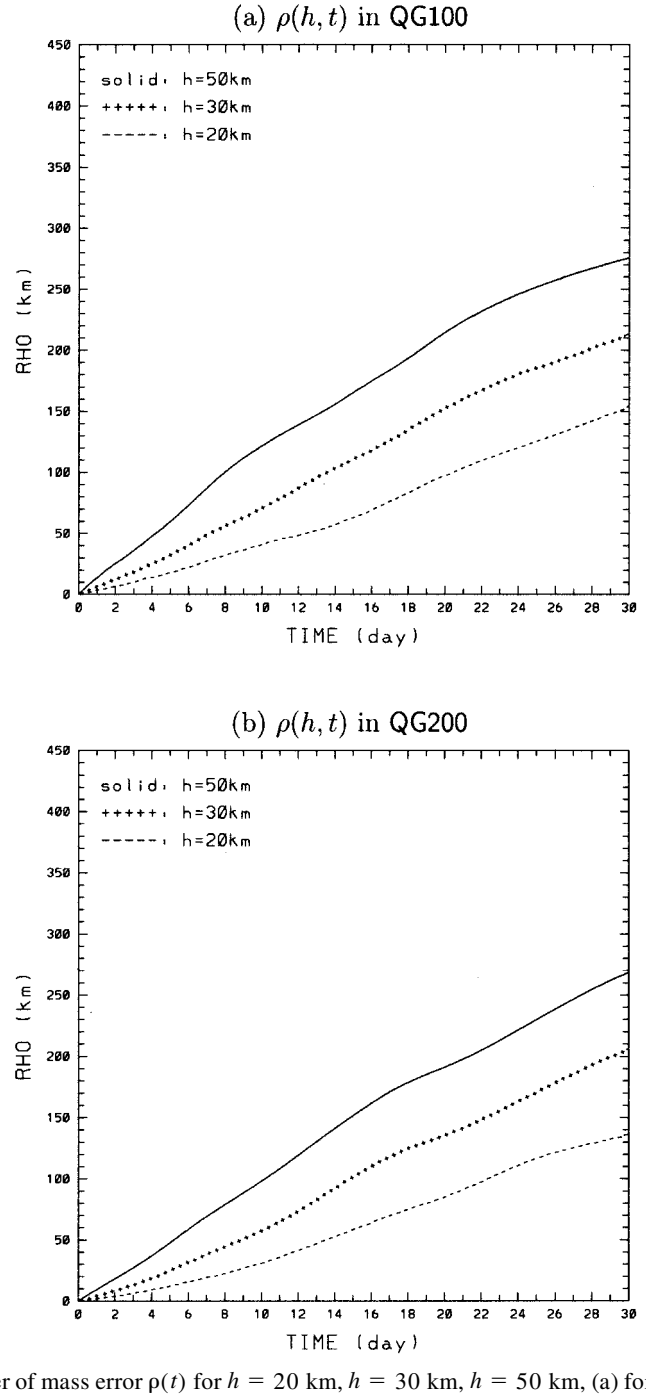

Figure 7. Center of mass error $\rho(t)$ for $h = 20$ km, $h = 30$ km, $h = 50$ km, (a) for the high energy QG experiment `QG100`, and (b) for the low energy QG experiment `QG200`.

Table 2. Comparison of center of mass error $\rho(t = 30$ days) from QG and LSM experiments.

Center of mass error (km)	QG100	LSM100 w/o **U**	LSM100 with **U**	QG200	LSM200 w/o **U**	LSM200 with **U**
$\rho(h = 20$ km; $t = 30$ days)	150	150	160	140	110	120
$\rho(h = 30$ km; $t = 30$ days)	210	240	230	200	175	190
$\rho(h = 50$ km; $t = 30$ days)	280	375	360	270	270	250

for $h = 20$ and 50 km, respectively. This suggests that particles, even at high h, are still correlated during the first few days, even though the error is quite high for any practical application.

The behavior of the scatter error $s(t)$ is shown in Figure 8 and summarized in Table 4. Despite the lack of robustness of the results (Fig. 13) which does not allow for a close quantitative assessment, the results appear significantly different from the ones for $\rho(t)$. The $s(t)$ values, in fact, are almost an order of magnitude smaller than $\rho(t)$, and they show a greater tendency to saturate in time, at least for $h = 20$–30 km. The smoothing, therefore, appears to strongly influence the behavior of the CM, while the scatter around it is only partially affected.

4. Results from LSM experiments

In this section, results from numerical experiments performed using the LSM (8) with parameters estimated from the QG experiments are presented and compared with the asymptotics presented in Section 2c.

The experiment parameters are summarized in Table 1. Two main sets of experiments are considered, denoted as LSM100 and LSM200, with parameter values for τ, σ_u^2, R estimated from the QG100 and QG200 results, respectively. As for the QG experiments, three values of $h = 20$, 30, 50 km are considered, and $L = 500$ cluster realizations are used to compute statistics. Three cluster sizes of $M = 1$, 10, 50 particles are considered, in order to explore the sensitivity of the results to the number of particles in the cluster. The initial cluster radius is given by $d_0 = 8$ km. The methodology is the same as for the QG

Table 3. Comparison of predictability time T_{pred} from QG, LSM experiments, and theory.

Predictability time (day)	QG100	LSM100 with **U**	Eq. (12)	QG200	LSM200 with **U**	Eq. (12)
$T_{pred}(h = 20$ km)	17	19	18.9	14	17	14.1
$T_{pred}(h = 30$ km)	10	15	13	10	12	7.9
$T_{pred}(h = 50$ km)	4	9	5.8	6	7	0

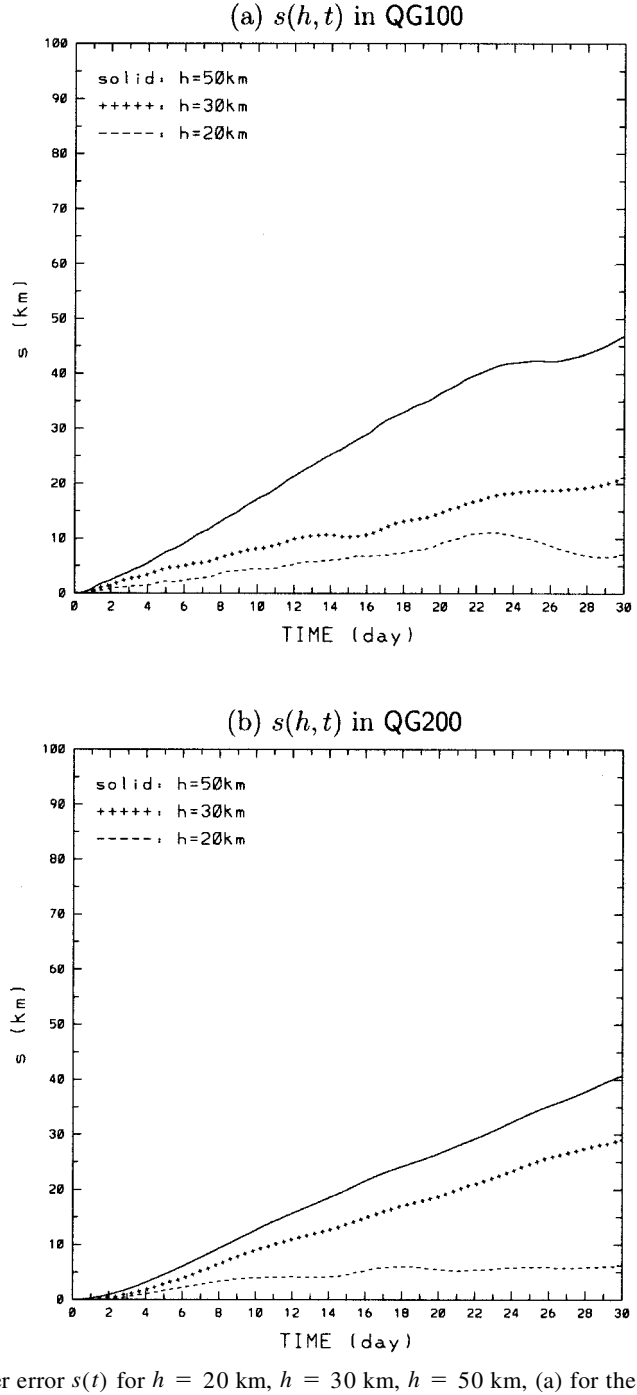

Figure 8. Scatter error $s(t)$ for $h = 20$ km, $h = 30$ km, $h = 50$ km, (a) for the high energy QG experiment QG100, and (b) for the low energy QG experiment QG200.

Table 4. Comparison of scatter error s (t = 30 days) from QG and LSM experiments.

Scatter error (km)	QG100	LSM100 with **U**	QG200	LSM200 with **U**
s (h = 20 km; t = 30 days)	6	28	6	30
s (h = 30 km; t = 30 days)	20	46	30	50
s (h = 50 km; t = 30 days)	45	100	40	100

experiments in Section 3. The true trajectories are computed integrating (8), while the model trajectories are obtained applying the smoothing kernel K_h on (8). Details on the numerical methodology are given in Piterbarg (2004).

The issue of how to compute estimates of τ, σ_u^2, R is very important for practical applications in the real ocean, since the applicability of LSM for prediction studies depends on the significance of the parameter estimates. In order to address this point, we have used simple estimation methods based on information that is available in the real ocean.

The parameters σ_u^2 and τ are evaluated from the Lagrangian autocovariance $R_u(t_{lag})$, as $R_u(0)$ and as e-folding time scale, respectively, using a simple method commonly used in oceanography (e.g. Bauer *et al.,* 2002). Historical Lagrangian data are now available in most of the world oceans, so that a similar approach can be easily persued in real ocean applications. Notice that the τ estimates are expected to be approximate in cases when the autocovariance is significantly different from exponential (e.g., Griffa, 1996). In real ocean data, exponential R_u is relatively common in regions away from high shears and coherent structures (Falco *et al.,* 2000; Bauer *et al.,* 2002), while in the presence of coherent structures, significant deviations from exponential can occur (Veneziani *et al.,* 2004). In our QG experiments, the flow is strongly dominated by coherent vortices and $R_u(\tau)$ (Figs. 5–6) is significantly different from exponential. As a consequence, this provides a good test, probably similar to the most "difficult" regions in the ocean. Parameter values estimated from Figure 6 suggest (Table 1): σ_u = 35 cm s^{-1} ($\approx$30 km day^{-1}), τ = 3.5 days for LSM100, and σ_u = 21 cm s^{-1} ($\approx$18 km day^{-1}), τ = 4.5 days for LSM200.

The parameter R is less straightforward to evaluate, especially from Lagrangian data alone. A rough indication can be obtained considering the ratio σ_u/τ, which suggest $R \approx$ 100–80 km for the QG experiments, in keeping with the size of the structures shown in Figure 1. Alternatively, as a first guess one can consider the value of the Rossby radius of deformation, which for the QG runs is of the order of 50 km. As can be seen, there is uncertainty in the R value. A similar situation can be envisioned for real ocean data, where satellite data might provide qualitative information on structure size and stratification data can provide Rossby radius estimate. Given the uncertainty of R, we have decided to test the sensitivity of the results considering different R values, ranging between 50 and 100 km (Table 1). In the following, we consider first the cases with R = 75 km for LSM100 and with R = 50 km for LSM200 (smaller values are chosen for LSM200 consistent with Fig. 1). Sensitivity to changes in R values is then discussed.

The shape of the $B(x)$ space autocovariance in (8) is taken as

$$B(x) \sim \exp(-x^2)$$

in all the experiments presented here. Sensitivity tests have been performed using different space covariances, e.g. with slower decay $B(x) \sim (1 + x^2)^{-3/2}$. It is found that as long as the values of σ_u, τ, and R hold, the quantitative behavior of ρ and s does not depend on $B(x)$.

For each set of experiments, two different choices of mean flow ($\mathbf{U}$) have been considered. The first choice is simply zero mean flow, $\mathbf{U}(\mathbf{r}) = 0$, while the second one qualitatively mimics the QG mean flow in Figure 2 and it is given by a meridional linear shear flow (Fig. 9) with velocity on the western boundary of $U \approx 35$ cm s^{-1} ($\approx$30 km day^{-1}), and $U = 0$ on the eastern boundary. The difference between the two choices will allow us to quantify the impact of the mean flow on the predictability of trajectories.

As for the case of QG, a sensitivity study has been carried out to check the robustness and convergence of the results in terms of ρ and s (see Appendix). The results are qualitatively similar to those obtained for the QG experiments; ρ appears stable at changing sampling size (L varying from 200 to 500) (Fig. 14a), and its standard deviation is of the order of 10–15% for $L = 500$ (Fig. 14c), whereas $s(t)$ is more unstable, in terms of both convergence at varying L and standard deviation error (Fig. 14b, d).

Results for $\rho(t)$ over 30 days are shown in Figure 10 and summarized in Table 2 for the two experiments `LSM100` and `LSM200`, considering both cases with no mean flow and with shear flow, and for cluster size of $M = 10$. The behavior of $\rho(t)$ with no mean flow is qualitatively similar for both `LSM100` and `LSM200` (Fig. 10a, b, respectively), and it shows the typical two-stage evolution depicted by (9) and (11), at least for $h = 20$, 30 km. The exponential stage lasts longer for smaller h, as expected. At a more quantitative level, `LSM100` shows a more pronounced growth, especially for higher h, reaching significantly higher values for $h = 50$ km at $t = 30$ days (see Table 2). Explanation is simple: the crucial role in the growth rate belongs to the velocity variance, which is higher in the high energy case. The presence of mean shear does not seem to impact significantly the results (Fig. 10c, d). The terminal values of $\rho(t)$ (Table 2) essentially coincide within the expected sampling error. Hence, the shear flow adds little uncertainty to the Lagrangian trajectories, while the velocity fluctuations appear to be primarily responsible for this uncertainty. Other mean flow shapes have also been tested, such as a periodic sinusoidal flow, and similar results are obtained.

The theoretical estimate of T_{pred} (12) has been computed and compared with experimental values (Table 3). T_{pred} (12) depends on the parameter Λ, which for the values of σ_u, τ and R in Table 1 is $\Lambda = 0.14$ and 0.13 day^{-1} for `LSM100` and `LSM200` respectively. The agreement between the T_{pred} estimates from (12) and from experiment results (Fig. 10) is quantitatively good for both `LSM100` and `LSM200` and for $h = 20$, 30 km. For $h = 50$ km, on the other hand, the theoretical values appear significantly different from the experimental one. This is not surprising given that (12) is valid for $h \ll R$ which is violated for $h = 50$ km.

The sensitivity of the results to the parameter R is illustrated in Table 5. Terminal values

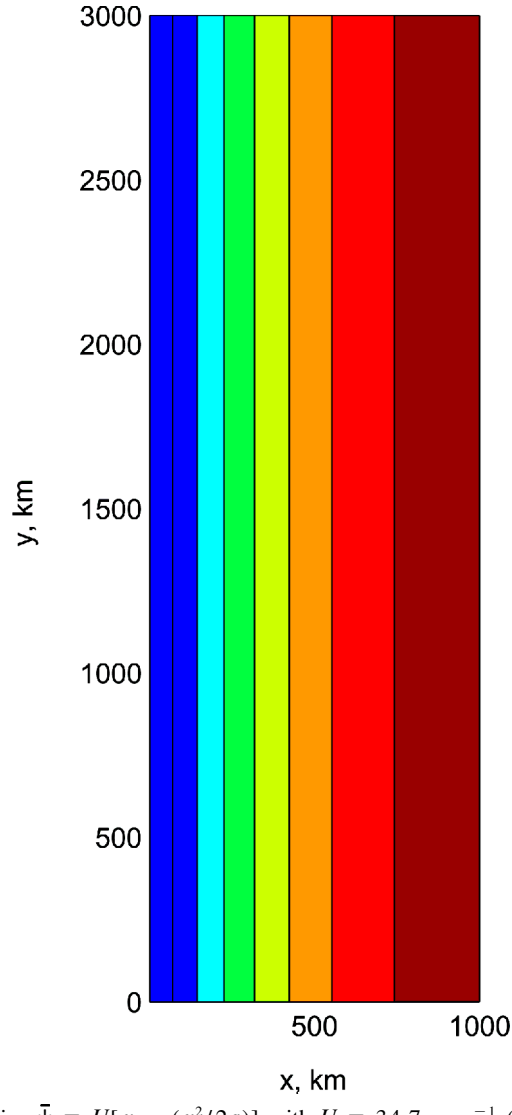

Figure 9. Streamfunction $\bar{\psi} = U[x - (x^2/2a)]$ with $U = 34.7$ cm s^{-1} (30 km day^{-1}) and $a = 1000$ km representing linear shear flow.

of $\rho(t)$ are shown for $R = 75$, 100 km for LSM100 and $R = 50$, 75 km for LSM200, considering $h = 20$, 50 km. For both experiments, the ρ values appear unaffected by the change in R for $h = 20$ km, while the results change significantly for $h = 50$ km. ρ is higher at smaller R, by a factor of 1.5 or more. This is probably due to the fact that $h = 50$ km is approximately of the order of R, so that even a relatively small change in R can determine a change in the particle regime and, therefore, a definite difference in ρ.

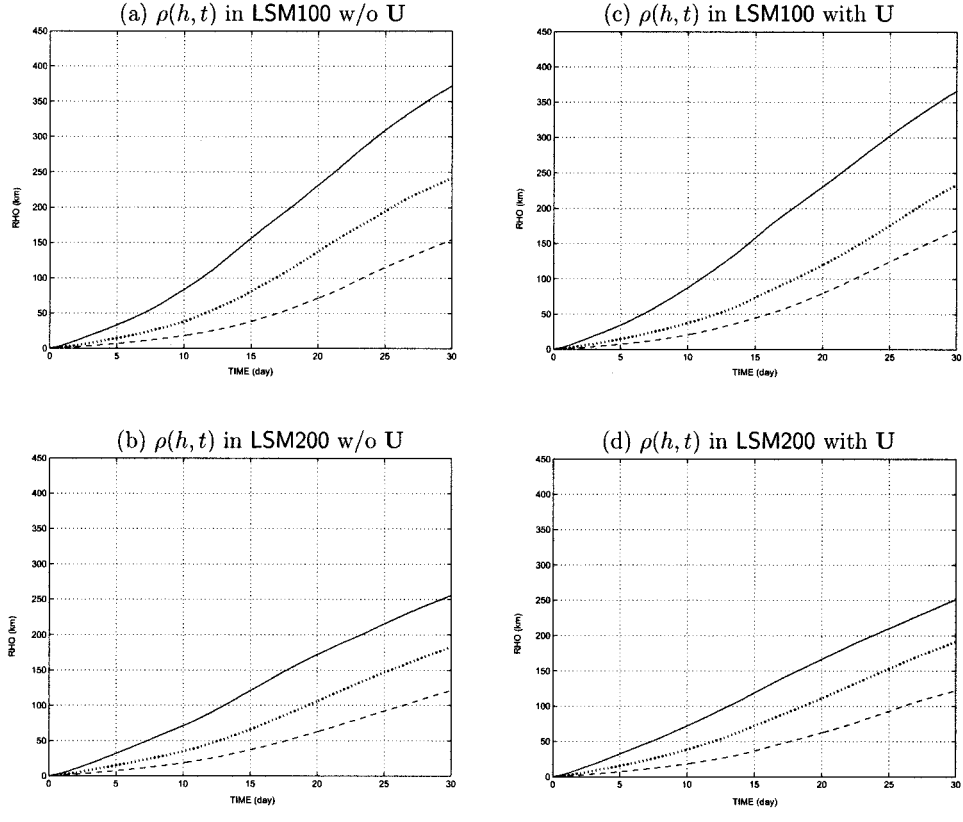

Figure 10. Center of mass error $\rho(t)$ for $h = 20$ km (dashed), $h = 30$ km (dotted), $h = 50$ km (solid), (a) for the high energy LSM experiment LSM100 without mean flow, (b) for the low energy LSM experiment LSM200 without mean flow, (c) for the high energy LSM experiment LSM100 with shear flow, (d) for the low energy LSM experiment LSM200 with shear flow.

Regarding the sensitivity to the other parameters, preliminary investigations changing τ by $\approx$50% suggest that the results are robust and more insensitive than with respect to R.

Finally, the sensitivity to cluster size M has been investigated, considering $\rho(t)$ for $M = 1, 50$. The results are shown in Figure 11 and Table 6, for both experiments in the shear flow case. The results can be directly compared with those in Figure 10 for $M = 10$. $\rho(t)$

Table 5. Sensitivity of center of mass error to selection of R in LSM experiments (with **U**)

Center of mass error (km)	LSM100 $R = 75$ km	LSM100 $R = 100$ km	LSM200 $R = 50$ km	LSM200 $R = 75$ km
ρ ($h = 20$ km; $t = 30$ days)	150	140	110	110
ρ ($h = 50$ km; $t = 30$ days)	375	200	270	180

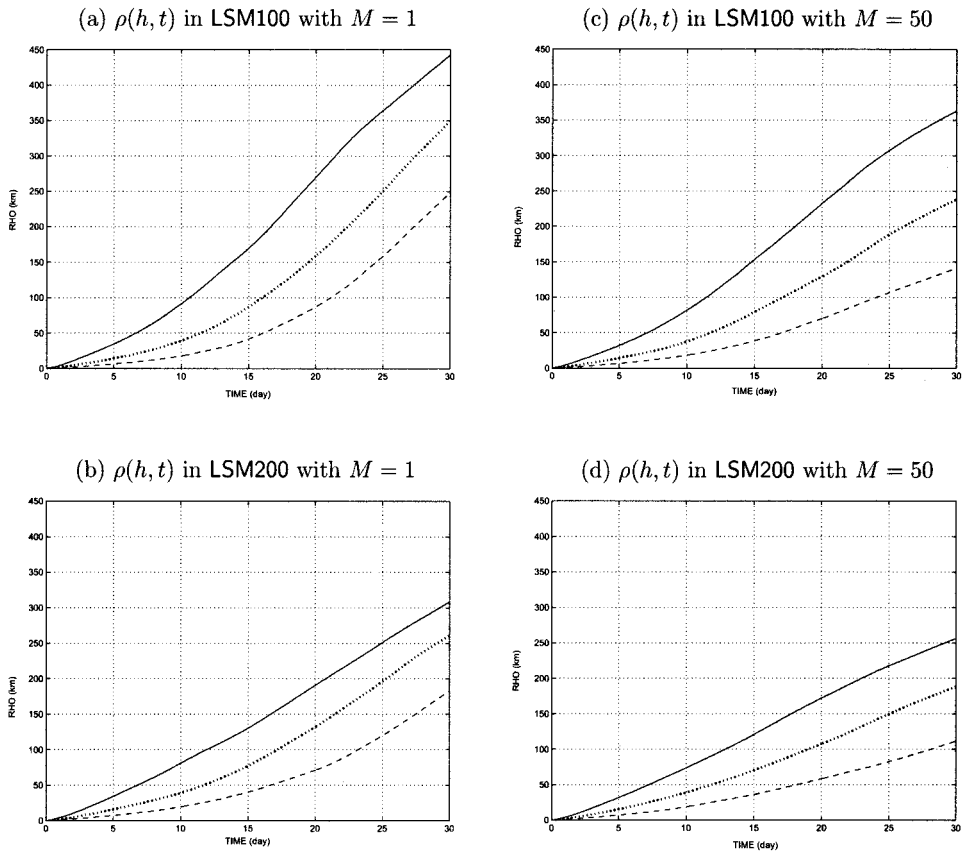

Figure 11. Sensitivity of center of mass error $\rho(t)$ to cluster size M. (a) LSM100 with $M = 1$, (b) LSM200 with $M = 1$, (c) LSM100 with $M = 50$, (d) LSM200 with $M = 50$.

appears significantly higher for $M = 1$ than for $M = 10, 50$, in qualitative agreement with (11). Intuitively, this indicates that the prediction error for a single particle is much greater than for the center of mass of a cluster. The results do not appear to change significantly between $M = 10$ and $M = 50$ over 30 days, even though they do differ for a longer run of 100 days (not shown), with $M = 10$ being significantly higher (a factor 1.5) than $M = 50$. The results then appear in qualitative agreement with (11), even though they do not agree in a quantitative sense. This is probably due to the asymptotic nature of the estimate (11).

The scattering difference, $s(t)$, is shown in Figure 12 and summarized in Table 4 for both experiments LSM100 and LSM200, in the case of shear flow. As for QG, the results appear significantly different from the ρ results (Fig. 10) in terms of values and, to some extent, in terms of shape. The $s(t)$ values are significantly smaller than $\rho(t)$, as shown also by the terminal values (Table 4), which are approximately $\frac{1}{3}$ to $\frac{1}{5}$ of the terminal $\rho(t)$ (Table 2).

Table 6. Sensitivity of center of mass error to selection of M, the number of drifters in each cluster, in LSM experiments (with **U**).

Center of mass error (km)	LSM100 $M = 1$	LSM100 $M = 10$	LSM100 $M = 50$	LSM200 $M = 1$	LSM200 $M = 10$	LSM200 $M = 50$
ρ ($h = 20$ km; $t = 30$ days)	250	160	130	180	120	120
ρ ($h = 30$ km; $t = 30$ days)	350	230	230	260	190	180
ρ ($h = 50$ km; $t = 30$ days)	440	360	370	310	250	250

This indicates that one can predict the scattering radius of the cluster with higher accuracy than the center of mass of the cluster.

In summary, the numerical experiments with LSM are in good qualitative agreement with the theoretical asymptotics derived from the same model. They also indicate that the fluctuation is the main source of uncertainty.

5. Comparison of LSM and QG experiments

In this section, results from QG experiments (Figs. 7–8) are compared with those from LSM, focusing for simplicity on the case with shear flow (Figs. 10c, d–12).

We focus mostly on the $\rho(t)$ metric which is more robust and has lower errors, therefore, allowing for a quantitative comparison. The qualitative behavior of the $\rho(t)$ curves is similar (Figs. 7, 10), especially for the lower energy experiments QG200 and LSM200. For the higher energy case, the initial exponential behavior which is present in LSM100 is not quite distinguishable in QG100 for $h = 50$ km. Terminal values of $\rho(t)$ for the QG and LSM experiments (Table 2) show quantitatively similar results (within the estimated sampling errors of the order of 10–15%) for $h = 20$–30 km. For $h = 50$ km, the behavior is different for the higher and lower energy experiments. For lower energy, QG200 and LSM200 show compatible results, while for higher energy QG100 and LSM100 are significantly different with LSM100 overestimating the QG ρ value of ≈30%.

A further comparison has been performed considering the values of T_{pred} (4) (Table 3). Again, the agreement appears quantitatively good overall, except for the high energy case with $h = 50$ km. In this case, LSM100 overestimates T_{pred} by almost 50% with respect to QG100.

Regarding the behavior of $s(t)$ (Figs. 8–12), the qualitative behavior of LSM and QG is similar and in both cases the values appear significantly lower than for $\rho(t)$. At a quantitative level, though, the LSM results overestimate the QG results (Table 4) for both experiments by a factor of 2 for $h = 30$, 50 km. For $h = 20$ km the discrepancy appears even higher even though it is difficult to quantify it given the high dependency of the estimates on sampling (Figs. 13–14).

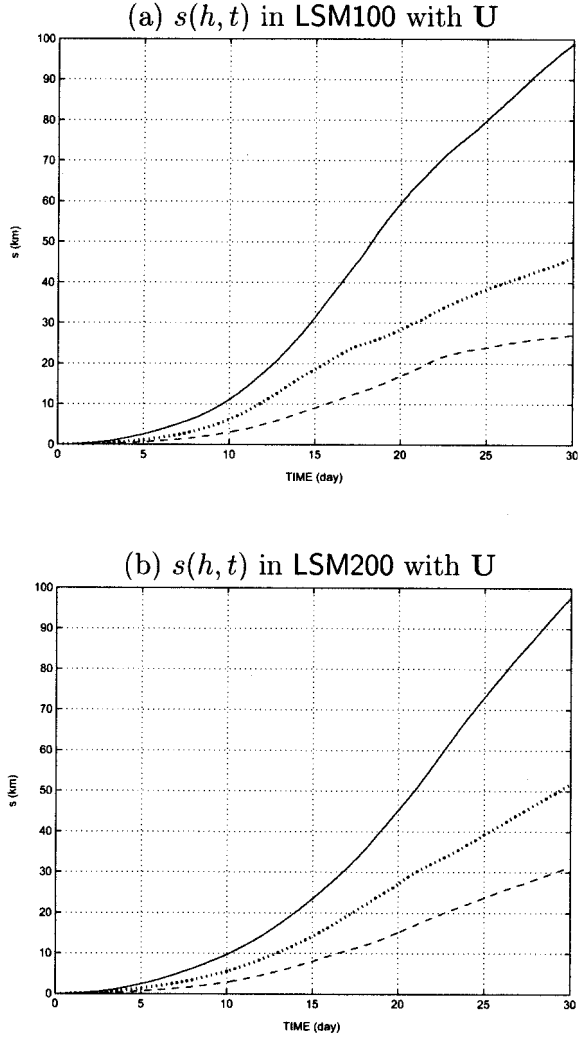

Figure 12. Scatter error $s(t)$ for $h = 20$ km (dashed), $h = 30$ km (dotted), $h = 50$ km (solid), (a) for the high energy LSM experiment `LSM100` with shear flow, and (b) for the low energy LSM experiment `LSM200` with shear flow.

In summary, the comparison shows a very good agreement between QG and LSM results for $\rho(t)$ indicating that LSM results can indeed be used to provide guidance on predictability estimates of CM for complex flow fields. The agreement is not only qualitative but also quantitative, except for the case of high energy and $h = 50$ km, which can be due to various reasons. The high energy case might be more difficult to simulate in terms of LSM, as also suggested by the difference in the curve shapes. The initial time exponential behavior predicted by the theory might not be present in the highly turbulent

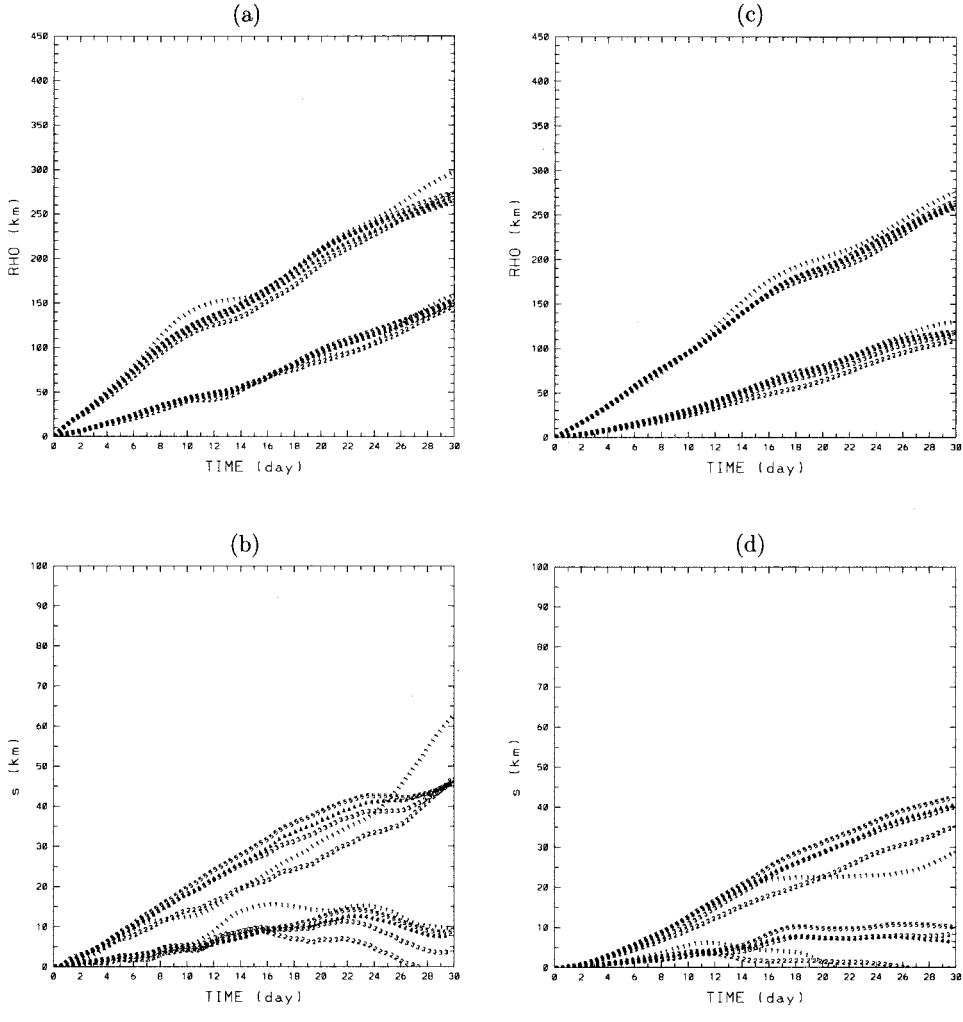

Figure 13. Sensitivity of QG results to sampling size, for number of clusters $L = 100$, 200, 300, 400, 500 (lines marked with 1, 2, 3, 4 and 5, respectively). (a) Center of mass error $\rho(t)$ and (b) scatter error $s(t)$ in QG100 for $h = 20$ km and $h = 50$ km. (c) $\rho(t)$ and (d) scatter error $s(t)$ in QG200 for $h = 20$ km and $h = 50$ km.

case. Also, it should be noted that for $h = 50$ km, the smoothing is performed at a scale close to the space scale R, so that the results might be more sensitive to the R parameter values, as suggested also by the sensitivity results in Section 4.

Regarding $s(t)$, the LSM results provide some qualitative information but they tend to overestimate significantly the s values, at least for the considered set of parameters. This might be due to the fact that the spreading of QG trajectories is constrained by the presence of vortices, which are not present in the LSM parameterization. In LSM, once the particles

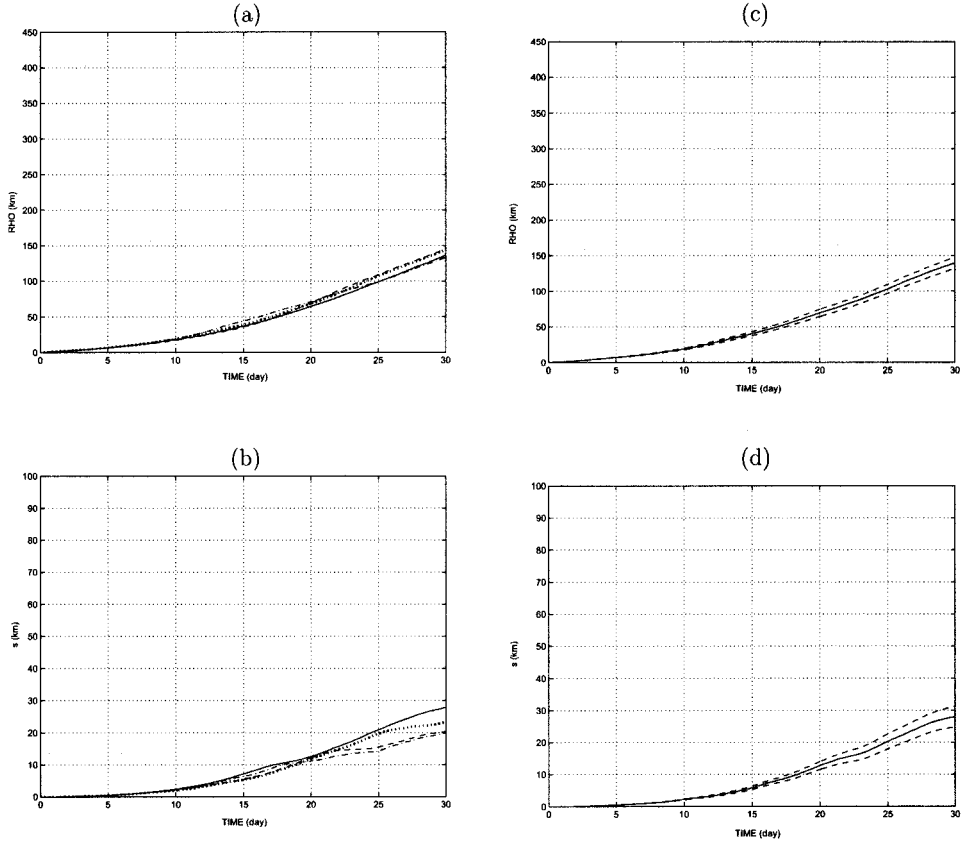

Figure 14. Sensitivity of LSM results to sampling size in `LSM100` for $h = 20$ km. (a) Center of mass error $\rho(t)$ for $L = 200$ (dash dot), 300 (dash), 400 (dot), 500 (solid), (b) scatter error $s(t)$ for $L = 100$, 200, 300, 400, 500, (c) confidence intervals of $\rho(t)$ for $L = 500$, (d) confidence intervals of $s(t)$ for $L = 500$.

have reached a distance R, they move independently, while QG particles tend to be effectively trapped by the vortices.

We underscore an important role of the spatial correlation in particle motion in the LSM used. In many previous studies, this correlation was ignored for a good reason: the goal was just to reconstruct the mean tracer field (e.g. Berloff and McWilliams, 2002). For such a goal the space correlation in Lagrangian motion does not matter. This is not the case in predictability problems like the one discussed here. On one side, any attempt to use uncorrelated particles for comparing LSM and QG performance leads to unrealistically large values of ρ compared with QG and common sense. On the other side, overestimated values of R result in large discrepancy between ρ in QG and LSM due to very small LSM values of ρ.

6. Summary and concluding remarks

This study is motivated by the effects of the discrepancy in scales between spatial discretization in ocean circulation models and available velocity data sets, and the forces acting on Lagrangian particles/drifters, typically 1 m, and their impact on the accuracy of the Lagrangian motion simulated by these modeled Eulerian fields. In numerical dynamical models, the parameterization of turbulent scales smaller than the model grid scale is typically modeled by Laplacian or other parameterization schemes, which are essentially smoothing operators. Even though small scales do not carry much energy, they may have high vorticity, due to high shears induced by wave reflection as revealed by aerial photographs and *in situ* moored current data near the northern California coast (Sheres *et al.,* 1985), or submesoscale eddies induced by episodic upwelling events revealed by recent high resolution surface radar observations off the Florida coast (Shay *et al.,* 2002; Peters *et al.,* 2002), and hence can exert significant impact on the accuracy of individual Lagrangian trajectories.

In this study, the impact of uncertainty due to smoothing errors in the Eulerian field on the prediction of particle trajectories is investigated. This is a first step toward the more complex problem of actually considering the effects of lowering resolution in a dynamical model. When changing resolution, the Reynolds number of the solution and, therefore, its degree of nonlinearity changes and other effects rather than simple smoothing can occur (e.g. Siegel *et al.,* 2001). The present results, then, should be considered as a first approximation to the problem of changing model resolution. On the other hand, the results are directly applicable to the case of an Eulerian velocity field computed using a "statistical model" (rather than a dynamical model), i.e. by smoothing and interpolating a set of discrete data point. This is a practical and important problem, for transport studies based on velocity fields obtained from HF radar measurements or coastal survey data (e.g. Paris *et al.,* 2002).

In the first part of the study, prediction uncertainty is quantified using numerical results from a QG model with stochastic boundary conditions. Clusters of particles are launched in the QG solution, considered as the "true" ocean, and compared with particles launched in smoothed versions of the solution, considered as the "model" ocean. This approach is conceptually similar to the so-called twin experiments in data assimilation, and it has the advantage that the truth is exactly known so that the uncertainty can be quantified. Two sets of experiments, characterized by different viscosity and different energy levels, are considered. For each experiment, various levels of smoothing are considered, characterized by the smoothing parameter h.

Sensitivity tests have been performed to test the robustness of the two metrics used to quantify the Lagrangian prediction error, i.e. the CM error ρ and the scatter error s. ρ is found to be robust and relatively insensitive to sample size, while s appears more sensitive to sample size and specific realizations. As a consequence, most of the quantitative analyses and comparisons are carried out considering the ρ metric, while $s(t)$ is used to provide qualitative information.

The QG experiment results show that the prediction of the cluster CM is highly

dependent on the Eulerian smoothing. Even when the smoothing scale is relatively small so that only minute changes in the flow field are induced, (i.e. for $h = 20$ km), the CM error $\rho(t)$ significantly increases from the first few days. At increasing h, $\rho(t)$ increases faster and starts at earlier times, reaching values of the order of 300 km after 30 days for $h = 50$ km. Similar results are obtained for both experiments, despite the different energy level.

Regarding the spreading error $s(t)$, a significant difference from the $\rho(t)$ behavior can be noticed. Values of $s(t)$ remain considerably lower (almost an order of magnitude) than the $\rho(t)$ values, reaching maximum values of $\approx$40 km for $h = 50$ km. The smoothing, therefore, appears to have a lower influence on the scatter around the CM than the behavior of the CM itself.

The QG results are then compared with results obtained using an LSM (Piterbarg, 2001a, b), in which the motion of multiple particles in a turbulent flow is parameterized. The parameters in the LSM, i.e. the velocity standard deviation σ_u, the Lagrangian time scale τ and the space scale R, are set considering the statistics of the QG solutions. The issue of how to estimate these parameters is crucial for practical applications in the real ocean. Here we have deliberately used very simple estimation methods based on information that can be realistically available in the real ocean. σ_u and τ are directly computed from Lagrangian autocovariances that can be computed from historical Lagrangian data sets now available in most of the world oceans. The R parameter is less straightforward to estimate, and various qualitative methods have been considered; for instance, based on Lagrangian information as well as on Rossby radius estimates or on visual information (e.g. from satellite data). These different methods provide results that, even though qualitatively similar, can easily differ by 25–50%. Given the uncertainty on the R estimate, a sensitivity study to R values has been performed.

Numerical experiments using the LSM have been performed and compared with the theoretical asymptotic results of Piterbarg (2004) and used to investigate sensitivity to various parameters, such as eddy size R, number of particles in the cluster M and mean flow $\mathbf{U}$ effects. It is found that $\rho(t)$ decreases with M, as suggested by the theory, even though the actual decrease rate is only in qualitative agreement with the asymptotic results. Results appear quite insensitive to the value of R when the smoothing parameter is relatively small, $h = 20$ km. For $h = 50$ km, a higher sensitivity is observed, with ρ varying by a factor of 1.5 or more. This is probably because $h = 50$ km is of the same order as R itself, so that even small changes can determine significant changes in the particle regime. A similar sensitivity at $h = 50$ km, is observed when comparing asymptotic and numerical results of the predictability time scale T_{pred}. Finally, the mean flow U is found to be only marginally responsible for Lagrangian uncertainty, which appears mostly due to the velocity fluctuations.

LSM results are then compared with QG results. Comparison of $\rho(t)$ results shows a very good agreement, not only qualitatively but also quantitatively. Results agree, within the estimated sampling error, except for the case of high energy and $h = 50$ km, in which the LSM results overestimate the QG results by $\approx$30%. Similar results hold also for T_{pred}. In summary, then, LSM results appear appropriate to provide guidance on predictability

estimates of CM in complex flow fields, especially when the smoothing scales are smaller than the eddy size.

Regarding $s(t)$, LSM results show significantly lower values than for $\rho(t)$, as for QG. At the quantitative level, though, LSM results appear significantly different from QG results, overestimating the s values. A possible reason for this difference is that the cluster spreading in QG is likely to be constrained by the presence of vortices, which are not present in the LSM.

The existence of quantitative differences between QG and LSM results is not surprising, given that the considered LSM is simple and it cannot reproduce completely the complex QG dynamics. The LSM, in fact, can be considered as a first step toward a more general hierarchy of stochastic processes (e.g. Griffa, 1996; Berloff and McWilliams, 2002). It corresponds to a one-dimensional first order process in time (Priestley, 1981), and for the case of a single particle, is characterized by a simple exponential autocovariance. Analysis of ocean Lagrangian data (e.g. Falco *et al.,* 2000; Bauer *et al.,* 2002) suggests that the exponential autocovariance is often adequate to describe surface mesoscale turbulence in the absence of persistent coherent features, but it is inadequate in describing regions dominated by strong waves or vortices. The QG solutions considered here are characterized by persistent vortices, and the corresponding autocovariances are quantitatively different from exponential, with negative lobes and a "wavy" behavior, that are more evident in the less energetic case. In a sense, the QG case considered here might be even more challenging for the LSM model than the real ocean, at least for what concerns surface flows away from strong shears.

In summary, the present results indicate that the use of LSM is a promising avenue to provide guidelines for predictability estimates in realistic ocean situations, at least for what concerns the position of cluster CM. This is quite advantageous since the LSM is simple to implement and it depends only on a few observable parameters that can be estimated from data. Especially for smoothing scale h smaller than the eddy size, LSM results appear precise and robust to parameter changes. For cluster spreading, instead, LSM results do not appear to provide reliable quantitative estimates. They appear to provide qualitative indications on an order of magnitude of expected $s(t)$.

In the future, LSM results could be further improved by considering more complex LSM, able to include more aspects of turbulent behavior, such as trapping due to vortices, Rossby wave effects and wall reflections, which are present in the QG runs. Some of these aspects have been recently considered in the framework of single-particle LSM (Berloff and McWilliams, 2002). In particular, it has been shown (Reynolds, 2002; Veneziani *et al.,* 2004) that the effects of coherent vortices can be successfully modeled by introducing a spin term in first order models, which appropriately couples the velocity components. These effects can, in principle, also be included in the multi-particle LSM model considered here.

Acknowledgments. The authors greatly appreciate the support of the Office of Naval Research under grants N00014-99-1-0049, N00014-03-1-0285 (Griffa and Özgökmen) and N00014-99-1-0042 (Piterbarg).

APPENDIX

The robustness of the statistics $\rho(t)$ (2) and $s(t)$ (3) has been tested in the framework of QG and LSM results.

For QG, the convergence of the results at increasing sample size has been studied by considering an increasing number of realizations: $L = 200, 300, 400, 500$. The results are shown in Figure 13 for the two experiments `QG100` and `QG200` and for two values of h, $h = 20, 50$ km.

The ρ results (Fig. 13a, c) show that the statistics are relatively insensitive to L, with a variability on the order of 10–15%. The s results (Fig. 13b, d), instead, show a much higher sensitivity to L, and a high variability depending on the specific realizations. The $s(t)$ values appear to vary by $\approx$100% for $h = 20$ km and $\approx$50% for $h = 50$ km.

Qualitatively similar results have also been obtained for the LSM experiments. Dependence on sample size has been studied as for QG, varying $L = 200, 300, 400, 500$. Also, standard deviations have been evaluated for $L = 500$. Typical results are shown in Figure 14 for the high energy experiment *LSM*100 and for $h = 20$ km.

The ρ results (Fig. 14a, c) show a confined variability at varying L, approximately of the same order of the standard deviation, i.e. $\approx$10%. The s sensitivity at changing L is considerably higher, as well as the standard deviation.

In summary, the ρ statistics appear robust and with relatively small variability, well suited for quantitative studies and comparisons. The s statistics, on the other hand, have higher error and higher sensitivity to specific realizations. This is especially true for the QG experiments, where the action of vortex trapping can highly influence the statistics.

REFERENCES

Arakawa, A. 1966. Computational design for long-term numerical integration of the equations of fluid motion: Two dimensional incompressible flow. Part I. J. Comput. Phys., *1,* 119–143.

Aref, H. 1984. Stirring by chaotic advection. J. Fluid Mech., *143,* 1–21.

Batchelor, G. K. 1967. An Introduction to Fluid Dynamics, Cambridge University Press, Cambridge, 635 pp.

Bauer, S. A., A. Griffa, M. Swenson, A. Mariano and K. Owens. 1998. Eddy-mean flow decomposition and eddy diffusivity estimates in the tropical Pacific Ocean. 1. Methodology. J. Geophys. Res., *103,* 30855–30871.

Bauer, S. A., M. Swenson and A. Griffa. 2002. Eddy-mean flow decomposition and eddy diffusivity estimates in the tropical Pacific Ocean. 2. Global estimates. J. Geophys. Res., *107(C10),* 3154–3174.

Berloff, P. S. and J. C. McWilliams. 2002. Material transport in oceanic gyres. Part II: Hierarchy of stochastic models. J. Phys. Oceanogr., *32,* 797–830.

Blanke, B., M. Arhan, G. Madec and S. Roche. 1999. Warm water paths in the equatorial Atlantic as diagnosed with a G.C.M. J. Phys. Oceanogr., *11,* 2753–2768.

Castellari, S., A. Griffa, T. M. Özgökmen and P.-M. Poulain. 2001. Prediction of particle trajectories in the Adriatic Sea using Lagrangian data assimilation. J. Mar. Sys., *29,* 33–50.

Chassignet, E. P. and Z. D. Garraffo. 2001. Viscosity parameterization and the Gulf Stream separation, *in* From Stirring to Mixing in a Stratified Ocean, Proceedings Aha Huliko'a Hawaiian Winter Workshop, U. of Hawaii, January 15–19, 2001, P. Muller and D. Henderson, eds., 37–41.

Coulliette, C. and S. Wiggins. 2000. Intergyre transport in a wind-driven, quasi-geostrophic double gyre: an application of lobe dynamics. Nonlinear Proc. Geoph., *7,* 59–85.

Dutkiewicz, S., A. Griffa and D. B. Olson. 1993. Particle diffusion in a meandering jet. J. Geophys. Res., *98,* 16487–16500.

Falco, P., A. Griffa, P.-M. Poulain and E. Zambianchi. 2000. Transport properties in the Adriatic Sea deduced from drifter data. J. Phys. Oceanogr., *30,* 2055–2071.

Garraffo, Z. D., A. J. Mariano, A. Griffa, C. Veneziani and E. P. Chassignet. 2001. Lagrangian data in a high-resolution numerical simulation of the North Atlantic. 1. Comparison with *in situ* drifter data. J. Mar. Sys., *29,* 157–176.

Gazdag, J. 1976. Time-differencing schemes and transform methods. J. Comput. Phys., *20,* 196–207.

Ghil, M. and P. Malanotte-Rizzoli. 1991. Data assimilation in meteorology and oceanography. Adv. Geophys., *33,* 141–266.

Griffa, A. 1996. Applications of stochastic particle models to oceanographic problems, *in* Stochastic Modelling in Physical Oceanography, R. Adler, P. Muller and B. Rozovskii, eds., Birkhauser, Boston, 113–128.

Kuznetsov, L., M. Toner, A. D. Kirwan, C. K. R. T. Jones, L. H. Kantha and J. Choi. 2002. The Loop Current and adjacent rings delineated by Lagrangian analysis of near-surface flow. J. Mar. Res., *60,* 405–429.

LaCasce, J. H. and A. Bower. 2000. Relative dispersion in the subsurface North Atlantic. J. Mar. Res., *58,* 863–894.

Leaman, K. D., E. Johns and T. Rossby. 1989. The average distribution of volume transport and potential vorticity with temperature at three sections across the Gulf Stream. J. Phys. Oceanogr., *19,* 36–51.

Lee, S. K. 2001. On the structure of supercritical western boundary currents. Dyn. Atmos. Oceans, *33,* 303–319.

Lorenc, A. C. 1986. Analysis methods for numerical weather prediction, Quart. J. R. Met. Soc., *112,* 1177–1194.

Mariano, A. J., T. M. Chin and T. M. Özgökmen. 2003. Stochastic boundary conditions for coastal flow modeling. Geophys. Res. Lett., *30,* doi:10.1029/2003GL016972.

Mariano, A. J., A. Griffa, T. M. Özgökmen and E. Zambianchi. 2002. Lagrangian analysis and predictability of coastal and ocean dynamics. 2000. J. Atmos. Ocean. Techn., *19,* 1114–1126.

McClean, J. L., P.-M. Poulain and J. W. Pelton. 2002. Eulerian and Lagrangian statistics from surface drifters and a high-resolution POP simulation in the North Atlantic. J. Phys. Oceanogr., *32,* 2472–2491.

Özgökmen, T. M. and Eric P. Chassignet. 1998. Emergence of inertial gyres in a two-layer quasigeostrophic ocean model. J. Phys. Oceanogr., *27,* 2460–2476.

Özgökmen, T. M., A. Griffa, L. I. Piterbarg and A. J. Mariano. 2000. On the predictability of Lagrangian trajectories in the ocean. J. Atmos. Ocean. Techn., *17,* 366–383.

Özgökmen, T. M., A. Molcard, T. M. Chin, L. Piterbarg and A. Griffa. 2003. Assimilation of drifter observations in primitive equation models of midlatitude ocean circulation. J. Geophys. Res. Oceans, *108(C7),* 3238, doi:10.1029/2002JC001719.

Özgökmen, T. M., L. I. Piterbarg, A. J. Mariano and E. H. Ryan. 2001. Predictability of drifter trajectories in the tropical Pacific Ocean. J. Phys. Oceanogr., *31,* 2691–2720.

Paris, C. B., R. K. Cowen, K. M. M. Lwiza, D. P. Wang and D. B. Olson. 2002. Multivariate objective analysis of the coastal circulation of Barbados, West Indies: implication for larval transport. Deep-Sea Res., I, *49,* 1363–1386.

Peters, H., L. Shay, A. Mariano and T. Cook. 2002. Current variability on a narrow shelf with large ambient vorticity. J. Geophys. Res., *107(C8),* 10.1029/2001JC000813.

Piterbarg, L. I. 2004. Uncertainty of Lagrangian trajectories versus uncertainty of forcing for a stochastic flow. SIAM J. Appl. Math. (Submitted) Preprint from: http://www-rcf.usc.edu/~lpiterba/preprints.html

——2001a. Short term prediction of Lagrangian trajectories. J. Atmos. Ocean. Tech., *18,* 1398–1410.

——2001b. The top Lyapunov exponent for a stochastic flow modeling the upper ocean turbulence. SIAM J. Appl. Math., *62,* 777–800.

Piterbarg, L. I. and T. M. Özgökmen. 2002. A simple prediction algorithm for the Lagrangian motion in two-dimensional turbulent flows. SIAM J. Appl. Math., *63,* 116–148.

Poje, A. C. and G. Haller. 1999. Geometry of cross-stream mixing in a double-gyre ocean model. J. Phys. Oceanogr., *29,* 1649–1665.

Priestley, M. B. 1981. Spectral Analysis and Time Series, Academic Press, London, 890 pp.

Reynolds, A. M. 2002. On Lagrangian stochastic modelling of material transport in oceanic gyres. Physica D, *172,* 124–138.

Risken, H. 1989. The Fokker-Planck Equation: Methods of Solutions and Applications, Springer-Verlag, Berlin, 472 pp.

Samelson, R. M. 1996. Chaotic transport by mesoscale motions, *in* Stochastic Modelling in Physical Oceanography, R. J. Adler, P. Müller and B. L. Rozovoskii, eds., Birkhäuser, 466 pp.

Sawford, B. L. 1993. Recent developments in the Lagrangian stochastic theory of turbulent dispersion. Boundary-layer Meteor., *62,* 197–215.

Shay, L. K., T. M. Cook, H. Peters, A. J. Mariano, R. Weisberg, P. E. An, A. Soloviev and M. Luther. 2002. Very high-frequency radar mapping of surface currents. IEEE J. Oceanic Eng., *27,* 155–169.

Sheres, D., K. E. Kenyon, R. L. Bernstein and R. C. Beardsley. 1985. Large horizontal surface velocity shears in the ocean obtained from images of refracting swell and *in situ* moored current data. J. Geophys. Res., *90(C3),* 4943–4950.

Siegel, A., J. B. Weiss, J. Toomre, J. C. McWilliams, P. S. Berloff and I. Yavneh. 2001. Eddies and vortices in ocean basin dynamics. Geophys. Res. Lett., *28,* 3183–3186.

Smith, R. D., M. E. Maltrud, F. O. Bryan and M. W. Hecht. 2000. Numerical simulation of the North Atlantic Ocean at 1/10°. J. Phys. Oceanogr., *30,* 1532–1561.

Stammer, D. and E. P. Chassignet. 2000. Ocean state estimation and prediction in support of oceanographic research. Oceanography, *13,* 51–56.

Swarztrauber, P. N. 1977. The methods of cyclic reduction, Fourier analysis and FACR algorithm for discrete solution of the Poisson equation an a rectangle. SIAM Rev., *19,* 490–501.

Thomson, D. J. 1990. A stochastic model for the motion of particle pairs in isotropic high-Reynolds-number turbulence, and its application to the problem of concentration variance. J. Fluid Mech., *210,* 113–153.

——1987. Criteria for the selection of stochastic models of particle trajectories in turbulent flows. J. Fluid Mech., *180,* 529–556.

Toner, M., A. C. Poje, A. D. Kirwan, C. K. R. T. Jones, B. L. Lipphardt and C. E. Grosch. 2001. Reconstructing basin-scale Eulerian velocity fields from simulated drifter data. J. Phys. Oceanogr., *31,* 1361–1376.

Veneziani, M., A. Griffa, A. M. Reynolds and A. J. Mariano. 2004. Oceanic turbulence and stochastic models from subsurface Lagrangian data for the northwest Atlantic Ocean. J. Phys. Oceanogr., (in press).

Wiggins, S. 1992. Chaotic Transport in Dynamical Systems, Springer, NY, 301 pp.

Zambianchi, Enrico and Annalisa Griffa. 1994. Effects of finite scales of turbulence on dispersion estimates. J. Mar. Res., *52,* 129–148.

Received: *21 July, 2003;* revised: *26 January, 2004.*