Predictability of Drifter Trajectories in the Tropical Pacific Ocean

TAMAY M. ÖZGÖKMEN

RSMAS/MPO, University of Miami, Miami, Florida

LEONID I. PITERBARG

Center for Applied Mathematical Sciences, University of Southern California, Los Angeles, California

ARTHUR J. MARIANO AND EDWARD H. RYAN

RSMAS/MPO, University of Miami, Miami, Florida

(Manuscript received 7 July 2000, in final form 11 January 2001)

ABSTRACT

Predictability of particle motion in the ocean over a timescale of one week is studied using three clusters of buoys consisting of 5–10 drifters deployed in the tropical Pacific Ocean. The analysis is conducted by using three techniques with increasing complexity: the center of mass of the cluster, advection by climatological currents, and a new technique that relies on the assimilation of both velocity and position data from the surrounding drifters into a Markov model for particle motion. A detailed mathematical description of the theory leading to this model is given.

The predictability of drifter motion in these clusters is characterized using the data density N_d, defined as the number of drifters over an area scaled by the mean diameter of the cluster. The data density N_d decreases along the drifter trajectories due to the tendency of particles to disperse by turbulent fluid motion. In the first regime, which corresponds to the period after the release of drifters in a tight cluster when $N_d \gg 1$ drifter/degree^2, the center of mass and the data assimilation methods perform nearly equally well, and both methods yield very accurate predictions of drifter positions with rms prediction errors $\leq$15 km up to 7 days. When a cluster starts to disperse, that is, in the regime where $N_d \geq 1$ drifter/degree^2, the data assimilation technique is the only method that gives accurate results. Finally, when $N_d \ll 1$ drifter/degree^2, no method investigated in this study is effective. Therefore, the data assimilation method performs better than relatively crude approaches of center of mass and mean flow field evolution in the intermediate regime, in which predictability is still possible. It is also found that advection by the climatological mean flow field is generally not an accurate indicator of drifter motion. Uncertainties in the knowledge of initial release positions and the frequency of data assimilation are found to have a strong impact on the prediction accuracy.

1. Introduction

In this study, predictability of Lagrangian motion in a large-scale oceanic flow field is investigated. Predicting particle trajectories in the ocean is of practical importance for problems such as searching for objects lost at sea, tracking floating mines, and studying ecological issues such as spreading of pollutants and fish larvae and designing oceanic observing systems. A number of authors (e.g., Aref 1984; Samelson 1996) have shown that prediction of particle motion is an intrinsically difficult problem because Lagrangian motion often exhibits chaotic behavior, even in regular and simple Eulerian flows. In the ocean, the combined effects of complex time dependence (Samelson 1992; Meyers 1994; Duan and Wiggins 1996) and three-dimensional structure (Yang and Liu 1997) are likely to induce chaotic transport. Chaos implies strong dependence on initial conditions, which are usually not known with great accuracy, so that the task of predicting particle motion is often extremely difficult. Also, velocity errors accumulate as errors of prediction position, further reducing the limits of predictability.

There has been a significant progress in the development of tools based on dynamical systems theory to address issues of Lagrangian mixing and transport in geophysical flows. These techniques focus on identifying the hyperbolic stagnation points, characterized by the intersection of a stable manifold (along which fluid particles are attracted toward the hyperbolic point) and an unstable manifold (along which fluid particles are repelled away from the hyperbolic point), from a ve-

Corresponding author address: Dr. Tamay M. Özgökmen, RSMAS/MPO, University of Miami, 4600 Rickenbacker Causeway, Miami, FL 33149-1098.
E-mail: tozgokmen@rsmas.miami.edu

TABLE 1. Experimental parameters for stochastic simulations of Lagrangian motion.

Number of drifters	Number of realizations	Initial diameter (km)	σ_v (cm s^{-1})	Δt (day)	T (day)	R (km)
6	50	60	20	0.5	3	250

locity field displaying complex time variability. One of these techniques is the so-called lobe dynamics (Wiggins 1992), which has been successfully applied for analyzing transport characteristics in a barotropic meandering jet (Rogerson et al. 1999) and quasigeostrophic double-gyre flow (Coulliette and Wiggins 2000). Haller and Poje (1997) focused on the transient stagnation points and finite-time analogs of stable and unstable manifolds, and extended the lobe analysis to the treatment of finite-time data. This method has then been used to analyze fluid particle pathways in an eddy-resolving, barotropic ocean model (Poje and Haller 1999). Therefore, concepts based on dynamical systems theory lead to accurate and efficient diagnostic tools to estimate and interpret transport and mixing characteristics from a *known* velocity field.

Another approach is to employ statistical dispersion models (e.g., Thomson 1986, 1987; Griffa 1996), which provide information about the probability of finding particles in a specific area, based on the known statistics of the velocity field, such as turbulent velocity variance and Lagrangian decorrelation timescale. Reasonable success has been achieved in reproducing the distribution of real drifters based on such models (e.g., Falco et al. 2000). Alternatively, ocean general circulation models can be integrated to generate the Eulerian flow field, which can be then used to predict the Lagrangian motion (e.g., Yang et al. 1999). However, it is unlikely that the precise location, size, and speed of major coherent structures such as jets and eddies in the ocean are captured with precision using general circulation models at the present time, and therefore this technique will probably offer better accuracy in the future when the assimilation of Eulerian data into ocean models becomes a routine operation.

A different approach is the use of models designed for the assimilation of Lagrangian data, which is a technique developed by Özgökmen et al. (2000). The philosophy behind this approach is that the Lagrangian data are the most natural choice to employ for the prediction of particle motion and that buoys can be released in situ for the practical problems mentioned above. In Özgökmen et al. (2000), drifter data in the vicinity of predicted trajectories are used to improve the accuracy of prediction. The prediction of particle motion is performed using a Gauss–Markov Lagrangian particle model, which relies on the estimate of climatological mean flow field, persistence of turbulence, and assimilation of velocity data from the surrounding drifters through a Kalman filtering technique. In essence, the method described in Özgökmen et al. (2000) is a means of reconstructing the velocity vector for the prediction particle from statistical considerations based on surrounding Lagrangian particles within a distance on the order of the deformation radius at a particular time. Theoretical error estimates for this prediction scheme have been derived. The effectiveness of this technique is verified with synthetic drifter trajectories generated in an idealized, three-layer, double-gyre ocean model. Further testing has been conducted (Castellari et al. 2001) with real drifter data collected in the Adriatic Sea (Poulain 1999), a semienclosed subbasin of the Mediterranean Sea, to show the validity of this method.

The primary purpose of the present study is to explore the effectiveness of this technique when applied to real drifters in large-scale ocean circulation. This will not only complement the previous modeling and semienclosed basin studies, but will also differ from them in several aspects. Perhaps most importantly, in the study by Castellari et al. (2001), the data density of only one of the drifter clusters was high enough for this technique to be effective. Therefore, the applicability of this particular technique to real drifters needs further investigation. Also, in Castellari et al. (2001) the drifters exhibited motion characterized by small time and space scales and particular boundary currents in the Adriatic Sea. Furthermore, the Adriatic drifters were surface drifters, making them very sensitive to changes in local wind forcing. Here, we work with mixed layer drifter clusters in large-scale, energetic, unbounded equatorial flows. These drifters are drogued at the 15-m depth level, making them relatively isolated from sudden changes in the local wind forcing.

The results of Özgökmen et al. (2000) correspond to the case where drifters are initially distributed uniformly over the entire subtropical Atlantic basin. Therefore, the primary parameter controlling the accuracy of the prediction, the data density, defined as the number of drifters within the characteristic Eulerian space scale, remains approximately constant over the prediction period. While trivial in a modeling study, a uniform distribution of drifters over a large area in the ocean is prohibitively expensive and difficult to accomplish in practice. At best, drifters can be released in clusters, which is defined as a number of buoys initially located from one another within the correlation distance of velocity fluctuations. The dispersion of particles, due to complex time variability in the underlying flow field and other effects, implies that the data density is not constant throughout the prediction period and, in contrast to the experimental setup with Özgökmen et al. (2000), the predictive capability decreases throughout the evolution of a cluster.

Finally, the assimilation method used in the present study is more general than the filter in Özgökmen et al. (2000) and in Castellari et al. (2001) in that the Kalman

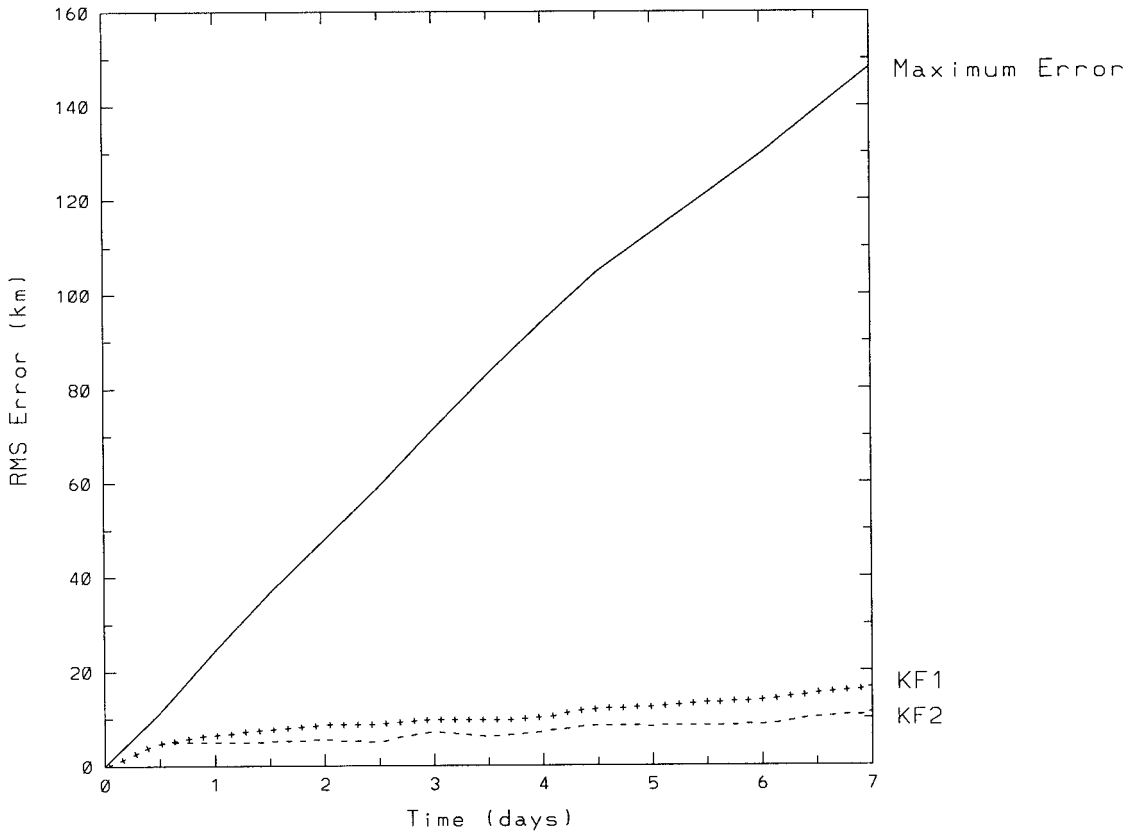

FIG. 1. Comparison of Kalman filters based on stochastic simulations of Lagrangian motion. The total distance traveled by particles or maximum error (solid line), the rms error of the predictions using the simplified Kalman filter corresponding to Eq. (9) (KF1, line with $+++$), and the rms error of the predictions using the improved Kalman filter corresponding to Eq. (8) (KF2, line with $---$). See Table 1 for the experimental parameters.

filtering is applied to drifter positions as well as the turbulent velocity components.

Here, we analyze the predictability of drifter clusters over a timescale of one week. Three drifter clusters, each consisting of 5–10 drifters, are selected from the World Ocean Circulation Experiment (WOCE) data. The trajectory of each particle, belonging to a specific cluster, is calculated using the other trajectories. The prediction error is computed as the difference between the true and predicted trajectories. The accuracy of the prediction is quantified by the root mean square (rms) error as a function of time and as the average of all cluster particles. The errors are systematically compared to the average distance traveled by the particles, the predictions obtained by using the climatological mean flow field, and by the center of mass of the cluster. The impact of the data density, which decreases throughout the evolution of the cluster, or equivalently along the trajectory of each drifter, on prediction accuracy is quantified. The effects of the knowledge about the mean flow field and the initial release positions on the predictions are investigated. Finally, the impact of data assimilation frequency on the predictions is explored.

The paper is organized as follows. In section 2, the model formulation and the Kalman filtering equations are introduced. In section 3, we compare the performance of our upgraded formulation to that of the previous method by means of a series of stochastic simulations. In section 4, the application of the method to real drifters in the tropical Pacific Ocean and the results are discussed. Finally, the conclusions from this study are summarized in section 5. Appendix A includes a time discretization procedure for the continuous stochastic differential equations describing the upper-ocean Lagrangian turbulence. A derivation of the adapted Kalman filter formulas is given in appendix B for the model introduced in section 2.

2. Theory

a. Mathematical problem statement

Let $\mathbf{u}(t, \mathbf{r})$ be a time-dependent, two-dimensional, horizontal Eulerian velocity field. Consider M Lagrangian particles starting at the same time $t = 0$ from different

positions $\mathbf{r}_1^0, \mathbf{r}_2^0, \ldots, \mathbf{r}_M^0$. Their motion is covered by the following system of $2M$ equations:

$$\frac{d\mathbf{r}_j}{dt} = \mathbf{u}(t, \mathbf{r}_j), \qquad \mathbf{r}_j(0) = \mathbf{r}_j^0,$$

$j = 1, \cdots, M$. Introduce the Lagrangian velocities in the usual manner

$$\mathbf{v}_j(t) = \mathbf{u}[t, \mathbf{r}_j(t)].$$

Assume that trajectories of the first $p = M - 1$ particles $\mathbf{r}_1(t), \mathbf{r}_2(t), \ldots, \mathbf{r}_p(t)$ are completely observed and the trajectory of the Mth particle $\mathbf{r}_M(t)$ is not observed during time interval $(0, t^o)$ where t^o is an observation time. We call the former particles *predictors* and the latter *predictand.* The problem is to find a best prediction, $\hat{\mathbf{r}}_M(t^o)$, of the unobserved particle position given the above predictor observations, a mean flow, the initial positions of the predictands, and finally statistics of the velocity fluctuations that will be specified below. The best prediction is defined here as the prediction minimizing the mean square deviation

$$E[|\hat{\mathbf{r}}_M(t^o) - \mathbf{r}_M(t^o)|^2] \to \min.$$

Here E is expectation or averaging over an ensemble. It is well known (e.g., Liptser and Shiryaev 1978) that the solution to this problem is given by the conditional expectation

$$\hat{\mathbf{r}}_M(t^o) = E(\mathbf{r}_M(t^o)|\mathbf{r}_1(t), \mathbf{r}_2(t), \cdots, \mathbf{r}_p(t), 0 \le t \le t^o),$$

given all the observations. Theoretically speaking, this conditional expectation is expressed in terms of the joint probability distribution of all M trajectories, which is a distribution in a functional space of $2M$-vector functions. If the Lagrangian displacement is a Markovian process, which is the case for a white noise velocity fluctuation, then the expectation depends on $\mathbf{r}_1(t^o), \mathbf{r}_2(t^o), \ldots, \mathbf{r}_p(t^o)$ only. Under some conditions a stochastic differential equation can be derived for the optimal prediction in this case (e.g., Liptser and Shiryaev 1978). Finally, if the joint distribution of $\mathbf{r}_1(t^o), \mathbf{r}_2(t^o), \ldots, \mathbf{r}_p(t^o)$ is Gaussian, then the dependence is linear and the corresponding coefficients are determined by the mean flow and statistics of the second order, such as the Lagrangian velocity covariance tensor. Unfortunately, Markovian and Gaussian approximations for the particle displacement are not satisfactory for real oceanic flows. As for the former, there is a strong indication (Griffa et al. 1995; Griffa 1996) that the Lagrangian velocity has a significant correlation time of 1–10 days and its time evolution can be well approximated by a stationary Markovian process. Thus, the displacement can be viewed as an integral of a Markovian process. Regarding the Gaussianity, note that even if the Eulerian velocity fluctuation is a Gaussian white noise in time, the joint distribution of a particle pair is not Gaussian. This is only so for the one-point distribution of the displacement. Thus, one should rule out any effort to get an exact expression for the optimal prediction. However, if the joint process of Lagrangian velocities $[\mathbf{v}_1(t), \mathbf{v}_2(t), \ldots, \mathbf{v}_M(t)]$ is Markovian, then so is the augmented process $[\mathbf{v}_1(t), \mathbf{r}_1(t), \mathbf{v}_2(t), \mathbf{r}_2(t), \ldots, \mathbf{v}_M(t), \mathbf{r}_M(t)]$ including both the velocities and positions, and one can try to apply an extended Kalman filter (EKF). There is a serious obstacle in the way: the diffusion matrix of the above velocity process depends on the positions of the particles. In such a case the EKF cannot be directly applied. Instead, we propose a prediction algorithm that is merely based on the EKF approach and can be viewed as an adaptive EKF. The method is not optimal in general; however, it performs perfectly in some simple cases raising hope that it would do so in more sophisticated cases. We check this with stochastic simulations.

First, we formulate a model for the M-particle motion, which is an extension of the well-known and often-used ''random flight'' model for the one-particle motion.

b. Model

Let

$$\mathbf{u}(t, \mathbf{r}) = \mathbf{U}(\mathbf{r}) + \mathbf{u}'(t, \mathbf{r})$$

be a decomposition of a 2D velocity field into a mean circulation and fluctuation. More exactly, $\mathbf{U}(\mathbf{r})$ is a deterministic velocity field and $\mathbf{u}'(t, \mathbf{r})$ is a random vector field with zero mean, $E[\mathbf{u}'(t, \mathbf{r})] = 0$.

Here we suggest a model for the stochastic flow determined by $\mathbf{u}'(t, \mathbf{r})$. In the framework of that model, the joint motion of M particles (M-particle motion) is described as $4M$-diffusion process (2D velocity and position for M particles) whose drift is determined by the mean flow, and its diffusion matrix is expressed in terms of some covariance matrix. Thus, the Lagrangian prediction problem reduces to filtering a multidimensional random process driven by a system of stochastic differential equations with space variable coefficients. Let $\xi(t, \mathbf{r})$ be a Gaussian vector random field, $\xi, \mathbf{r} \in R^2$ with zero mean, which is a stationary white noise in time

$$E[\xi(t, \mathbf{r})] = 0,$$

$$E[\xi(t_1, \mathbf{r}_1)\xi(t_2, \mathbf{r}_2)^*] = \delta(t_1 - t_2)\mathbf{B}(\mathbf{r}_1, \mathbf{r}_2),$$

where the vectors are considered as column vectors, the asterisk means transposition, $\delta(t)$ is the Dirac delta function, and $\mathbf{B}(\mathbf{r}_1, \mathbf{r}_2)$ is a space covariance tensor. We assume that the Lagrangian velocity and position are described by the following system of stochastic differential equations

$$\dot{\mathbf{v}} = -\mathbf{\Lambda}\mathbf{v} + \xi(t, \mathbf{r}), \qquad \dot{\mathbf{r}} = \mathbf{U}(\mathbf{r}) + \mathbf{v}, \qquad (1)$$

where the dot means the time derivative, $\mathbf{\Lambda}$ is a constant positive definite 2×2 matrix. In our applications we usually assume that

$$\mathbf{\Lambda} = \begin{pmatrix} 1/T_1 & 0 \\ 0 & 1/T_2 \end{pmatrix},$$

where T_1 and T_2 are the Lagrangian correlation timescales

for the zonal and meridional components, respectively, or even $T_1 = T_2 = T$. However, we derive prediction formulas for a general matrix $\mathbf{\Lambda}$ to provide a basis for including other effects such as the Coriolis acceleration and a mean flow shear. Note that the system (1) determines the flow in full; that is, the state of any M particles is determined by their initial positions and velocities. Namely, introduce the state and drift vectors by

$$\mathbf{z} = \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{r}_1 \\ \mathbf{v}_2 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{v}_M \\ \mathbf{r}_M \end{pmatrix} \qquad \mathbf{A}(\mathbf{z}) = \begin{pmatrix} -\mathbf{\Lambda}\mathbf{v}_1 \\ \mathbf{U}(\mathbf{r}_1) + \mathbf{v}_1 \\ -\mathbf{\Lambda}\mathbf{v}_2 \\ \mathbf{U}(\mathbf{r}_2) + \mathbf{v}_2 \\ \vdots \\ -\mathbf{\Lambda}\mathbf{v}_M \\ \mathbf{U}(\mathbf{r}_M) + \mathbf{v}_M \end{pmatrix}.$$

The drift vector $\mathbf{A}(\mathbf{z})$ contains both the large-scale mean flow and a Markov model of mesoscale flow.

Equation (1) is equivalent to the following: $\mathbf{z}(t)$ is a $4M$-dimensional diffusion process with the drift vector $\mathbf{A}(\mathbf{z})$ and diffusion matrix $\mathbf{D}(\mathbf{z})$ given by

$$\mathbf{D}(\mathbf{z}) = [\mathbf{D}_{ij}(\mathbf{z})],$$

with 4×4 blocks

$$\mathbf{D}_{ij} = \begin{pmatrix} \mathbf{B}(\mathbf{r}_i, \mathbf{r}_j) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

More accurately,

$$d\mathbf{z} = \mathbf{A}(\mathbf{z})dt + \mathbf{D}(\mathbf{z})^{1/2}d\mathbf{W}, \tag{2}$$

where $\mathbf{W}(t)$ is a standard Wiener process in $4M$ dimensions. The model (2) describing M-particle motion is a starting point for the prediction algorithm, but we need to adjust it for discrete observations (see next section). In particular, the one-particle motion equations in the homogeneous case $\mathbf{B}(\mathbf{r}_i, \mathbf{r}_j) = \mathbf{B}(\mathbf{r}_i - \mathbf{r}_j)$ with diagonal $\mathbf{B}(\mathbf{0})$

$$\mathbf{B}(\mathbf{0}) = \begin{pmatrix} 2\sigma_1^2/T_1 & 0 \\ 0 & 2\sigma_2^2/T_2 \end{pmatrix}$$

become

$$dv_i = -(v_i/T_i)dt + \sigma_i\sqrt{2/T_i}\,dw_i$$
$$dx_i = (U_i + v_i)dt, \qquad i = 1, 2,$$

where v_1 and v_2 are zonal and meridional Lagrangian velocity fluctuations, x_1 and x_2 are the particle zonal and meridional displacements, U_1 and U_2 are the mean flow zonal and meridional components, and σ_1^2 and σ_2^2 are variances of the zonal and meridional fluctuation components. Thus, the well-known random flight model for the one-particle motion studied in Thomson (1986) and Griffa (1996) is a special case of our more general model.

Note that the generic model (1) can be derived from an Euler equation for the Eulerian velocity field with a random forcing (Piterbarg 2001).

c. Discretization

In practice, we encounter measurements separated by a finite timescale Δt, which can be of the same order as a characteristic variability scale of the Lagrangian velocity. Let us define

$$\mathbf{z}(n) \equiv \mathbf{z}(n\Delta t).$$

It is shown in appendix A that the approximation of (2) to order $(\Delta t)^{3/2}$ results in the following first-order Markov model

$$\mathbf{z}(n) = \mathbf{\Phi}(\mathbf{z}(n-1)) + \mathbf{q}(n, \mathbf{r}(n-1)), \tag{3}$$

where

$$\mathbf{r} = \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_M \end{pmatrix}, \qquad \mathbf{\Phi}(\mathbf{z}) = \begin{pmatrix} (\mathbf{I} - \Delta t\mathbf{\Lambda})\mathbf{v}_1 \\ \mathbf{r}_1 + \Delta t[\mathbf{U}(\mathbf{r}_1) + \mathbf{v}_1] \\ (\mathbf{I} - \Delta t\mathbf{\Lambda})\mathbf{v}_2 \\ \mathbf{r}_2 + \Delta t[\mathbf{U}(\mathbf{r}_2) + \mathbf{v}_2] \\ \vdots \\ (\mathbf{I} - \Delta t\mathbf{\Lambda})\mathbf{v}_M \\ \mathbf{r}_M + \Delta t[\mathbf{U}(\mathbf{r}_M) + \mathbf{v}_M] \end{pmatrix}, \tag{4}$$

$\mathbf{I}$ is the 2×2 unit matrix and the covariance matrix of the stochastic part is expressed in the form

$$E[\mathbf{q}(n, \mathbf{r})\mathbf{q}(n, \mathbf{r})^*] \equiv \mathbf{Q}(\mathbf{r}) = (\mathbf{Q}_{ij}),$$

where 4×4 blocks $\mathbf{Q}_{ij}$ are given by

$$\mathbf{Q}_{ij} = \frac{1}{6}\begin{pmatrix} 6\Delta t\mathbf{B}(\mathbf{r}_i, \mathbf{r}_j) & 3(\Delta t)^2\mathbf{B}(\mathbf{r}_i, \mathbf{r}_j) \\ 3(\Delta t)^2\mathbf{B}(\mathbf{r}_i, \mathbf{r}_j) & 2(\Delta t)^3\mathbf{B}(\mathbf{r}_i, \mathbf{r}_j) \end{pmatrix}. \tag{5}$$

Note that the straightforward first difference approximation of (1),

$$\mathbf{v}(n) = (\mathbf{I} - \Delta t\mathbf{\Lambda})\mathbf{v}(n-1) + (\Delta t)^{1/2}\varepsilon(n, \mathbf{r}(n-1)),$$
$$\mathbf{r}(n) = \mathbf{r}(n-1) + [\mathbf{U}(\mathbf{r}(n-1)) + \mathbf{v}(n-1)]\Delta t, \tag{6}$$

where $\varepsilon(n, \mathbf{r})$ is a sequence of independent random fields, leads to a system similar to (3) but with a degenerate covariance matrix of the stochastic part. This fact would prevent an implementation of the EKF approach. The second drawback of (6) is that it does not give any explicit relation between $\varepsilon(n, \mathbf{r})$ and $\xi(t, \mathbf{r})$. Our approach to the discretization is different: We first linearize the original time continuous system at the neighborhood of the current state. We then solve the equations exactly at an interval of duration Δt, and, finally, truncate the terms of order higher than $(\Delta t)^{3/2}$. See appendix A for details.

d. Prediction algorithm

Let us augment the state equation (3) with the observation equation

TABLE 2. Characteristics of the clusters and experimental parameters.

Cluster	Number of drifters	Initial date	Analysis time (day)	σ_v (cm s^{-1})	T (day)	R (km)
I	5	25 Oct 1993	100	24	3	250
II	7	22 Aug 1990	160	31	3	250
III	10	2 Dec 1990	60	54	3	250

$$\mathbf{y}(n) = \mathbf{H}\mathbf{z}(n), \tag{7}$$

where the observational $4p \times 4M$ matrix is

$$\mathbf{H} = (\mathbf{I}_p \qquad \mathbf{0})$$

and $\mathbf{I}_p$ is $4p \times 4p$ unit matrix. The above structure of the observation matrix implies that the first p components of the state vector are fully observed and the last one is not observed at all.

The distinguished feature of the system (3) is that the stochastic part depends on the state. If it was not the case, that is, $\mathbf{q}(n, \mathbf{r}) \equiv \mathbf{q}(n)$, then one could use the extended Kalman filter for the system (3) and (7). Alas, the standard Kalman filter approach is not applicable for handling the considered case. Moreover, systems with noise depending on the state are poorly studied in mathematics. The difficulty here is that even though the sequence $\{\mathbf{q}(n, \mathbf{r})|_{n=1,2,\ldots}\}$ itself is a sequence of independent Gaussian random fields, it looses this property after substituting $\mathbf{r}(n-1)$ instead of $\mathbf{r}$. More exactly, the sequence $\{\mathbf{q}(n, \mathbf{z}(n-1))|_{n=1,2,}\}$ is not a sequence of independent random vectors. For the above reasons we suggest a prediction algorithm that inherits only the idea of the Kalman filter and is not optimal. The suggested algorithm can be viewed as an adaptive Kalman filter since its coefficients depend on the previous observations. Here we give the computational formulas only for the case of small Δt. The general case as well as the derivation are brought to appendix B. Namely, the adaptive Kalman filter formulas take the following form:

$$\mathbf{v}_M^a(n) = \mathbf{v}_M^f(n) + \sum_{j=1}^{p} \mathbf{K}_j[\mathbf{v}_j(n) - \mathbf{v}_j^f(n)],$$

$$\mathbf{r}_M^a(n) = \mathbf{r}_M^f(n) + \sum_{j=1}^{p} \mathbf{K}_j[\mathbf{r}_j(n) - \mathbf{r}_j^f(n)], \tag{8}$$

where

$$\mathbf{v}_M^f(n) = (\mathbf{I} - \Delta t\mathbf{\Lambda})\mathbf{v}_M^a(n-1),$$

$$\mathbf{r}_M^f(n) = \mathbf{r}_M^a(n-1) + \{\mathbf{U}[\mathbf{r}_M^a(n-1)] + \mathbf{v}_M^a(n-1)\}\Delta t \quad \text{and}$$

$$\mathbf{K}_j = \sum_{k=1}^{p} \mathbf{B}[\mathbf{r}_M^a(n-1), \mathbf{r}_k(n-1)]\mathbf{B}_{kj}^{-1}(n-1), \qquad j = 1, \cdots, p.$$

The superscripts a and f stand for ''analyzed'' and ''forecasted,'' respectively, in accordance with the common Kalman filter notation. $\mathbf{B}_{ij}^{-1}(n)$ are the 2×2 entries of $(\mathbf{B}(\mathbf{r}_i(n), \mathbf{r}_j(n)))^{-1}$. Note that for the observed particles, analyzed variables coincide with observational ones $\mathbf{v}_i^a(n) = \mathbf{v}_i(n)$, $\mathbf{r}_i^a(n) = \mathbf{r}_i(n)$, $i = 1, \cdots, p$.

Using the ''naive'' approximation (6) of order Δt gives

$$\mathbf{v}_M^a(n) = \mathbf{v}_M^f(n) + \sum_{k=1}^{p} \mathbf{K}_j[\mathbf{v}_j(n) - \mathbf{v}_j^f(n)],$$

$$\mathbf{r}_M^a(n) = \mathbf{r}_M^f(n). \tag{9}$$

Equation (9) was used in Özgökmen et al. (2000) and differs from (8) by absence of the correcting term in the position prediction. For applications of (8) and (9), we use

$$\mathbf{B}(\mathbf{r}_1, \mathbf{r}_2) = \begin{pmatrix} \exp[-(\mathbf{r}_1 - \mathbf{r}_2)^2/R^2] & 0 \\ 0 & \exp[-(\mathbf{r}_1 - \mathbf{r}_2)^2/R^2] \end{pmatrix}, \qquad \mathbf{\Lambda} = \begin{pmatrix} 1/T & 0 \\ 0 & 1/T \end{pmatrix}, \tag{10}$$

where R is the space correlation scale estimated from data and T is the Lagrangian correlation time. Thus, besides the mean flow, only two parameters are needed to implement the algorithm. Note that the coefficient before the covariance does not matter since the weighing matrices $\mathbf{K}_j$ are determined by ratios of covariances. For this reason we set the coefficient equal to one.

In our previous paper (Özgökmen et al. 2000), we developed an error analysis for the prediction (9) in terms of N_R—the number of the predictors in the R neighborhood of the predictand divided by the area of the neighborhood. Such a mapping is not appropriate for a cluster prediction analysis and we suggest another approach in section 4. A theoretical error analysis of the prediction formulas (8) for a cluster is a matter of our current studies and is based on the Lyapunov exponent theory for the model (1).

e. Prediction by center of mass

It is interesting to compare the Kalman filter (KF) prediction with the position of the center of mass of the cluster. In fact, the time when the error of KF becomes equal to error using the center of mass, can be treated

as a predictability limit for the KF. Now we find the asymptotical distance between the center of mass and the predictand for the above model. Define the center of mass by

$$\mathbf{r}_c(t) = \frac{1}{p}\sum_{j=1}^{p}\mathbf{r}_j(t) \tag{11}$$

and set

$$s_c^2(t) = E[|\mathbf{r}_M(t) - \mathbf{r}_c(t)|^2].$$

The difference between displacement of two particles released simultaneously is called the separation process. Introduce the variance of the separation process by

$$\rho_{ij}(t) = E[|\mathbf{r}_i(t) - \mathbf{r}_j(t)|^2].$$

Then

$$s_c^2(t) = \frac{1}{p^2}E\left[\left|\sum_{j=1}^{p}(\mathbf{r}_M(t) - \mathbf{r}_j(t))\right|^2\right]$$

$$= \frac{1}{p^2}E\left[\sum_{j=1,i=1}^{p}(\mathbf{r}_M(t) - \mathbf{r}_j(t))(\mathbf{r}_M(t) - \mathbf{r}_i(t))\right].$$

Recalling that for any two vectors $\mathbf{ab} = \frac{1}{2}(\mathbf{a}^2 + \mathbf{b}^2 - (\mathbf{a} - \mathbf{b})^2)$ and substituting for $\mathbf{a} = \mathbf{r}_M(t) - \mathbf{r}_j(t)$ and $\mathbf{b} = \mathbf{r}_M(t) - \mathbf{r}_i(t)$ in the above formula, we get

$$s_c^2(t) = \frac{1}{2p^2}E\left[\sum_{j=1,i=1}^{p}\{(\mathbf{r}_M(t) - \mathbf{r}_j(t))^2 + (\mathbf{r}_M(t) - \mathbf{r}_i(t))^2 - (\mathbf{r}_j(t) - \mathbf{r}_i(t))^2\}\right],$$

which can then be rewritten as

$$s_c^2(t) = \frac{1}{p}\sum_{j=1}^{p}\rho_{jM}(t) - \frac{1}{2p^2}\sum_{i,j=1}^{p}\rho_{ij}(t).$$

It can be shown (Piterbarg 1998; Zirbel and Cinlar 1996) for an incompressible flow in an unbounded domain that

$$\rho_{ij}(t) \sim 2Dt,$$

as $t \to \infty$ where D is the diffusivity due to the velocity fluctuations. For the above model with constant mean flow and statistically homogeneous velocity fluctuations (e.g., Griffa 1996)

$$D = 2(\sigma_1^2 T_1 + \sigma_2^2 T_2).$$

Thus, for large t

$$s_c^2(t) \sim \left(1 + \frac{1}{p}\right)Dt$$

with D given above.

3. Results from simulated Lagrangian motion

Prior to conducting a predictability analysis of drifter clusters in the real ocean, it is desirable to quantify the difference in prediction error from algorithm (9) [used before in Özgökmen et al. (2000)] and the mathematically more complete formula (8). This is best accomplished by conducting a systematic set of simulations of Lagrangian motion using the stochastic model described in section 2b. This will also allow us to test the validity of the EKF prediction error in the ideal situation when one knows exactly the residual covariance function and all the model parameters. Regarding the performance of two Kalman filter formulations [(8) and (9)], note that the additional term in (8) is of order Δt^2 and is probably not very important in the case $\Delta t \ll T$. Nonetheless, a quantitative analysis is needed.

In these simulations, a constant mean flow and divergence-free isotropic velocity fluctuations are assumed. The analytical ground for the simulations is Eq. (2) describing the M-particle motion in the continuous case. Its simplest difference form used in our simulation is given by

$$\mathbf{v}(n) = \exp(\Delta t/T)\mathbf{v}(n-1) + \sigma_v\mathbf{B}[\mathbf{r}(n-1)]^{1/2}\sqrt{2dt/T}\,\varepsilon_n,$$

$$\mathbf{r}(n) = \mathbf{r}(n-1) + [\mathbf{U} + \mathbf{v}(n)]dt, \tag{12}$$

where $\mathbf{v}(n)$, $\mathbf{r}(n)$ are the $2M$-vectors of the particle velocities and positions respectively, $\mathbf{B}(\mathbf{r}) = [\mathbf{B}(\mathbf{r}_i - \mathbf{r}_j)]$ is $2M \times 2M$ matrix, and its root square is found by the Cholesky decomposition, $\{\varepsilon_n\}$ is a sequence of independent Gaussian $2M$-vectors with zero mean and unit covariance matrix. In all experiments we assumed the constant mean flow $\mathbf{U}$, the covariance $\mathbf{B}(\mathbf{r}_i - \mathbf{r}_j)$ given by (10), and $dt = 1$ h. The rest of parameter values are chosen consistent with the following analysis of Pacific drifters: Lagrangian correlation timescale $T = 3$ days, space correlation scale $R = 250$ km, the mean square velocity $\sigma_v = 20$ cm s^{-1}, $\Delta t = 0.5$ days, and the number of drifters $M = 6$. Thus, we first run (12) for 7 days and then apply (8), (9) for the predictions. Initially, the predictors are located at the vorticies of a right pentagon with radius 30 km and the predictand is located at the distance 15 km from the pentagon center. The prediction errors are calculated by averaging through 50 experiments using a random number generator for simulating the velocity fluctuations. The experimental parameters are summarized in Table 1.

The result from this set of simulations is illustrated in Fig. 1, which shows the rms distance traveled by all simulated drifters with respect to the release position (dispersion or maximum error) and prediction errors by two Kalman filtering formulations, respectively, over a period of 7 days. With respect to the total dispersion of particles (148 km at 7 days), both Kalman filtering algorithms yield order of magnitude gains in accuracy (17 and 11 km, respectively, at 7 days). The improved Kalman filter given by (8) performs visibly better compared to the simplified one (9). The maximum difference in this particular parameter range is about 5–6 km or ap-

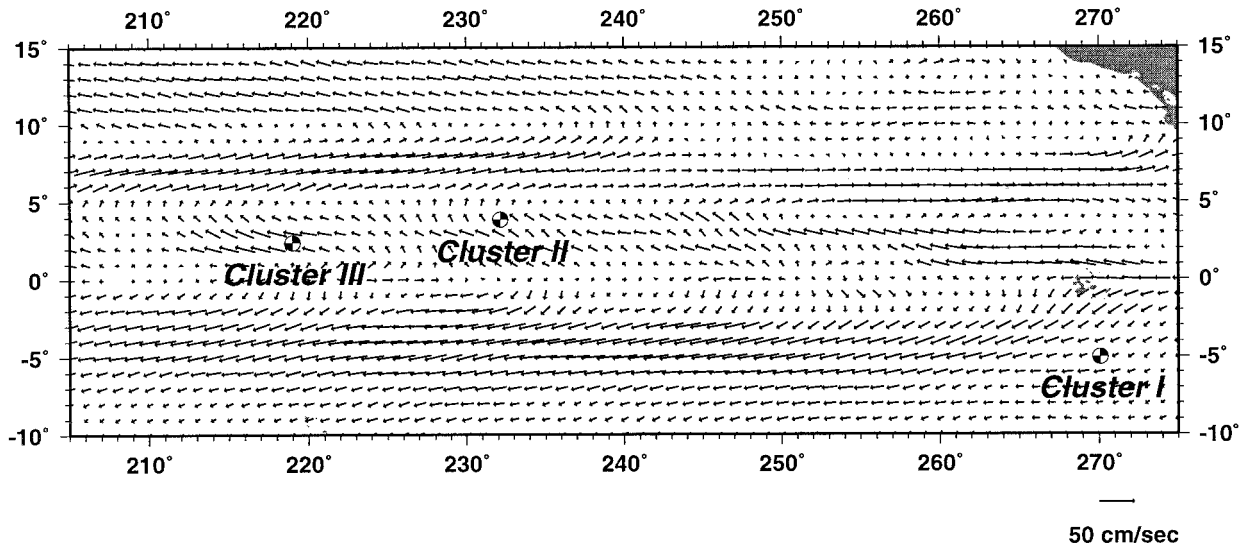

FIG. 2. The mean flow field **U** used in the prediction experiments. The circles mark the initial positions of Clusters I, II, and III.

proximately 30% improvement over a timescale of one week.

Numerous other experiments (not shown) indicate that the additional term does not make a substantial difference for other realistic values of the parameters. However, in all the experiments the prediction algorithm (8) gives lower prediction error than (9). Therefore, the improved Kalman filtering formulation (8) is used in the following study of the Pacific drifter clusters.

4. Pacific drifter clusters

a. Methodology

Three clusters, each consisting of 5–10 drifters, are selected from the 1988–96 WOCE data obtained from the NOAA Atlantic Oceanographic and Meteorological Laboratory, Global Drifter Center. The drifter design has been calibrated to an accuracy of 1–2 cm s^{-1} for conditions that prevail in the tropical Pacific (Niiler et al. 1995). The raw drifter data then undergo standard quality control procedures (Hansen and Poulain 1996) and are posted in the so-called krigged form, in which positions and velocities are given at 6 h intervals. We have searched the entire 1988–96 dataset in the tropical Pacific by plotting annual maps of release positions. The clusters are defined by a group of at least 5 drifters released within a time interval less than or equal to decorrelation timescale (3 days) and/or within the Eulerian decorrelation space scale, or approximately the radius of deformation (~250 km in the equatorial Pacific; Cushman-Roisin 1994). We found that this criterion leads to only about a half dozen clusters. This is because drifter clusters cannot be formed from a fixed or random deployment pattern from ships of opportunity, but require specific preparation. In the specific cases that the drifters have been deployed on tight grids,

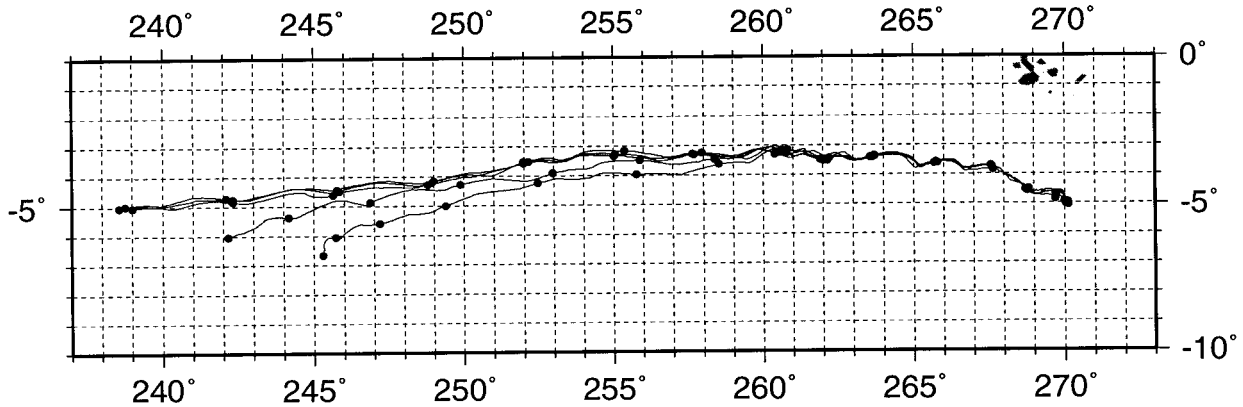

FIG. 3. Drifter trajectories in Cluster I (buoy id numbers: 20 024, 20 025, 20 026, 20 027, and 20 028) for approximately 100 days. The circles mark 7-day intervals.

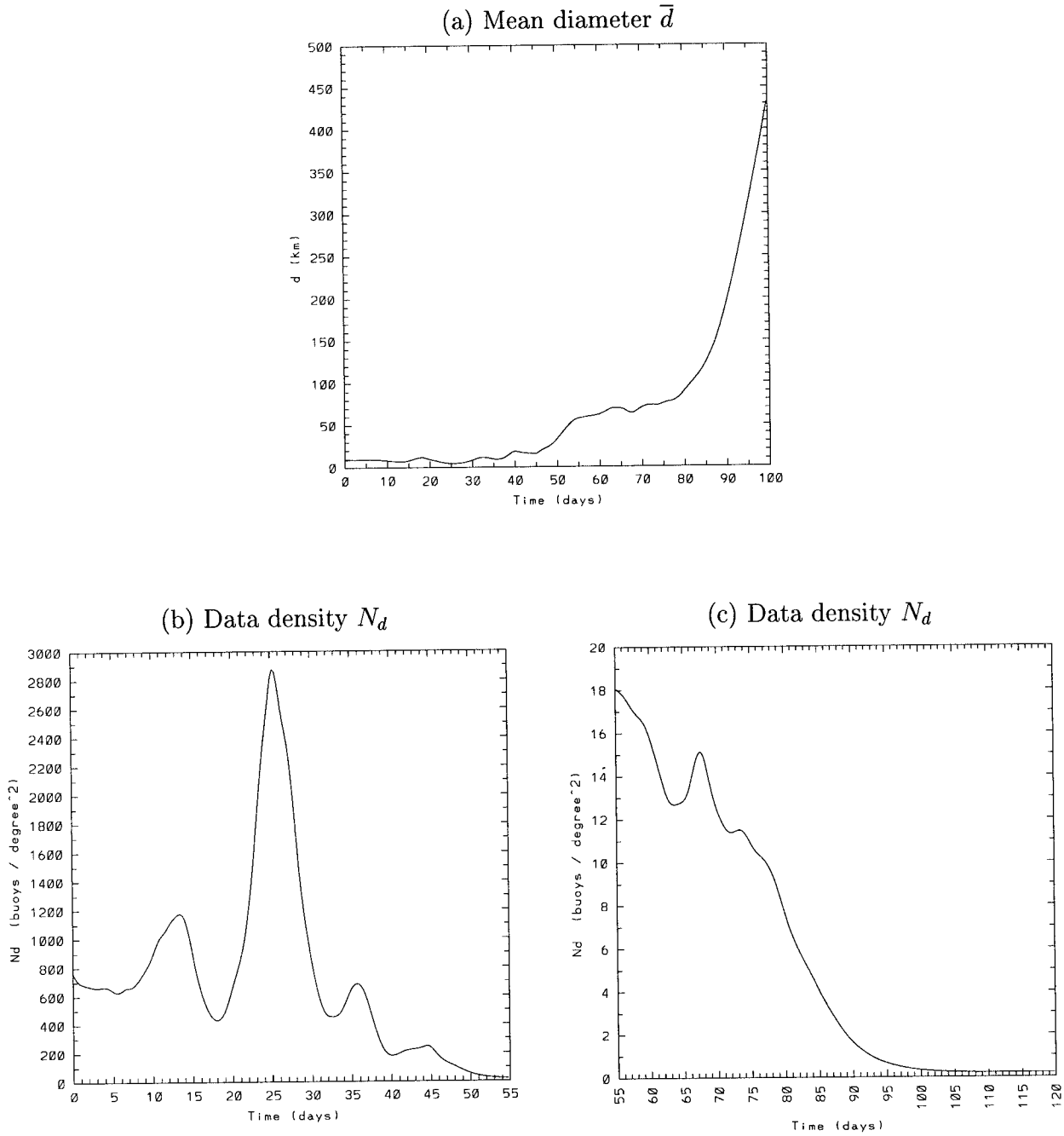

FIG. 4. (a) Mean diameter $\overline{d}$ (km) for $0 \leq t \leq 100$ days, (b) the cluster data density N_d (drifter/degree^2) for $0 \leq t \leq 55$ days, and (c) N_d for $55 \leq t \leq 120$ days for Cluster I.

a real-time analysis of a variety of datasets including current meter profilers and satellite images have been necessary for a detailed dynamical analysis (e.g., Flament et al. 1996) due to strong currents in this region.

The prediction model outlined in detail in section 2 requires the mean flow field **U** and two parameters, the Lagrangian correlation timescale T and the space correlation scale R, which can be thought of as the space correlation scale of the corresponding Eulerian velocity field.

The separation of the Lagrangian particle velocity into a mean component (deterministic) and fluctuating component (stochastic) is at the basis of random flight models for one-particle motion (Thomson 1986; Griffa 1996). However, it is not clear whether a clear separation is applicable to oceanic flows, even though the Rossby

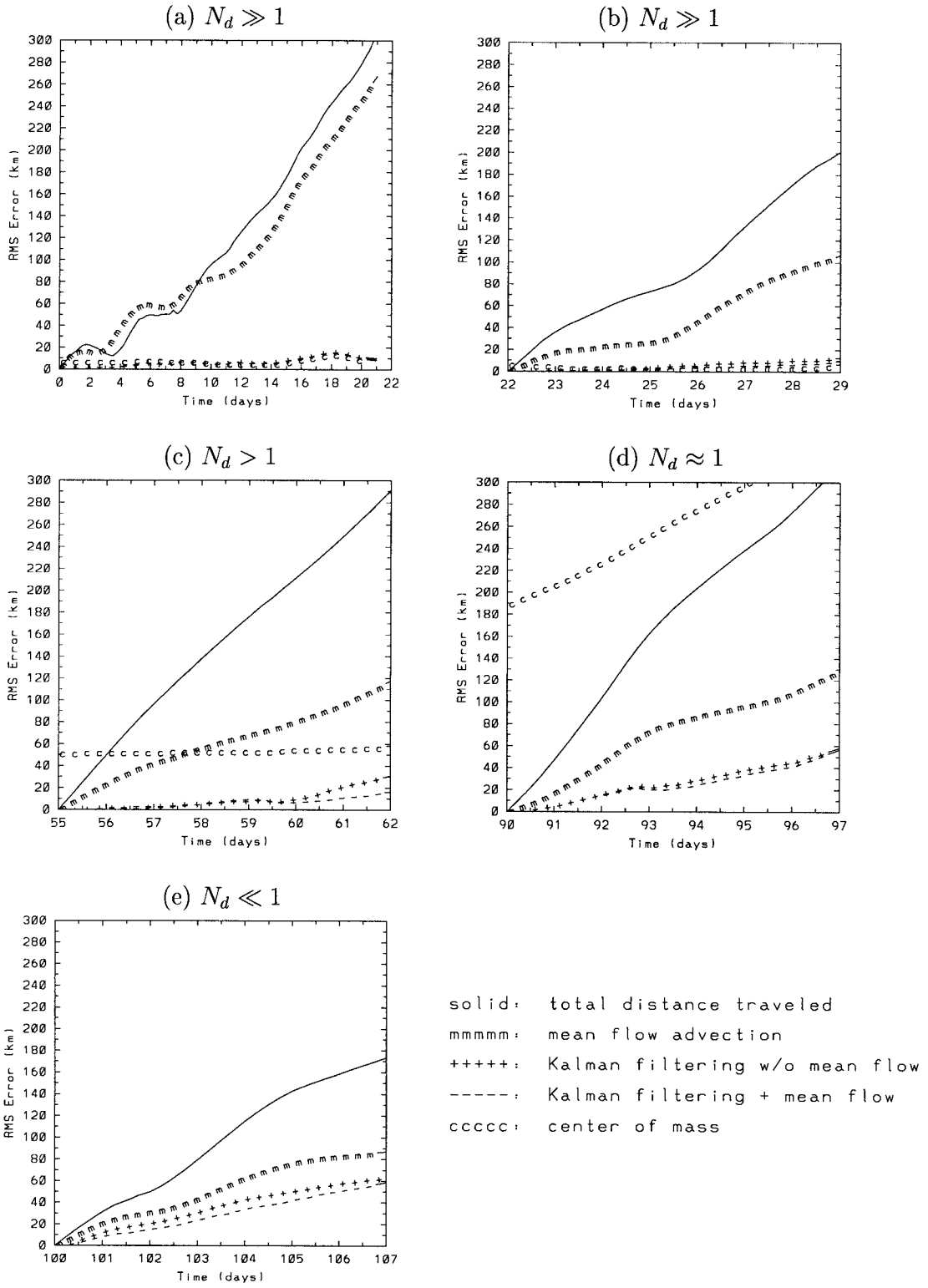

FIG. 5. Total distance traveled s, prediction error with mean flow advection s_m, prediction error based on the center of mass s_c, prediction error using data assimilation s_a with and without mean flow **U** for Cluster I during (a) $0 \le t \le 21$ days ($N_d \gg 1$ drifter/degree^2), (b) $22 \le t \le 29$ days ($N_d \gg 1$ drifter/degree^2), (c) $55 \le t \le 62$ days ($N_d > 1$ drifter/degree^2), (d) $90 \le t \le 97$ days ($N_d \approx 1$ drifter/degree^2), and (e) $100 \le t \le 107$ days ($N_d \ll 1$ drifter/degree^2).

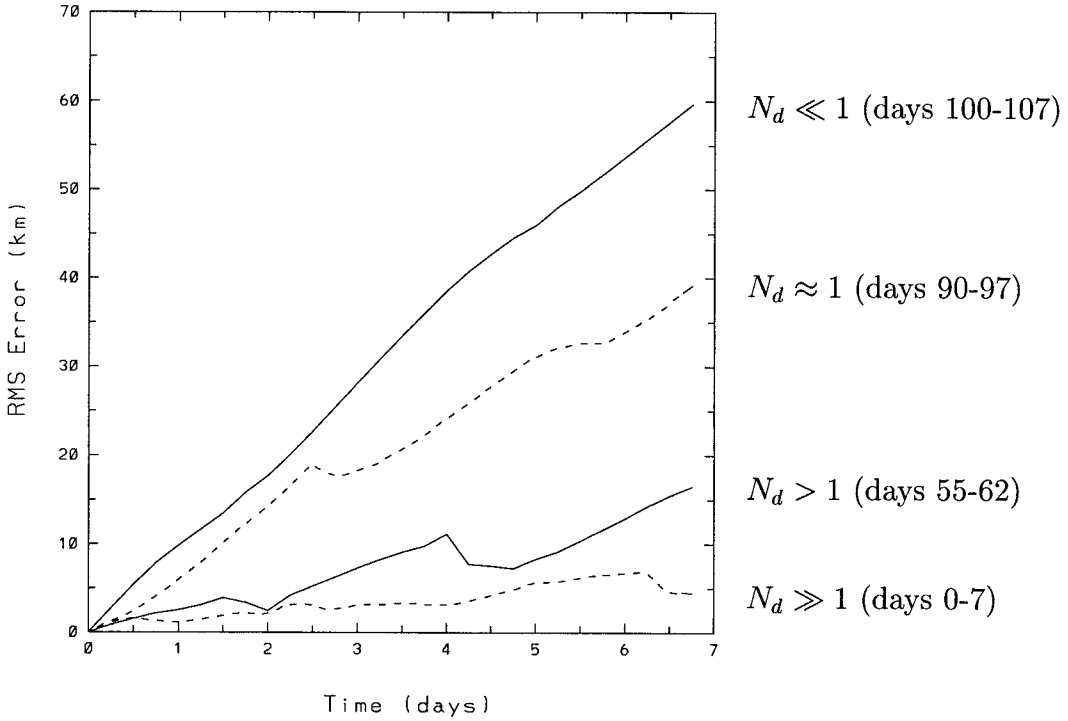

FIG. 6. The rms prediction error using data assimilation s_a (with **U**) in different regimes of data density along the drifter trajectories in Cluster I.

deformation radius suggests a scale for coherent structures of turbulence. In general, the fluctuating component, as defined in this study, represents turbulence at smaller scales not captured by the larger scale mean flow. It was shown that the performance of random flight models is not too sensitive to the assumption of scale separation between the mean flow and fluctuation field (Zambianchi and Griffa 1994), and hence on the precise definition these components. Reasonable success has been achieved in reproducing real drifter motion based on such models (e.g., Falco et al. 2000). In the present study, the mean flow is calculated in the *climatological* sense, representing basin-scale currents and therefore the fluctuating component is representative primarily of mesoscale turbulence, even though smaller-scale turbulence is also included.

The climatological mean field is calculated from the 1988–96 WOCE drifter dataset by using a least squares bicubic spline interpolation technique developed by Bauer et al. (1998). When compared to the traditional binning technique, the spline interpolation technique minimizes the energy in the fluctuation field and maintains a degree of smoothness for the mean flow even in regions of strong currents and strong shears, such as the equatorial Pacific. The mean flow field is calculated from $-10°$ to 15° (or from 10°S to 15°N) in the meridional direction and from 205° to 275° (or from 155° to 85°W) in the zonal direction, the general area covered by the three clusters of drifters during the period of interest (Fig. 2). The well-known major currents are clearly visible in the mean field: the westward North Equatorial Current (NEC) north of about 10°N, the eastward North Equatorial Countercurrent (NECC) between approximately 4° and 9° and the westward South Equatorial Current (SEC), which extends across the equator to 10°S. The reader is referred to Reverdin et al. (1994) for the temporal variability of the large-scale velocity, derived from drifter data. Drifters in the first cluster (Cluster I) have been released in the SEC, whereas the others in Clusters II and III have been launched just south of the NECC (Fig. 2). The other model parameters are chosen as follows. The Lagrangian timescale T is taken as 3 days for all clusters, in agreement with the values computed by Bauer et al. (1998). As explained above, since the mean flow field is calculated as climatology, the fluctuating velocity field is primarily representative of mesoscale turbulence, and therefore the Eulerian space scale R corresponds approximately to the radius of deformation, and R is taken as 250 km in this study, the deformation radius in the equatorial Pacific. The results presented in the following are not sensitive to the specific value of T and R in the range of $\pm 25\%$ of the above cited numbers. The model parameters are listed in Table 2.

Next we define some quantities used in the following analysis of predictability:

$$s(t) = \langle (\mathbf{r}(t) - \mathbf{r}(0))^2 \rangle^{1/2}, \tag{13}$$

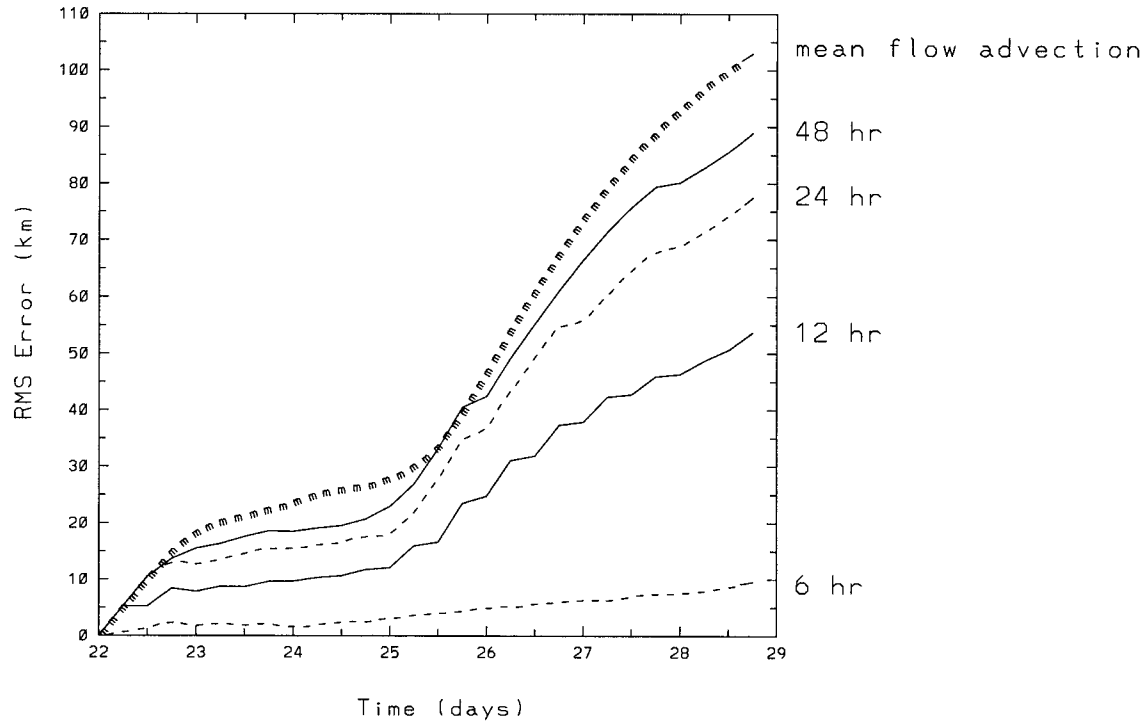

FIG. 7. The rms prediction error using assimilation s_a (with **U**) for different data assimilation frequencies: every 6 h (dashed line), 12 h (solid line), 24 h (dashed line), and 48 h (solid line) during $22 \leq t \leq 29$ days for Cluster I. The line marked with *mmm* shows the prediction error using mean flow advection (i.e., when data are not assimilated).

where the angle brackets denote ensemble averaging by assigning the role of predictand to each drifter in a cluster. Therefore, s provides a measure of the total distance traveled by the particles from their original starting position during the period of interest t, and it can be therefore considered as an upper bound for error for any reasonable prediction algorithm. Here

$$s_m(t) = \left\langle \left(\int_0^t \mathbf{U}[\mathbf{r}(\tau)]\, d\tau - \mathbf{r}(t) \right)^2 \right\rangle^{1/2} \tag{14}$$

is the error for trajectory prediction based only on the advection by the mean flow field shown in Fig. 2. In this case, no fluctuating component is calculated and utilized. Here

$$s_c(t) = \langle (\mathbf{r}_c(t) - \mathbf{r}(t))^2 \rangle^{1/2} \tag{15}$$

is the rms position error with respect to the center of mass defined by (11). Finally, the rms prediction error with assimilation of surrounding data [i.e., using the complete algorithm (8)] is defined as

$$s_a(t) = \langle (\mathbf{r}^a(t) - \mathbf{r}(t))^2 \rangle^{1/2}. \tag{16}$$

In order to quantify the behavior of the cluster, a mean diameter $\overline{d}$ of the cluster is defined as

$$\overline{d}(t) = \frac{1}{Mp} \sum_{i=1}^{M} \sum_{i \neq j}^{M} |\mathbf{r}_i(t) - \mathbf{r}_j(t)|. \tag{17}$$

Next, we define the cluster data density N_d as

$$N_d \equiv \frac{M}{\overline{d}^2}. \tag{18}$$

The above definition of data density differs from that in Özgökmen et al. (2000), that is, $N_R \equiv p/\pi R^2$, the number of particles within a circle of radius R. Since $\overline{d}(t)$ is a true measure of cluster behavior and $R = \text{const}$ is an approximate scale, N_d is a more accurate definition of data density than N_R. For instance, changes in data density cannot be detected using N_R unless predictors are farther apart than R, while, as shown below, such changes are important for the prediction accuracy, especially for small t.

In the following analysis, the effect of data assimilation is quantified by comparing s_a to the total distance traveled s, the rms error of prediction using the mean flow field alone s_m, and the estimate from the center of mass of the cluster s_c. Three important uncertainties, the turbulent initial velocity, the exact knowledge of the mean flow, and the initial positions are addressed as follows. While computing s_a, the initial turbulent velocity is estimated from the surrounding data using (8). The impact of the mean flow field on the accuracy of the prediction algorithm is quantified by calculating s_a with $\mathbf{U} = 0$ and by comparing the results to those obtained with the mean flow included in the algorithm. Finally, given that for some practical applications listed above (search and rescue missions, etc.) the exact initial location of the object is not likely to be known, we

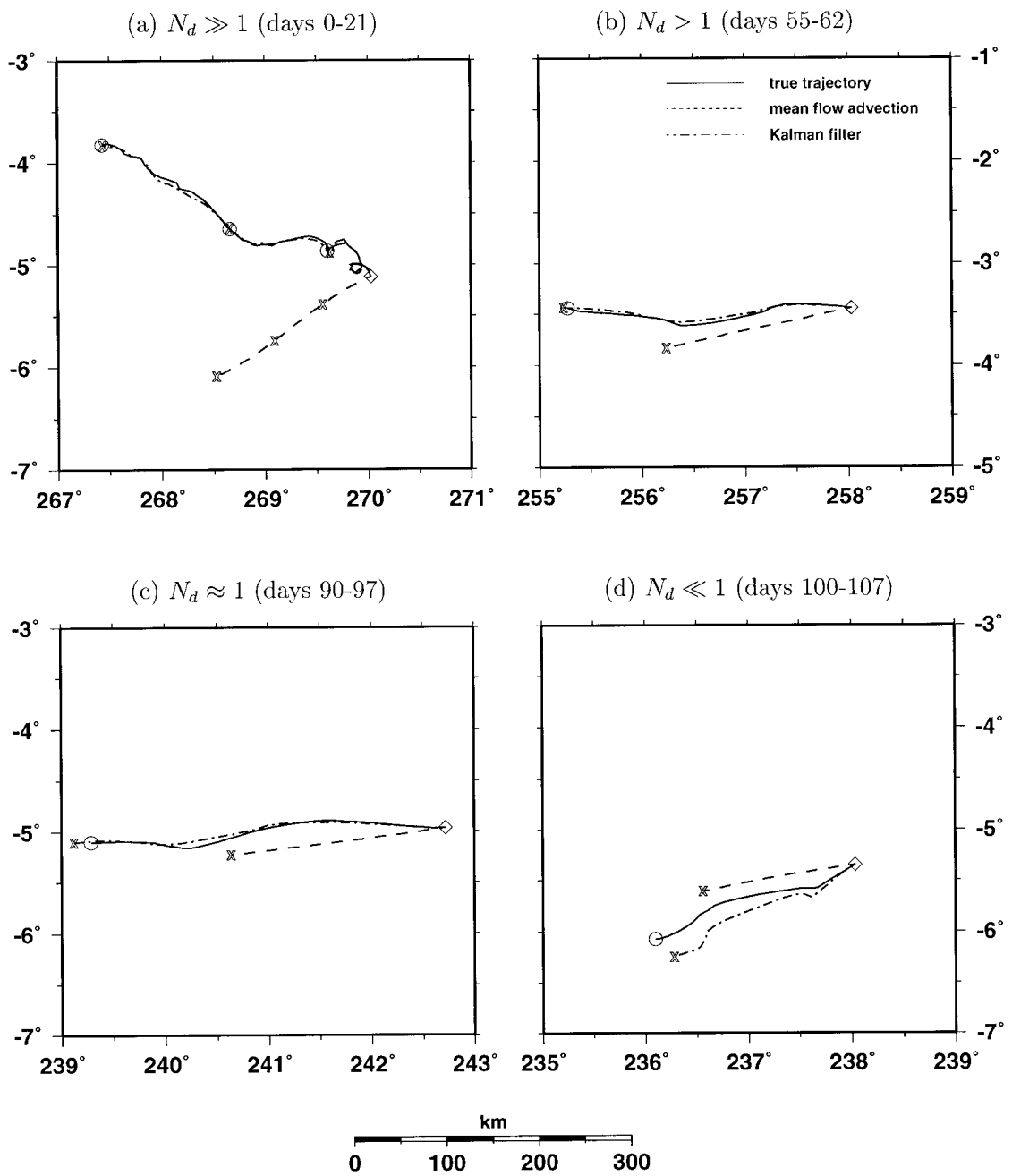

FIG. 8. The true (solid line) and predicted trajectories using mean flow advection (dashed line) and data assimilation (dash–dotted line) methods for buoy 20 026 in Cluster I during (a) $0 \leq t \leq 21$ days ($N_d \gg 1$ drifter/degree^2), (b) $55 \leq t \leq 62$ days ($N_d > 1$ drifter/degree^2), (c) $90 \leq t \leq 97$ days ($N_d \approx 1$ drifter/degree^2), and (d) $100 \leq t \leq 107$ days ($N_d \ll 1$ drifter/degree^2). Every 7 days are marked.

introduce errors in the initial position, and observe their effect on the prediction error.

b. Results for Cluster I

Cluster I consists of 5 drifters that have been released on 25 October 1993 in a tight grid. The trajectories of these drifters over the following 100 days are shown in Fig. 3, which illustrates that they remain surprisingly close to each other for roughly 3000 km. The velocity variance $\sigma_v = 24$ cm s^{-1} indicates that while being the least energetic of the three clusters analyzed here (Table 1), this cluster is far more energetic than the Adriatic drifters investigated previously (Poulain 1999) with σ_v ranging between 10 and 15 cm s^{-1}.

The mean diameter $\overline{d}$ for Cluster I is plotted in Fig.

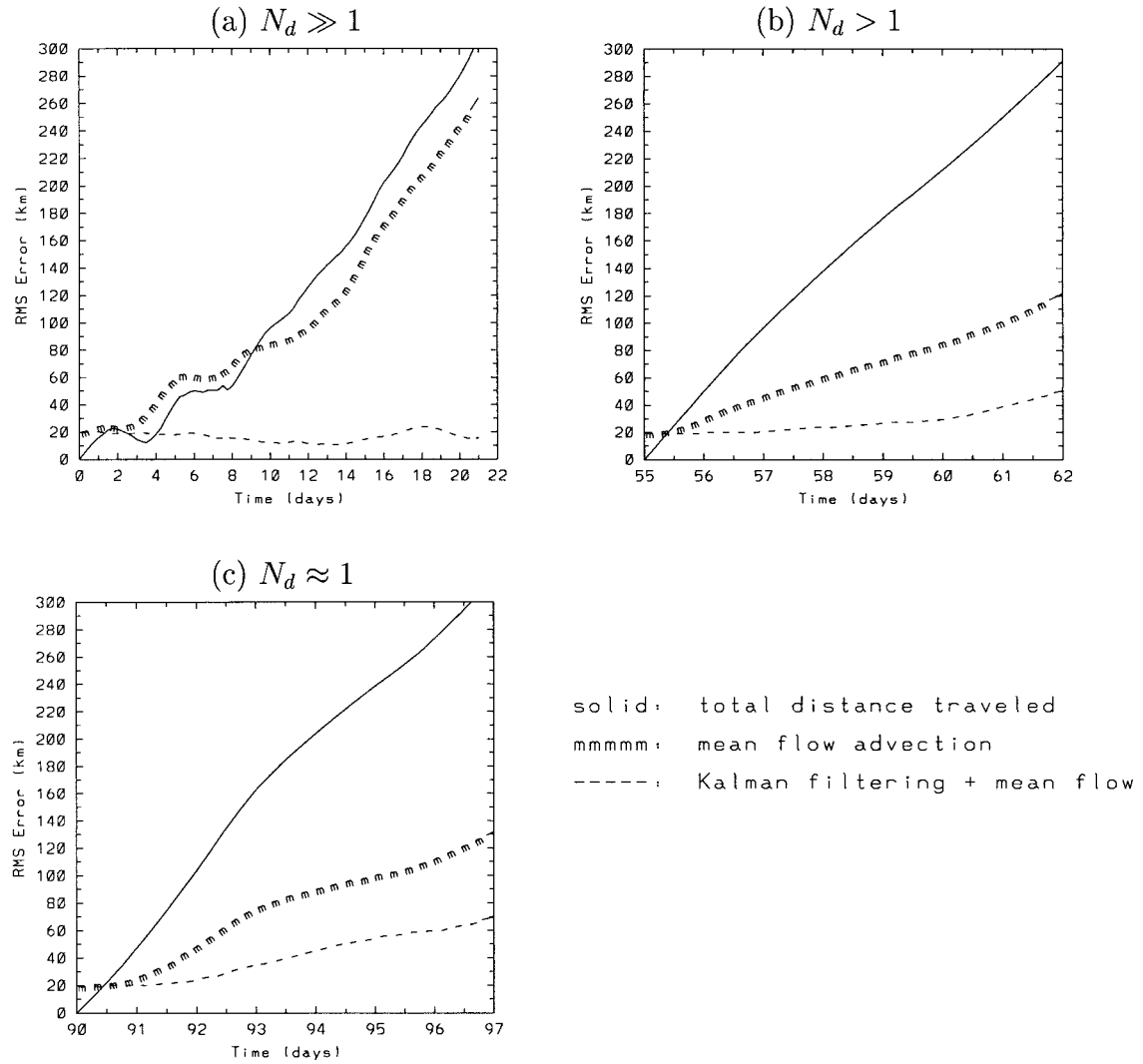

FIG. 9. Prediction errors with mean flow advection s_m and by using data assimilation s_a (with mean flow **U**) for Cluster I subject to 20-km initial position errors, during (a) $0 \le t \le 21$ days ($N_d \gg 1$ drifter/degree^2), (b) $55 \le t \le 62$ days ($N_d > 1$ drifter/degree^2), and (c) $90 \le t \le 97$ days ($N_d \approx 1$ drifter/degree^2).

4a for 100 days, which indicates that during this period, the cluster goes through three distinct regimes. For the first 45 days, the particles remain very close with $\overline{d} \approx 10$ km. Then a transition to a different regime takes place, characterized by $\overline{d} \approx 70$ km between $t = 55$–80 days. Afterward, particles disperse rapidly and $\overline{d}$ exceeds 400 km at $t = 100$ days, at which point in time we assume that the existence of the cluster expires (primarily due to two drifters separating from the cluster) and we focus our attention to the first 100 days of motion.

The cluster data density N_d for Cluster I is plotted in Figs. 4b,c. Based on this figure, five different regimes are characterized. Regime (i): $0 \le t \le 21$ days, $N_d > 500$ drifters/degree^2; Regime (ii): $22 \le t \le 29$ days, $1000 \le N_d < 3000$ drifters/degree^2; Regime (iii): $55 \le t \le 80$ days, $N_d > 1$ drifter/degree^2; Regime (iv): $90 \le t \le 97$ days, $N_d \approx 1$ drifter/degree^2; and Regime (v): $t > 100$ days, $N_d \ll 1$ drifter/degree^2.

Predictability analysis is performed for these five regimes separately, by calculating the distance traveled s, the rms prediction errors using mean flow advection s_m, the center of mass s_c, and using data assimilation s_a with and without mean flow **U**. For s_m and s_a, the particles are initialized from the exact positions along their true trajectory in these five sets of integrations. Figure 5a shows the cluster-averaged results for the first regime $0 \le t \le 21$ days, where $N_d > 500$ drifters/degree^2. Over this period, the drifters travel an average distance of approximately 300 km. The use of mean flow advection for the purpose of prediction is not effective, $s_m \approx 260$ km at $t = 21$ days. Data assimilation technique, however, gives very accurate results; $s_a < 10$ km throughout the entire three week period. As obvious from Figs. 3

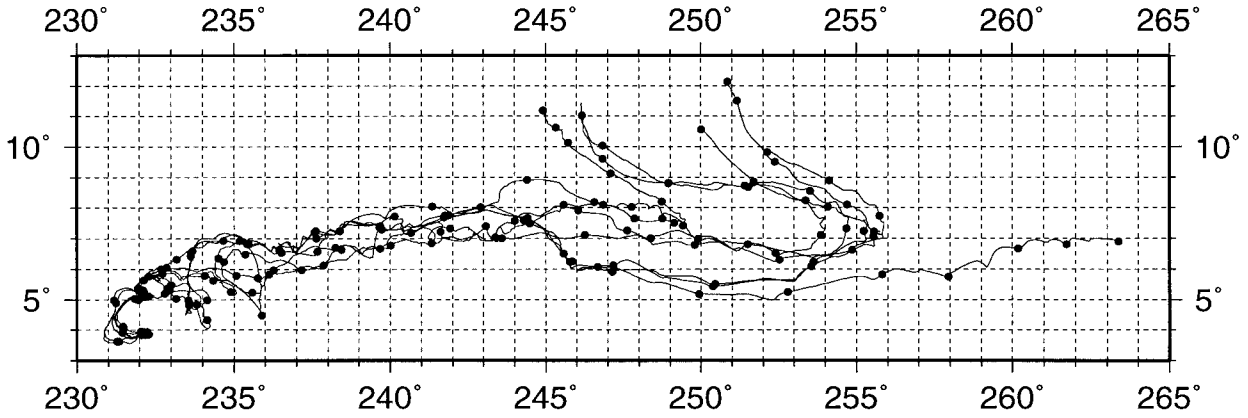

FIG. 10. Drifter trajectories in Cluster II (buoy id numbers: 11 663, 11 672, 11 678, 11 682, 11 685, 11 688, 11 689) for approximately 160 days. The circles mark 7-day intervals.

and 4a, this accuracy is due to the fact that all drifters stay together during this period. Therefore, a far simpler technique, the center of mass, is as accurate as the data assimilation method for $t > 7$ days. The performance of the data assimilation method is better only for the initial few days, which can be of high importance in practical applications (Schneider 1998): at $t = 2$ days, $s_c = 8.5$ km while $s_a = 1.9$ km. The effect of the mean flow **U** in the calculation of s_a is negligible; assimilation of high density data compensates for the lack of knowledge of the mean flow, as concluded in Özgökmen et al. (2000).

In the second regime investigated, $22 \leq t \leq 29$ days, the data density is even higher than the first one. However, there is no qualitative and quantitative difference in s_a and s_c between the first and second regimes (Fig. 5b).

Predictions for the third regime when $N_d > 1$ drifter/degree^2 are conducted by sampling the first week of this time interval, $55 \leq t \leq 62$ days. The results are shown in Fig. 5c. The drifters travel nearly 300 km after 7 days. The rms error with mean flow advection yields $s_m \approx 120$ km at $t = 62$ days. The performance of the mean flow advection method depends on the variability of the flow. The fluctuation velocity variance is $\sigma_v \approx 24$ (cm s^{-1}) when calculated over $0 \leq t \leq 100$ days, but during the third regime, fluctuations are much smaller $\sigma_v \approx 9$ cm s^{-1} for $55 \leq t \leq 62$ days) leading to an improved performance for the mean flow advection method. As indicated in Fig. 4a, the particles spread apart during this time interval, and with $s_c \approx 50$ km, center of mass of is not a satisfactory predictor for drifter positions anymore. However, the data assimilation model remains very accurate, $s_a = 2.5$ ($s_a = 16.5$, with mean flow) km at $t = 57$ ($t = 62$) days.

In the fourth regime, $90 \leq t \leq 97$ days, when $N_d \approx 1$ drifter/degree^2, $s > 300$ km after 7 days (Fig. 5d), similar to that in the third regime. The particles are so far apart, the center of mass method is not a satisfactory means of prediction, with s_c exceeding 180 km even at the beginning ($t = 90$ days). Despite low data density, the data assimilation method performs satisfactorily, $s_a = 15.6$ ($s_a = 56.8$, with mean flow) km when compared to the mean flow method, $s_m = 44.6$ ($s_m = 129.2$) km at $t = 92$ ($t = 97$) days, respectively.

Finally, examples of predictions during the fifth regime when $N_d \ll 1$ drifter/degree^2 are shown in Fig. 5e for the period $100 \leq t \leq 107$ days. The primary result is that the data assimilation method does not lead to much improvement over the mean flow advection technique, simply because the drifters are too far apart, and their motion is not as correlated as before.

An important conclusion from these experiments is that the utility of the data assimilation method appears to be when $N_d \geq 1$ drifter/degree^2 (Figs. 5c,d). In this regime, this technique outperforms the others. When $N_d \gg 1$ drifter/degree^2 (Figs. 5a,b), complex data assimilation scheme is unnecessary, since the simple center of mass technique is nearly equally accurate. When $N_d \ll 1$ drifter/degree^2, no technique investigated in this study is found to give accurate predictions. Probably more complete dynamical models such as those including rotation and in situ wind forcing will be necessary for accurate prediction.

The effect of data density N_d on the prediction accuracy of the data assimilation method is summarized by replotting s_a for the regimes discussed above in Fig. 6: $N_d \ll 1$, $N_d \approx 1$, $N_d > 1$, and $N_d \gg 1$ drifter/degree^2 for a one week sampling period from each regime. This figure indicates that there is not a substantial difference in prediction accuracy between the case when $N_d > 1$ and $N_d \gg 1$ drifter/degree^2. This is encouraging since it implies that the data assimilation technique appears to be ''cost effective''; that is, it can give accurate results without the need of releasing a very high number of drifters in practical applications. The generality of this conclusion is investigated in the following by repeating the above procedure for other drifter clusters, with more complex patterns of motion.

Figures 5a,b indicate that there appears to be an upper

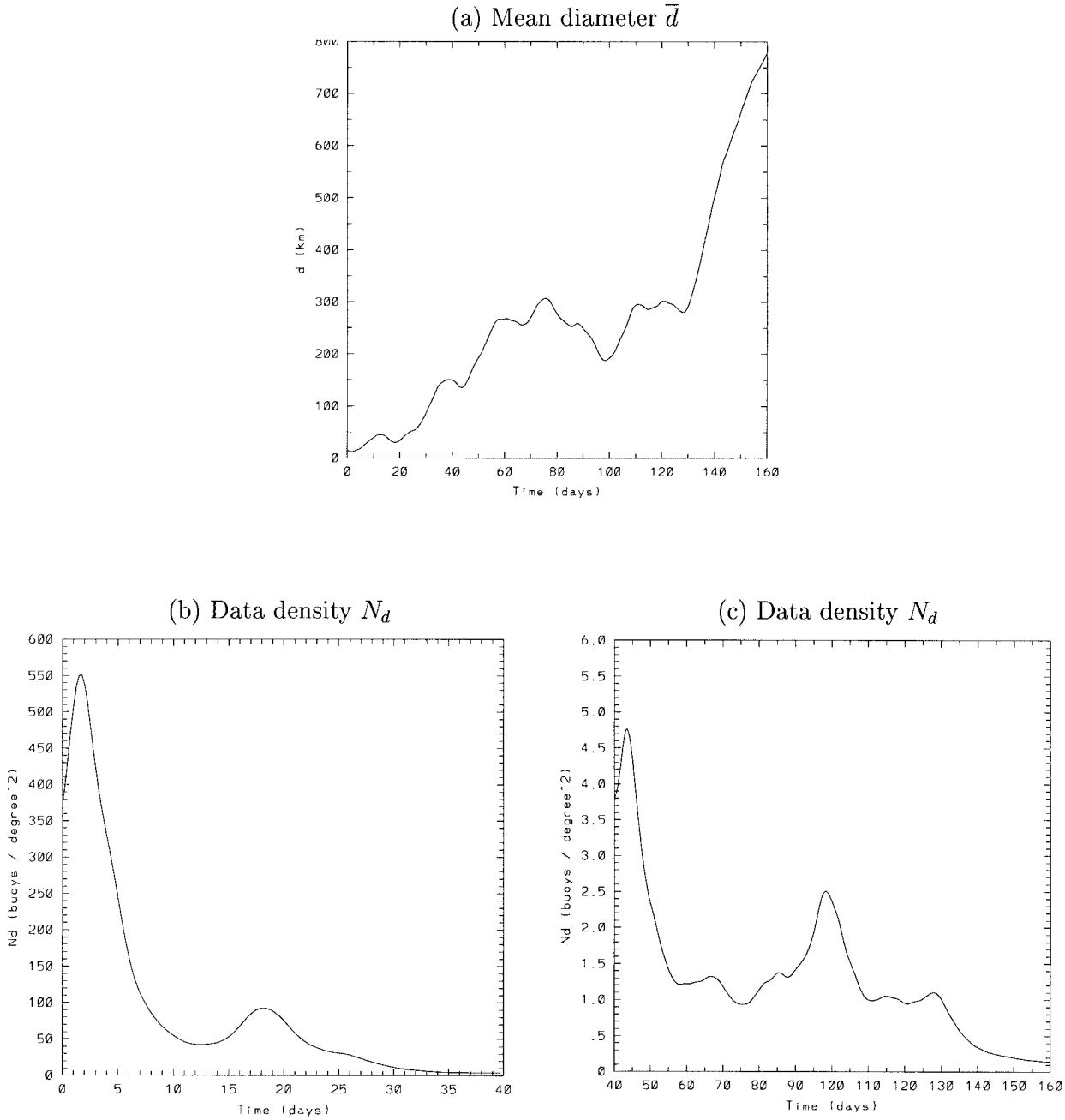

FIG. 11. (a) Mean diameter $\overline{d}$ (km) for $0 \leq t \leq 160$ days, (b) the cluster data density N_d (drifter/degree^2) for $0 \leq t \leq 40$ days, and (c) N_d for $40 \leq t \leq 160$ days for Cluster II.

bound for accuracy that our model cannot exceed, that is, there is no improvement in accuracy despite the fact that N_d is higher in the second regime than the first. Due to the standard of the Global Drifter Center, the drifter positions and velocities are assimilated every 6 h, and in between the model is marched forward based on the remaining simple dynamics, the mean flow field, and the memory term of the turbulent velocity. For instance, over the one week period when N_d is highest ($22 \leq t \leq 29$ days), the drifters travel approximately 200 km, or $\approx$7 km during each 6-h assimilation period, a distance comparable to rms error s_a. It appears that when the data density is very high, the sampling frequency (or assimilation frequency of data) and/or the advection errors introduced by such a crude dynamical model become important for improving the accuracy of

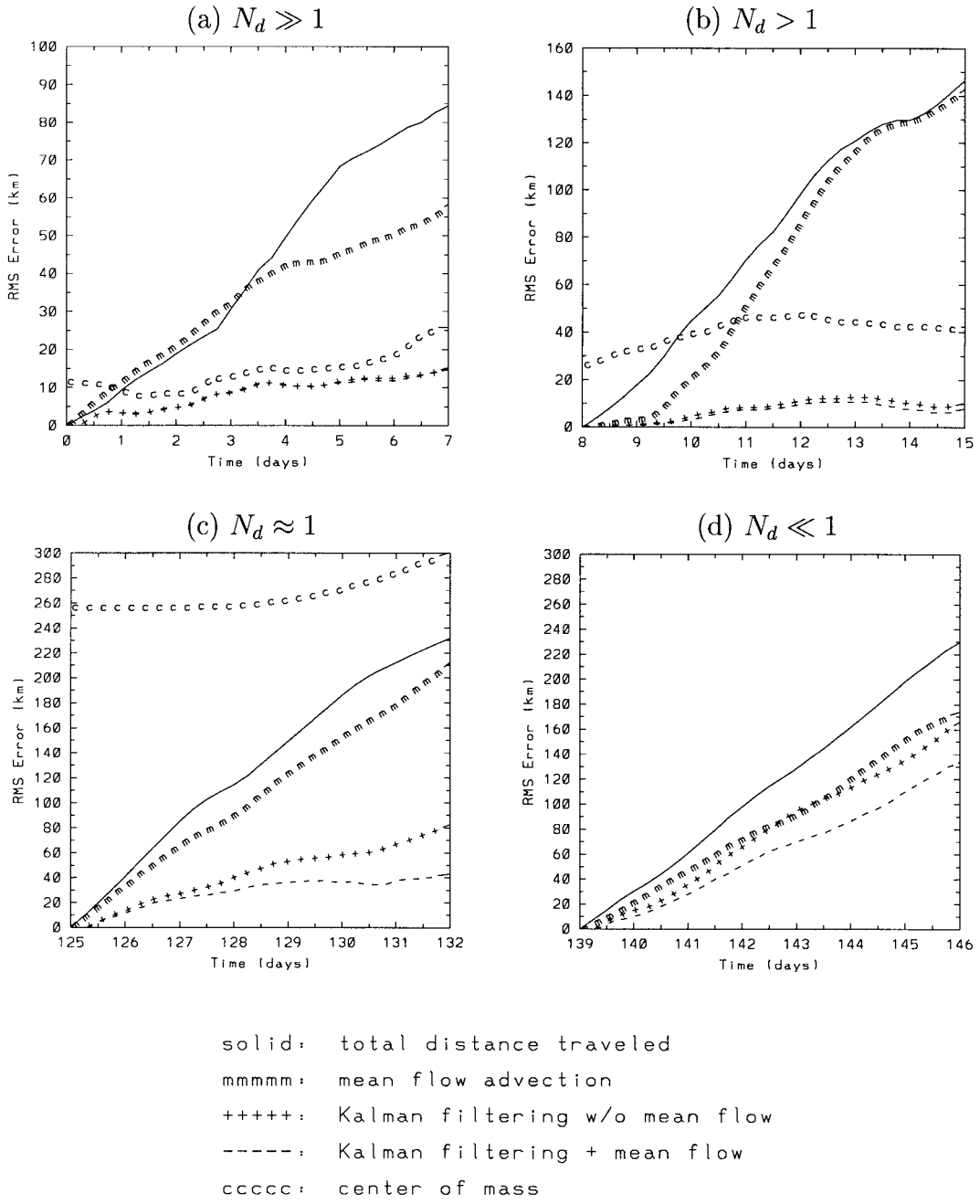

FIG. 12. Total distance traveled s, prediction error with mean flow advection s_m, prediction error based on the center of mass s_c, prediction error using data assimilation s_a with and without mean flow **U** for Cluster II during (a) $0 \leq t \leq 7$ days ($N_d \gg 1$ drifter/degree^2), (b) $8 \leq t \leq 15$ days ($N_d > 1$ drifter/degree^2), (c) $125 \leq t \leq 132$ days ($N_d \approx 1$ drifter/degree^2), and (d) $139 \leq t \leq 146$ days ($N_d \ll 1$ drifter/degree^2).

prediction. Here, we test the validity of the former statement, that is, the effect of the data assimilation frequency on the prediction error. Since we do not have data more frequently than every 6 h, the effect of reducing the assimilation frequency is investigated. Figure 7 shows the rms prediction error s_a with mean flow **U** obtained by assimilating surrounding drifter data every 6 h (as before Fig. 5a), every 12 h, every 24 h and every 48 h for the period $22 \leq t \leq 29$ days (when N_d is highest). Clearly, there is a significant reduction in prediction accuracy when the data assimilation frequency is reduced from every 6 h to every 12 h; with s_a increasing from $\approx$10 km to $\approx$54 km at $t = 29$ days. As anticipated, as the data assimilation frequency is re-

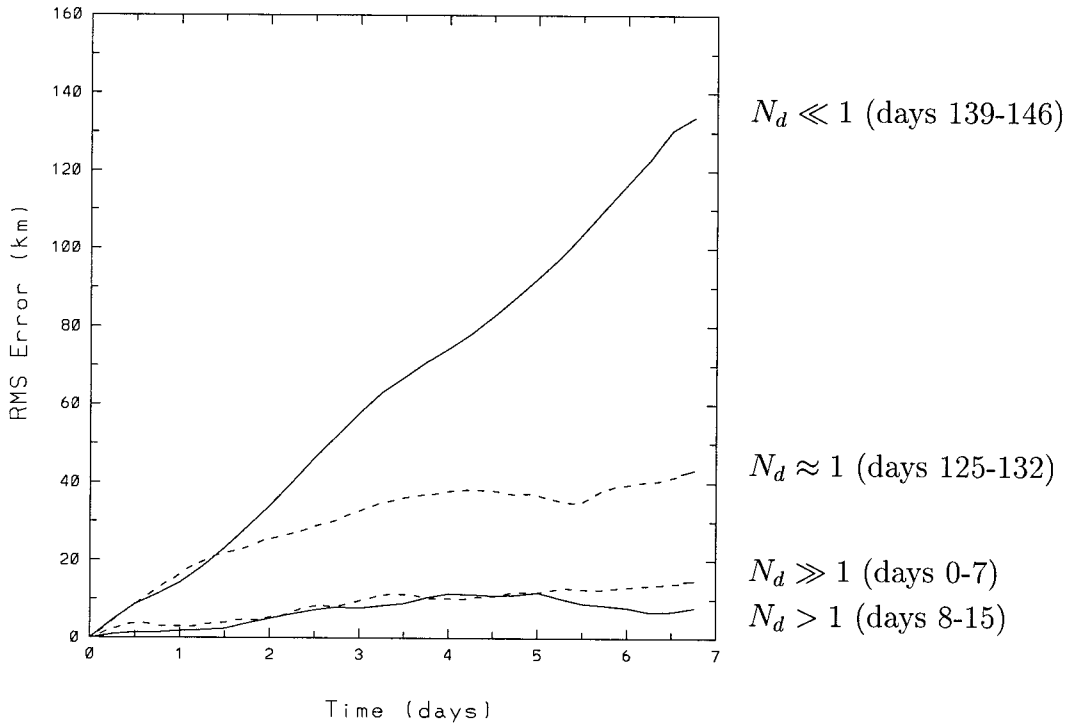

FIG. 13. The rms prediction error using data assimilation s_a (with **U**) in different regimes of data density along the drifter trajectories in Cluster II.

duced further, the rms error s_a converges to s_m, which is obtained without assimilation of data. The result illustrated in Fig. 7 indicates the importance of data assimilation frequency in regions characterized by strong currents.

Since the error plots presented so far do not necessarily give a clear idea about the predicted trajectories, true and predicted trajectories for one of the drifters are plotted in Fig. 8. The results are qualitatively similar for the other drifters. When $N_d \gg 1$ drifters/degree^2 and $N_d > 1$ drifter/degree^2 (Fig. 8a,b), very accurate predictions of the true trajectory are possible with the data assimilation method. As the data density decreases, the accuracy of predicted trajectories deteriorates gradually, as expected.

Finally, the effect of the knowledge of the initial release positions, which is not likely to be known exactly in practical applications, is investigated by introducing 20-km position errors in the eastward, westward, northward, and southward directions, respectively. The error curves s_a (with **U**) are then averaged and plotted for different data density regimes in Fig. 9. The same is also done for s_m for completeness. There appears to be a tendency for the initial error to be reduced when $N_d \gg 1$ drifter/degree^2 since s_a decreases from 20 km to approximately 10 km during the second week (Fig. 9a). However, this behavior is likely to be due to the convergence of drifters during this period (Fig. 4b). In the third week of the prediction, s_a behaves as before (i.e., as in Fig. 5a). For $N_d > 1$ drifter/degree^2 (Fig. 9b), the initial error is not recovered; s_a remains roughly at 20 km for the first 5 days and then it starts increasing at the same rate as that in Fig. 5c. Therefore, the initial position errors appear to be preserved. For lower data densities (Fig. 9c), s_a and s_m differ from their counterparts with exact initial position only during the period required to reach an rms error of 20 km (typically for the first 1–1.5 days).

c. Results for Cluster II

Cluster II consists of 7 drifters that have been released on 21 and 22 August 1990 at approximately 4°N, just south of the NECC. The trajectories of these drifters (Fig. 10) illustrate that while they propagate westward during the first week, all drifters turn northward in the second week and then they are entrained into the NECC and follow an eastward course for over 100 days (and approximately 2500 km). During the last few weeks of the 160-day period shown in Fig. 10, most of them appear to turn westward again, probably due to the effect of NEC. The average velocity variance over the entire period is $\sigma_v = 31$ cm s^{-1}, indicating that the drifters in this cluster are somewhat more energetic than those in Cluster I.

The variations in the mean distance $\overline{d}$ and the data density N_d over the 160-day period following the release of drifters are shown in Figs. 11a and 11b,c, respec-

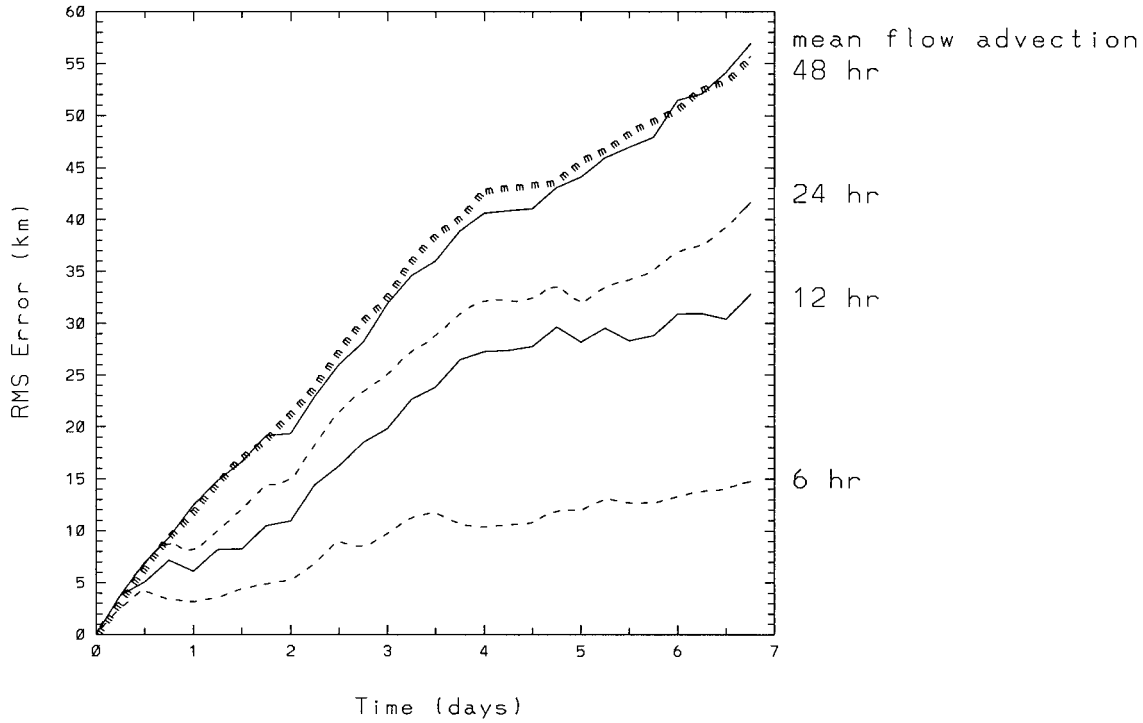

FIG. 14. The rms prediction error using data assimilation s_a (with **U**) for different data assimilation frequencies: every 6 h (dashed line), 12 h (solid line), 24 h (dashed line), and 48 h (solid line) during $0 \leq t \leq 7$ days for Cluster II. The line marked with *mmm* shows the prediction error using mean flow advection (i.e., when data are not assimilated).

tively. Based on these plots, the following one-week regimes are selected to test the general validity of the results discussed for Cluster I: Regime (i): $0 \leq t \leq 7$ days where $N_d \gg 1$ drifter/degree^2; Regime (ii): $8 \leq t \leq 15$ days where $N_d > 1$ drifter/degree^2; Regime (iii): $125 \leq t \leq 132$ days where $N_d \approx 1$ drifter/degree^2; and Regime (iv): $139 \leq t \leq 146$ days where $N_d \ll 1$ drifter/degree^2.

In the first regime, when $N_d \gg 1$ drifter/degree^2 the prediction errors are $s_m = 24$ ($s_m = 58.4$), $s_c = 9.4$ ($s_c = 25.8$), and $s_a = 5.3$ ($s_a = 14.8$, with mean flow) km at $t = 2$ ($t = 7$) days (Fig. 12a). As in the case of Cluster I, the accuracy of the center of mass technique is approximately the same as the data assimilation technique (within a factor of 2). In the second regime, when $N_d > 1$ drifters/degree^2 ($8 \leq t \leq 15$ days), the particles are far apart, and the center of mass technique yielding $s_c \approx 27$–49 km is not satisfactory (Fig. 12b). This is the regime, as postulated in section 4b, where the advantage of the data assimilation method over the others is greatest: $s_a = 3.7$ ($s_a = 7.8$, with mean flow) km as compared to $s_m = 21.3$ ($s_m = 143$) km at $t = 10$ ($t = 15$) days. As before, there is not a significant difference in s_a for $N_d > 1$ and $N_d \gg 1$ drifter/degree^2. In the third regime, characterized by a data density $N_d \approx 1$ drifter/degree^2 ($125 \leq t \leq 132$ days), the center of mass technique is useless with $s_c > 257$ km (Fig. 12c), which is greater than the total distance traveled s, the upper bound of error for any meaningful prediction method. While not as accurate as for higher data densities, the data assimilation technique offers significant gains with respect to the mean flow method; $s_a = 26$ ($s_a = 43.3$, with mean flow) km as opposed to $s_m = 74.3$ ($s_m = 212$) km at $t = 127$ ($t = 132$) days (i.e., roughly 300%–500% more accurate for this particular case). Finally, when $N_d \ll 1$ drifter/degree^2, the improvement in prediction accuracy of the data assimilation technique with respect to the mean flow advection method disappears (Fig. 12d).

The effect of data density on the predictive error of the data assimilation scheme s_a is summarized in Fig. 13. While the distance covered by the drifters s is not the same in the four different regimes plotted in this figure, it is still useful to illustrate that the availability of surrounding data leads to substantial gains over the predictive capability of the basic dynamical model based on mean flow advection (compare the cases $N_d \ll 1$ and $N_d \approx 1$ drifter/degree^2), and prediction accuracy quickly reaches a saturation level (i.e., errors for cases $N_d > 1$ and $N_d \gg 1$ drifter/degree^2).

The issue of the prediction accuracy reaching a saturation level at high data densities is in part related to the data assimilation frequency. Figure 14 illustrates s_a (with mean flow) from assimilating data every 6 h (as before), and then every 12, 24, and 48 h for the period of $0 \leq t \leq 7$ days when the data density is the highest. As for Cluster I, there appears to be a substantial difference in prediction accuracy between 6 h and 12 h

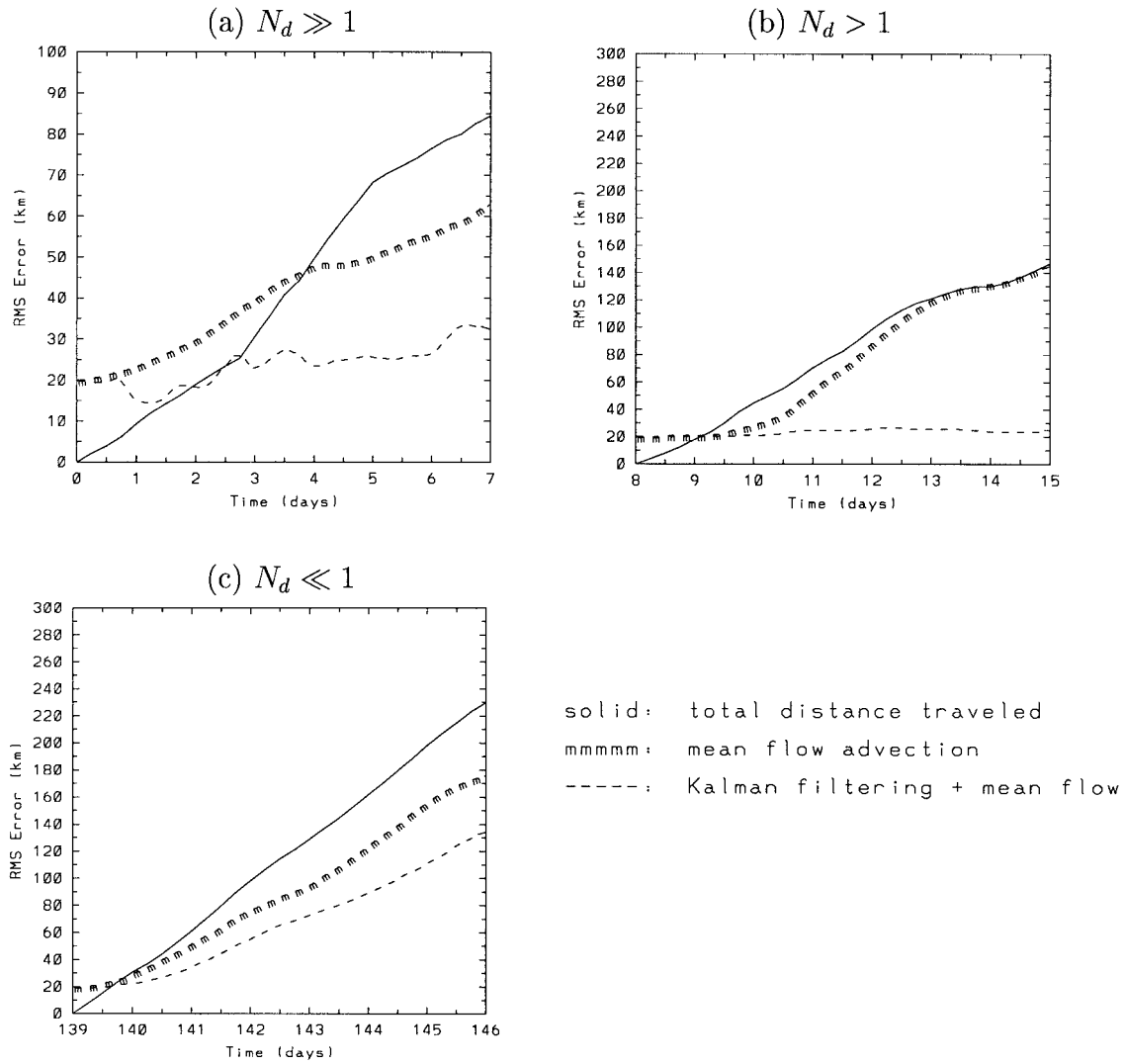

FIG. 15. Prediction errors using mean flow advection s_m and using data assimilation s_a (with mean flow **U**) for Cluster II subject to 20-km initial position errors, during (a) $0 \le t \le 7$ days ($N_d \gg 1$ drifter/degree^2), (b) $8 \le t \le 15$ days ($N_d > 1$ drifter/degree^2), and (c) $139 \le t \le 146$ days ($N_d \approx 1$ drifter/degree^2).

cases with $s_a \approx 15$ km and $s_a \approx 33$ km, respectively, at $t = 7$ days. As the assimilation frequency is reduced further, s_a converges to s_m; for the 48-h case, s_a is nearly identical to s_m.

As for Cluster I, the effect of an initial position error of 20 km (averaged in four directions) on the prediction error is tested for Cluster II. When $N_d \gg 1$ drifter/degree^2, there appears to a reduction of s_a from 20 km to approximately 14 km after three data assimilation sequences (Fig. 15a), most probably due to the convergence of trajectories shortly after drifters have been launched (Fig. 11b). The recovery in error is short lived, and $s_a \approx 33$ km at $t = 7$ days, nearly 20 km higher with respect to the case starting from exact initial positions (Fig. 12a). In the case when $N_d > 1$ drifter/degree^2 (Fig. 15b), there is no reduction of the initial error, but an error of approximately 20 km is maintained throughout the integration period. When $N_d \ll 1$ drifter/degree^2 (Fig. 15c), the prediction error curves, s_a and s_m, are different from those shown in Fig. 12d only for the first 1.5 days, which is the time to reach an rms error of 20 km in this case.

d. Results for Cluster III

Cluster III consists of 10 drifters that have been released on 1 and 2 December 1990 south of the NECC. All buoys drift westward in the first week and turn northward in the second and third weeks (Fig. 16). All drifters exhibit a sudden change in course when they reach the latitude of the NECC at approximately 5°N. Later, the drifter trajectories bifurcate, with 4 drifters recirculating

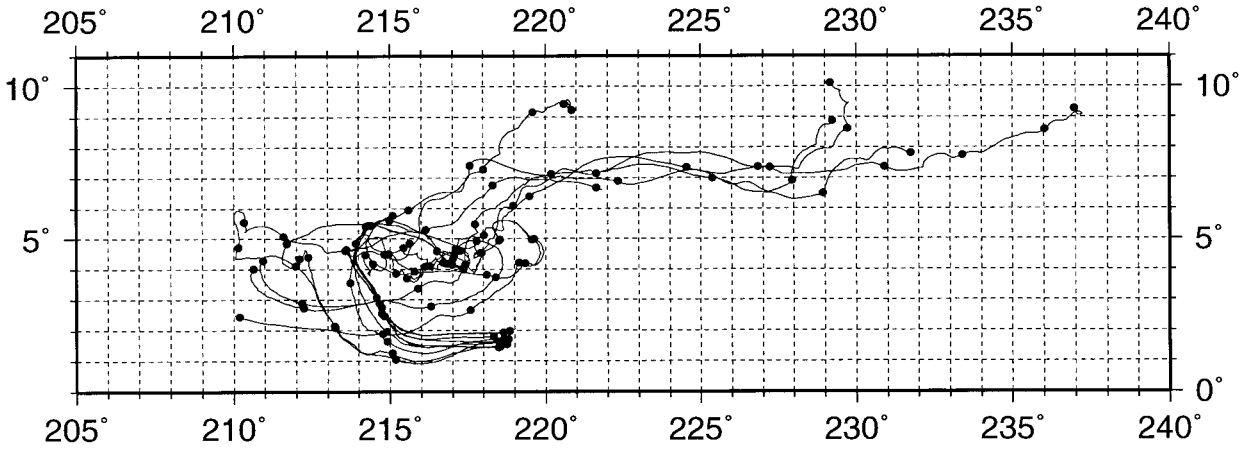

FIG. 16. Drifter trajectories in Cluster III (buoy id numbers: 11 662, 11 664, 11 666, 11 669, 11 674, 11 683, 11 684, 11 693, 11 699, 11 704) for approximately 70 days. The circles mark 7-day intervals.

westward and the other 6 continuing along the NECC. The average velocity variance over the 70-day period shown in Fig. 16 is $\sigma_v = 54$ cm s^{-1}, which is roughly twice as high as that for clusters I and II (Table 1). The analytical error in Özgökmen et al. (2000) indicates that the rate of growth in time of the rms prediction error is proportional to σ_v. While those derivations are strictly applicable for a homogeneous distribution of drifters (i.e., constant data density), one can anticipate that the rate of error growth will be much faster for Cluster III than that for the other clusters. Indeed, when the mean distance $\overline{d}$ for Cluster III is plotted in Fig. 17a, a fast rate of dispersion is observed, with $\overline{d}$ exceeding 1000 km after 60 days, which corresponds to the same distance in 40% (35%) of the time when compared to that in Cluster I (Cluster II). Based on the data density plots (Figs. 17b,c) the following regimes of a period of one week are chosen to apply the prediction methods: Regime (i): $0 \leq t \leq 7$ days where $N_d \gg 1$ drifter/degree^2; Regime (ii): $7 \leq t \leq 14$ days where $N_d > 1$ drifter/degree^2; Regime (iii): $50 \leq t \leq 57$ days where $N_d \ll 1$ drifter/degree^2.

In the first week, when $N_d \gg 1$ drifter/degree^2 the prediction errors are $s_m = 100$ ($s_m = 256$), $s_c = 37$ ($s_c = 92$), and $s_a = 10.8$ ($s_a = 61.7$, with or without mean flow) km at $t = 2$ ($t = 7$) days (Fig. 18a). As before, substantial gains in accuracy are obtained with the data assimilation technique in short period of prediction, roughly 10 times (4 times) better at $t = 2$ days with respect to that using mean flow (center of mass) method. However, due to the fast dispersion of this cluster, the errors s_a and s_c are rather high at $t = 7$ days. When $N_d > 1$ drifter/degree^2 ($7 \leq t \leq 14$ days), $s_c \approx 92$–168 km, hence center of mass is not a good indicator of drifter position. But, as for other clusters, the data assimilation method offers good accuracy: $s_a = 14$ ($s_a = 61$, with mean flow) km as compared to $s_m = 81$ ($s_m = 235.9$) km at $t = 9$ ($t = 14$) days. Again, s_a gradually converges with s_m for $N_d \ll 1$ drifters/degree^2 (Fig. 18c).

In agreement with the conclusions for the other clusters, significant improvement in s_a with respect to s_m is obtained provided that $N_d > 1$ drifter/degree^2 (Fig. 19). As before, s_a does not decrease continually as N_d increases. As shown before, assimilation frequency may play an important role in prediction accuracy as the dispersion rate of the cluster increases. This is demonstrated for Cluster III in Fig. 20, which illustrates s_a (with mean flow) from assimilating data every 6 h, and then every 12, 24, and 48 h for the period of $0 \leq t \leq 7$ days. The results are essentially the same as for the other two clusters: there is a significant deterioration in prediction accuracy between the 6- and 12-h cases, and s_a converges to s_m for the 48-h case.

5. Summary and conclusions

This study is conducted to improve oceanographers' ability to perform Lagrangian prediction, and to enhance our understanding of the predictability limits of particle motion in the real ocean. Lagrangian prediction problem is not only an outstanding conceptual problem due to the tendency of Lagrangian motion to be chaotic, but it also has important practical applications such as searching for objects lost at sea, ecological issues like spreading of pollutants or fish larvea, and design of observing systems.

The present study complements the two previous investigations conducted in the context of synthetic drifter data generated using a three-layer, idealized ocean model (Özgökmen et al. 2000) and real surface drifter data in a semienclosed sea, the Adriatic Sea (Castellari et al. 2001). In this study, three drifter clusters consisting of 5–10 drifters are selected from the 1988–96 WOCE dataset in the tropical Pacific Ocean. Therefore, the results of this study are reflective of predictability of particle motion in unbounded regions characterized by strong ocean currents and high shears.

Predictability of drifters is investigated by using three

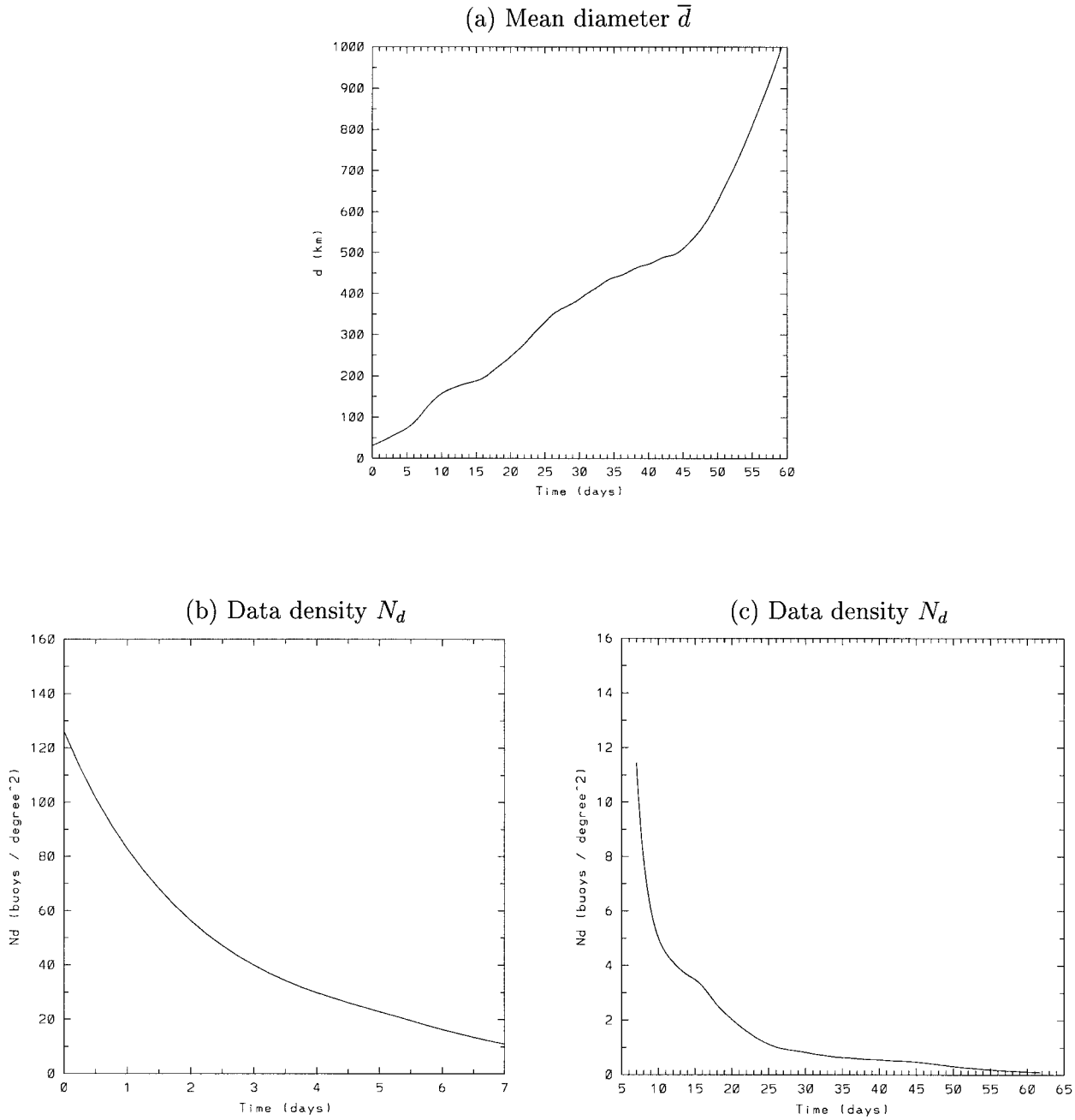

FIG. 17. (a) Mean diameter $\overline{d}$ (km) for $0 \leq t \leq 60$ days, (b) the cluster data density N_d (drifter/degree^2) for $0 \leq t \leq 7$ days, and (c) N_d for $7 \leq t \leq 60$ days for Cluster III.

methods with increasing complexity: the center of mass of the drifter cluster, advection by the climatological currents, and a new technique, which relies on the assimilation of data from the surrounding drifters into a Gauss–Markov model for particle motion. The assimilation is conducted using a Kalman filter, which differs from that in the aforementioned studies in that drifter positions as well as turbulent velocities are used in the assimilation. The mean flow field is calculated from the entire 1988–96 WOCE drifter dataset in the tropical Pacific by using a spline interpolation technique described by Bauer et al. (1998).

First, the difference in prediction accuracy of the present Kalman filtering scheme and the one used in the previous studies is quantified from a series of stochastic simulations of Lagrangian motion. We conclude that while the additional term for position assimilation does not make in general a substantial difference, improve-

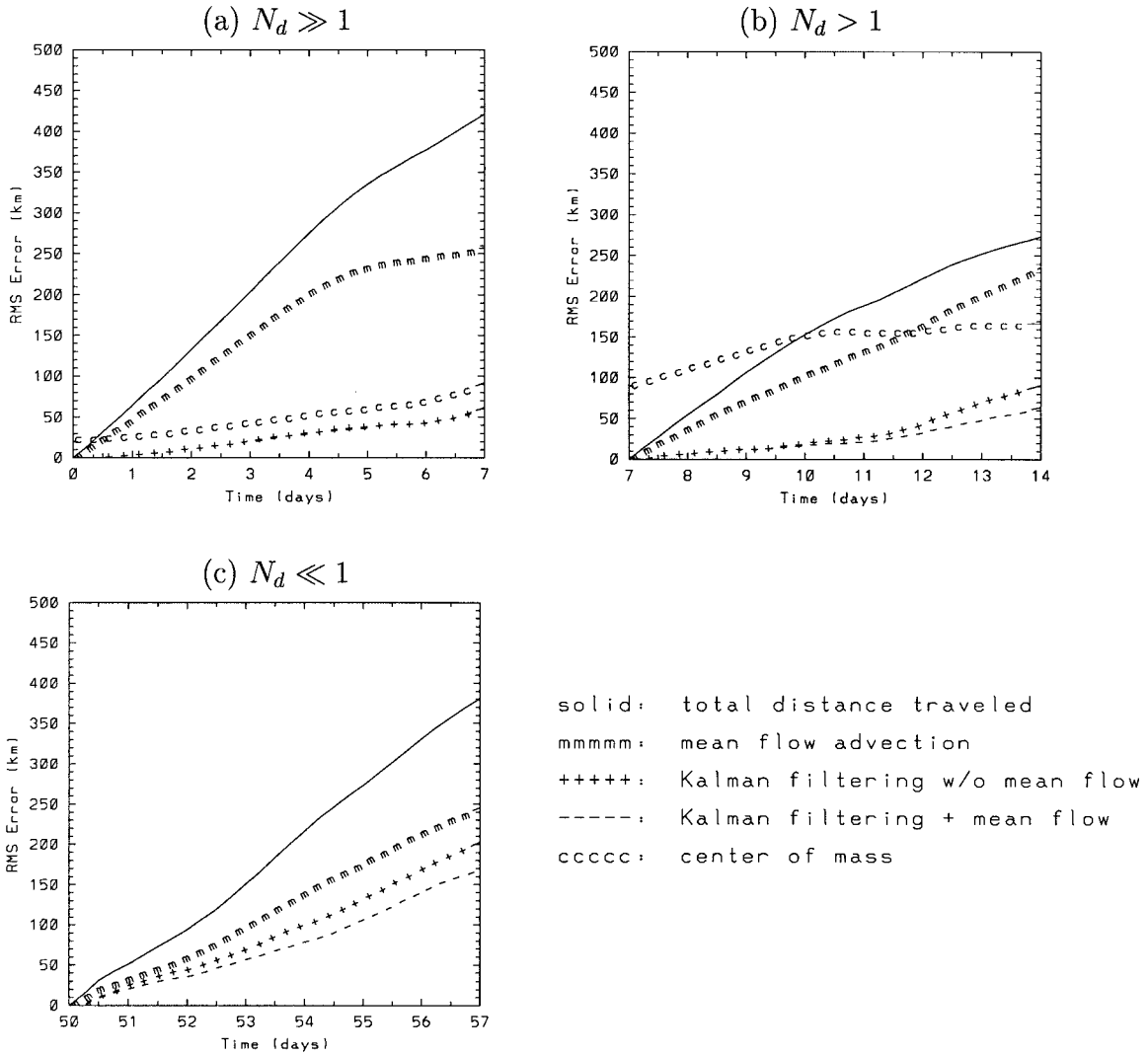

FIG. 18. Total distance traveled s, prediction error with mean flow advection s_m, prediction error based on the center of mass s_c, prediction error using data assimilation s_a with and without mean flow **U** for Cluster III during (a) $0 \leq t \leq 7$ days ($N_d \gg 1$ drifter/degree^2), (b) $7 \leq t \leq 14$ days ($N_d > 1$ drifter/degree^2), and (c) $50 \leq t \leq 57$ days ($N_d \ll 1$ drifter/degree^2).

ments up to 30% over a timescale of one week are calculated. Therefore, the new Kalman filtering formulation, which is described in detail, is used in the second part of the study, for the predictability analysis of oceanic drifters.

The primary accomplishment of the present study is to quantify the range of effectiveness of a new technique when applied to oceanic drifter clusters. The results indicate that three general regimes can be characterized using the cluster data density N_d, defined as the number of drifters over an area scaled by the mean diameter of the cluster. This study is not aimed at quantifying the effects of small-scale turbulence and time variability on particle dispersion and predictability. Such effects are implicitly represented by N_d, which then trivially controls the effectiveness of the technique. In the first regime, which corresponds to the period after the release of drifters in a tight cluster when $N_d \gg 1$ drifter/degree^2, the center of mass technique performs nearly as well as the data assimilation technique, and both methods yield very accurate predictions with rms errors $\leq$15 km over a timescale of a few days: about 7 days for clusters I and II with $\sigma_v = 24$–31 cm s^{-1}, and about 3 days for Cluster III with $\sigma_v = 54$ cm s^{-1}. When the drifters start to disperse, that is, in the regime where $N_d \geq 1$ drifter/degree^2, the data assimilation technique is the only method that gives accurate results. This result shows the utility of the data assimilation technique and implies that this technique is cost effective; that is, it can give accurate results without the need of releasing a very high number of drifters in practical applications. Finally, when $N_d \ll 1$ drifter/degree^2, no method investigated in this study is effective. Probably more complete dynamical models including the effect of rotation and in

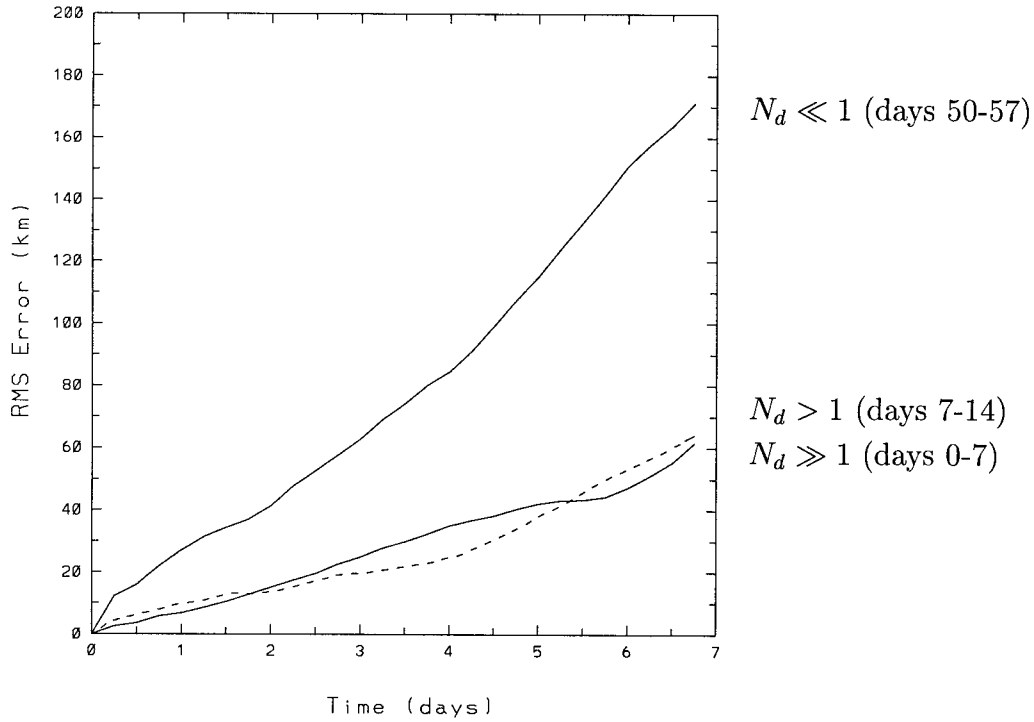

FIG. 19. The rms prediction error using data assimilation s_a (with **U**) in different regimes of data density along the drifter trajectories in Cluster III.

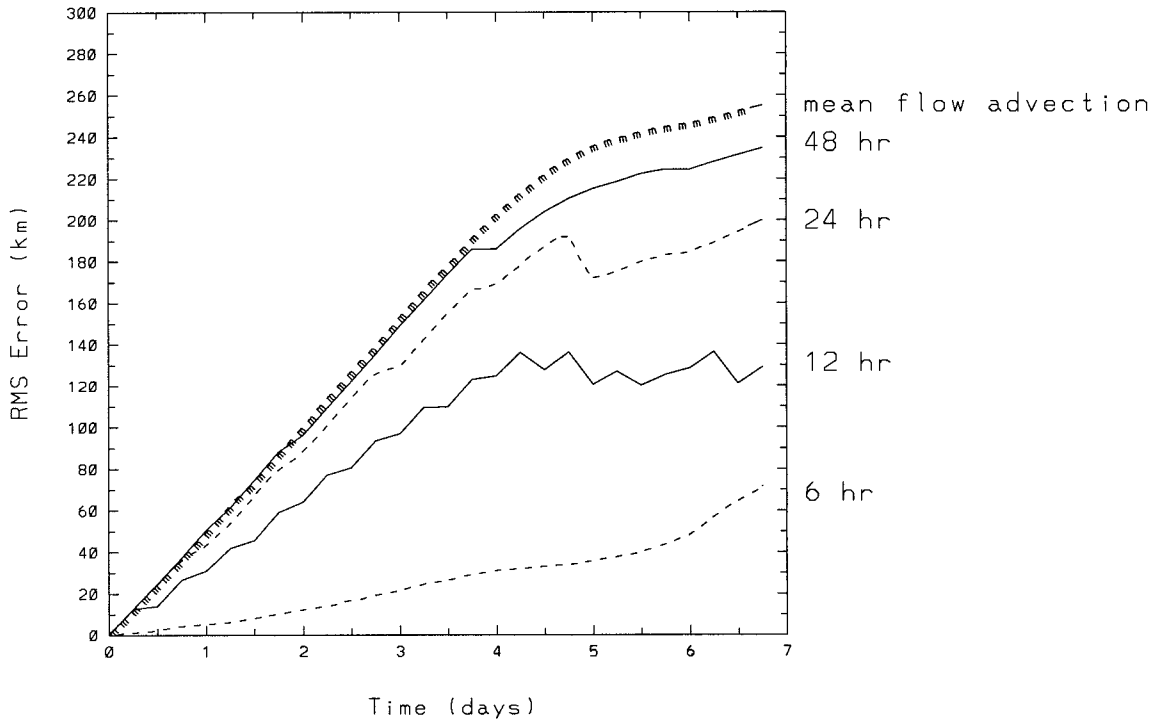

FIG. 20. The rms prediction error using data assimilation s_a (with **U**) for different data assimilation frequencies: every 6 h (dashed line), 12 h (solid line), 24 h (dashed line), and 48 h (solid line) during $0 \leq t \leq 7$ days for Cluster III. The line marked with *mmm* shows the prediction error using mean flow advection (i.e., when data are not assimilated).

situ wind forcing can be useful. We also find, in general, that advection by the mean flow field is not a good indicator of drifter motion. Given that this area has relatively large amounts of drifter data to calculate the mean flow, data assimilation methods and dynamical data will be necessary for most of the world's oceans.

The effects of uncertainties in the knowledge of the mean flow field and initial release positions on the predictions by the data assimilation method are investigated. It is found that the effect of the mean flow on the data assimilation method is negligible; assimilation of high density data compensates for the lack of knowledge of the mean flow, as concluded in Özgökmen et al. (2000). It appears in general that the initial position errors are not recovered. Therefore, the knowledge of initial positions seems to be an important factor for practical applications.

We note that the data assimilation method does not show a significant difference between the regimes $N_d \gg 1$ and $N_d > 1$ drifter/degree^2. While this can be interpreted that the model accuracy is reaching a degree of saturation, a series of experiments conducted with assimilation periods of 6, 12, 24, and 48 h indicate that the data-availability or assimilation frequency is likely to be an important factor for improving prediction accuracy in regions of strong currents. It is shown that assimilating surrounding drifter data every 6 h gives significantly better predictions (i.e., reducing the rms error by half or more) when compared to that from carrying out the assimilation procedure every 12 h. Consequently, 3-day sampling periods used by a number of new drifters may lead to unsatisfactory results using the data assimilation techniques described in this study.

As final remarks, we point out that the deterministic part of the present model consists only of advection by the climatological flow field. There is not sufficient information for accurate prediction capability in the absence of data assimilation, that is, once the drifters in the cluster have dispersed or between the assimilation moments. Therefore, improvement of deterministic model dynamics and employing other data (e.g., in situ winds) can be considered as a next step. Finally, given the practical importance of the first few days after the launch in practical applications and the accuracy of the center of mass technique in this regime, new formulations based on the center of mass of the cluster will be considered in future studies.

Acknowledgments. The authors greatly appreciate the support of the Office of Naval Research under Grants N00014-97-1-0620 and N00014-99-1-0049 (T. M. Özgökmen), N00014-99-1-0042 (L. I. Piterbarg), and N00014-95-1-0257 and N00014-99-1-0049 (A. J. Mariano and E. Ryan). The authors are grateful to A. Griffa for many useful discussions and she has contributed conceptually to this paper as much as the authors. The authors thank S. Bauer for supplying the code for mean flow calculation from the drifter dataset. Finally, we thank the two reviewers for their constructive comments, which led to an improved manuscript.

APPENDIX A

Discretization of the Particle Stochastic Model

For fixed t_0, Δt linearize (1) at the point $\mathbf{r}(t_0)$ by setting for all $t \in (t_0, t_0 + \Delta t)$

$$\mathbf{U}(\mathbf{r}(t)) = \mathbf{U}(\mathbf{r}(t_0)) + \nabla\mathbf{U}(\mathbf{r}(t_0))(\mathbf{r}(t) - (\mathbf{r}(t_0)),$$

$$\xi(t, \mathbf{r}(t)) = \xi(t, \mathbf{r}_0(t)).$$

Denote

$$\mathbf{U} = \mathbf{U}(\mathbf{r}(t_0)), \qquad \mathbf{G} = \nabla\mathbf{U}(\mathbf{r}(t_0)), \qquad \xi(t) = \xi(t, \mathbf{r}(t_0))$$

and write (1) in linearized form

$$\dot{\mathbf{v}} = -\Lambda\mathbf{v} + \xi(t), \qquad \dot{\mathbf{r}} = \mathbf{U} + \mathbf{G}\mathbf{r} + \mathbf{v}.$$

Solving this linear system with constant coefficients on $(t_0, t_0 + \Delta t)$ we get

$$\mathbf{v}(t) = \exp[-(t - t_0)\Lambda]\left\{\mathbf{v}(t_0) + \int_{t_0}^{t} \exp[-(t - s)\Lambda]\xi(s)\, ds\right\},$$

$$\begin{aligned}
\mathbf{r}(t) &= \exp[(t - t_0)\mathbf{G}]\left\{\mathbf{r}(t_0) + \int_{t_0}^{t} \exp[(t - s)\mathbf{G}][\mathbf{U} + \mathbf{v}(s)]\, ds\right\} \\
&= \exp[(t - t_0)\mathbf{G}]\left\{\mathbf{r}(t_0) + \int_{t_0}^{t} \left\{\exp[(t - s)\mathbf{G}](\mathbf{U} + \exp[-(s - t_0)\Lambda]\mathbf{v}(t_0) + \int_{t_0}^{s} \exp[(s - x)\Lambda)]\xi(x)\, dx\right\} ds\right\} \\
&= \exp[(t - t_0)\mathbf{G}]\mathbf{r}(t_0) + \int_{t_0}^{t} \exp[(t - s)\mathbf{G}]\, ds\mathbf{U} + \int_{t_0}^{t} \exp[(t - s)\mathbf{G}]\, \exp[-(s - t_0)\Lambda]\, ds\mathbf{v}(t_0) \\
&\quad + \int_{t_0}^{t}\int_{t_0}^{s} \exp[(t - s)\mathbf{G}]\, \exp[-(s - x)\Lambda]\xi(x)\, dx\, ds.
\end{aligned}$$

Then we set $t_0 = (n-1)\Delta t$ and $t = n\Delta t$, change the variables by $s = t_0 + u$, $x = t_0 + v$ in the integrals, and arrive at

$$\mathbf{v}(n) = \mathbf{C}\mathbf{v}(n-1) + \xi(n),$$
$$\mathbf{r}(n) = \mathbf{D}(n-1)\mathbf{r}(n-1) + \mathbf{F}(n-1)\mathbf{U}(n-1) + \mathbf{S}(n-1)\mathbf{v}(n-1) + \eta(n),$$

where

$$\mathbf{U}(n) = \mathbf{U}(\mathbf{r}(n))$$

is the mean velocity at moment n at position $\mathbf{r}(n)$. The matrix coefficients are given by

$$\mathbf{C} = \exp(-\Delta t\boldsymbol{\Lambda}), \qquad \mathbf{D}(n) = \exp(\Delta t\mathbf{G}(n)),$$
$$\mathbf{F}(n) = \int_0^{\Delta t} \exp(u\mathbf{G}(n))\, du,$$
$$\mathbf{S}(n) = \mathbf{D}(n)\int_0^{\Delta t} \exp(-u\mathbf{G}(n))\exp(-u\boldsymbol{\Lambda})du,$$

where

$$\mathbf{G}(n) = \nabla\mathbf{U}(\mathbf{r}(n)).$$

Random vectors $\xi(n)$, $\eta(n)$ are expressed in terms of the original stochastic forcing as

$$\xi(n) = \int_0^{\Delta t} \exp(-u\boldsymbol{\Lambda})\xi[(n-1)\Delta t + u, \mathbf{r}(n-1)]\, du,$$
$$\eta(n) = \int_0^{\Delta t} \mathbf{T}(n, u)\xi[(n-1)\Delta t + u, \mathbf{r}(n-1)]\, du,$$

where

$$\mathbf{T}(n, u) = \mathbf{D}(n)\int_u^{\Delta t} \exp(-v\mathbf{G}(n))\exp(-v\boldsymbol{\Lambda})\, dv \times \exp(-u\boldsymbol{\Lambda}).$$

Notice that for the same Lagrangian time T for both components and invertible $\mathbf{G}$ the above formulas are essentially simplified

$$\mathbf{v}(n) = \exp(-\Delta t/T)\mathbf{v}(n-1) + \xi(n)$$
$$\mathbf{r}(n) = \exp(\Delta t\mathbf{G})\mathbf{r}(n-1) + \mathbf{G}^{-1}[\exp(\Delta t\mathbf{G}) - \mathbf{I}] \times \mathbf{U}(n-1) + [\exp(\Delta t\mathbf{G}) - \exp(-\Delta t/T)\mathbf{I}] \times (\mathbf{G} + T^{-1}\mathbf{I})^{-1}\mathbf{v}(n-1) + \eta,$$

with

$$\xi(n) = \int_0^{\Delta t} \exp(-u/T)\xi[(n-1)\Delta t + u, \mathbf{r}(n-1)]\, du,$$
$$\eta(n) = \int_0^{\Delta t} [\exp(u\mathbf{G}) - \exp(-u/T)\mathbf{I}](\mathbf{G} + T^{-1}\mathbf{I})^{-1} \times \xi[(n-1)\Delta t + u, \mathbf{r}(n-1)]\, du.$$

So far we dealt with the linearization and discretization for the one particle motion. Now consider the M-particle motion

$$\dot{\mathbf{v}}_1 = -\boldsymbol{\Lambda}\mathbf{v}_1 + \xi(t, \mathbf{r}_1),$$
$$\dot{\mathbf{r}}_1 = \mathbf{U}(\mathbf{r}_1) + \mathbf{v}_1,$$
$$\dot{\mathbf{v}}_2 = -\boldsymbol{\Lambda}\mathbf{v}_2 + \xi(t, \mathbf{r}_2),$$
$$\dot{\mathbf{r}}_2 = \mathbf{U}(\mathbf{r}_2) + \mathbf{v}_2,$$
$$\vdots$$
$$\dot{\mathbf{v}}_M = -\boldsymbol{\Lambda}\mathbf{v}_M + \xi(t, \mathbf{r}_M),$$
$$\dot{\mathbf{r}}_M = \mathbf{U}(\mathbf{r}_M) + \mathbf{v}_M.$$

The important observation is that all M systems are related only through the correlations in the noise. Thus, we can carry out the linearization and discretization for each pair of equations exactly same way as before and take into account the space correlation of the noise. The result is

$$\mathbf{v}_1(n) = \mathbf{C}\mathbf{v}_1(n-1) + \xi_1(n)$$
$$\mathbf{r}_1(n) = \mathbf{D}_1\mathbf{r}_1(n-1) + \mathbf{F}_1(n)\mathbf{U}_1(n-1) + \mathbf{S}_1(n)\mathbf{v}_1(n-1) + \eta_1(n)$$
$$\mathbf{v}_2(n) = \mathbf{C}\mathbf{v}_2(n-1) + \xi_2(n)$$
$$\mathbf{r}_2(n) = \mathbf{D}_2\mathbf{r}_2(n-1) + \mathbf{F}_2(n)\mathbf{U}_2(n-1) + \mathbf{S}_2(n)\mathbf{v}_2(n-1) + \eta_2(n)$$
$$\vdots$$
$$\mathbf{v}_M(n) = \mathbf{C}\mathbf{v}_M(n-1) + \xi_M(n)$$
$$\mathbf{r}_M(n) = \mathbf{D}_M\mathbf{r}_M(n-1) + \mathbf{F}_M(n)\mathbf{U}_M(n-1) + \mathbf{S}_M(n)\mathbf{v}_M(n-1) + \eta_M(n),$$

where

$$\mathbf{C} = \exp(-\Delta t\boldsymbol{\Lambda}), \qquad \mathbf{D}_j(n) = \exp(\Delta t\mathbf{G}_j(n)),$$
$$\mathbf{F}_j(n) = \int_0^{\Delta t} \exp(u\mathbf{G}_j(n))\, du$$
$$\mathbf{S}_j(n) = \mathbf{D}_j(n)\int_0^{\Delta t} \exp(-u\mathbf{G}_j(n))\exp(-u\boldsymbol{\Lambda})\, du$$

and

$$\xi_j(n) = \int_0^{\Delta t} \exp(-u\boldsymbol{\Lambda})\xi[(n-1)\Delta t + u, \mathbf{r}_j(n-1)]\, du,$$
$$\eta_j(n) = \int_0^{\Delta t} \mathbf{T}_j(n-1, u) \times \xi[(n-1)\Delta t + u, \mathbf{r}_j(n-1)]\, du.$$

The conditional expectations given the positions at the previous moment are expressed in the form

$$E_{n-1}\{\xi_i(n)\xi_j(n)^*\} = \int_0^{\Delta t} \exp(-u\mathbf{\Lambda})\mathbf{B}[\mathbf{r}_i(n-1), \mathbf{r}_j(n-1)] \exp(-u\mathbf{\Lambda}^*)\, du,$$

$$E_{n-1}\{\eta_i(n)\eta_j(n)^*\} = \int_0^{\Delta t} \mathbf{T}_i(n, u)\mathbf{B}[\mathbf{r}_i(n-1), \mathbf{r}_j(n-1)]\mathbf{T}(n, u)^*\, du,$$

$$E_{n-1}\{\xi_i(n)\eta_j(n)^*\} = \int_0^{\Delta t} \exp(-u\mathbf{\Lambda})\mathbf{B}[\mathbf{r}_i(n-1), \mathbf{r}_j(n-1)]\mathbf{T}_j(n, u)^*\, du.$$

In particular the ''one-and-one-half'' order approximation gives

$$\begin{aligned}
\mathbf{v}_1(n) &= \mathbf{v}_1(n-1) - \Delta t\mathbf{\Lambda}\mathbf{v}_1(n-1) + \xi_1(n),\\
\mathbf{r}_1(n) &= \mathbf{r}_1(n-1) + [\mathbf{U}_1(n-1) + \mathbf{v}_1(n-1)]\Delta t + \eta_1(n)\\
\mathbf{v}_2(n) &= \mathbf{v}_2(n-1) - \Delta t\mathbf{\Lambda}\mathbf{v}_2(n-1) + \xi_2(n),\\
\mathbf{r}_2(n) &= \mathbf{r}_2(n-1) + [\mathbf{U}_2(n-1) + \mathbf{v}_2(n-1)]\Delta t + \eta_2(n)\\
&\vdots\\
\mathbf{v}_M(n) &= \mathbf{v}_M(n-1) - \Delta t\mathbf{\Lambda}\mathbf{v}_M(n-1) + \xi_M(n),\\
\mathbf{r}_M(n) &= \mathbf{r}_M(n-1) + [\mathbf{U}_M(n-1) + \mathbf{v}_M(n-1)]\Delta t + \eta_M(n),
\end{aligned}$$

with

$$E_{n-1}\{\xi_i(n)\xi_j(n)^*\} = \Delta t\mathbf{B}_{ij}(n-1),$$

$$E_{n-1}\{\eta_i(n)\eta_j(n)^*\} = \frac{(\Delta t)^3}{3}\mathbf{B}_{ij}(n-1),$$

$$E_{n-1}\{\xi_i(n)\eta_j(n)^*\} = \frac{(\Delta t)^2}{2}\mathbf{B}_{ij}(n-1),$$

where $\mathbf{B}_{ij}(n) = \mathbf{B}(\mathbf{r}_i(n), \mathbf{r}_j(n))$. Notice that the latter is expressible in the form

$$E_{n-1}\{\zeta_i(n)\zeta_j(n)^*\} = \frac{1}{6}\begin{pmatrix} 6\Delta t\mathbf{B}_{ij} & 3(\Delta t)^2\mathbf{B}_{ij} \\ 3(\Delta t)^2\mathbf{B}_{ij} & 2(\Delta t)^3\mathbf{B}_{ij}\end{pmatrix},$$

where

$$\zeta_i(n) = \begin{pmatrix}\xi_i(n)\\ \eta_i(n)\end{pmatrix}.$$

APPENDIX B

Derivation of the Adaptive EKF Formulas

We sum up the results of appendix A as

$$\mathbf{z}(n) = \mathbf{\Phi}(\mathbf{z}(n-1)) + \mathbf{q}(n, \mathbf{r}(n-1)), \qquad \text{(B1)}$$

where

$$\mathbf{z} = \begin{pmatrix}\mathbf{v}_1\\ \mathbf{r}_1\\ \mathbf{v}_2\\ \mathbf{r}_2\\ \vdots\\ \mathbf{v}_M\\ \mathbf{r}_M\end{pmatrix}, \qquad \mathbf{\Phi}(\mathbf{z}) = \begin{pmatrix}\mathbf{C}\mathbf{v}_1\\ \mathbf{D}_1\mathbf{r}_1 + \mathbf{F}_1\mathbf{U}(\mathbf{r}_1) + \mathbf{S}_1\mathbf{v}_1\\ \mathbf{C}\mathbf{v}_2\\ \mathbf{D}_2\mathbf{r}_2 + \mathbf{F}_2\mathbf{U}(\mathbf{r}_2) + \mathbf{S}_2\mathbf{v}_2\\ \vdots\\ \mathbf{C}\mathbf{v}_M\\ \mathbf{D}_M\mathbf{r}_M + \mathbf{F}_M\mathbf{U}(\mathbf{r}_M) + \mathbf{S}_M\mathbf{v}_M\end{pmatrix},$$

$$\mathbf{q} = \begin{pmatrix}\xi_1\\ \eta_1\\ \xi_2\\ \eta_2\\ \vdots\\ \xi_M\\ \eta_M\end{pmatrix}$$

and the $4M \times 4M$ covariance matrix of the noise is built with 4×4 blocks given by

$$\mathbf{Q}_{ij} = \begin{pmatrix} \int_0^{\Delta t} \exp(-u\mathbf{A})\mathbf{B}(\mathbf{r}_i, \mathbf{r}_j) \exp(-u\mathbf{A}^*)\, du & \int_0^{\Delta t} \exp(-u\mathbf{A})\mathbf{B}(\mathbf{r}_i, \mathbf{r}_j)\mathbf{T}_j(n, u)^*\, du \\ \int_0^{\Delta t} \mathbf{T}(n, u)\mathbf{B}(\mathbf{r}_j, \mathbf{r}_i) \exp(-u\mathbf{A}^*)\, du & \int_0^{\Delta t} \mathbf{T}(n, u)\mathbf{B}(\mathbf{r}_i, \mathbf{r}_j)\mathbf{T}_j(n, u)^*\, du \end{pmatrix}.$$

For small $\Delta t(\Delta t\|\mathbf{\Lambda}\| + \Delta t\|\mathbf{G}\| \ll 1)$ one gets the basic model (3–5). To derive the prediction formulas, first split the state vector into observed, $\mathbf{z}^{(p)}$, and unobserved, $\mathbf{z}^{(M)}$, parts

$$\mathbf{z}^{(p)} = \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{r}_1 \\ \mathbf{v}_2 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{v}_p \\ \mathbf{r}_p \end{pmatrix}, \qquad \mathbf{z}_M = \begin{pmatrix} \mathbf{v}_M \\ \mathbf{r}_M \end{pmatrix}.$$

Rewrite (B1) as

$$\begin{aligned} \mathbf{z}^{(p)}(n) &= \mathbf{\Phi}_p(\mathbf{z}^{(p)}(n-1)) \\ &\quad + \mathbf{q}^{(p)}(n, \mathbf{r}^{(p)}(n-1), \mathbf{r}_M(n-1)) \\ \mathbf{z}_M(n) &= \mathbf{\Phi}_M(\mathbf{z}_M(n-1)) \\ &\quad + \mathbf{q}_M(n, \mathbf{r}^{(p)}(n-1), \mathbf{r}_M(n-1)), \end{aligned}$$

where

$$\mathbf{\Phi}_p(\mathbf{z}^{(p)}) = \begin{pmatrix} \mathbf{C}\mathbf{v}_1 \\ \mathbf{D}_1\mathbf{r}_1 + \mathbf{F}_1\mathbf{U}(\mathbf{r}_1) + \mathbf{H}_1\mathbf{v}_1 \\ \mathbf{C}\mathbf{v}_2 \\ \mathbf{D}_2\mathbf{r}_2 + \mathbf{F}_2\mathbf{U}(\mathbf{r}_2) + \mathbf{H}_2\mathbf{v}_2 \\ \vdots \\ \mathbf{C}\mathbf{v}_M \\ \mathbf{D}_p\mathbf{r}_p + \mathbf{F}_p\mathbf{U}(\mathbf{r}_p) + \mathbf{H}_p\mathbf{v}_p \end{pmatrix},$$

$$\mathbf{\Phi}_M(\mathbf{z}_M) = \begin{pmatrix} \mathbf{C}\mathbf{v}_M \\ \mathbf{D}_M\mathbf{r}_M + \mathbf{F}_M\mathbf{U}(\mathbf{r}_M) + \mathbf{H}_M\mathbf{v}_M \end{pmatrix},$$

$$\mathbf{r}^{(p)} = \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_p \end{pmatrix}, \qquad \mathbf{q}^{(p)} = \begin{pmatrix} \mathbf{q}_1 \\ \mathbf{q}_2 \\ \vdots \\ \mathbf{q}_p \end{pmatrix}.$$

Accordingly, the covariance matrix is represented as

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q}_{pp} & \mathbf{Q}_{pM} \\ \mathbf{Q}_{Mp} & \mathbf{Q}_{MM} \end{pmatrix},$$

where

$$\mathbf{Q}_{pp} = E\mathbf{q}^{(p)}\mathbf{q}^{(p)*}, \qquad \mathbf{Q}_{pM} = E\mathbf{q}^{(p)}\mathbf{q}_M^*,$$
$$\mathbf{Q}_{Mp} = E\mathbf{q}_M\mathbf{q}^{(p)*}, \qquad \mathbf{Q}_{MM} = E\mathbf{q}_M\mathbf{q}_M^*.$$

Suppose that the predictor observations are given until moment n

$$\mathbf{z}^{(p)}(1) = \mathbf{x}_1, \qquad \mathbf{z}^{(p)}(2) = \mathbf{x}_2, , \qquad \mathbf{z}^{(p)}(n) = \mathbf{x}_n.$$

Assume for a while that the predictand is also observed until moment $n - 1$

$$\mathbf{z}_M(1) = \mathbf{y}_1, \qquad \mathbf{z}_M(2) = \mathbf{y}_2, \cdots,$$
$$\mathbf{z}_M(n-1) = \mathbf{y}_{n-1}.$$

Since all the positions moment $n - 1$ are given, $\mathbf{z}^{(p)}(n)$ and $\mathbf{z}_M(n)$ become Gaussian random vectors given by

$$\mathbf{z}^{(p)}(n) = \mathbf{\Phi}_p(\mathbf{x}_{n-1}) + \mathbf{q}^{(p)}(n, \mathbf{x}_{n-1})$$
$$\mathbf{z}_M(n) = \mathbf{\Phi}_M(\mathbf{y}_{n-1}) + \mathbf{q}_M(n, \mathbf{y}_{n-1}).$$

Hence, we can find

$$\hat{\mathbf{z}}_M(n) = E(\mathbf{z}_M(n) \,|\, \text{observations})$$

using the Gauss–Markov theorem

$$\begin{aligned} \hat{\mathbf{z}}_M(n) &= \mathbf{\Phi}_M(\mathbf{y}_{n-1}) + E(\mathbf{q}_M(n, \mathbf{y}_{n-1}) \,|\, \text{observations}) \\ &= \mathbf{\Phi}_M(\mathbf{y}_{n-1}) + \mathbf{Q}_{Mp}(\mathbf{y}_{n-1}, \mathbf{x}_{n-1}) \\ &\quad \times \mathbf{Q}_{pp}^{-1}(\mathbf{x}_{n-1}, \mathbf{x}_{n-1})\mathbf{q}^{(p)}(n, \mathbf{x}_{n-1}) \\ &= \mathbf{\Phi}_M(\mathbf{y}_{n-1}) + \mathbf{Q}_{Mp}(\mathbf{y}_{n-1}, \mathbf{x}_{n-1}) \\ &\quad \times \mathbf{Q}_{pp}^{-1}(\mathbf{x}_{n-1}, \mathbf{x}_{n-1})(\mathbf{x}_n - \mathbf{\Phi}(\mathbf{x}_{n-1})). \end{aligned}$$

The right-hand side includes the quantity $\mathbf{y}_{n-1} = \mathbf{z}_M(n - 1)$, which is in fact unobservable. We simply replace it by the prediction at the previous step to suggest the following prediction formulas

$$\begin{aligned} \mathbf{z}_M^a(n) &= \mathbf{z}_M^f(n) + \mathbf{Q}_{Mp}(\mathbf{z}_M^a(n-1), \mathbf{z}^{(p)}(n-1)) \\ &\quad \times \mathbf{Q}_{pp}^{-1}(\mathbf{z}^{(p)}(n-1), \mathbf{z}^{(p)}(n-1)) \\ &\quad \times (\mathbf{z}^{(p)}(n) - \mathbf{z}_p^f(n)) \end{aligned} \tag{B2}$$

with

$$\mathbf{z}_M^f(n) = \mathbf{\Phi}_M(\mathbf{z}_M^a(n-1)),$$
$$\mathbf{z}_p^f(n) = \mathbf{\Phi}_p(\mathbf{z}_p^a(n-1)).$$

Note that the formula before (B2) is an exact expression for the conditional expectation while (B2) itself is just a reasonable prediction formula that is not optimal in general. Indeed, it is optimal only if the noise covariance does not depend on the state vector and the dynamics is linear. We dropped the superscript ''*a*'' (''analyzed'') in the second formula since the analyzed and observed values for the predictors coincide. To finish, we first express the ''regression coefficient'' $\mathbf{Q}_{Mp}\mathbf{Q}_{pp}^{-1}$ in terms of the covariance $\mathbf{B}_{ij} = \mathbf{B}(\mathbf{r}_i, \mathbf{r}_j)$ and then separate the prediction formulas for the velocities and positions. Let

$$\widetilde{\mathbf{B}}_{ij} = \begin{pmatrix} \mathbf{B}_{ij} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{ij} \end{pmatrix}$$

be the 4×4 matrix with $\mathbf{B}_{ij}$ along the main diagonal. Notice that (5) is written as

$$\mathbf{Q}_{ij} = \mathbf{L}\begin{pmatrix} \mathbf{B}(\mathbf{r}_i, \mathbf{r}_j) & \mathbf{0} \\ \mathbf{0} & \mathbf{B}(\mathbf{r}_i, \mathbf{r}_j) \end{pmatrix}$$

$$= \begin{pmatrix} \mathbf{B}(\mathbf{r}_i, \mathbf{r}_j) & \mathbf{0} \\ \mathbf{0} & \mathbf{B}(\mathbf{r}_i, \mathbf{r}_j) \end{pmatrix}\mathbf{L}$$

with

$$\mathbf{L} = \frac{1}{6}\begin{pmatrix} 6\Delta t\mathbf{I} & 3(\Delta t)^2\mathbf{I} \\ 3(\Delta t)^2\mathbf{I} & 2(\Delta t)^3\mathbf{I} \end{pmatrix}$$

where $\mathbf{I}$ is the 2×2 unit matrix. Then

$$\mathbf{Q} = (\widetilde{\mathbf{B}_{ij}}L) \equiv \begin{pmatrix} \mathbf{B}_{ij}\mathbf{L} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{ij}\mathbf{L} \end{pmatrix}.$$

Thus,

$$\mathbf{Q}_{pp} = (\widetilde{\mathbf{B}_{ij}}L)_{i=1,\cdots,p;\, j=1,\cdots,p}, \qquad \mathbf{Q}_{pp} \equiv (\widetilde{\mathbf{B}_{iM}}L)_{i=1,\cdots,p}$$

and then

$$\mathbf{Q}_{pp}^{-1} = (\mathbf{L}^{-1}\widetilde{\mathbf{B}}_{ij}^{-1})_{i=1,\cdots,p;\, j=1,\cdots,p},$$

$$\mathbf{Q}_{Mp} \equiv (\mathbf{L}^{-1}\widetilde{\mathbf{B}_{iM}}^{-1})_{i=1,\cdots,p},$$

where

$$\widetilde{\mathbf{B}}_{ij}^{-1} = \begin{pmatrix} \mathbf{B}_{ij}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{ij}^{-1} \end{pmatrix}.$$

where $\mathbf{B}_{ij}^{-1}$ are the entries of $(\mathbf{B}_{ij})^{-1}$. Hence $\mathbf{Q}_{Mp}\mathbf{Q}_{pp}^{-1}$ also has a block structure. The explicit expression for each block is given by

$$(\mathbf{Q}_{Mp}\mathbf{Q}_{pp}^{-1})_{ij} = \sum_{k=1}^{p} (\mathbf{Q}_{Mp})_{ik}(\mathbf{Q}_{pp}^{-1})_{kj}$$

$$= \sum_{k=1}^{p} \widetilde{\mathbf{B}}_{ik}\mathbf{L}\mathbf{L}^{-1}\widetilde{\mathbf{B}}_{kj}^{-1} = \sum_{k=1}^{p} \widetilde{\mathbf{B}}_{ik}\widetilde{\mathbf{B}}_{kj}^{-1}$$

$$= \begin{pmatrix} \sum_{k=1}^{p} \mathbf{B}_{ik}\mathbf{B}_{kj}^{-1} & \mathbf{0} \\ \mathbf{0} & \sum_{k=1}^{p} \mathbf{B}_{ik}\mathbf{B}_{kj}^{-1} \end{pmatrix}.$$

Thus, each block of $\mathbf{Q}_{Mp}\mathbf{Q}_{pp}^{-1}$ has a diagonal form with the same diagonal entries. This means that the weighing coefficients for the velocity and position prediction are the same and given by

$$\mathbf{K}_k = \sum_{k=1}^{p} \mathbf{B}(\mathbf{r}_i, \mathbf{r}_k)\mathbf{B}_{kj}^{-1}, \qquad j = 1, \cdots, p.$$

Thus, we arrive at (8).

REFERENCES

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

Bauer, S., M. S. Swenson, A. Griffa, A. J. Mariano, and K. Owens, 1998: Eddy–mean flow decomposition and eddy-diffusivity estimates in the tropical Pacific Ocean. Part 1: Methodology. *J. Geophys. Res.*, **103,** 30 855–30 871.

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. Syst.*, **29,** 1–4, 32–50.

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

Cushman-Roisin, B., 1994: *Introduction to Geophysical Fluid Dynamics.* Prentice Hall, 285 pp.

Duan, J., and S. Wiggins, 1996: Fluid exchange across a meandering jet with quasiperiodic variability. *J. Phys. Oceanogr.*, **26,** 1176–1188.

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.

Flament, P. J., S. C. Kennan, R. A. Knox, P. P. Niiler, and R. L. Bernstein, 1996: The three-dimensional structure of an upper ocean vortex in the tropical Pacific Ocean. *Nature,* **383,** 610–613.

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

——, K. Owens, L. Piterbarg, and B. Rozovskii, 1995: Estimates of turbulence parameters from Lagrangian data using a stochastic particle model. *J. Mar. Res.*, **53,** 212–234.

Haller, G., and A. C. Poje, 1997: Eddy growth and mixing in mesoscale oceanic flows. *Nonlinear Proc. Geophys.*, **4,** 223–235.

Hansen, D. V., and P.-M. Poulain, 1996: Quality control and interpolations of WOCE/TOGA drifter data. *J. Atmos. Oceanic Technol.*, **13,** 900–909.

Liptser, R. S., and A. N. Shiryaev, 1978: *Statistics of Random Processes.* Vol. 2. Springer-Verlag, 422 pp.

Meyers, S., 1994: Cross-frontal mixing in a meandering jet. *J. Phys. Oceanogr.*, **24,** 1641–1646.

Niiler, P. P., A. S. Sybrandy, K. Bi, P.-M. Poulain, and D. S. Bittermam, 1995: Measurements of the water-following capability of holey-sock and TRISTART drifters. *Deep-Sea Res.*, **42,** 1951–1964.

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

Piterbarg, L. I., 1998: Drift estimation for Brownian flows. *Stochastic Proc. Appl.*, **79,** 132–149.

——, 2001: The top Lyapunov exponent for a stochastic flow modeling the upper ocean turbulence. *SIAM J. Appl. Math.*, in press.

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

Poulain, P.-M., 1999: Drifter observations of surface circulation in the Adriatic Sea between December 1994 and March 1996. *J. Mar. Syst.*, **20,** 231–253.

Reverdin, G., C. Frankignoul, and E. Kestenare, 1994: Seasonal variability in the surface currents of the equatorial Pacific. *J. Geophys. Res.*, **99,** 20 323–20 344.

Rogerson, A. M., P. D. Miller, L. J. Pratt, and C. K. R. T. Jones, 1999: Lagrangian motion and fluid exchange in a barotropic meandering jet. *J. Phys. Oceanogr.*, **29,** 2635–2655.

Samelson, R. M., 1992: Fluid exchange across a meandering jet. *J. Phys. Oceanogr.*, **22,** 431–440.

——, 1996: Chaotic transport by mesoscale motions. *Stochastic Modelling in Physical Oceanography,* R. J. Adler, P. Müller, and B. L. Rozovoskii, Eds., Birkhauser, 423–438.

Schneider, T., 1998: Lagrangian drifter models as search and rescue tools. M.S. thesis, Department of Meteorology and Physical Oceanography, University of Miami, 96 pp.

Thomson, D. J., 1986: A random walk model of dispersion in tur-

bulent flows and its application to dispersion in a valley. *Quart. J. Roy. Meteor. Soc.*, **112,** 511–529.
——, 1987: Criteria for the selection of stochastic models of particle trajectories in turbulent flows. *J. Fluid Mech.*, **180,** 529–556.
Wiggins, S., 1992: *Chaotic Transport in Dynamical Systems.* Springer, 301 pp.
Yang, H., and Z. Liu, 1997: The three-dimensional chaotic transport and the great ocean barrier. *J. Phys. Oceanogr.*, **27,** 1258–1273.
——, R. H. Weisberg, P. P. Niiler, W. Sturges, and W. Johnson, 1999: Lagrangian circulation and forbidden zone on the West Florida Shelf. *Contin. Shelf Res.*, **19,** 1221–1245.
Zambianchi, E., and A. Griffa, 1994: Effects of finite scales of turbulence on dispersion estimates. *J. Mar. Res.*, **52,** 129–148.
Zirbel, C. L., and E. Cinlar, 1996: Dispersion of particle systems in Brownian flows. *Adv. Appl. Prob.*, **28,** 53–74.