Journal of Marine Research, 53, 371–401, 1995

Estimates of turbulence parameters from Lagrangian data using a stochastic particle model

by Annalisa Griffa[1], Kenneth Owens[2,3], Leonid Piterbarg[2] and Boris Rozovskii[2]

ABSTRACT

A new parametric approach for the study of Lagrangian data is presented. It provides parameter estimates for velocity and transport components and is based on a stochastic model for single particle motion. The main advantage of this approach is that it provides more accurate parameter estimates than existing methods by using the *a-priori* knowledge of the model. Also, it provides a complete error analysis of the estimates and is valid in presence of observation errors. Unlike nonparametric methods (e.g. Davis, 1991b), our technique depends on *a-priori* assumptions which require that the model validity be checked in order to obtain reliable estimates. The model used here is the simplest one in a hierarchy of "random flight" models (e.g. Thomson, 1987), and it describes the turbulent velocity as a linear Markov process, characterized by an exponential autocorrelation. Experimental and numerical estimates show that the model is appropriate for mesoscale turbulent flows in homogeneous regions of the upper ocean. More complex models, valid under more general conditions, are presently under study.

Estimates of the mean flow, variance, turbulent time scale and diffusivity are obtained. The properties of the estimates are discussed in terms of biases and sampling errors, both analytically and using numerical experiments. Optimal sampling for the measurements is studied and an example application to drifter data from the Brazil/Malvinas extension is presented.

1. Introduction

Lagrangian data (i.e. data provided by current-following instruments such as drifting buoys) have become increasingly common in the last two decades, providing good coverage of extensive regions in the ocean (Davis, 1991a). In the mesoscale range, Lagrangian data provide a satisfactory measurement of the horizontal motion of ideal water particles (Davis, 1991b), and are therefore particularly suitable for statistical studies of the transport of passively advected substances in the ocean. In this context, they can be considered complementary to more direct transport studies

1. Rosenstiel School of Marine and Atmospheric Sciences, University of Miami, 4300 Rickenbacker Causeway, Miami, Florida, 33149, U.S.A.

2. Center for Applied Mathematics, University of Southern California, Los Angeles, California, 90089, U.S.A.

3. Present address: NASA, Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, California, 91109, U.S.A.

based on observations of geochemical tracers. The simplest statistical description of transport phenomena is provided by the average concentration of a passive tracer $\langle\theta\rangle$. Ideally, $\langle\theta\rangle$ can be computed from the single particle probability function P_1 (e.g. Davis, 1983) that can be reconstructed given enough Lagrangian data in a certain region. In real applications, though, the number of data necessary to compute P_1 is unrealistically large, so that the computation of $\langle\theta\rangle$ from Lagrangian data relies on the use of indirect methods. In what follows, we briefly summarize the methods in the literature, and then introduce a new and alternative approach.

The most common way to compute $\langle\theta\rangle$ is based on its Eulerian evolution equation, using a simple closure form to parameterize the turbulent transport term. The simplest and most commonly used closure is the eddy-diffusion parameterization, for which the equation for $\langle\theta\rangle$ reduces to the advection-diffusion equation

$$\partial\langle\theta\rangle/\partial t = -\mathbf{U}\cdot\nabla\langle\theta\rangle + \nabla(\mathbf{K}\nabla\langle\theta\rangle) \tag{1}$$

where $\mathbf{U}$ is the mean velocity and $\mathbf{K}$ is the eddy-diffusivity. The parameters $\mathbf{U}$ and $\mathbf{K}$ can be estimated, at least in principle, from the data and inserted into (1) to compute $\langle\theta\rangle$. Notice that Eq. (1) is valid under restrictive conditions (e.g. Holloway, 1989; Zambianchi and Griffa, 1994a) so that its applications are limited to processes with time scales larger than the Lagrangian time scale $\mathbf{T}$ (and corresponding space scale), which are found in selected regions of the ocean where the flow is approximately homogeneous. These restrictions can be relaxed by considering generalized forms of (1) (generalized "K-models"). An example of this type of model is the "elaborated advection-diffusion equation" derived by Davis (1987), which is written in terms of a time-dependent $\mathbf{K}$ representing the particle history. These models are not frequently used in practical applications due to their complexity.

One problem with the approach based on (1) (or on its generalizations) is the difficulty in estimating the turbulent parameter $\mathbf{K}$ from the Lagrangian data. The basic reason for this is the dependence of $\mathbf{K}$ on the lowest frequency of variability which introduces an inherent problem in the estimates. This can be easily seen in the case of a homogeneous flow (Taylor, 1921), where $\mathbf{K}$ is defined as

$$\mathbf{K} = \sigma^2 \int_0^\infty \mathbf{R}(\tau)d\tau$$

where $\mathbf{R}(\tau)$ is the Lagrangian velocity autocorrelation and σ^2 is the velocity variance. As it is evident $\mathbf{K}$ depends on the asymptotic behavior of the particles so that very long records must be used to correctly compute it. On the other hand, as shown by Davis (1991b), the sampling error grows with time, so that reliable estimates are very hard to obtain.

The approach proposed in this paper is different from the one outlined above, insofar as it assumes, in the range of scales of interest, that the Lagrangian velocity field obeys a known statistical model, so that the velocity autocorrelation has a known

shape. The model depends on several parameters (i.e. the Lagrangian time scale **T**, the velocity variance σ^2 and the mean flow **U**), which can be estimated from the data. The main advantage of this approach (which is called "parametric") is that it gives accurate parameter estimates by using the *a-priori* knowlege of the model. In particular, in the application presented here, all the estimates turn out to be independent of the lowest frequencies, but rather depend on measurements at scales on the order of **T**, resulting in greater estimate reliability and other benefits to be discussed. On the other hand, our approach has stronger *a-priori* constraints than nonparametric methods (e.g. Davis, 1991b) and thus the validity of the model must be checked in order to assure reliability of the estimates. Details on the choice of the model and on its validity are given in the following. The analysis presented here is only a first step in the use of the general method, and it is restricted to the case of energetic flows, such as flows in the upper ocean, occurring in approximately homogeneous and stationary regions. The attention is focused on the role of the mesoscale in the transport processes, whereas the impact of long time fluctuation is not included. Notice that in the presence of homogeneity and stationarity, a simple relationship exists between the model parameters σ^2, **T** and the turbulent parameter **K** in (1) ($\mathbf{K} = \sigma^2\mathbf{T}$), so that estimates of **T** and σ^2 automatically provide estimates for **K**.

The model that we use belongs to the general class of the "random flight" models, which have a long history in the literature (e.g. Chandrasekhar, 1943), and it has been used in a number of applications in the physics of the atmosphere and of the ocean (e.g. Thomson, 1986; Dutkiewicz *et al.,* 1993). It describes the motion of single independent particles in a turbulent flow, such as a random mesoscale eddy field, by using a system of ordinary stochastic differential equations. The particles can be thought of as belonging to a tracer so that their concentration corresponds to the mean concentration of the tracer itself (Csanady, 1980). The basic assumption of the model is that the turbulent velocity sampled by the particles during their motion is a Markov process, characterized in the case of homogeneous turbulence, by an exponential autocorrelation. The exponential autocorrelation is an approximation of the real particle autocorrelation valid for flows where the time scales of the acceleration can be considered small with respect to the other velocity scales. Experimental (e.g. Colin de Verdiere, 1983; Davis, 1985; Krauss and Boning, 1987) and numerical (e.g. Verron and Nguyen, 1989; Davis, 1991b; Figueroa and Olson, 1994) evidence shows that the approximation is appropriate for turbulent flows at the surface or in the upper portion of the ocean. The occurrence of the exponential shape in data taken in different geographical regions and in numerical simulations obtained with different wavenumber spectra suggests that it is indeed a "relatively universal" (Davis, 1991b) form for this type of flow.

In Section 2 we introduce and discuss the Lagrangian model. In Section 3, the estimates for the parameters are derived using the method of moments. A discussion

on biases and sampling errors of the estimates is given in Section 4. The results of Sections 3 and 4 are tested using numerical simulations in Section 5, and an application to drifter data is presented in Section 6. A summary and a discussion are given in Section 7.

2. Stochastic model for particle motion

In this section we present the "random flight" stochastic model which describes the motion of single particles in a flow with a mean flow **U** and a turbulent velocity u. Each particle is considered as independently launched in a different realization of u, and the statistics are computed by averaging over the ensemble of all the particles. An important consequence of the assumption of independence of the particles is that u is considered as a purely time-dependent process representing the turbulent velocity as sampled by the particle during its motion. This allows for a substantial simplification with respect to the fully space dependent representation of u necessary to study two-particle or higher order statistics.

In the following, we briefly summarize the basic characteristics of the model in a simple form, valid under the following assumptions: (1) the velocity field is 2-dimensional (on the free surface or on isopycnal surfaces), (2) the velocity field is homogeneous and stationary with constant **U**, and (3) the two components of the velocity are independent. Even though more general forms of the model are available (e.g. Pasquill and Smith (1983), Thomson (1987)), we think that the present assumptions, while providing useful simplifications in the analytical calculations, provide a valid starting point for discussion. Assumption (1) is usually valid for mesoscale phenomena (Davis, 1991b), and assumption (3) can be met, at least approximately, in selected regions of the ocean. Assumption (2) is more difficult to satisfy, especially for what concerns the stationarity since the ocean circulation is essentially a red spectrum process. The presence of low-frequency variability in the ocean is of course a problem not only for our analysis, but in general for any statistical analysis concerned with the description of average quantities or "typical" aspects of the ocean circulation and transport. As suggested by Davis (1991b), this difficulty might be overcome with a pragmatic approach, i.e. by substituting the notion of a true mean by that of "an average over some finite time, representative of a particular ocean climate." In this sense, our analysis can be considered as focusing on the "typical" role of the mesoscale in the transport processes over the time interval of the available measurements, while the role of the long term fluctuations is not directly considered.

We use assumption (3), which implies that each velocity component can be treated separately, to write the model for each component in 1-dimensional form as follows

$$dx = vdt = (U + u)dt \tag{2}$$

$$du = -\theta u dt + \sigma\sqrt{2\theta}dW \tag{3}$$

where

$$\theta = 1/T. \tag{4}$$

In the above expressions, dx is the displacement, in the considered component, of the particle during the time dt, v is the velocity, U is the mean velocity, u is the turbulent velocity, T is the turbulent time scale, σ is the square root of the turbulent velocity variance σ^2, and dW is a random increment from a normal distribution with zero mean and second order moment $\langle dW \cdot dW \rangle = dt$ where the increments corresponding to non-overlapping time intervals are statistically independent. The above properties of the continuous stochastic process imply that $W(t)$ is a Brownian motion and that Eq. (3) should be understood in the sense of Ito calculus. Its solution is usually referred to as the Ornstein-Uhlenbeck process where we assume that the probability law of the initial condition u_0 is the steady state distribution (invariant measure). The Ornstein-Uhlenbeck process is a stationary Markov process.

Physically, Eq. (3) states that the turbulent velocity u along a particle trajectory is a Markovian process. This means that at each time step as the particle moves through the fluid it loses a fraction of its momentum u_T^{dt} and in turn receives a random impulse proportional to dW due to the random turbulent interactions. As a consequence, the particle progressively "loses memory" of its initial turbulent velocity. The characteristic time scale over which the particle still "remembers" is given by T.

Eq. (3) can be solved exactly, and its solution is well known (e.g. Yaglom, 1962). It is characterized by the exponential autocorrelation

$$R(\tau) = \frac{1}{\sigma_2} \langle u(t)u(t+\tau) \rangle = e^{-\theta\tau}. \tag{5}$$

From (5), $T = 1/\theta$, turns out to be the Lagrangian integral time of the turbulent velocity.

Notice that, although the model is physically much more acceptable than a Markovian model for the particle position $x(t)$ which would imply a discontinuity in the velocity field (e.g. Zambianchi and Griffa, 1994a), it still is not completely realistic since it implies a discontinuity in the acceleration. This in turn implies that the scales of the acceleration are assumed to be negligible in the model and they are not explicitly resolved. A more realistic model, which would resolve the acceleration scales, would be characterized by a smoother autocorrelation than the exponential (5) at $\tau = 0$ (eg. Pope, 1994). Lagrangian observations and simulations (e.g. Colin de Verdiere, 1983; Davis, 1991b) show that the "exponential-like" behavior is a good approximation for mid-latitude mesoscale turbulent flows in the upper ocean. This can be understood by considering that for these flows the scale of the velocity T is of the order of 2–10 days, whereas the scale of the acceleration is of the order of one day or less. Since motions of the order of one day are characterized by high frequency

dynamics, such as inertial and tidal oscillations, they are usually filtered out and neglected in mesoscale studies. In the mesoscale range, then, the scales of the acceleration are not considered and the exponential approximation holds. Notice that the situation can be different for deep flows, where the time scales of the evolution are longer. This is suggested by deep float observations and simulations (e.g. Riser and Rossby, 1983) which show a smoother autocorrelation at the origin. In these cases, a second-order modification should be added to the model (3) (Sawford, 1991).

As a final remark, we notice that the model (2)–(3) can be generalized to include the effects of inhomogeneity by allowing the turbulence parameters to vary in space and by adding in (3) a term proportional to $\partial\sigma^2/\partial x$ (Van Dop *et al.,* 1985). For this modified model, the turbulent velocity autocorrelation is not necessarily exponential because of the inhomogeneous turbulence.

3. Parameter estimation

In this section estimates for the model parameters θ, σ and U are derived starting from the Lagrangian observations of particle positions and using the method of moments. We choose this method instead of the maximum likelihood method because, even when noise-free observations are available, the maximum likelihood approach does not yield a tractable analytic expression for the likelihood ratio, making the maximum likelihood estimates difficult to study.

Lagrangian instruments such as drifting buoys are designed to move with the currents and to report their position at discrete times either acoustically or via satellite. In the range of scales we are interested in, the instruments follow the currents with good accuracy (Davis, 1991b) so that their motion satisfactorily approximates the motion of ideal particles. This allows us to write the discrete position measurements, s_n, provided by drifting buoys as the sum of the true particle positions, x_n, plus a random instrument error due to the acoustic or satellite transmission,

$$s_n = x_n + \epsilon\xi_n = \int_0^{t_n} v(t)dt + \epsilon\xi_n \tag{6}$$

where v is the particle velocity, which we assume to obey the model (2)–(3), ϵ^2 is the variance of the observation errors, and $\{\xi_n,\ n = 1 \ldots, N\}$ is a standard Gaussian sequence of independent observation errors with mean zero and variance one. In the following we assume that ϵ is an unknown parameter, which will be estimated together with the model parameters and that measurements are available at regular intervals $0 = t_0 < t_1 < \cdots < t_N = T_{\text{obs}}$, with $t_n - t_{(n-1)} = h$. The assumption of

regularity can be partially relaxed assuming irregular but comparable intervals with average h, as it has been shown by numerical experiments.

From the position measurements (6), a sequence of finite difference approximations for the velocity can be computed as

$$z_n = \frac{s_n - s_{n-1}}{h}. \tag{7}$$

The sequence of velocity approximations (7) can be written as a function of the particle velocity using (2),

$$z_n = U + \bar{u}_n + \epsilon \frac{\xi_n - \xi_{n-1}}{h}. \tag{8}$$

where

$$\bar{u}_n = \frac{1}{h} \int_{t_{n-1}}^{t_n} u(t)dt \tag{9}$$

represents the mean value of the turbulent component of the velocity on the time interval $t_{n-1} \leq t \leq t_n$.

The method of moments is applied using the following procedure. First we compute theoretically four moments of z_n using the known statistics of the model (2)–(3). Details on the theoretical computation of the moments are shown in Appendix A. Here we only present the final results, which turn out to be functions of the model parameters θ, σ, U and of the observation error parameter ϵ,

$$\begin{aligned}
\langle z_n \rangle &= U \\
r_0 = \langle z_n^2 \rangle - \langle z_n \rangle^2 &= \left\{ \frac{2\epsilon^2}{h^2\sigma^2} + \frac{2}{h^2\theta^2}(e^{-\theta h} - 1 + \theta h) \right\} \sigma^2 \\
r_1 = \langle z_n z_{n+1} \rangle - \langle z_n \rangle^2 &= \left\{ -\frac{\epsilon^2}{h^2\sigma^2} + \frac{1}{h^2\theta^2}(1 - e^{-\theta h})^2 \right\} \sigma^2 \\
r_2 = \langle z_n z_{n+2} \rangle - \langle z_n \rangle^2 &= \frac{\sigma^2}{h^2\theta^2}(1 - e^{-\theta h})^2 e^{-\theta h}.
\end{aligned} \tag{10}$$

We now observe that the moments r_0, r_1, r_2 can be approximately computed from the data using the fact that the process z_n is stationary and ergodic (as evidenced by the exponential autocorrelation) so that the average $\langle\ \rangle$ can be approximated as a time

average over measurements from one particle trajectory or from a set of independent trajectories,

$$\begin{aligned}
\hat{U} &= \frac{1}{N}\sum_{n=1}^{N} z_n \\
\hat{r}_0 &= \frac{1}{N}\sum_{n=1}^{N} z_n^2 - \hat{U}^2 \\
\hat{r}_1 &= \frac{1}{N}\sum_{n=1}^{N} z_n z_{n+1} - \hat{U}^2 \\
\hat{r}_2 &= \frac{1}{N}\sum_{n=1}^{N} z_n z_{n+2} - \hat{U}^2.
\end{aligned} \tag{11}$$

A system of approximate equations can then be obtained by equating the moment estimates (11) to the r.h.s. of the theoretical expressions in (10). The solutions of this system give the estimates of the parameters, $\hat{\theta}$, $\hat{\sigma}$, $\hat{U}$, $\hat{\epsilon}$. After a simple manipulation, the system can be rewritten as

$$\begin{aligned}
\hat{U} &= \frac{1}{N}\sum_{n=1}^{N} z_n \\
\frac{2\hat{r}_1 + \hat{r}_0}{\hat{r}_2} &= \frac{2(1 - e^{-\hat{\theta}h})^2 + 2(e^{-\hat{\theta}h} - 1 + \hat{\theta}h)}{(1 - e^{-\hat{\theta}h})^2 e^{-\hat{\theta}h}} \\
\hat{\sigma} &= \frac{\sqrt{\hat{r}_2} e^{\hat{\theta}h/2}\hat{\theta}h}{1 - e^{-\hat{\theta}h}} \\
\hat{\epsilon} &= h\sqrt{\frac{\hat{\sigma}^2}{(\hat{\theta}h)^2}(1 - e^{-\hat{\theta}h})^2 - \hat{r}_1}.
\end{aligned} \tag{12}$$

The system (12) is quite simple to implement in practical applications. Eq. (12b) can be solved numerically for $\hat{\theta}$ and the other estimates are obtained by substitution. Notice that an estimate of the diffusion parameter K can also be computed as

$$\hat{K} = \frac{\hat{\sigma}^2}{\hat{\theta}}.$$

Before proceeding to the analysis of consistency and accuracy of the estimates that will be presented in the next section, let us make some remarks on the basic characteristics of the estimates as they appear in (12). A first remark is that the estimates depend only on the moments of the approximated velocities z_n computed at time lags less than or equal to $2h$, in contrast to the estimates for T (and K) commonly used in the literature (e.g. Krauss and Boning, 1987; Figueroa and Olson, 1989),

which are based on the definition

$$T = \frac{1}{\sigma^2} \int_0^\infty R(\tau) d\tau \tag{13}$$

and which depend on asymptotically long lags. This is due to the fact that our estimates are derived from the assumption of an exponential autocorrelation (5). Conceptually then, the estimate of $\theta = 1/T$ in (12) depends on the slope of the autocorrelation computed at the time lag h, and it is therefore easier to compute from the data than the asymptotic estimate (13). A second remark is that the estimate for σ in (12c) is different from the intuitive estimate $\hat{\sigma} = \sqrt{\hat{r}_0}$ which is commonly used in applications. As it will be explained in detail in the next section, the estimate (12c) has better properties than $\hat{\sigma} = \sqrt{\hat{r}_0}$. The reason for this is inherent in the nature of the Lagrangian measurements which are position and not velocity. The velocity is only approximately computed from the discrete position observation so that the statistics of the approximated velocity do not necessarily correspond to the statistics of the true velocity unless the time step h is sufficiently small. In this limit the estimates (12c) and $\hat{\sigma} = \sqrt{\hat{r}_0}$ coincide.

4. Consistency and sampling errors of the estimates

In this section we analytically investigate the basic statistical properties of the model parameter estimates, $\hat{U}$, $\hat{\theta}$, $\hat{\sigma}$, $\hat{\epsilon}$ as given by (12). Also, for comparison with other results in the literature, we include the analysis of the estimate $\hat{K}$, even though the parameter K does not enter in our model and its estimate does not play a direct role in our approach.

a. Consistency properties

The property of consistency is fundamental for acceptable estimates. An estimate $\hat{A}_N$ of an arbitrary parameter A is said to be consistent if $\hat{A}_N$ converges to A with probability one as $N \to \infty$. Notice that consistency implies asymptotic unbiasness, $\lim_{K\to\infty} \langle \hat{A}_{N_K} \rangle = A$ where $\hat{A}_{N_K}$ is a subsequence taken from $\hat{A}_N$.

The proof of consistency for the estimates in (12) is straightforward. Given our assumption of stationarity and ergodicity, the estimates of the moments $\hat{U}$, $\hat{r}_0$, $\hat{r}_1$, $\hat{r}_2$ defined by (11) converge to the true moments as $N \to \infty$. As a consequence, the solutions of the system (12) converge to the true parameters θ, σ, U, ϵ since the system has a unique solution.

Notice that the commonly used estimate for σ, $\hat{\sigma} = \sqrt{\hat{r}_0}$, is not consistent even in the ideal case of $\epsilon = 0$. $\sqrt{\hat{r}_0}$, in fact, converges to $\sqrt{r_0}$ for large N, but $\sqrt{r_0}$ is the square root variance of the approximated velocity and it does not necessarily correspond to the square root variance of the real velocity, σ (this is shown also by the relationship 10b). Only if $h \to 0$, does $\sqrt{r_0} \to \sigma$, because the approximated velocities z_n converge to

the real velocities. For Lagrangian estimates, where the velocity is not directly measured and the limit as $h \to 0$ is not always approximated, the estimate $\hat{\sigma} = \sqrt{\hat{r}_0}$ can give misleading results. The estimates in (12) are instead specifically derived assuming that only approximate velocities are known, and they guarantee consistency even in presence of a finite time step h and observation errors.

As a last remark, we notice that the consistency property does not guarantee that the estimates are also unbiased for a finite number of sampling N. The estimates (11)–(12) are actually biased for finite N, $\langle \hat{A}_N \rangle = A + C/N$, where C is a constant which depends on the values of the parameters. The bias term C/N, though, can be considered negligible because, as it is shown in Section 4b, it is of lower order than the accuracy of the estimates $(1/\sqrt{N})$.

b. Sampling errors

Here we discuss the sampling mean squared error, defined as

$$Er^2(A) = \frac{\langle (\hat{A} - A) \rangle^2}{(A^2)},$$

for the estimates $\hat{U}$, $\hat{\theta}$, $\hat{\sigma}$, $\hat{\epsilon}$ and $\hat{K}$. The analytical computations are performed using the definitions of the estimates given by (12) and the statistics of the model (2)–(3). The results are obtained in the asymptotic limit where the number of observations N is large and the sampling interval h is nonzero. The amplitude ϵ of the observation error is arbitrary. In order to isolate the role of the observation error, in the following we discuss first the results for $\epsilon = 0$, and then for $\epsilon \neq 0$. The details of the calculations are quite lengthy and are reported in Appendix B. Only the final results are presented here and discussed.

We remark that, for completeness, the theoretical analysis is carried out on a range including very small values of h and $h\theta$, the sampling rate relative to the autocorrelation scale T. It is important to keep in mind, though, that in practical applications for mesoscale motions, $h\theta$ is always > 0.1. $T = 1/\theta$, in fact, is of the order of 2–10 days and h is of the order of 1 day. Even when more data are actually available (possibly five or more data points per day), they are usually filtered at a one day interval, because higher frequencies obey different, inertial and subinertial, dynamics.

i. Zero observation error, $\epsilon = 0$. We start by considering the sampling mean squared error for the mean velocity estimate $\hat{U}$, $Er^2(U)$. As shown in Appendix B (B1), the expression for large N is

$$Er^2(U) = \frac{2\sigma^2}{U^2 N h \theta}. \tag{14}$$

The expression (14) states that the error in estimating U is directly proportional to the ratio between fluctuation and mean flow, σ^2/U^2, and inversely proportional to the number of independent samples,

$$N^* = Nh\theta.$$

This coincides with the expression commonly used in the literature (e.g. Flierl and McWilliams, 1977), derived assuming direct velocity measurements.

We now discuss the asymptotic mean squared errors for $\hat{\theta}$, $\hat{\sigma}$ and $\hat{K}$. The analytic expressions, given in details in Appendix B (B2–B4), have a general structure of the form

$$Er^2(\theta) = \frac{f_1(h\theta)}{Nh\theta}, \tag{15}$$

$$Er^2(\sigma) = \frac{f_2(h\theta)}{Nh\theta}, \tag{16}$$

$$Er^2(K) = \frac{f_3(h\theta)}{Nh\theta}. \tag{17}$$

The expressions (15)–(17) indicate that the errors for the three estimates are inversely proportional to N^*, as expected, and that at given N^* there is a nontrivial dependence on $h\theta$. The dependence of the errors on $h\theta$ at fixed values of N^* is shown in Figure 1 for the three estimates, and it is discussed in the following. The values of N^* chosen are $N^* = 50$, 100, 200, 400, and 1000 which correspond to the range of value found in the literature for Lagrangian statistical studies.

Figure 1a shows the error for $\hat{\sigma}$, $Er^2(\sigma)$. As is evident, the error tends to a constant value for small sampling intervals, $h\theta \ll 1$, whereas it increases at larger sampling intervals, $h\theta \approx 1$. Intuitively this result can be explained as follows. When the number of independent observations N^* is held fixed, reducing h to values much smaller than the autocorrelation scale, T, corresponds to over sampling the process and does not improve the quality of the estimate. On the other hand if h is increased to values of the order of T or larger, the scale of the variability is not resolved anymore, and the estimate deteriorates.

The error for $\hat{\theta}$, shown in Figure 1b, has a structure similar to $Er^2(\sigma)$. Similarly to $Er^2(\sigma)$, $Er^2(\theta)$ also converges to constant values for small values of $h\theta$, but the absolute values of the errors are higher. This indicates that the estimate of θ is more sensitive and less accurate than the estimate of σ. This is not surprising considering that conceptually θ depends on the autocorrelation slope, whereas σ depends on its amplitude. In order to make a simple quantitative comparison, consider the value of the error for the two estimates at the lowest value of $N^* = 50$. $Er^2(\sigma)$ converges for small $h\theta$ to an approximate value 0.01, indicating a relative mean error on σ of 10%.

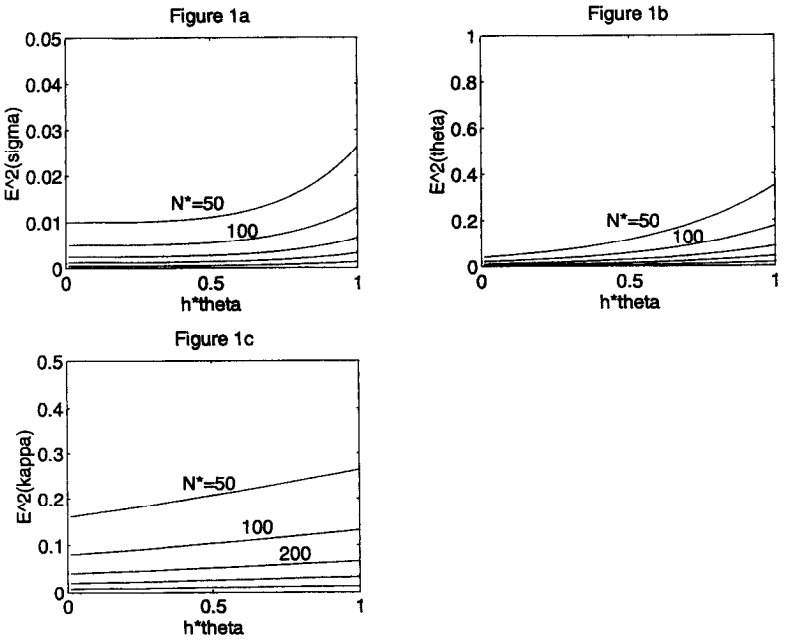

Figure 1. Theoretical mean square sampling errors Er^2 versus $h\theta$ at fixed number of independent measurements $N^* = 50, 100, 200, 400, 1000$ computed for $\epsilon = 0$. (a) $Er^2(\sigma)$; (b) $Er^2(\theta)$; (c) $Er^2(K)$.

$Er^2(\theta)$, instead converges to a value of approximately 0.04, with a relative mean error of 20%.

In Figure 1c the error for $\hat{K}$, $Er^2(K)$, is shown. The dependence on $h\theta$ shows some different characteristics than $Er^2(\sigma)$ and $Er^2(\theta)$. The basic difference is that the estimates $\hat{\sigma}$ and $\hat{\theta}$ deteriorate rapidly when $h\theta$ grows, (and the scale of the variability is not well resolved) while the estimate $\hat{K}$ is more stable. This can be understood intuitively by considering that K is a parameter characteristic of the asympototic behavior of the system, and therefore can be estimated satisfactorily even for relatively large $h\theta$, whereas σ and θ, which are parameters of the variability, cannot. On the other hand, the estimate $\hat{K}$ does not improve as much as the other estimates for small $h\theta$. For the same value $N^* = 50$ considered above, $Er^2(K)$ converges to approximately 0.16, with a relative mean error of 40%.

We remark that the results (15)–(17) can be presented in an equivalent but alternative way. As an example, in some situations it is useful to consider the curves when the number of measurements N is fixed, instead of N^*. These curves differ from the ones in Figure 1 only by a factor of $1/h\theta$, but their direct visualization is useful in highlighting some interesting features of the error structure. Typically, this kind of curve is useful when planning an experiment, since the number of measurements N is approximately fixed a-*priori* (for instance by the life of the instrument batteries),

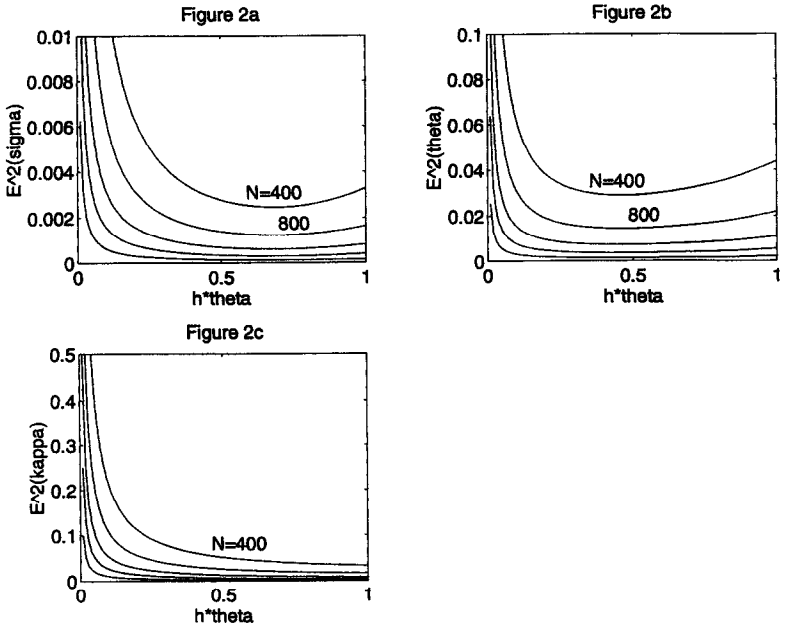

Figure 2. Theoretical mean square sampling errors Er^2 versus $h\theta$ at fixed number of measurements $N = 400, 800, 1600, 3200, 8000$, computed for $\epsilon = 0$. (a) $Er^2(\sigma)$; (b) $Er^2(\theta)$; (c) $Er^2(K)$.

whereas the sampling interval must be chosen in order to optimize the estimate quality. The curves for fixed N, are shown in Figure 2.

Figure 2a and 2b show that both $Er^2(\sigma)$ and $Er^2(\theta)$ are characterized by curves with finite minimum values, occurring at $h\theta \approx 0.7$ and 0.5 respectively. Intuitively this result can be explained as follows. When h becomes small at given N and θ, the total time of observation, $T_{obs} = Nh$, becomes also small, reducing the number of independent measurements N^* which are required for a good estimate. When h becomes too large, on the other hand, the scale $T = 1/\theta$ is not resolved, and the parameters cannot be satisfactorily estimated. In between these two tendencies there is an optimal range of sampling intervals.

The behavior of $Er^2(K)$, shown in Figure 2c, is different in so far as it does not show a minimum value, but rather it tends to slightly decrease at increasing $h\theta$, (even for the range $h\theta > 1$, which is not shown in the Fig. 2c). As previously discussed with reference to Figure 1c, this is conceptually a consequence of the fact that K is an asymptotic parameter. Notice that the result suggests that, at least in principle, K can be estimated satisfactorily using long sampling intervals. On the other hand, it should be considered that the assumption of homogeneity in our analysis is at best only a local property in the ocean. Buoys launched in a region with certain local characteristics will tend to move out in a finite time sampling other regions, so that the use of long sampling intervals, which imply long T_{obs}, is usually not appropriate.

ii. Finite observation error, $\epsilon \neq 0$. The influence of the observation error ϵ on the quality of the estimates can be directly evaluated (at least in the asymptotic limit of large N and finite h), using the complete expressions for the sampling errors derived in Appendix B. The results of Appendix B show that only $Er^2(U)$ does not depend on ϵ(B1), whereas $Er^2(\theta)$, $Er^2(\sigma)$, $Er^2(K)$ (B2–B4) conserve the same structure shown in (15–17), but with f_i ($i = 1, 2, 3$) depending on $\epsilon/h\sigma$ as well as $h\theta$. $\epsilon/h\sigma$ is the ratio between the typical error size ϵ and the typical particle displacement due to turbulent motion during the time interval h, $h\sigma$.

The general behavior of the error is the following. For values of h greater than $h^* = \epsilon/\sigma$ (i.e. the sampling for which the turbulent displacement has the same size as the observation error), the sampling errors are the same as for $\epsilon = 0$. For smaller values of h, instead, the errors increase rapidly as h decreases, making the estimates unreliable. The reasons for this can be understood intuitively by considering the fact that the estimates are based on the velocity approximations (8). When the sampling interval h becomes small, the terms U and $\overline{u}$ on (8) tend to constant values, whereas the term depending on the observation error, $\epsilon\,(\xi_n - \xi_{n-1})/h$, increases as $1/h$. As a consequence, the approximated velocity becomes more and more influenced by the observation error, and the parameters of the flow become progressively more difficult to estimate. Only the estimate of U is not influenced by the observation errors, because they average out.

A specific example of the behavior of the sampling errors for a given set of parameters is given in Figures 3a-b-c, with $Er^2(\sigma)$, $Er^2(\theta)$, $Er^2(K)$ plotted as functions of $h\theta$ at constant N^*. The flow parameters, representative for oceanographical values, are $\sigma = 15$ km/day, $\theta = 1/4$ day^{-1}, $U = 12$ km/day, whereas ϵ is chosen to be 0.6 km, a conservative estimate for satellite tracked instruments. A comparison with the curves of Figure 1, corresponding to the case $\epsilon = 0$, shows that the observation error influences the quality of the estimates only for $h\theta \ll 0.1$, as expected since $h^* = \epsilon/\sigma = 0.04$ and therefre $h^*\theta = 0.01$. This suggests that for realistic values of the parameters and of the sampling, the observation error can be neglected in the sampling error analysis.

A formula similar to the ones for $\hat{\sigma}$, $\hat{\theta}$ and $\hat{K}$ is valid also for the sampling error of ϵ, $Er^2(\epsilon) = f_4(h\theta, \epsilon/h\sigma)/Nh\theta$, where f_4 is given in (B5). In Figure 3d, the mean squared sampling error for $\hat{\epsilon}$ is shown for the same parameters as before. The error appears to be small only for $h\theta < 0.1$ or 0.2, depending on N^*, and increases dramatically as $h\theta$ increases. This suggests that the parameter ϵ is actually very hard to estimate in practical situations because of the large estimation error.

5. Numerical simulations

In Section 4 a theoretical analysis of the sampling errors for $\hat{U}$, $\hat{\theta}$, $\hat{\sigma}$, $\hat{\epsilon}$ and $\hat{K}$ was performed, valid in the asymptotic range of large N and finite h. In this section,

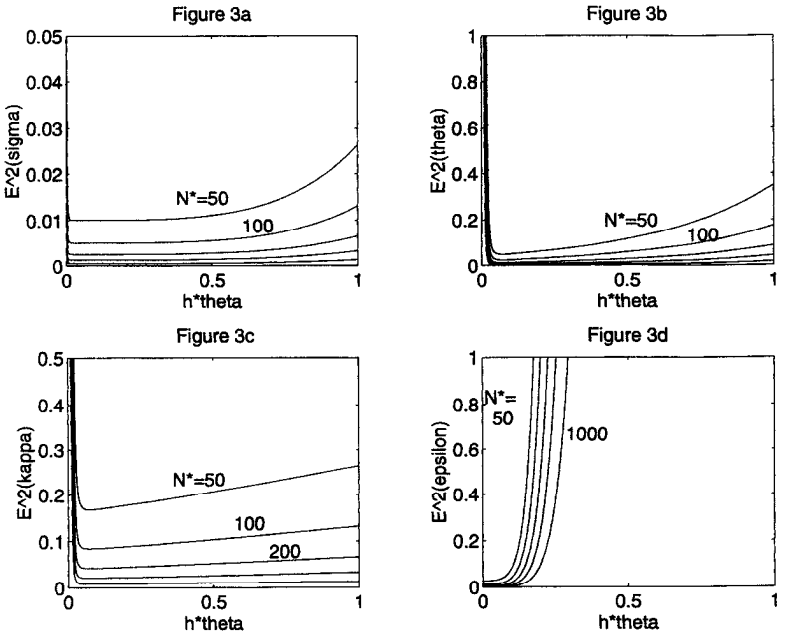

Figure 3. Theoretical mean square sampling errors Er^2 versus $h\theta$ at fixed number of independent measures $N^* = 50, 100, 200, 400, 1000$ computed for $\epsilon = 0.6$ km. (a) $Er^2(\sigma)$; (b) $Er^2(\theta)$; (c) $Er^2(K)$; (d) $Er^2(\epsilon)$.

results from numerical simulations are presented, aimed at verifying the theoretical asymptotic results.

The numerical simulations are performed simulating the succession of approximate velocities (8),

$$z_n = U + \bar{u}_n + \epsilon \frac{\xi_n - \xi_{n-1}}{h}$$

where

$$\bar{u}_n = \frac{1}{h} \int_{t_{n-1}}^{t_n} u(t)dt.$$

$\bar{u}_n$ is simulated exactly without introducing finite differences or other types of approximations. This is done using the fact that, as shown in Appendix C, the process $\bar{u}_n$ can be represented as an autoregressive moving average process that satisfies

$$\bar{u}_n = \alpha \bar{u}_{n-1} + p\eta_n + q\eta_{n-1}$$

where $\{\eta_n, n = 1 \ldots, N\}$ is a sequence of independent standard Gaussian random variables with mean zero and variance one. The parameters α, p, q (defined in

Appendix C) depend on the parameters σ, θ and on the time interval h, and the initial condition $\bar{u}_0$ can be computed using the statistics of $\bar{u}_n$ also derived in Appendix C.

The experiments have been performed in the following way. For each set of parameter values (U, θ, σ and ϵ) and for each value of N^*, we have considered 60 values of $h\theta$ equally spaced in the range $0 < h\theta < 1$. For each value of $h\theta$ we have run 500 independent simulations characterized by different initial conditions and we have estimated the parameters using (12). We have then computed an approximate value of the mean square error by averaging over simulations. Thus we have the error expression,

$$\hat{Er}^2(A) = \frac{1}{M}\sum_{n=1}^{M}\frac{\langle(\hat{A} - A)\rangle^2}{(A^2)}$$

where $M = 500$. The values have been plotted as functions of $h\theta$ at given N^*, and compared with the theoretical curves.

We discuss here two sets of experiments, EXP1 and EXP2, that are representative of the general results. The two sets differ by the value of the parameter ϵ, which is $\epsilon = 0.0$ km in EXP1 and $\epsilon = 0.6$ km in EXP2. The values of the other parameters are the same in the two sets, and correspond to the ones used in Section 4b(ii) $\sigma = 15$ km/day, $\theta = 1/4$ day^{-1}, $U = 12$ km/day.

The results of EXP1 are shown in Figure 4 as curves of the estimated mean squared errors for the parameters σ, θ and K. A comparison of these curves with the equivalent theoretical curves in Figure 1 shows that the trends predicted by the theory agree with the numerical results. The errors for σ and θ tend to constant values for small $h\theta$ and increase with increasing $h\theta$ as expected, whereas the errors for K are much less sensitive to the value of $h\theta$. At a more quantitative level, for all the estimates, the theory appears to predict with good accuracy the values of the errors for small $h\theta$. At $h\theta \geq 0.5$, instead, the theory seems to underestimate in some cases the errors for σ and θ with regard to the simulations. This effect is more evident for small values of N^* ($N^* = 50, 100$) and it might be due to the fact that the number of actual observations N decreases at increasing $h\theta$ and fixed N^*. The theoretical results, based on the assumption of large N, thus become less reliable in this range.

The results of the second set of experiments, EXP2, are shown in Figure 5. In agreement with the theoretical results in Figure 3, Figures 5a–5c show that the errors for σ, θ and K are altered by the presence of the observation error ϵ only at small $h\theta < 0.1$. In this range of $h\theta$, the quantitative agreement with the theory is very good. The error in ϵ is shown in Figure 5d. With respect to the theoretical curves in Figure 3d, the numerical curve is in good agreement. In conclusion, the numerical

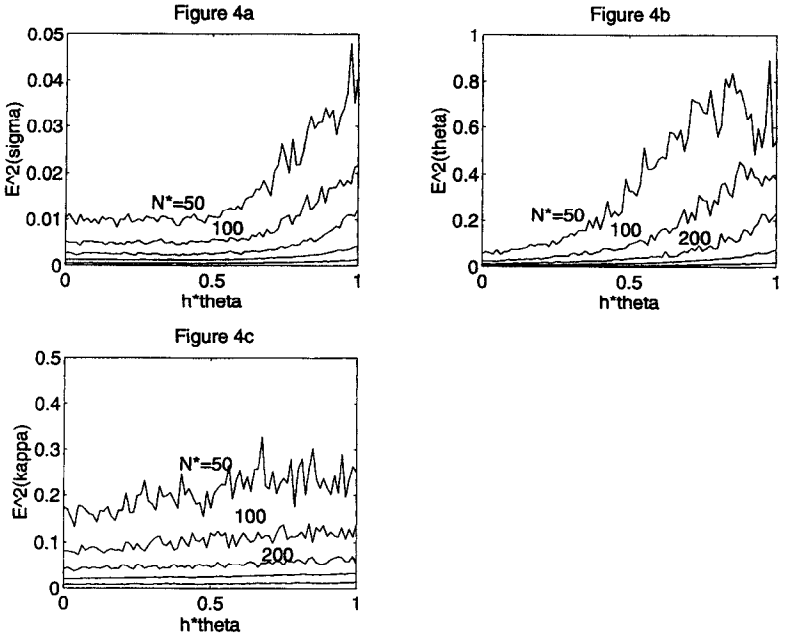

Figure 4. Estimates of mean square sampling errors from numerical simulations $\hat{E}r^2$ versus $h\theta$ at fixed number of independent measurements $N^* = 50, 100, 200, 400, 1000$ computed for $\epsilon = 0$. (a) $\hat{E}r^2(o)$; (b) $\hat{E}r^2(\theta)$; (c) $\hat{E}r^2(K)$.

results confirm the fundamental results of the theory, and suggest some quantiative corrections in the ranges of small N and small h.

a. Comparison with other estimators

Here we compare the properties of our estimates with some of the other estimates used in the literature. The comparison will be restricted to the case $\epsilon = 0$, since the results for the other estimates are usually obtained assuming that the velocity is measured directly and exactly, without an observation error.

As remarked in Section 4, the expression for the sampling errors in U (14) is the same as the one known in the literature, which is derived assuming direct velocity measurements. The error in σ (16) converges asymptotically for small $h\theta$ to the known expression given by Flierl and McWilliams (1977) derived for direct velocity measurements. At values of $h\theta \leq 0.5$, though, the error in (16) increases, due to the peculiar property of the Lagrangian measurements which cannot be captured by the analysis of Flierl and McWilliams. As h increases, in fact, the velocity approximations from the positions become less satisfactorily, degrading the estimate. As shown by the numerical simulations, this feature might be even more pronounced than that suggested by (16), at least for N^* small.

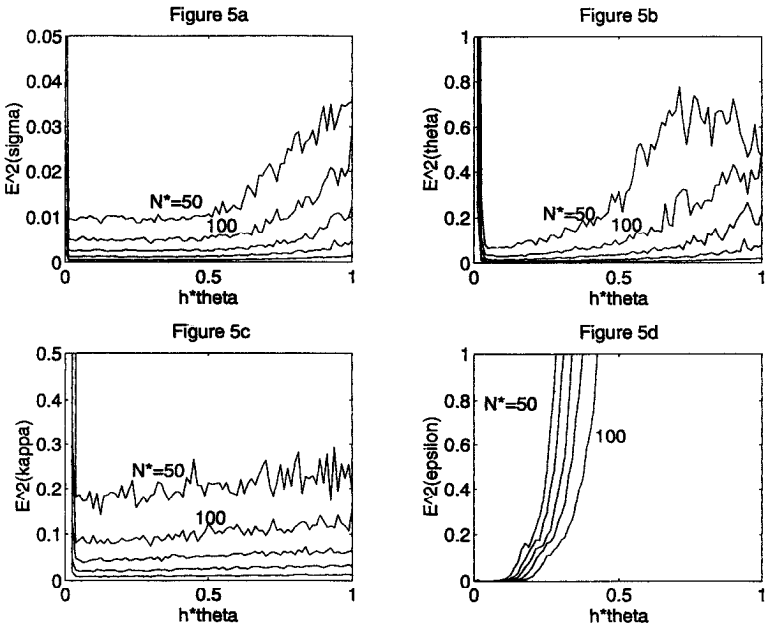

Figure 5. Estimates of mean square sampling errors from numerical simulations $\hat{E}r^2$ versus $h\theta$ at fixed number of independent measures $N^* = 50, 100, 200, 400, 1000$ computed for $\epsilon = 0.6$ km. (a) $\hat{E}r^2(\sigma)$; (b) $\hat{E}r^2(\theta)$; (c) $\hat{E}r^2(K)$; (d) $\hat{E}r^2(\epsilon)$.

As remarked in Section 3, the estimate $\hat{\theta}$ (12) is very different from the estimates of the corresponding parameter $T = 1/\theta$ which are usually found in the literature, and which are based on the equation (13). Eq. (13) depends on asymptotic time lags of the autocorrelation, and T is usually approximated by integrating up to a truncation time $T_{tr} < T_{obs}$. The choice of T_{tr} plays a very important role because it can influence the value of T up to order one, and unfortunately there is no theoretical procedure for determining T_{tr} (a common choice is the first zero crossing of the autocorrelation, e.g. Krauss and Boning, 1987). The uncertainty in the choice of T_{tr} introduces a strong element of subjectivity in the estimates, and the values of T computed in the literature are often not provided with error bounds. In comparison, our estimate $\hat{\theta}$ is based on the a-priori assumption of the stochastic model (2)–(3). With respect to this assumption, the estimate is "objective" and the error analysis is consistent.

A similar discussion applies also to the estimates of K, which are often obtained directly from the estimates of T using the relationship $K = \sigma^2 T$. Other, more general estimates of K have also been used in the literature. As an example, Davis (1991b) has proposed a time-dependent estimate of K valid for inhomogeneous flows. This estimate has the advantage of being more general than ours, but on the other hand it has the same fundamental problem of depending on the long time behavior of the system like the estimates based on (13).

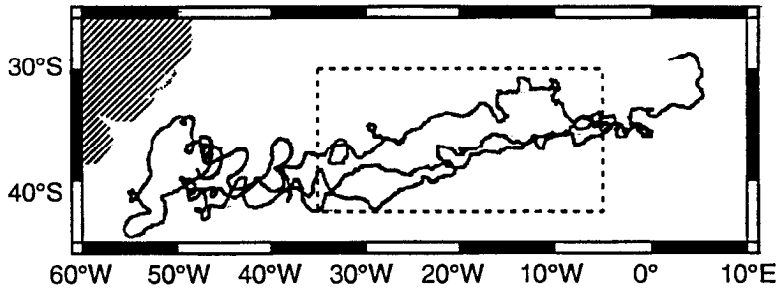

Figure 6. Trajectories of the three SOS drifters 85, 86, 87. The dashed box correspond to the region where the estimates are computed.

6. Application to drifter data in the Brazil-Malvinas extension

a. The data

In this section we present an application of the parameter estimation (12) based on a subset of data taken from a set of surface drifting buoys deployed in 1984 in the Brazil/Malvinas extension as part of the ONR sponsored SOS (Southern Ocean Studies) initiative. The total SOS data set has been analyzed by Figueroa and Olson (1989, FO in the following) in conjunction with another data set (FGGE, i.e. First Garp Global Experiments) to give a statistical description of several regions in the South Atlantic. The aim of the following analysis is not to add new oceanographical results to the ones of FO, but rather to provide a simple example of how our theoretical results can be applied in practice and of how they compare with results obtained with different methods.

The SOS subset considered gives position data for three drifters (85, 86 and 87) in the region extending from 35W to 5W in longitude and from 43S to 30S in latitude, for a total of 800 days. See Figure 6. The region lies in the open ocean extension of the Braxil/Malvinas current. As shown by FO, even though still highly turbulent, this region is characterized by less variability with respect to the more western region of the confluence, and has a well defined northeastern average current. The original data, provided by the Argos system, are taken at irregular time intervals with an average of about 5 fixes per day. Before being used in the parameter estimation, they have been filtered using a triangular filter to obtain one value a day (Zambianchi and Griffa, 1994b, ZG in the following).

b. Model applicability

In order to apply the estimation procedure (12) to the data, the applicability of the model (2)–(3) in the region of interest must be first verified. A preliminary analysis on the compatibility of the model with the data is presented in ZG. Here we summarize the main results of ZG and present some new evidence of the model's applicability. First of all, we recall that the model (2)–(3) is based on the assumptions

that the flow is homogeneous and that the two velocity components are independent. Let us indicate with z_{EW} and z_{NS} the East-West and North-South approximations of the velocity computed from the data according to (7). The assumption of independence has been verified by ZG calculating the mean coherence between z_{EW} and z_{NS}. Regarding the homogeneity a first, rough test has been performed by ZG computing the parameters U and σ from the data for both components, first over the whole region and then over two subregions, from 35 to 20W and from 20 to 5W respectively. The results are found to agree within the limit of the sampling error. This simple test is far from being conclusive, given the relatively small number of data available, but at least it indicates that there are no obvious discrepancies with the basic assumptions.

A more significant way of testing the model consists in verifying the exponential form of the assumed autocorrelation, or analogously, its spectral form. The spectral shape has been studied by ZG. Here we concentrate on the form of the autocorrelation. The autocorrelations computed from the data for the two velocity components z_{EW}, z_{NS} are shown in Figure 7a–7b (dashed lines). At a first inspection, it appears evident that the shape of the two autocorrelations for short time lags is, at least qualitatively, in agreement with the exponential shape assumed by the model, indicating that the scales of the acceleration can be safely neglected. At longer time lags, though, the shape of the autocorrelations significantly deviates from exponential, showing zero crossings, negative lobes and oscillations. The point that we want to investigate is whether these effects are due to different dynamics than the one of the simple model (2)–(3), or whether the data can actually be considered as a particular realization of (2)–(3) which deviates from the exponential shape only because of the relatively small number of data points. In order to address this question, we have used the model to generate a set of realizations of simulated velocities, obtained by setting the parameters to the values estimated from the data (see Table 1). The number of measurements N and the sampling interval h for the simulations are taken to be the same as the ones of the data: $N = 800, h = 1$ day. We have then compared the data and the simulations by plotting together the autocorrelation functions (examples for three realizations are shown in Figure 7a–7b, continuous lines). The results strongly suggest that the data are indeed consistent with the model, since data and simulations appear to be essentially indistinguishable, with the data autocorrelation falling right into the envelope of the simulated ones. Even though this test cannot prove positively that the model is completely accurate, it certainly indicates that it is a reasonable and valid starting point.

c. Parameter estimation and error analysis

Once the applicability of the model has been verified, the estimate of the parameters U, σ, θ and K can be evaluated in a straightforward way from the data, computing the moments (11) and then solving the system (12). The results for the

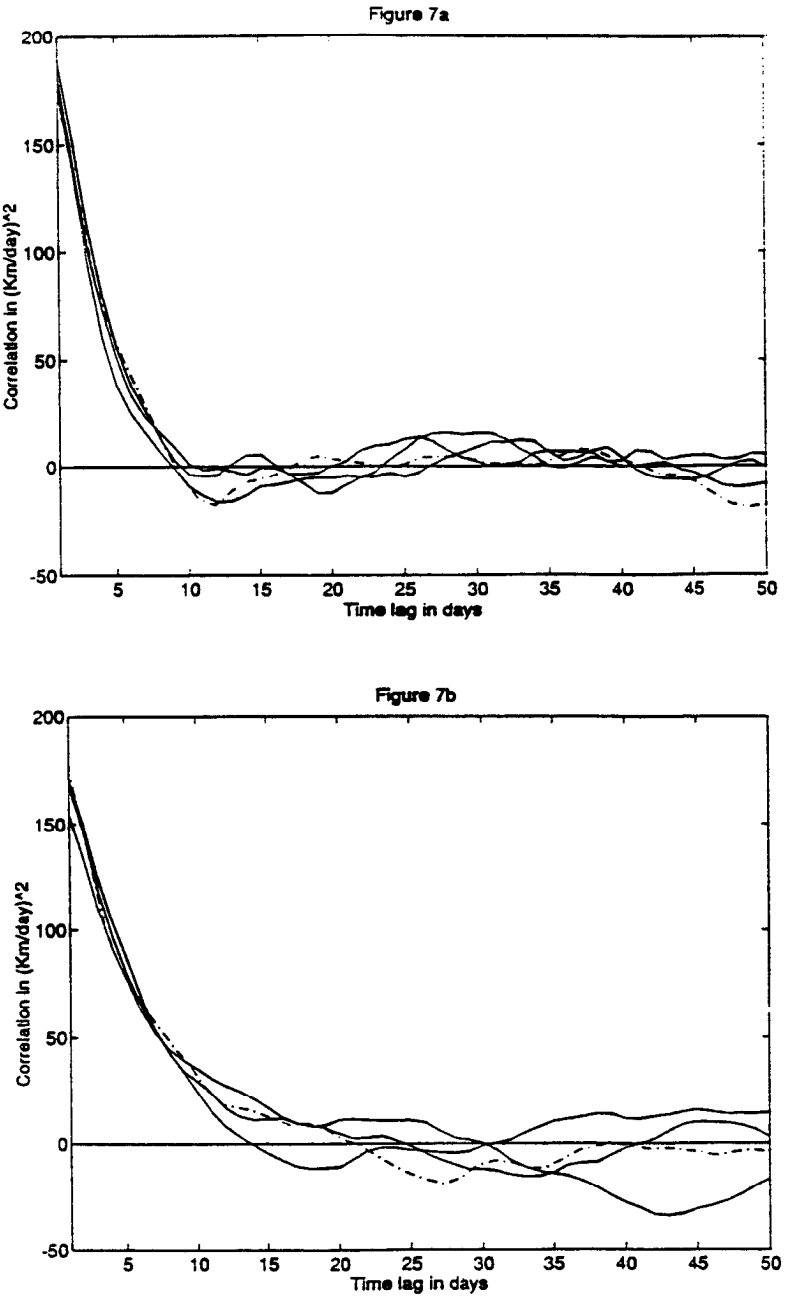

Figure 7. Non-normalized autocorrelations for the North-South (a) and East-West (b) velocity components. Dashed lines correspond to velocities computed from the data. Continuous lines correspond to velocities computed from numerical simulations.

Table 1

	$\hat{U}$ *K* m/day	$\hat{\sigma}$ *K* m/day	$1/\hat{\theta}$ days	$\hat{K}$ *K* m²/day
E.W. Component	7.5	13	5	845
N.S. Component	2	14	2.8	560

two velocity components are shown in Table 1. As expected, they show that there is a well defined eastward average motion, with a small northward component. The turbulent motion appears to be quite energetic, and anisotropic in the time scales.

In order to compute the errors in the estimates, the values of $N^* = Nh\theta$ and $h\theta$ for the two components are first evaluated using the estimate of θ in Table 1. They turn out to be $N^* = 160$ and $h\theta = 0.2$ for the East-West component and $N^* = 280$ and $h\theta = 0.35$ for the North-South component. Notice that for these values the theoretical estimates of the errors are accurate (compare Fig. 1–2 and 3–4), so that no corrections from the numerical results are needed. The values of the mean square errors for the four estimates are computed from (14)–(17), and the corresponding percentage errors are reported in Table 2. A very high error (88%) is found for the estimate of the North-South velocity. This is not surprising, since the fluctuation is much higher than the mean for this component. The errors in the variance σ are quite small for both components (4–5%), and the errors in θ are higher but still acceptable (10–15%). Notice that the error analysis confirms that the difference in the East-West and North-South time scales is significant, so that the anisotropy suggested by the value in Table 1 is real.

We conclude with a brief comparison of our results with the time scale estimates of T (and of the diffusivity K) given by FO. FO choose a truncation time T_{tr} coinciding with 50% of the longest velocity time series in the region. The arbitrary value of 50% is chosen in order to guarantee the stability of the estimate and, as shown by Figure 3 in FO, plays an important role in determining the value of the estimates. The values obtained by FO in two regions overlapping with our region (see Fig. 2 FO) are qualitatively similar to ours, even though they show an overestimate of T in the East-West direction and an underestimate in the North-South direction. This is likely to be an effect of considering the individual "bumps" in the autocorrelations near the zero crossing, which are eliminated in our estimates.

7. Summary and discussion

A new parametric approach to the study of Lagrangian data is presented, based on a stochastic model for particle motion. The model assumes an exponential autocorre-

Table 2

	N^*	$h\theta$	$Er(U)$	$Er(\sigma)$	$Er(\theta)$	$Er(K)$
E.W. Component	160	0.2	20%	5.6%	14%	23%
N.S. Component	280	0.35	88%	4.3%	12%	18%

lation for the turbulent velocity and is appropriate for mid-latitude mesoscale flows in homogeneous regions of the upper ocean where the scales of the acceleration do not play an important role. Estimates of the mean velocity and of the turbulent parameters are derived from particle position data. A complete error analysis of the estimates is presented which specifically takes into account the Lagrangian nature of the data and considers the dependence of the errors on the sampling interval. The results are derived in presence of an arbitrary measurement error. For satellite tracked drifters and for realistic parameters and sampling, the quality of the estimates does not appear to be influenced by the observation error.

If compared to nonparametric methods where less stringent assumptions are necessary (e.g. Davis, 1991b), our method has the disadvantage of more restricted applicability. Prior to using the method, the compatibility of the model with the data must be studied, as shown in the example of the Brazil/Malvinas data considered here. On the other hand, the method provides relevant advantages. The effects of the observation error are directly included in the estimates and in the error analysis, and the accuracy of the estimates is enhanced. In particular, the estimates of T and K depend on the slope of the autocorrelation and are computed using only the first two time lags. With respect to the other estimates in the literature which depend on subjective procedures of truncation of the long time lags of the correlations, they are "objective" and have a well defined associated error.

Even though the results presented here are valid only for the simple homogeneous model, the methodology is very general and can be applied to more complete and less restrictive models. Work is presently under way on a generalization of the model which will include dependent velocity components, mean shear and simple turbulence inhomogeneity. Also, a model which includes deep ocean motion by explicitly resolving the acceleration scales, is under consideration.

Acknowledgments. The authors would like to thank S. Molchanov, D. Olson, C. Rooth and M. Swenson for interesting comments and discussions. The drifter data, gathered in the framework of ONR-sponsored SOS initiative, were made available by D. Olson. This work was supported by the Office of Naval Research (Grant N00014-91-J1346 and N-00014-91-J1526).

APPENDIX A

Method of moments

Recall that the velocity approximations satisfy (8)–(9),

$$z_n = U + \bar{u}_n + \epsilon \frac{\xi_n - \xi_{n-1}}{h}. \qquad [A1]$$

We will now compute the moments of z_n and we will often use the following facts in the derivation. First, the average value of cross terms involving $\bar{u}_n$ and ξ_k is zero since they are independent and $\langle \xi_k \rangle = 0$. Second, $\langle \xi_i \xi_j \rangle = \delta_{ij}$ where δ_{ij} is the Kronecker

delta. Using [A1] we see that,

$$\langle z_n - U\rangle = \langle \bar{u}_n\rangle + \epsilon\left\langle \frac{\xi_n - \xi_{n-1}}{h}\right\rangle = \langle \bar{u}_n\rangle$$

$$\langle (z_n - U)^2\rangle = \langle \bar{u}_n^2\rangle + \frac{\epsilon^2}{h^2}\left(\langle \xi_n^2\rangle - 2\langle \xi_n \xi_{n-1}\rangle + \langle \xi_{n-1}^2\rangle\right)$$

$$= \langle \bar{u}_n^2\rangle + 2\frac{\epsilon^2}{h^2}$$

$$\langle (z_n - U)(z_{n+1} - U)\rangle = \langle \bar{u}_n \bar{u}_{n+1}\rangle + \frac{\epsilon^2}{h^2}\left(\langle \xi_n \xi_{n+1}\rangle - \langle \xi_n \xi_n\rangle\right.$$

$$\left. - \langle \xi_{n-1}\xi_{n+1}\rangle + \langle \xi_{n-1}\xi_n\rangle\right) = \langle \bar{u}_n \bar{u}_{n+1}\rangle - \frac{\epsilon^2}{h^2}$$

For $j \geq 2$ $\langle (z_n - U)(z_{n+j} - U)\rangle = \langle \bar{u}_n \bar{u}_{n+j}\rangle$.

Hence, we have reduced the problem of calculating the moments of the velocity observations z_n to the problem of calculating the moments of $\bar{u}_n$.

The calculation of the moments of $\bar{u}_n$ is as follows.

$$\langle \bar{u}_n\rangle = \frac{1}{h}\int_{t_{n-1}}^{t_n} \langle u(s)\rangle ds = 0$$

$$\langle \bar{u}_n^2\rangle = \frac{1}{h^2}\int_{t_{n-1}}^{t_n}\int_{t_{n-1}}^{t_n} \langle u(t)u(s)\rangle dt\, ds = \frac{1}{h^2}\int_{t_{n-1}}^{t_n}\int_{t_{n-1}}^{t_n} \sigma^2 e^{-\theta|t-s|} dt\, ds$$

$$= \frac{2}{h^2}\int_{t_{n-1}}^{t_n}\int_{s}^{t_n} \sigma^2 e^{-\theta(t-s)} dt\, ds = \frac{2\sigma^2}{h^2\theta^2}(\theta h - 1 + e^{-\theta h}).$$

In a similar fashion we calculate for $j \geq 1$,

$$\langle \bar{u}_{n+j}\bar{u}_n\rangle = \frac{\sigma^2}{h^2\theta^2} e^{-(j-1)\theta h}(e^{-\theta h} - 1)^2.$$

Substituting these expressions for the moments of $\bar{u}_n$ we find

$$\langle z_n\rangle = U \qquad \text{[A2]}$$

$$r_0 = \langle z_n^2\rangle - \langle z_n\rangle^2 = \left\{\frac{2\epsilon^2}{h^2\sigma^2} + \frac{2}{h^2\theta^2}(e^{-\theta h} - 1 + \theta h)\right\}\sigma^2 \qquad \text{[A3]}$$

$$r_1 = \langle z_n z_{n+1}\rangle - \langle z_n\rangle^2 = \left\{-\frac{\epsilon^2}{h^2\sigma^2} + \frac{1}{h^2\theta^2}(1 - e^{-\theta h})^2\right\}\sigma^2 \qquad (18)$$

$$r_2 = \langle z_n z_{n+2}\rangle - \langle z_n\rangle^2 = \frac{\sigma^2}{h^2\theta^2}(1 - e^{-\theta h})^2 e^{-\theta h}.$$

We can approximate U, r_0, r_1 and r_2 with the estimates $\hat{U}, \hat{r}_0, \hat{r}_1$ and $\hat{r}_2$ computed from the data as indicated in (11). Thus, the parameters will satisfy the approximate equations

$$\begin{aligned}
\hat{r}_0 &= \langle z_n^2 \rangle - \langle z_n \rangle^2 = \left\{ \frac{2\epsilon^2}{h^2\sigma^2} + \frac{2}{h^2\theta^2}(e^{-\theta h} - 1 + \theta h) \right\} \sigma^2 \\
\hat{r}_1 &= \langle z_n z_{n+1} \rangle - \langle z_n \rangle^2 = \left\{ -\frac{\epsilon^2}{h^2\sigma^2} + \frac{1}{h^2\theta^2}(1 - e^{-\theta h})^2 \right\} \sigma^2 \\
\hat{r}_2 &= \langle z_n z_{n+2} \rangle - \langle z_n \rangle^2 = \frac{\sigma^2}{h^2\theta^2}(1 - e^{-\theta h})^2 e^{-\theta h}
\end{aligned} \tag{19}$$

Notice that,

$$\frac{2\hat{r}_1 + \hat{r}_0}{\hat{r}_2} = f(h\hat{\theta}) \tag{A4}$$

where

$$f(x) = \frac{2(1 - e^{-x})^2 + 2(e^{-x} - 1 + x)}{(1 - e^{-x})^2 e^{-x}}.$$

We solve Eq. (A4) numerically using the bisection method and substitute this result back into the system to obtain the estimates for the parameters given in the text.

APPENDIX B

Sampling mean squared error formulas

First we will compute the error in U. We will then demonstrate the general procedure for estimating the mean squared error for the rest of the papameters by deriving in detail the mean squared error in θ.

$$\begin{aligned}
\langle (\hat{U} - U)^2 \rangle &= \langle \hat{U}^2 \rangle - U^2 = \frac{1}{N^2} \sum_{i,j=1}^{N} (\langle z_i z_j \rangle - U^2) \\
&= \frac{1}{N^2} \sum_{i,j=1}^{N} r_{i-j} = \frac{1}{N^2} \sum_{k=-N}^{N} (N - |k|) r_k \\
&\approx \frac{1}{N} \sum_{k=-\infty}^{\infty} r_k = \frac{1}{N} \left(r_0 + 2r_1 + 2r_2 \frac{1}{1 - \alpha} \right) = \frac{2\sigma^2}{Nh\theta}
\end{aligned} \tag{B1}$$

where $r_k = \langle (z_i - U)(z_{i-k} - U) \rangle$, $\alpha = e^{-h\theta}$ and we used formulas [A2] and [A3] from Appendix A for r_0, r_1 and r_2.

Now we proceed to the error estimates for the other parameters. Recall that the

parameter estimate for θ is

$$\hat{\theta} = \frac{1}{h} f^{-1}\left(\frac{2\hat{r}_1 + \hat{r}_0}{\hat{r}_2}\right)$$

where $f(\cdot)$ is a known function specified in Appendix A, [A4]. Let $F(a, b, c) = 1/h f^{-1}((a + 2b)/c)$ then

$$F(r_0, r_1, r_2) = \theta \text{ and } F(\hat{r}_0, \hat{r}_1, \hat{r}_2) = \hat{\theta}.$$

The mean squared error is given by $\langle(\hat{\theta} - \theta)^2\rangle$ and using the above definition we can write $\langle(\hat{\theta} - \theta)^2\rangle = \langle(F(\hat{r}_0, \hat{r}_1, \hat{r}_2) - F(r_0, r_1, r_2))^2\rangle$. This expression can be approximated as

$$\langle(\hat{\theta} - \theta)^2\rangle \approx \left\langle\left(\frac{\partial F}{\partial r_0}(\hat{r}_0 - r_0) + \frac{\partial F}{\partial r_1}(\hat{r}_1 - r_1) + \frac{\partial F}{\partial r_2}(\hat{r}_2 - r_2)\right)^2\right\rangle = \left(\frac{\partial F}{\partial r_0}\right)^2 \gamma_{00} + \left(\frac{\partial F}{\partial r_1}\right)^2 \gamma_{11}$$
$$+ \left(\frac{\partial F}{\partial r_2}\right)^2 \gamma_{22} + 2\frac{\partial F}{\partial r_0}\frac{\partial F}{\partial r_1}\gamma_{01} + 2\frac{\partial F}{\partial r_0}\frac{\partial F}{\partial r_2}\gamma_{02} + 2\frac{\partial F}{\partial r_1}\frac{\partial F}{\partial r_2}\gamma_{12}$$

where $\gamma_{p,q} = \langle(\hat{r}_p - r_p)(\hat{r}_q - r_q)\rangle$. Using the fact that $\hat{r}_p \to r_p$ with probability one as N goes to infinity, we can write

$$\gamma_{p,q} = \langle\hat{r}_p\hat{r}_q\rangle - r_p r_q$$

where you will recall that $r_p = \langle z_n z_{n-p}\rangle$ and $\hat{r}_p = 1/(N-p)\sum_{n=1}^{N-p} z_{n+p}z_n$. Hence,

$$\langle\hat{r}_p\hat{r}_q\rangle = \frac{1}{(N-p)(N-q)}\sum_{i=1}^{N-p}\sum_{j=1}^{N-q}\langle z_i z_{i+p} z_j z_{j+q}\rangle.$$

Further it can be shown that z_n is Gaussian and thus the latter expression can be simplified using the following standard fact about Gaussian random variables. If ξ_1, ξ_2, ξ_3, ξ_4 are Gaussian random variables then

$$\langle\xi_1\xi_2\xi_3\xi_4\rangle = \langle\xi_1\xi_2\rangle\langle\xi_3\xi_4\rangle + \langle\xi_1\xi_3\rangle\langle\xi_2\xi_4\rangle + \langle\xi_1\xi_4\rangle\langle\xi_2\xi_3\rangle.$$

Using the fact that $\langle z_i z_{i+p}\rangle\langle z_j z_{j+q}\rangle = r_p r_q$ we can thus write

$$\gamma_{p,q} = \frac{1}{(N-p)(N-q)}\sum_{i=1}^{N-p}\sum_{j=1}^{N-q}(r_{i-j}r_{i-j+p-q} + r_{i-j-q}r_{i-j+p}).$$

Since the subscripts in the above expression depend only on the i-j for large N we can write

$$\gamma_{p,q} = \frac{1}{N^2}\sum_{l=-N}^{N}(N - |l|)(r_l r_{l+p-q} + r_{l-q}r_{l+p}) + o\left(\frac{1}{N}\right).$$

Expanding this expression and keeping terms to $o(1/N)$ we get

$$\gamma_{p,q} = \frac{1}{N}\sum_{l=-N}^{N}(r_l r_{l+p-q} + r_{l-q}r_{l+p}) + o\left(\frac{1}{N}\right)$$

Since the series is convergent we can write

$$\gamma_{p,q} = \frac{1}{N}\sum_{l=-\infty}^{\infty}(r_l r_{l+p-q} + r_{l-q}r_{l+p}) + o\left(\frac{1}{N}\right).$$

This formula is the basis for further calculations. In particular if $r_k = r_2\alpha^{k-2}$ for $k \geq 2$ (as we have in our model) we obtain by summation of the above infinite series the following formulas for the first moments

$$\tilde{\gamma}_{00} = 2\left(r_0^2 + 2r_1^2 + \frac{2r_2^2}{1-\alpha^2}\right)$$

$$\tilde{\gamma}_{10} = 4\left(r_0 r_1 + r_1 r_2 + \frac{\alpha r_2^2}{1-\alpha^2}\right)$$

$$\tilde{\gamma}_{20} = 2\left(r_1^2 + 2r_0 r_2 + 2r_1 r_2\alpha + \frac{2r_2^2\sigma^2}{1-\alpha^2}\right)$$

$$\tilde{\gamma}_{11} = \frac{\tilde{\gamma}_{00} + \tilde{\gamma}_{20}}{2}$$

$$\tilde{\gamma}_{12} = \frac{\tilde{\gamma}_{10}}{2} + 2\left(r_0 r_2\alpha + r_1 r_2(1+\alpha^2) + \frac{r_2^2\alpha^3}{1-\alpha^2}\right)$$

$$\tilde{\gamma}_{22} = \frac{\tilde{\gamma}_{00}}{2} + 2r_1 r_2\alpha(1+\alpha^2) + 2r_0 r_2\alpha^2 + \frac{r_2^2(1-\alpha^2+2\alpha^4}{1-\alpha^2}$$

where $\tilde{\gamma}_{pq} = N\tilde{\gamma}_{pq}$.
Substituting these results into the mean squared error expression we find

$$\frac{\langle(\hat{\theta}-\theta)^2\rangle}{\theta^2} = \frac{f_1(h\theta)}{Nh\theta} \qquad \text{[B2]}$$

where

$$f_1(x) = \frac{1}{x(f'(x))^2 r_2^2(x)}(\tilde{\gamma}_{00} + 4\tilde{\gamma}_{01} + f^2(x)\tilde{\gamma}_{22} - 2f(x)(\tilde{\gamma}_{02} + 2\tilde{\gamma}_{12})).$$

Using a similar procedure we obtain the following mean squared error estimates for

the other parameters

$$\frac{\langle(\hat{\sigma} - \sigma)^2\rangle}{\sigma^2} = \frac{f_2(h\theta)}{Nh\theta} \qquad \text{[B3]}$$

where

$$f_2(x) = \frac{\tilde{\gamma}_{22}x}{4r_2^2} + \psi_0^2(x)x^2 f_1(x) + \frac{\psi_0(x)x}{r_2 f'(x)\left(\frac{\tilde{\gamma}_{02}}{r_2} + \frac{2\tilde{\gamma}_{12}}{r_2} - \frac{(r_0 + 2r_1)\tilde{\gamma}_{22}}{r_2^2}\right)}$$

$$\text{and } \psi_0(x) = \frac{1}{2} + \frac{1}{x} - \frac{e^{-x}}{1 - e^{-x}}; \qquad \text{[B4]}$$

$$\frac{\langle(\hat{K} - K)^2\rangle}{K^2} = \frac{f_3(h\theta)}{Nh\theta}$$

where

$$f_3(x) = \frac{\tilde{\gamma}_{22}x}{r_2^2} + \psi^2(x)x^2 f_1(x) + \frac{2\psi(x)x}{r_2 f'(x)\left(\frac{\tilde{\gamma}_{02}}{r_2} + \frac{2\tilde{\gamma}_{12}}{r_2} - \frac{(r_0 + 2r_1)\tilde{\gamma}_{22}}{r_2^2}\right)}$$

$$\text{and } \psi(x) = 2\psi_0(x) - \frac{1}{x}; \qquad \text{[B5]}$$

$$\frac{\langle(\hat{\epsilon} - \epsilon)^2\rangle}{\epsilon^2} = \frac{f_4(h\theta)}{Nh\theta}$$

where

$$f_4(x) = \frac{1}{4(r_2e^x - r_1)^2}\left[r_2^2e^{2x}xf_1(x) + \tilde{\gamma}_{11}\left(1 - \frac{4e^x}{f'(x)}\right) + \tilde{\gamma}_{22}e^{2x}\left(1 - \frac{2f(x)}{f'(x)}\right)\right.$$
$$\left. + 2\tilde{\gamma}_{12}e^x\left(\frac{(f(x) + 2e^x)}{f'(x)} - 1\right) + \tilde{\gamma}_{20}\frac{2e^{2x}}{f'(x)} - \tilde{\gamma}_{10}\frac{2e^x}{f'(x)}\right].$$

Note that $f_i(x)$ depend only on x when $\epsilon = 0$. However, when $\epsilon > 0$ we must keep in mind that $f_i(x) = f_i(x;\epsilon/h\sigma)$ because r_0, r_1 and γ_{pq} depend on $\epsilon/h\sigma$ as well.

APPENDIX C

Autoregressive property

Recall that $\bar{u}_n = 1/h \int_{t_{n-1}}^{t_n} u(t)dt$ and that in Appendix A we derived the following

$$\langle \bar{u}_n \bar{u}_{n+j} \rangle = \begin{cases} \frac{\sigma^2}{h^2\theta^2}(1 - e^{-h\theta})^2 e^{-(j-1)h\theta} & j \geq 1 \\ \frac{2\sigma^2}{h^2\theta^2}(e^{-h\theta} - 1 + h\theta) & j = 0 \end{cases}$$

We will show that the process $\bar{u}_n$ can be represented as an autoregressive moving average process that satisfies,

$$\bar{u}_n = \alpha \bar{u}_{n-1} + p\eta_n + q\eta_{n-1} \qquad \text{[C1]}$$

where $\{\eta_n, n = 1 \ldots, N\}$ is an independent sequence of standard Gaussian random variables with zero mean and variance one, $\alpha = e^{-h\theta}$, and

$$p = \frac{\sigma}{\sqrt{2h\theta}}\left(1 - e^{-h\theta} + \sqrt{(1 + e^{-h\theta})^2 - \frac{2(1 - e^{-2h\theta})}{h\theta}}\right)$$

$$q = \frac{\sigma}{\sqrt{2h\theta}}\left(1 - e^{-h\theta} - \sqrt{(1 + e^{-h\theta})^2 - \frac{2(1 - e^{-2h\theta})}{h\theta}}\right).$$

Let x_n be an arbitrary process that satisfies

$$x_n = \alpha x_{n-1} + p\eta_n + q\eta_{n-1}. \qquad \text{[C2]}$$

Since x_n and $\bar{u}_n$ are both Gaussian showing that they have the same correlation function is equivalent to showing that they are the same process. So, let us begin by computing the correlation function for x_n. Let x_0 be equal to a zero mean Gaussian random variable and let us adopt a compact notation for the correlation function $\langle x_n x_{n+j}\rangle$. Let $\langle x_n x_{n+j}\rangle = r_j$ for $j \geq 0$. Then taking the first moment of x_n we find

$$\langle x_n\rangle = \langle \alpha x_{n-1}\rangle.$$

This equation implies that $\langle x_n\rangle = 0$ for all $n \geq 0$, since $\langle x_0\rangle = 0$ by assumption. Taking second moments yields

$$\langle x_n x_n\rangle = \langle \alpha^2 x_{n-1}^2\rangle + 2\alpha\,\langle x_{n-1}(p\eta_n + q\eta_{n-1})\rangle + p^2 + q^2.$$

Using [C2] we get

$$\langle x_n x_n\rangle = \langle \alpha^2 x_{n-1}^2\rangle + 2\alpha\,\langle x_{n-1}(x_n - \alpha x_{n-1})\rangle + p^2 + q^2.$$

Using our compact notation this becomes

$$r_0 = \alpha^2 r_0 + 2\alpha(r_1 - \alpha r_0) + p^2 + q^2. \qquad \text{[C3]}$$

Computing $\langle x_n x_{n-1}\rangle$ and using [C2] we get

$$\langle x_n x_{n-1}\rangle = \langle[\alpha x_{n-1} + p\eta_n + q\eta_{n-1}][\alpha x_{n-2} + p\eta_{n-1} + q\eta_{n-2}]\rangle = \alpha^2\langle x_{n-1}x_{n-2}\rangle$$
$$+ \alpha\,\langle x_{n-1}(x_{n-1} - \alpha x_{n-2})\rangle + \alpha\langle x_{n-2}(x_n - \alpha x_{n-1})\rangle + \langle(p\eta_n + q\eta_{n-1})(p\eta_{n-1} + q\eta_{n-2})\rangle$$

or in a simpler notation

$$r_1 = \alpha^2 r_1 + \alpha(r_0 - \alpha r_1) + \alpha(r_2 - \alpha r_1) + pq \qquad \text{[C4]}$$

which implies

$$(1 + \alpha^2)r_1 - \alpha r_2 - \alpha r_0 = pq.$$

Performing similar calculations we find for $j \geq 2$

$$(1 + \alpha^2)r_j = \alpha(r_{j-1} + r_{j+1}). \qquad [C5]$$

Now let us check that $\bar{u}_n$ satisfies these moment conditions. Let us define $\langle \bar{u}_0^2 \rangle = \rho_0$ and notice that $\langle \bar{u}_n \bar{u}_{n-j} \rangle = \rho\alpha^{j-1}$ where $\rho = \sigma^2/h^2\theta^2(1 - e^{-h\theta})^2$. Upon substituting the moments $\bar{u}_n$ for $j \geq 2$ into [C5] we see that

$$(1 + \alpha^2)\rho\alpha^{j-1} = \alpha(\rho\alpha^{j-2} + \rho\alpha^j).$$

This implies that $\alpha^{j-1} + \alpha^{j+1} = \alpha^{j-1} + \alpha^{j+1}$; which is an identity. Thus, all of the moments of $\bar{u}_n$ for $j \geq 2$ agree with those of the autoregressive solution x_n. Thus we must check that the first and second moments of $\bar{u}_n$ and x_n agree. To do this, substitute the first and second moments of $\bar{u}_n$ into [C3] and [C4]. We see that

$$\rho_0(1 + \alpha^2) - 2\alpha\rho = p^2 + q^2$$

$$\rho(1 + \alpha^2) - \alpha^2\rho - \alpha\rho_0 = pq.$$

After substituting the value of ρ, ρ_0, p and q it can be shown that this system of equations holds. Thus, $\bar{u}_n$ and x_n have the same correlation function. Since $\bar{u}_n$ and x_n have the same correlation function and are both Gaussian, the random vectors $\mathbf{X} = (x_1, x_2, \ldots, x_N)$ and $Y = (\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_N)$ have the same multivariate Gaussian distributions. Therefore, $\bar{u}_n$ can be represented as a solution of the above autoregressive process x_n provided that we set $\langle x_0^2 \rangle = \langle \bar{u}_0^2 \rangle$.

REFERENCES

Chandrasekhar, S. 1943. Stochastic problems in physics and astronomy. Rev. Modern Physics, *15,* 1–89.

Colin de Verdiere, A. C. 1983. Lagrangian eddy statistics from surface drifters in the eastern North Atlantic. J. Mar. Res., *41,* 375–398.

Csanady, G. T. 1980. Turbulent Diffusion in the Environment, Dordrecht, D. Reidel Publ. Co., 248 pp.

Davis, R. E. 1983. Oceanic property transport, Lagrangian particle statistics, and their prediction. J. Mar. Res., *41,* 163–194.

—— 1985. Drifter observations of coastal surface currents during CODE: The statistical and dynamical views. J. Geophys. Res., *90,* 4756–4772.

—— 1987. Modelling eddy transport of passive tracers. J. Mar. Res., *45,* 635–666.

—— 1991a. Lagrangian ocean studies. Ann. Rev. Fluid Mech., *23,* 43–64.

—— 1991b. Observing the general circulation with floats. Deep Sea Res., *38,* S531–S571.

Dutkiewicz, S., A. Griffa and D. B. Olson. 1993. Particle diffusion in a meandering jet. J. Geophys. Res., *93,* 16,487–16,500.

Figueroa, H. A. and D. B. Olson. 1989. Lagrangian statistics in the South Atlantic as derived from SOS and FGGE drifters. J. Mar. Res., *47,* 525–546.

—— 1994. Eddy resolution vs. eddy diffusion in a double gyre GCM Part I: The Lagrangian and Eulerian description. J. Phys. Oceanogr., *24,* 371–386.

Flierl, G. R. and J. C. McWilliams. 1977. On the sampling requirements for measuring moments of eddy variability. J. Mar. Res., *35,* 797–820.

Holloway, G. 1989. Subgridscale representation in Oceanic Circulation Models: Combining Data and Dynamics. L. T. Anderson and J. Willebrand, eds., Kluwer, NATO ASI Series C 284, 513–587.

Krauss, W. and C. W. Boning. 1987. Lagrangian properties of eddy fields in the northern North Atlantic as deduced from satellite-tracked buoys. J. Mar. Res., *45,* 259–291.

Pasquill, F. and F. B. Smith. 1983. Atmospheric diffusion, 3rd Ed. Halsted Press, New York, 437 pp.

Pope, S. B. 1994. Lagrangian PDF methods for turbulent flows. Ann. Rev. Fluid Mech., *26,* 23–63.

Riser, S. C. and H. T. Rossby. 1983. Quasi-Lagrangian structure and variability of the subtropical western North Atlantic circulation. J. Mar. Res., *41,* 127–162.

Sawford, B. L. 1991. Reynolds number effects in Lagrangian stochastic models of turbulent dispersion. Phys. Fluids A, *3*(6), 1577–1586.

Taylor, G. I. 1921. Diffusion by continuous movements. Proc. London Math. Soc., *20,* 196–212.

Thomson, D. J. 1986. A random walk model of dispersion in turbulent flows and its application to dispersion in a valley. Q.J.R. Meteor. Soc., *112,* 511–530.

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

Van Dop, H., F. T. M. Niewstand and J. C. R. Hunt. 1985. Random walk models for particle displacements in inhomogeneous unsteady turbulent flows. Phys. Fluids, *28,* 1639–1653.

Verron, J. and K.-D. Nguyen. 1989. Lagrangian diffusivity estimates from a gyre-scale numerical experiment on float tracking. Oceanologica Acta, *11,* 167–176.

Yaglom, A. M. 1962. Stationary Random Functions. Dover 69–71.

Zambianchi, E. and A. Griffa. 1994a. Effects of finite scales of turbulence on dispersion estimates. J. Mar. Res., *52,* 129–149.

—— 1994b. A comparison between a stochastic model for particle motion and drifter data in the Brazil/Malvinas extension. Annales Istituto Universitario Navale.

Received: *26 April, 1994;* revised: *14 November, 1994.*