SIAM J. APPL. MATH.
Vol. 63, No. 1, pp. 116–148

A SIMPLE PREDICTION ALGORITHM FOR THE LAGRANGIAN MOTION IN TWO-DIMENSIONAL TURBULENT FLOWS*

LEONID I. PITERBARG† AND TAMAY M. ÖZGÖKMEN‡

Abstract. A new algorithm is suggested for prediction of a Lagrangian particle position in a stochastic flow, given observations of other particles. The algorithm is based on linearization of the motion equations and appears to be efficient for an initial tight cluster and small prediction time. A theoretical error analysis is given for the Brownian flow and a stochastic flow with memory. The asymptotic formulas are compared with simulation results to establish their applicability limits. Monte Carlo simulations are carried out to compare the new algorithm with two others: the center-of-mass prediction and a Kalman filter–type method. The algorithm is also tested on real data in the tropical Pacific.

Key words. stochastic flow, Lagrangian motion, prediction, stochastic simulations, oceanographic applications

AMS subject classifications. 76F25, 76F55, 86A05, 62M20

PII. S003613990139194X

1. Introduction.

The problem discussed is motivated by applications to rescue and search operations in the sea. An important part of such operations is to properly narrow the search area based on the best possible prediction of the position of a lost object, given its approximate initial position (Schneider (1998)). It is hard to make a reasonable prediction based only on knowledge of the mean current, because of strong velocity fluctuations drifting the object away from the path indicated by the mean velocity field. One can expect more realistic help from other floating objects in the same area, like debris or drifters (human-made floats), which can be observed from planes or satellites. We consider here a simplified model of such a situation, as follows. Several current following floats are released simultaneously at different known positions in a stochastic flow. One of the floats, called the predictand, is unobservable, while the remaining floats, called the predictors, are observed. The problem is to predict the position of the unobservable float, given the above observations. In addition to practical needs, this problem is of great importance from a theoretical viewpoint since it addresses the predictability issue for the Lagrangian motion in turbulent flows. Here, by turbulent flows, we mean velocity fields with fluctuations described by stochastic differential equations. Thus, a kinematic approach is employed: given flow statistics one should conclude with the mean square error of a prediction algorithm. The mathematical framework we set up here is as follows.

Let $\mathbf{u}(t, \mathbf{r})$ be a random velocity field varying in time. By the method of applications we consider only the two-dimensional case: $\mathbf{u}, \mathbf{r} \in R^2$. Consider $M > 1$ Lagrangian (current following) particles starting at time $t = 0$ from different positions

*Received by the editors July 6, 2001; accepted for publication (in revised form) January 28, 2002; published electronically August 28, 2002. This research was supported by Office of Naval Research grants N00014-99-0042 and N00014-99-1-0049.

http://www.siam.org/journals/siap/63-1/39194.html

†Center for Applied Mathematical Sciences, University of Southern California, Los Angeles, CA 90089 (piter@math.usc.edu).

‡Division of Meteorology and Physical Oceanography, Rosenstiel School of Marine and Atmospheric Science, University of Miami, 4600 Rickenbacker Causeway, Miami, FL 33149.

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

$$\frac{d\mathbf{r}_j}{dt} = \mathbf{u}(t, \mathbf{r}_j) \ , \qquad \mathbf{r}_j(0) = \mathbf{r}_j^0, \tag{1}$$

$j = 1, \ldots, M$. 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 during time interval $(0, T)$, while the trajectory of the last one, $\mathbf{r}_M(t)$, is not observed. The problem is to find a reasonable prediction of the position of the unobserved particle, given the above predictor observations and the initial predictand position. The optimal prediction in the mean square sense,

$$E|\widehat{\mathbf{r}}_M(T) - \mathbf{r}_M(T)|^2 \to \min,$$

is given by the conditional expectation (e.g., Liptser and Shiryaev (1978))

$$\hat{\mathbf{r}}_M(T) = E\left(\mathbf{r}_j(T) \mid \mathbf{r}_1(t), \mathbf{r}_2(t), \ldots, \mathbf{r}_p(t), \quad 0 \le t \le T\right),$$

based on all the observations. Alas, this general formula gives too little in the considered situation. Normally, this conditional expectation cannot be explicitly found even in the simplest case when the velocity fluctuations are delta-correlated in time. However, it can be approximated (with an uncertain accuracy) for some Markov models of stochastic flows (Piterbarg (2001b)). This approximation resulted in a Kalman filter–type prediction algorithm which was tested on synthetic (Özgökmen et al. (2000), Piterbarg (2001b)) and real (Castellari et al. (2001), Özgökmen et al. (2001)) data. In general, that algorithm, called henceforth the KF algorithm, performs well, but its essential drawback is that it requires knowledge of some statistics of the underlying stochastic flow such as the Lagrangian correlation time and the space correlation radius of the Eulerian velocity field.

The goal of this paper is to introduce and investigate a new model-independent prediction algorithm. At first glance, the suggested prediction method looks a little bit naive. Roughly, we linearize (1) in a vicinity of the initial cluster, obtain a linear regression model where regressors are the initial particle positions, then estimate the regression coefficients at any given moment based on the observations of the predictor positions, and, finally, use them for predicting the unknown particle. The idea for the new algorithm emerged when we found from real data that the position of the cluster center of mass is a not bad alternative to the KF algorithm. The trouble is that the center-of-mass algorithm (CM) performs poorly at the initial stage if the predictor is located far from the cluster center of mass. In fact, the suggested algorithm, called here the regression algorithm (RA), can be viewed as a CM algorithm adjusted to the initial position of the predictand. As it will be shown, the RA performs very well at the initial stage if the cluster diameter is essentially less than the space correlation radius of the velocity fluctuations and performs as well as the CM algorithm in the long term. The good predictive skill of RA demonstrated in real data processing has had an impact on development of theoretical and Monte Carlo error analysis for RA. Such an analysis is based on investigating the second moment $\rho(t, \mathbf{r}_0)$ of the difference between positions of two particles initially separated by $\mathbf{r}_0$, called the separation process. The quantity $\rho(t, \mathbf{r}_0)$ can be effectively studied for two important models: the well-known Brownian flow and a stochastic model with memory recently developed in (Piterbarg (2001a)).

The Brownian stochastic flow arises when the Eulerian velocity field is delta-correlated in time. In this case a closed partial differential equation for $\rho(t, \mathbf{r}_0)$ is

readily written and the asymptotical prediction error is obtained for a tight initial cluster by expanding the equation solution in $\mathbf{r}_0$. It is interesting that the prediction error is not determined by the second Lyapunov moment, but rather by the fourth order term in $\mathbf{r}_0$. Asymptotical behavior of the error for large t is determined by the top Lyapunov exponent of the stochastic flow. In the case of a positive Lyapunov exponent, the behavior of the mean square error is close to that of the dispersion, $\sqrt{2Dt}$, where D is an effective diffusivity. In fact, it is even a little bit worse due to the bias between the predictand initial position and the cluster center of mass. For a negative Lyapunov exponent the growth order is $\sqrt{t/\log t}$. It is worthwhile to notice that our algorithm is not designed for long-term prediction.

Being an important mathematical example, the Brownian stochastic flow is not a realistic model for upper ocean turbulence. We focus instead on the second model in which a joint vector of the positions and velocities form a Markov process. For this reason it is called the first-order Markov model to distinguish it from the zero-order model determined by the Brownian flow. The first-order model implies an additional important parameter, the Lagrangian correlation time τ, which can be estimated well from real data (Griffa et al. (1995)). In the framework of the model, the Lagrangian velocity of a single particle is an Ornstein–Uhlenbeck process with parameter τ. In this case the variance of the separation process $\rho(t, \mathbf{r}_0, \mathbf{v}_0)$ also depends on the difference of the initial velocities. As a consequence, the prediction error for close initial positions and velocities is determined by the expansion coefficient of ρ at $\mathbf{v}_0^2$. The equation for ρ in the first-order model is a standard Kolmogorov equation. An expansion of the solution in both $\mathbf{r}_0$ and $\mathbf{v}_0$ results in an approximate mean square error for the prediction. The proposed approximation is in good agreement with simulations. A special focus is on a linear shear mean flow determined by the stretch and rotation parameters γ and ω, respectively. It is shown that the relative prediction error decreases as γ and ω increase. Comparison with the KF algorithm shows that the RA performs essentially better in the presence of a deterministic linear shear flow, while for a pure stochastic flow they are equivalent or KF is better.

The main points of this work are (1) formulation of the new prediction algorithm (section 2); (2) formulas for the prediction error which are in very good agreement with stochastic simulations (sections 3–5); (3) comparative analysis of RA and KF performance based on synthetic and real data (sections 6 and 7).

The main investigation tools used are stochastic simulations, together with standard diffusion process analytic techniques. For the simulations we take real values of model parameters and show the error in dimension units to give an idea of the usefulness of the real prediction.

2. Prediction formula. Assume the following classical regression model for motion of M particles:

$$\mathbf{r}_i(t) = \mathbf{A}(t)\mathbf{r}_i(0) + \mathbf{b}(t) + \mathbf{y}_i(t), \tag{2}$$

where $\mathbf{A}(t)$ and $\mathbf{b}(t)$ are an unknown, random in general, 2×2 matrix and 2-vector, respectively, and $\mathbf{y}_j(t)$ are stochastic processes with zero mean uncorrelated for any fixed t. Notice that this model does not follow from the model (1) in the general case. Moreover, it even contradicts (1) for a nonlinear velocity field. The idea is to construct a prediction algorithm based on (2) and then forget (2) and investigate the algorithm performance for some important particular cases of the model (1). The reason to expect a good performance is that the system (1) can be linearized on short times, and then the obtained formula would be useful for the short-term prediction.

Recall that the first $p = M - 1$ particles (predictors) are supposed to be observed and the Mth one is to be predicted. To identify six unknown parameters at each time (four entries of $\mathbf{A}$ and two entries of $\mathbf{b}$) one should have $p \geq 3$. We accept this assumption for the rest of the paper. The underdetermined situation $p = 2$ is of practical interest as well, but it requires a special consideration which is outside the scope of this paper. The least square estimators of $\mathbf{A}(t)$ and $\mathbf{b}(t)$ based on the observed particles at the moment t are given by

$$\hat{\mathbf{A}}(t) = \mathbf{S}(t)\mathbf{S}(0)^{-1}, \quad \hat{\mathbf{b}}(t) = \mathbf{r}_c(t) - \hat{\mathbf{A}}(t)\mathbf{r}_c(0),$$

where

$$\mathbf{r}_c(t) = \frac{1}{p}\sum_{i=1}^{p} \mathbf{r}_i(t)$$

is the center of mass of the predictor cluster and

$$\mathbf{S}(t) = \sum_{i=1}^{p} (\mathbf{r}_i(t) - \mathbf{r}_c(t))(\mathbf{r}_i(0) - \mathbf{r}_c(0))^T,$$

the superscript T stands for transposition, the vectors mean column-vectors, and it is assumed that $p > 2$ to have nondegenerate matrix $S(0)$. The obtained estimators then are used to predict the unobservable particle

$$\hat{\mathbf{r}}_M(t) = \mathbf{r}_c(t) + \mathbf{S}(t)\mathbf{S}(0)^{-1}(\mathbf{r}_M(0) - \mathbf{r}_c(0)). \tag{3}$$

This prediction formula is optimal in the framework of the model (2) if $\mathbf{A}$ and $\mathbf{b}$ are supposed to be deterministic. Further we reject the regression model and study its performance for some specific models of the velocity field $\mathbf{u}(t, \mathbf{r})$ appearing in (1).

If the velocity field is smooth enough in time, then it is worthwhile to include the initial velocities as regressors as well. We do not do that in the present paper for two reasons: first, this does not make any sense when considering the Brownian flow since it implies infinite velocities, and second, determining initial velocities in practice is a very hard problem. However, a study of an initial velocity–based formula is of theoretical interest and will be considered in a further work.

Once again we underscore that the prediction formula (3) does not include any parameters except the initial particle positions. Of course, it is not always a strength. Including well-known parameters would probably improve prediction essentially, but the problem is that statistical estimates of mean currents and turbulence parameters are often not reliable in oceanic conditions. Therefore, apparently, sometimes it is better to use a rough prediction algorithm than fine algorithms with misspecified parameters. Further we try to evaluate limits of this "roughness."

3. General error analysis. Let

$$s^2(t) = E|\hat{\mathbf{r}}_M(t) - \mathbf{r}_M(t)|^2 \tag{4}$$

be the mean square error of the prediction (3). Introduce the variance of the separation process by

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

For the sake of brevity we call ρ_{ij} the separation. Then (see appendix)

(5)

$$s^2(t) = \frac{1}{p}\sum_{k=1}^{p}\rho_{kM}(t) - \frac{1}{2p^2}\sum_{k,l=1}^{p}\rho_{kl}(t) - \frac{1}{p}\sum_{k,l=1}^{p} b_k\rho_{kl}(t) + \sum_{k=1}^{p} b_k\rho_{kM}(t) - \frac{1}{2}\sum_{k,l=1}^{p} b_k b_l\rho_{kl}(t),$$

where coefficients

$$b_i = (\mathbf{r}_i(0) - \mathbf{r}_c(0))^T \mathbf{S}(0)^{-1}(\mathbf{r}_M(0) - \mathbf{r}_c(0)) \tag{6}$$

are defined by the initial positions only, while the behavior of $\rho_{ij}(t)$ also depends on the flow properties. Notice that the deviation $E|\mathbf{r}_c(t) - \mathbf{r}_M(t)|^2$ of the predictand from the cluster center of mass is determined by the first two terms in (5).

In the next two sections we will consider two different asymptotics of (4): first, small initial distances between particles and, second, the long time asymptotic. As one will see, in the first case the problem is reduced to a factorized separation

$$\rho_{ij}(t) \sim \rho_0(t)c_{ij},$$

where ρ_0 is independent on the initial configuration and c_{ij} are completely determined by the initial conditions. In the second case under common conditions, the separation is independent of the initial conditions:

$$\rho_{ij}(t) \sim \rho(t)(1 - \delta_{ij}),$$

where δ_{ij} is the Kronecker delta. Hence, in the first case (5) becomes

$$s^2(t) \sim C_0\rho_0(t), \tag{7}$$

where

$$C_0 = \frac{1}{p}\sum_{k=1}^{p} c_{kM} - \frac{1}{2p^2}\sum_{k,l=1}^{p} c_{kl} - \frac{1}{p}\sum_{k,l=1}^{p} b_k c_{kl} + \sum_{k=1}^{p} b_k c_{kM} - \frac{1}{2}\sum_{k,l=1}^{p} b_k b_l c_{kl} \tag{8}$$

is a function of the initial conditions only. A similar formula appears in the second case:

$$s^2(t) \sim C\rho(t), \tag{9}$$

with

$$C = \frac{1}{2} + \frac{1}{2p} + \frac{1}{2}(\mathbf{r}_M(0) - \mathbf{r}_c(0))^T \mathbf{S}(0)^{-1}(\mathbf{r}_M(0 - \mathbf{r}_c(0)). \tag{10}$$

Notice that this algorithm is not designed for long time prediction and the formula (9) is of theoretical interest only.

4. Brownian flow. Assume that the velocity field is decomposed into mean circulation and fluctuation:

$$\mathbf{u}(t, \mathbf{r}) = \mathbf{U}(t, \mathbf{r}) + \mathbf{u}'(t, \mathbf{r}), \tag{11}$$

where $\mathbf{U}_0(t,\mathbf{r})$ is a deterministic velocity field and $\mathbf{u}'(t,\mathbf{r})$ is a random vector field with zero mean, $E\mathbf{u}'(t,\mathbf{r})=0$. The Brownian flow is determined by the assumption that the velocity fluctuation $\mathbf{u}'(t,\mathbf{r})$ is a Gaussian white noise in t, i.e.,

$$E\mathbf{u}'(t,\mathbf{r})=0,\quad E\mathbf{u}'(t_1,\mathbf{r}_1)\mathbf{u}'(t_2,\mathbf{r}_2)^T=\delta(t_1-t_2)\mathbf{B}(\mathbf{r}_1,\mathbf{r}_2)\ , \tag{12}$$

where $\delta(t)$ is the Dirac delta-function and $\mathbf{B}(\mathbf{r}_1,\mathbf{r}_2)$ is the spatial covariance tensor of the velocity field. Introduce the state, drift, and noise vectors:

$$\mathbf{z}=\begin{pmatrix}\mathbf{r}_1\\ \mathbf{r}_2\\ \dots\\ \mathbf{r}_M\end{pmatrix},\quad \mathbf{A}(t,\mathbf{z})=\begin{pmatrix}\mathbf{U}(t,\mathbf{r}_1)\\ \mathbf{U}(t,\mathbf{r}_2)\\ \dots\\ \mathbf{U}(t,\mathbf{r}_M)\end{pmatrix},\quad \mathbf{W}(t)=\begin{pmatrix}\mathbf{w}_1\\ \mathbf{w}_2\\ \dots\\ \mathbf{w}_M\end{pmatrix},$$

where $\mathbf{r}_j(t)$ are the positions of M particles at time t starting from different locations and $\mathbf{w}_j(t)$ are independent standard Wiener processes. Then a rigorous interpretation of (11), (12) is as follows. The process $\mathbf{z}(t)$ is a $2M$-dimensional Markov process satisfying the stochastic Ito differential equation (Kunita (1990))

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

where the $2M\times 2M$ diffusion matrix is given by

$$\mathbf{D}(\mathbf{z})=(\mathbf{B}(\mathbf{r}_i,\mathbf{r}_j)).$$

Another equivalent formulation of this model is as follows: $\mathbf{z}(t)$ is a Markov process with the generator given by

$$L=\mathbf{U}(t,\mathbf{r}_i)\cdot\nabla_{\mathbf{r}_i}+\frac{1}{2}\nabla_{\mathbf{r}_i}\cdot\mathbf{B}(\mathbf{r}_i,\mathbf{r}_j)\nabla_{\mathbf{r}_j}.$$

In the homogeneous case characterized by the assumptions that the mean flow is constant $\mathbf{U}(t,\mathbf{r})\equiv\mathbf{U}$ and that the covariance is a function of the position difference $\mathbf{B}(\mathbf{r}_1,\mathbf{r}_2)=\mathbf{B}(\mathbf{r}_1-\mathbf{r}_2)$, the separation process $\mathbf{r}(t)=\mathbf{r}_1(t)-\mathbf{r}_2(t)$ (the difference between displacements of two different particles) is also a Markov process with the generator

$$L_s=\nabla_{\mathbf{r}}\cdot(\mathbf{B}(\mathbf{0})-\mathbf{B}(\mathbf{r})))\nabla_{\mathbf{r}}.$$

Further we assume that the velocity field is isotropic (that is, $\mathbf{U}\equiv 0$) and the entries $b_{ij}(\mathbf{r})$ of $\mathbf{B}(\mathbf{r})$ are given by (Monin and Yaglom (1975))

$$b_{ij}(\mathbf{r})=b_N(r)\delta_{ij}+\frac{x_ix_j}{r^2}(b_L(r)-b_N(r)),$$

where $r=|\mathbf{r}|$, $\mathbf{r}=(x_1,x_2)$. Assume that the longitudinal and normal covariances are four times continuously differentiable:

$$\begin{aligned}b_L(r)&=D-\tfrac{1}{2}\beta_Lr^2+\tfrac{1}{2}\gamma_Lr^4+O(r^6),\\ b_N(r)&=D-\tfrac{1}{2}\beta_Nr^2+\tfrac{1}{2}\gamma_Nr^4+O(r^6),\end{aligned} \tag{13}$$

where $D,\beta_L,\gamma_L,\beta_N,\gamma_N$ are positive parameters whose physical meaning is explained below. For the isotropic Brownian flow the squared dispersion is expressed as

$$d^2(t)\equiv E(\mathbf{r}(t)-\mathbf{r}(0))^2=2Dt; \tag{14}$$

hence, D means a diffusivity. Introduce the space correlation radius of the velocity field by

$$R^2 = D/\beta_L$$

and set $\mathbf{r}_{ij} = \mathbf{r}_i(0) - \mathbf{r}_j(0)$, $r_{ij} = |\mathbf{r}_{ij}|$. First, consider the "tight cluster" asymptotic characterized by

$$r_{ij} \ll R, \quad i, j = 1, M.$$

Under this approximation each separation $\rho_{ij}(t) = \rho(t, r_{ij})$ is expressible in form

$$\rho(t, r) = \rho_1(t) r^2 - \rho_0(t) r^4 + O(r^6), \tag{15}$$

where (see appendix)

$$\rho_1(t) = \exp(\overline{\beta} t), \quad \rho_0(t) = K(\exp(\beta_0 t) - \exp(\overline{\beta} t)), \tag{16}$$

where

$$\beta_0 = 2\beta_N + 6\beta_L, \quad \overline{\beta} = \beta_L + \beta_N, \quad K = \frac{\overline{\gamma}}{5\beta_L + \beta_N}, \quad \overline{\gamma} = \gamma_L + \gamma_N.$$

After substitution of the expansion (15) into (5) the terms quadratic in r_{ij} disappear, and we get

$$s^2(t) \sim C_0 \rho_0(t), \tag{17}$$

where C_0 is given by (8) with

$$c_{kl} = |\mathbf{r}_k(0) - \mathbf{r}_l(0)|^4$$

and $\rho_0(t)$ is given in (16)

For the incompressible Brownian flow, characterized by $b_N(r) = (d/dr)(r b_L(r))$, with the longitudinal covariance

$$b_L(r) = D \exp(-r^2/R^2),$$

we have

$$\rho_1(t) = \exp(8Dt/R^2), \quad \rho_0(t) = \frac{3}{8R^2}(\exp(24Dt/R^2) - \exp(8Dt/R^2)).$$

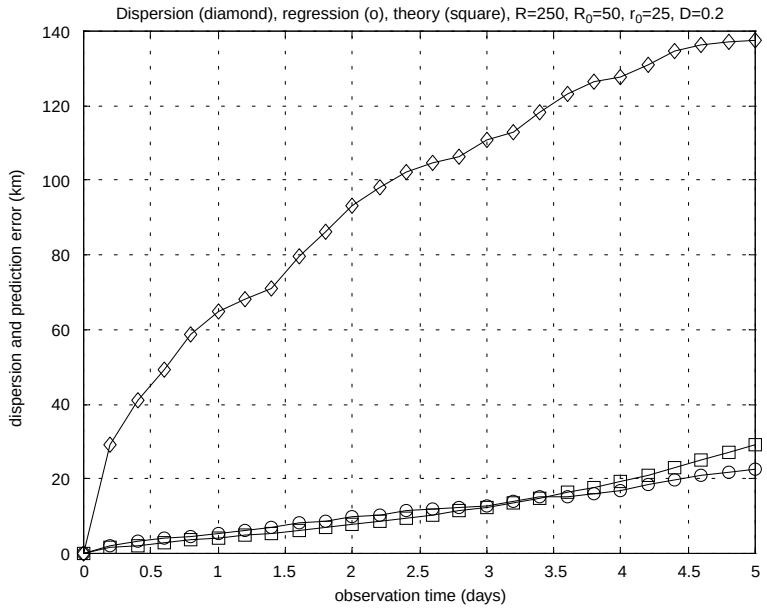
Assume that initially the predictors are located at the vertices of a right polygon at a distance R_0 from the center and the predictand is at a distance r_0 from the center. We call such an initial configuration perfect. In this case (17) becomes

$$s^2(t) \sim (3r_0^4 + 2R_0^4 - 4r_0^2 R_0^2)\rho_0(t) \tag{18}$$

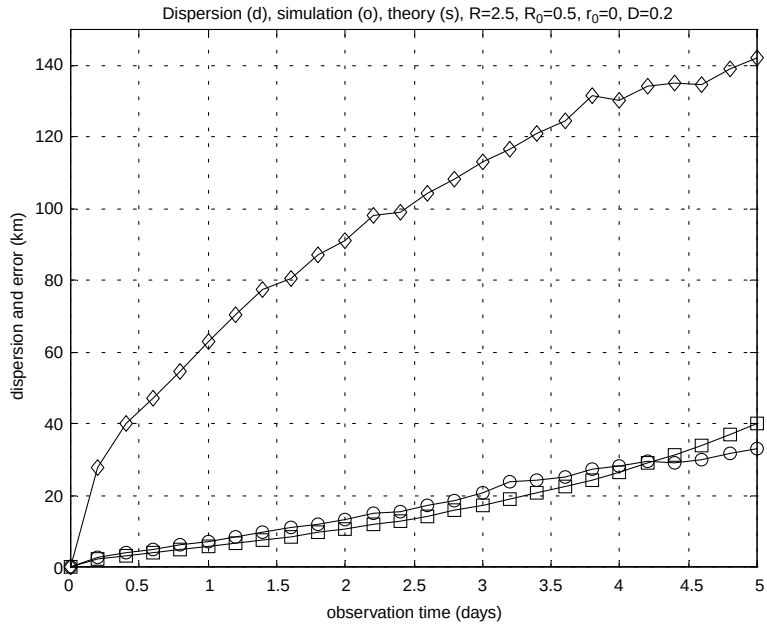
for $p > 3$ and

$$s^2(t) \sim (3r_0^4 + 2R_0^4 - 3r_0^2 R_0^2 - 2R_0 r_0^3 \cos(3\alpha))\rho_0(t)$$

for $p = 3$, where α is the angle between the directions from the center to the predictand and from the center to one of the predictors. The approximation (18) was checked via simulations, and the results are presented in Figure 1(a) and (b). For the simulation



(a)



(b)

FIG. 1. (a) *Dependence of the dispersion,* $d(t)$ (*diamonds*) *and prediction error,* $s(t)$, *obtained from simulations* (*circles*) *and from theory* ((18), *squares*) *on the observation time for the Brownian stochastic flow.* 100 *independent runs are used for* $d(t)$ *and* $s(t)$. *The diffusivity* $D = 2 \times 10^3\, km^2/day$, *number of predictors* $p = 6$, *initial hexagon radius* $R_0 = 50\, km$, *velocity space correlation radius* $R = 250\, km$, *distance of the predictand from the hexagon center* $r_0 = 25$ *km.* (b) *Same as in* (a) *with* $r_0 = 0$.

it was assumed that $R = 250$ km, $R_0 = 50$ km, and $D = 2000$ km^2/day. In the first case the predictor is distanced by $r_0 = 25$ km from the center, and in the second case it is initially located at the center. The dispersion $d(t)$ (diamonds) and experimental prediction error $s(t)$ (circles) are obtained from the simulations by averaging over 100 independent runs. A modest sample size is used on purpose to illustrate graphically how large the noise is under a moderate sample volume typical in oceanographic measurements. The line marked by squares expresses the suggested error formula (18). This approximation performs pretty well up to $T = 5$ days. After that the theoretical curve sharply diverges from the experimental one. The simulated dispersion is in good agreement with formula (14): it behaves as $\sqrt{t}$ and the value $d(5) \approx 139$ is very close to that of given by (14), $d(5) = 100\sqrt{2}$. In the second case (Figure 1(b)) the prediction is slightly worse in full agreement with (18). Notice that the relative prediction error for small t is approximately constant:

$$s_r(t) \equiv \frac{s(t)}{d(t)} \sim \frac{\sqrt{3(3r_0^4 + 2R_0^4 - 4r_0^2R_0^2)}}{2R^2}.$$

As for large t, we have two different situations depending of the sign of the Lyapunov exponent for the underlying flow. The Lyapunov exponent, λ, characterizes the exponential divergence (convergence) of initially close particles. It can be expressed in terms of the flow parameters (Baxendale and Harris (1986)):

$$\lambda = (\beta_N - \beta_L)/2.$$

If $\lambda > 0$, then for large t the difference between the positions of two particles goes to infinity with probability 1, and the mean square distance between them grows as

$$\rho(t) \sim 4Dt.$$

From (9), (10) it follows that the relative error is also approximately constant,

$$s_r(t) \sim \sqrt{1 + \frac{1}{p} + (\mathbf{r}_M(0) - \mathbf{r}_c(0))^T \mathbf{S}(0)^{-1} (\mathbf{r}_M(0 - \mathbf{r}_c(0))},$$

and greater than one.

In the opposite case $\lambda < 0$, the picture is more sophisticated: the difference goes to zero with probability 1; however, the mean square distance still grows, but at a lower rate (see Zirbel and Cinlar (1996)):

$$\rho(t) \sim \frac{ct}{\log t}$$

with constant c depending on the initial distance. Thus, the relative error goes to zero slowly as t goes to infinity:

$$s_r(t) \sim C\sqrt{\frac{1}{\log t}}$$

with constant C depending on the initial cluster configuration.

5. Stochastic flow with memory. As we already noticed, the Brownian flow is not an appropriate model for the upper ocean turbulence, since it is based on the white noise assumption for Lagrangian velocity. In fact, numerous observations clearly

demonstrate that Lagrangian velocity is well approximated by the first-order Markov process (Thomson (1986), Griffa (1996)). The following model of multiparticle motion suggested by Piterbarg (2001a), (2001b) generalizes the above experimental fact.

In addition to the decomposition (11) assume that there is a deterministic acceleration $\mathbf{a}(\mathbf{v},\mathbf{r})$ depending in general on the particle velocity and position such that the motion equations take the form

$$\begin{aligned} d\mathbf{r} &= (\mathbf{U}(t,\mathbf{r})+\mathbf{v})dt, \\ d\mathbf{v} &= \mathbf{a}(\mathbf{v},\mathbf{r})dt + d\mathbf{w}(t,\mathbf{r}), \end{aligned} \tag{19}$$

where

$$E\mathbf{w}(t,\mathbf{r}) = 0, \qquad E\mathbf{w}(t_1,\mathbf{r}_1)\mathbf{w}(t_2,\mathbf{r}_2)^T = \min(t_1,t_2)\mathbf{B}(\mathbf{r}_1,\mathbf{r}_2)\ .$$

In other words, now the velocity field is not a white noise itself but rather is driven by a white noise with a space covariance structure determined by tensor $\mathbf{B}(\mathbf{r}_1,\mathbf{r}_2)$. A rigorous formulation of (19) is given in the above-mentioned references. For a cluster of M particles, introduce a state vector containing the particle velocities as well as positions and a drift vector:

$$\mathbf{z} = \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{r}_1 \\ \mathbf{v}_2 \\ \mathbf{r}_2 \\ \cdots \\ \mathbf{v}_M \\ \mathbf{r}_M \end{pmatrix}, \qquad \mathbf{A}(t,\mathbf{z}) = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{U}_1+\mathbf{v}_1 \\ \mathbf{a}_2 \\ \mathbf{U}_2+\mathbf{v}_2 \\ \cdots \\ \mathbf{a}_M \\ \mathbf{U}_M+\mathbf{v}_M \end{pmatrix},$$

where $\mathbf{U}_m = \mathbf{U}(t,\mathbf{r}_m)$, $\mathbf{a}_m = \mathbf{a}(\mathbf{v}_m,\mathbf{r}_m)$. The model (19) implies that the motion of any M particles is a Markov process in $4M$ dimensions described by a stochastic differential equation. Namely,

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

where $\mathbf{W}(t)$ is a standard Wiener process in $4M$ dimensions and the diffusion matrix $\mathbf{D}(\mathbf{z})$ is given by

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

with 4×4 blocks

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

Recall that now $\mathbf{B}(\mathbf{r}_1,\mathbf{r}_2)$ is the covariance tensor of the forcing, not the Eulerian velocity field itself. The equivalent formulation is given by the generator

$$L = (\mathbf{U}(t,\mathbf{r}_i)+\mathbf{v}_i)\cdot\nabla_{\mathbf{r}_i} + \mathbf{a}_i\cdot\nabla_{\mathbf{v}_i} + \frac{1}{2}\nabla_{\mathbf{v}_i}\cdot\mathbf{B}(\mathbf{r}_i,\mathbf{r}_j)\nabla_{\mathbf{v}_j}.$$

For our purposes the following homogeneous case is most important:

$$\mathbf{a}(\mathbf{v},\mathbf{r}) = -\tau^{-1}\mathbf{v}, \quad \mathbf{U}(\mathbf{r}) = \mathbf{U}+\mathbf{G}\mathbf{r}, \quad \mathbf{B}(\mathbf{r}_1,\mathbf{r}_2) = \mathbf{B}(\mathbf{r}_1-\mathbf{r}_2),$$

$$\mathbf{G} = \begin{pmatrix} \gamma & \omega \\ -\omega & -\gamma \end{pmatrix},$$

where τ is the Lagrangian correlation time and the mean velocity field is a divergence-free linear shear flow characterized by constant drift $\mathbf{U}$, stretching parameter γ, and rotation parameter ω. The one-particle motion in this case is described by the well-known Langevin equation for the Lagrangian velocity and the standard motion equation for the displacement

$$\begin{aligned} d\mathbf{v} &= -\tau^{-1}\mathbf{v}dt + \sigma_v\tau^{-1/2}d\mathbf{w}(t), \\ d\mathbf{r} &= (\mathbf{U} + \mathbf{G}\mathbf{r} + \mathbf{v})dt\ , \end{aligned} \tag{20}$$

where $\mathbf{w}(t)$ is a two-dimensional Brownian motion and $\sigma_v^2 = E\mathbf{v}^2$ is the velocity variance. To obtain (20) we assumed that $\mathbf{B}(\mathbf{0}) = (\sigma_v^2/2\tau)\mathbf{I}$. In particular, for the dispersion we get (see appendix)

(21)
$$d^2(t) \equiv\ E(\mathbf{r}(t) - \mathbf{r}(0))^2 = d_0^2(t) + \sigma_v^2 \int_0^t \int_0^t \cosh(\sqrt{\gamma^2(2t - s_1 - s_2)^2 - \omega^2(s_1 - s_2)^2}) \exp(-|s_1 - s_2|/\tau) ds_1 ds_2,$$

where $d_0(t)$ is determined by the mean flow and initial position $\mathbf{r}_0$ only. An explicit expression is given in the appendix. Further we assume that $\mathbf{U} = 0$ and $\mathbf{r}_0 = 0$, which results in $d_0^2(t) = 0$. Notice two partial cases of (21): if $\omega = 0$, then

$$d^2(t) = \frac{\sigma_v^2\tau}{1 - \gamma^2\tau^2}(\gamma^{-1}\ \sinh(2\gamma t) - \tau\cosh(2\gamma t) - \tau + 2\tau\cosh(\gamma t)\exp(-t/\tau)); \tag{22}$$

if $\gamma = 0$, then

$$d^2(t) = \frac{\sigma_v^2\tau}{1 + \omega^2\tau^2}\left(t - \frac{\tau(1 - \omega^2\tau^2)}{1 + \omega^2\tau^2}\ (1 - \exp(-t/\tau)\cos(\omega t)) - \frac{2\omega\tau^2}{1 + \omega^2\tau^2}\exp(-t/\tau)\sin(\omega t)\right).$$

Finally, for the zero shear (Zambianchi and Griffa (1994))

$$d^2(t) = \sigma_v^2\tau\,(t - \tau\ (1 - \exp(-t/\tau)))\,. \tag{23}$$

The stochastic equations for the separation process,

$$\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2, \quad \mathbf{v} = \mathbf{v}_1 - \mathbf{v}_2,$$

take the form

$$\begin{aligned} d\mathbf{v} &= -\tau^{-1}\mathbf{v}dt + (2(\mathbf{B}(\mathbf{0}) - \mathbf{B}(\mathbf{r})))^{1/2}d\mathbf{w}(t), \\ d\mathbf{r} &= (\mathbf{G}\mathbf{r} + \mathbf{v})dt. \end{aligned}$$

In other words, the generator of the separation process is

$$L_s = (\mathbf{G}\mathbf{r} + \mathbf{v}) \cdot \nabla_{\mathbf{r}} - \tau^{-1}\mathbf{v} \cdot \nabla_{\mathbf{v}} + \nabla_{\mathbf{v}} \cdot (\mathbf{B}(\mathbf{0}) - \mathbf{B}(\mathbf{r})))\nabla_{\mathbf{v}}. \tag{24}$$

Assume that the forcing is isotropic, i.e.,

$$b_{ij}(\mathbf{r}) = b_N(r)\delta_{ij} + \frac{x_i x_j}{r^2}(b_L(r) - b_N(r)),$$

with twice differentiable covariances

$$b_L(r) = \frac{\sigma_v^2}{2\tau} - \frac{1}{2}\beta_L r^2 + O(r^4), \quad b_N(r) = \frac{\sigma_v^2}{2\tau} - \frac{1}{2}\beta_N r^2 + O(r^4). \tag{25}$$

Let $\rho = \rho(t,u,v,x,y) = E\mathbf{r}(t)^2$, where $\mathbf{r}(0) = (x,y), \mathbf{v}(0) = (u,v)$. For small x, y, u, v expand

$$\rho = a_1x^2 + 2a_2xy + a_3y^2 + a_4u^2 + 2a_5uv + a_6v^2 + a_7xu + a_8xv + a_9yu + a_{10}yv \tag{26}$$

and set $\mathbf{a} = (a_1, a_2, \ldots, a_{10})$. It is shown in the appendix that

$$\frac{d\mathbf{a}}{dt} = \mathbf{A}\mathbf{a}, \quad \mathbf{a}|_{t=0} = \mathbf{a}_0, \tag{27}$$

where matrix $\mathbf{A}$ and vector $\mathbf{a}_0$ are given. In particular, for the zero shear

$$\rho(t) = \rho(t, \mathbf{r}, \mathbf{v}) = \rho_1(t)\mathbf{r}^2 + \rho_{10}(t)(\mathbf{r}\cdot\mathbf{v}) + \rho_0(t)\mathbf{v}^2, \tag{28}$$

where

$$\begin{aligned} \frac{d\rho_0(t)}{dt} &= -2\tau^{-1}\rho_0(t) + \rho_{10}(t), \\ \frac{d\rho_1(t)}{dt} &= 2\overline{\beta}\rho_0(t), \\ \frac{d\rho_{10}(t)}{dt} &= -\tau^{-1}\rho_{10}(t) + 2\rho_1(t), \end{aligned} \tag{29}$$

where $\overline{\beta} = \beta_L + \beta_N$. Assume that the initial velocities are Gaussian random values with zero mean and independent components with the same variance σ_0^2. Additionally assume that the velocities are independent for different particles and are independent of the forcing. Averaging (28) over the ensemble of initial values gives

$$\rho_{ij}(t) = \rho(t, \mathbf{r}_{ij}, \mathbf{v}_{ij}) = \rho_1(t)r_{ij}^2 + 2\rho_0(t)\sigma_0^2(1-\delta_{ij}). \tag{30}$$

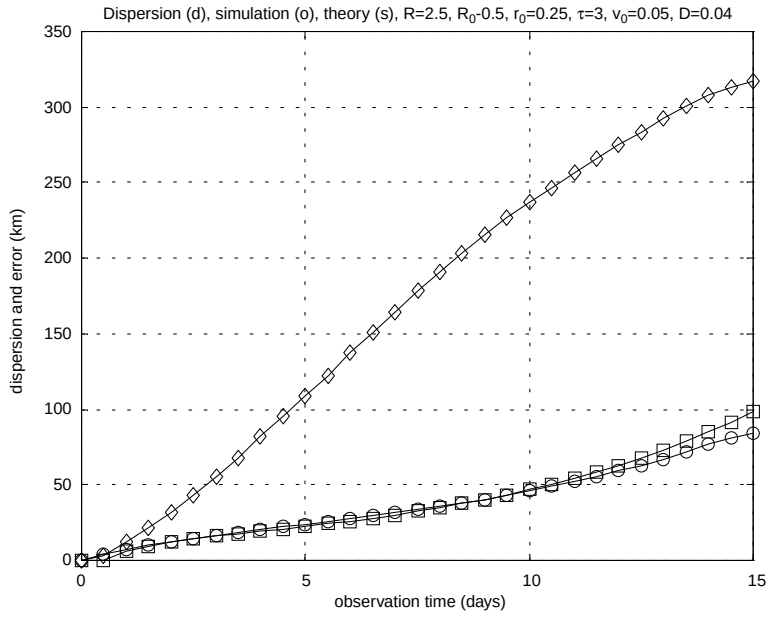
After substituting (30) into (5), the terms containing the distances disappear, and for the "tight cluster" asymptotic in the case of the perfect predictor, we get the initial configuration

$$s^2(t) \sim 2\sigma_0^2\left(1 + \frac{1}{p} + \frac{4r_0^2}{pR_0^2}\right)\rho_0(t), \tag{31}$$

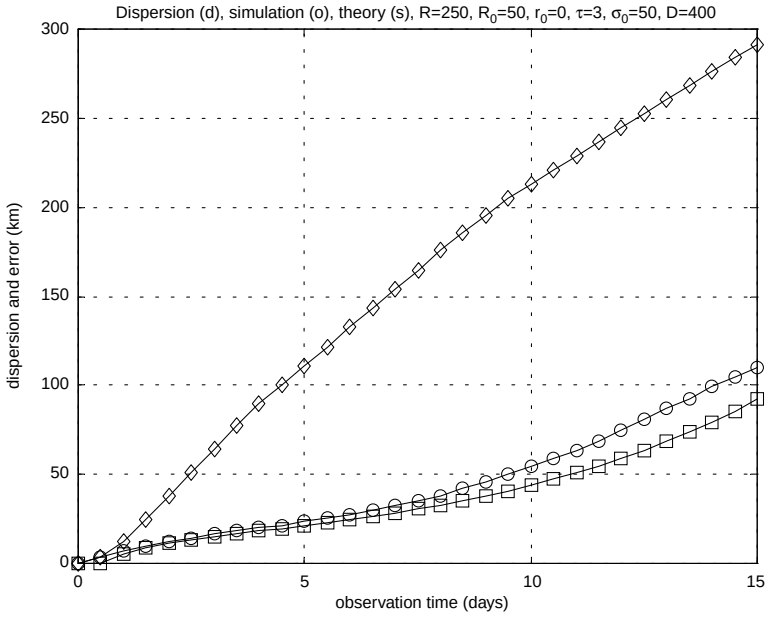
where $\rho_0(t)$ is obtained from (29). In the presence of the mean shear flow we get a similar formula:

$$s^2(t) \sim \sigma_0^2\left(1 + \frac{1}{p} + \frac{4r_0^2}{pR_0^2}\right)(a_4(t) + a_6(t)), \tag{32}$$

where $a_4(t), a_6(t)$ are obtained from (27). The asymptotic (31) is compared with simulations in Figure 2(a) and (b). For the simulations we used Lagrangian correlation time $\tau = 3$ days, the velocity variance $\sigma_v^2 = 0.12 \times 10^4$ km^2/day^2, initial velocity variance $\sigma_0^2 = 0.25 \times 10^2$ km^2/day^2, number of predictors $p = 6$, initial hexagon radius $R_0 = 50$ km, velocity space correlation radius $R = 250$ km. In Figure 2(a) the



(a)



(b)

FIG. 2. (a) *Dependence of the dispersion,* $d(t)$ (*diamonds*) *and prediction error,* $s(t)$*, obtained from simulations* (*circles*) *and from theory* ((31), *squares*) *on the observation time, for the stochastic flow with memory. The Lagrangian correlation time* $\tau = 3\,days$*, the velocity variance* $\sigma_v^2 = 12 \times 10^2\,km^2/day^2$*, initial velocity variance* $v_0^2 = 0.25 \times 10^2\,km^2/day^2$*, number of predictors* $p = 6$*, initial hexagon radius* $R_0 = 50\,km$*, velocity space correlation radius* $R = 250\,km$*, distance of the predictand from the hexagon center* $r_0 = 25$ *km.* (b) *Same as in* (a) *with* $r_0 = 0$.

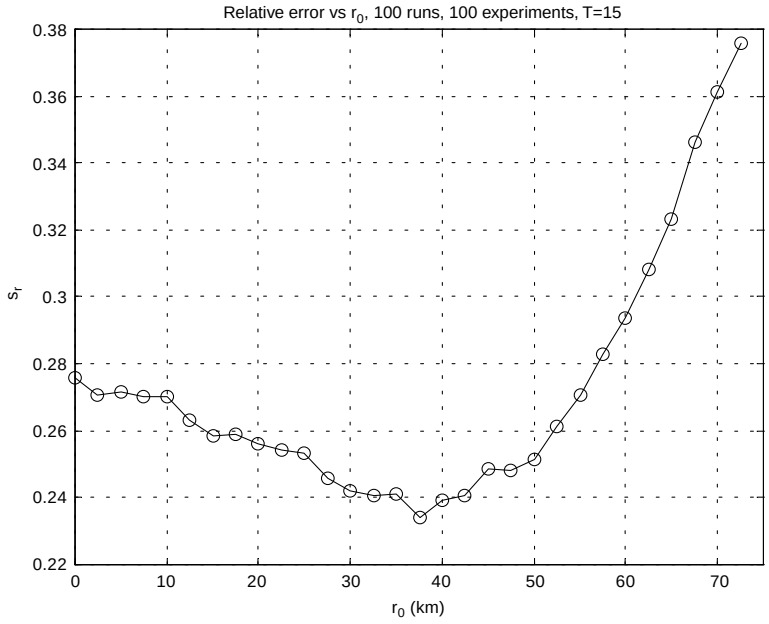


FIG. 3. *Dependence of the relative error s_r on r_0 for the observation time $T = 15$ days. The remaining parameters are the same as those as in Figure* 2(a).

predictand is distanced from the hexagon center by $r_0 = 25$ km, while in Figure 2(b) it is located exactly at the center. We take σ_0 much less than σ_v because the approximation (31) requires both small initial distances and small initial velocity differences. The dispersion $d(t)$ is shown as diamonds; the prediction error, $s(t)$, obtained from simulations, as circles; and the prediction error asymptotic (31) as squares. First, notice a good qualitative agreement of the simulated dispersion (100 runs) with the theoretical formula (23): for small t we have the ballistic regime ($d \sim t$), and for larger time the diffusion regime ($d \sim \sqrt{t}$). The quantitative agreement is also satisfactory. Then these figures show that the theoretical error formula works well for the ratio cluster radius/velocity correlation radius of less than 5 and for prediction periods of fewer than 15 days. The agreement is clearly better in the first case. In this regard, notice that unlike the Brownian flow case, the suggested approximation (31) does not give a correct dependence of the prediction error on the initial distance r_0 from the center. Indeed, in accordance with (31) the error in the second case should be less, but the simulations show the opposite. To get a correct dependence one should account for terms of higher order in the expansion of ρ. We do not do that here but instead study the dependence of the relative prediction error s_r on r_0 by the Monte Carlo method. Figure 3 demonstrates this dependence. The curve was obtained by averaging over 100 experiments with the same parameters $\tau = 3$ days, $\sigma_v^2 = 0.12 \times 10^4$ km^2/day^2, $\sigma_0^2 = 0.25 \times 10^2$ km^2/day^2, $p = 6$, $R_0 = 50$ km, $R = 250$ km, while each experiment included 100 runs to obtain s_r. This figure supports the previous observations that the error first decreases as r_0 increases, then assumes a minimum between 0 and R_0, and finally increases approaching R_0. For $r_0 > R_0$ the prediction worsens drastically. The obtained curve is affected by sampling variability, and the exact dependence s_r on r_0 is still to be investigated. It is interesting that the analytical dependence of s_r on the ratio $x = r_0^2/R_0^2$ obtained for the zero-order model

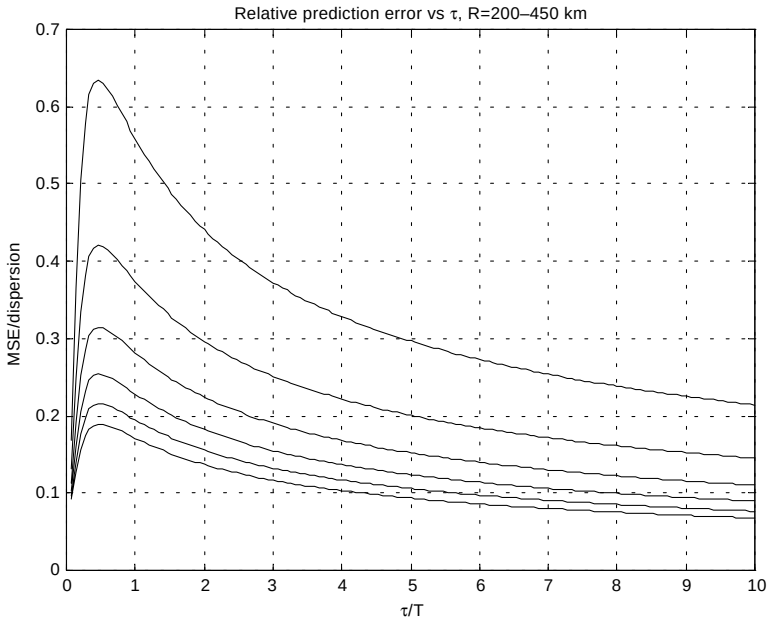
(19) describes pretty well the experimental curve for the first-order model. Indeed, from (19) $f(x) \equiv s(x)/s(0) = 1.5x^2 - 2x + 1$ assumes $f(0.5) = 0.375, f(2/3) = 1/3$ (minimum point), and $f(1) = 0.5$. Approximately the same values for that ratio follow from the curve in Figure 3. The reason is that for $T \gg \tau$ the first-order model approaches the Brownian flow.

Having sufficiently good agreement between (31) and the simulations, at least for values of r_0 close to $R_0/2$, we investigate the dependence of the prediction error on the model parameters using the analytical formulas (31) and (32). Figure 4(a) illustrates the dependence of the relative error $s_r(T)$ on τ for $R = 200, 250, 300, 350, 400, 450$ km with the zero mean flow and $T = 15$ days obtained from (21), (29), (31). The curves line up with R: the larger the R, the better the prediction. As for the Lagrangian correlation time, the error decreases with τ whenever $\tau/T > 0.5$. The effect of the error increasing for small value of τ is due to the regime changing in the dispersion behavior from ballistic to diffusive. As before, $R_0 = 50$ km and $p = 6$.

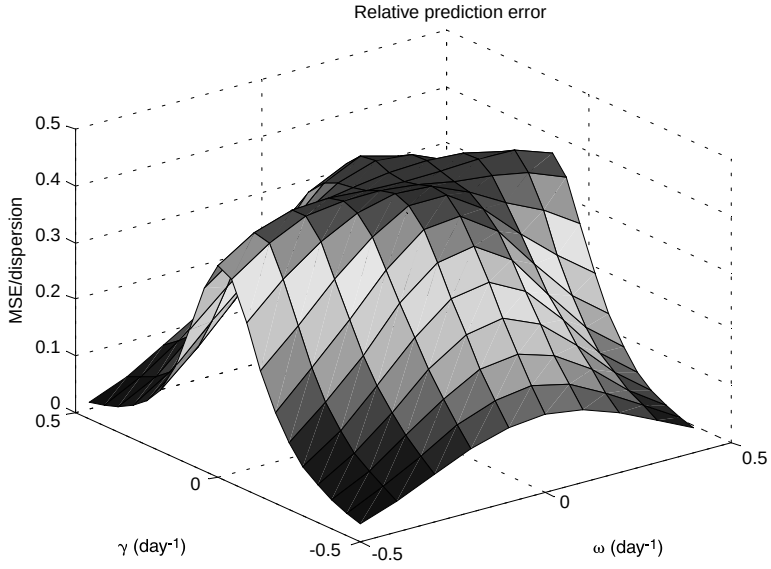
Figure 4(b) shows the dependence of s_r on γ and ω for fixed $\tau = 3$ days, $T = 15$ days, and $R = 250$ km, obtained from (21), (27), (32). The values of other parameters are indicated in the figure captions. Obviously, $s_r(\gamma, \omega)$ as a function of the shear parameters is even in both of them. The maximum relative error corresponds to $\gamma = 0$ because of a strong growth of the dispersion with γ (21), (22).

In the next series of experiments with zero mean flow we try a different initial configuration of predictors and a different velocity initialization to determine how the initialization affects the prediction skill. An eventual goal is to find an initial predictor configuration ensuring the best prediction. This problem is very complex and is beyond the scope of this paper. The goal of the present experiments with randomly distributed predictors is to understand the extent of the prediction error's dependence on the initial configuration and velocities. Namely, the alternative configuration we consider is a random initial configuration of predictors with the uniform distribution in a square with side of a. Now we compare four cases, the first two of which were discussed before: (1) perfect configuration ($R_0 = 50$ km) with the predictand $r_0 = 25$ km from the cluster center; (2) perfect configuration with the predictand at the center ($r_0 = 0$); (3) random configuration with $a = 2R_0$ and predictand $r_0 = 25$ km from the square center; (4) random configuration with $a = 2R_0$ and predictand at the square center ($r_0 = 0$). First, we consider the dependence of the error on the number of predictors for these four cases (Figure 5). As one can see there is not much difference in the algorithm performance when the number of predictors is 6 or more. It is worth noting that the random case is slightly worse than the perfect case for low p, but they quickly converge at $p = 6$ and do not change much as p increases. The error does not decrease significantly as p grows for all the initializations. This is in agreement with the theoretical formula (31).

Next we compare the statistical moments and histograms of the prediction error for 6 predictors and prediction time of 15 days (Table 1 and Figure 6(a)). The table gives the statistical moments of the relative deviation $\xi = |r_M - \widehat{r_M}|/d$, and the histograms are histograms of ξ. Thus, $s_r = \sqrt{E\xi^2}$, and it is also given in the table even though it can be found from the first two columns. First, it can be seen that for the perfect predictor configuration (series 1a, 2a) the initial location of the predictand is essential. Approaching the predictand to the predictors ($r_0 = 25$ km, series 1a) diminishes the mean from 0.275 to 0.249. This is in agreement with the previous experiments shown in Figures 2, 3, and 5. As for the random distribution of predictors, the initial position of the predictand almost does not make a difference,



(a)



(b)

FIG. 4. (a) *Dependence of the relative prediction error s_τ on the Lagrangian correlation time τ for different values of the Eulerian velocity correlation radius R obtained from the asymptotic formula* (31) *for observation time $T = 15$ days. Initial hexagon radius $R_0 = 50$ km, distance of the predictand from the hexagon center $r_0 = 25$ km, R varies from* 450 *km (lower curve) to* 200 *km (upper curve) with step* 50 *km.* (b) *Dependence of the relative prediction error s_τ on the shear parameters γ and ω obtained from the asymptotic formula* (32) *for observation time $T = 15$ days. Lagrangian correlation time $\tau = 3$ days, Eulerian velocity correlation radius $R = 250$ km, initial hexagon radius $R_0 = 50$ km, distance of the predictand from the hexagon center $r_0 = 25$ km.*

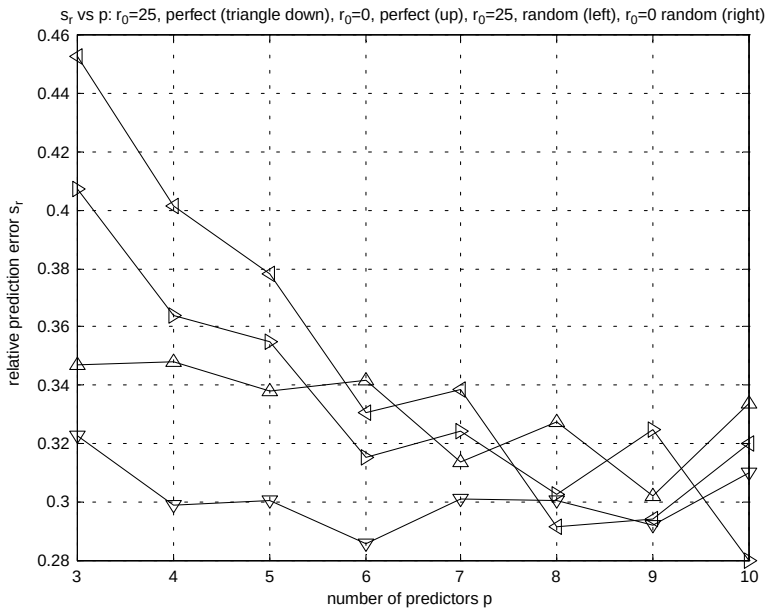


FIG. 5. *Dependence of the relative prediction error on the number of predictors (simulation) for different initial configurations.* (1) *Perfect configuration with the biased predictand (triangle down),* (2) *perfect configuration with the predictand at the center (triangle up),* (3) *uniformly distributed predictors with the biased predictand (triangle left),* (4) *uniformly distributed predictors with the predictand at the center (triangle right).*

TABLE 1

	Mean	STD	Median	Error
Series 1a	0.249	0.1947	0.203	0.3161
Series 2a	0.2751	0.2032	0.2294	0.342
Series 3a	0.2601	0.2088	0.2117	0.3335
Series 4a	0.2716	0.1981	0.2175	0.3362
Series 1b	0.2443	0.1923	0.198	0.3109
Series 2b	0.2792	0.2008	0.2329	0.3439
Series 3b	0.2525	0.2133	0.1998	0.3305
Series 4b	0.2601	0.202	0.2155	0.3293

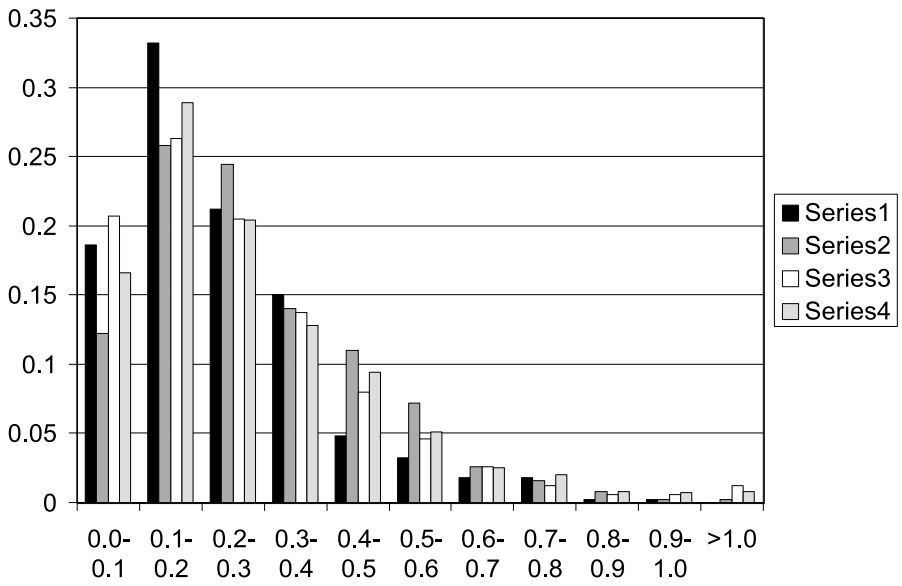
as can also be seen from the histogram (series 3 and 4). In contrast, the histograms in the case of perfect configuration (series 1 and series 2) look quite different. The mode of the first distribution is essentially higher, and the tail decays much faster.

Now introduce the alternative method of the velocity initialization as follows:

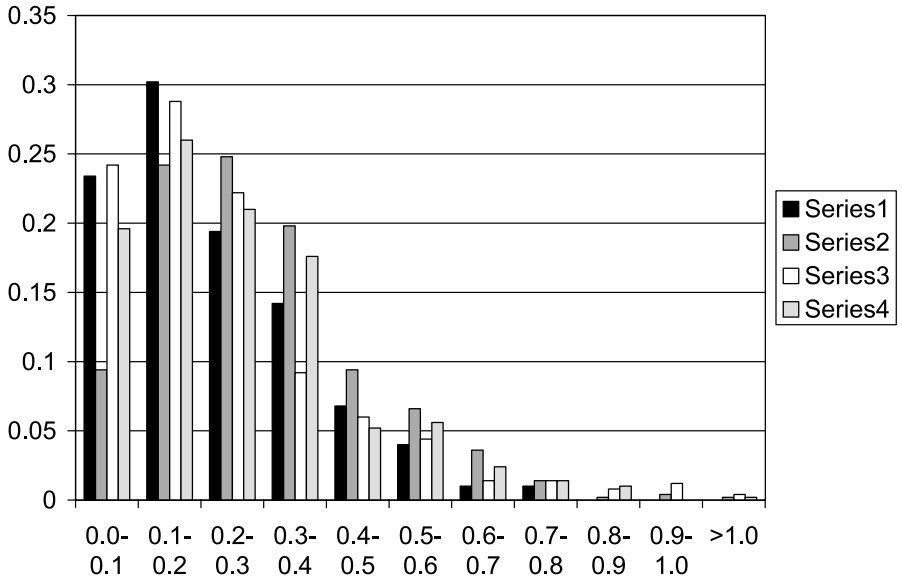
$$\mathbf{v}_j(0) = k\mathbf{r}_j(0), \tag{33}$$

where k is a constant independent of $j = 1, 2, \ldots, M$.

For the experiments we took $k = 0.1$ day^{-1} to have the same order of the initial velocities as in the case of the random initialization. Table 1 and Figure 6(b) demonstrate the statistical moments and histograms for the four cases discussed above, corresponding to the new velocity initialization (33). Regarding the predictand location, the conclusion is as before: for the perfect configuration of the predictors it is better to distance the predictor from the center, and for the random configuration it does not matter.



(a)



(b)

FIG. 6. (a) *Histograms of the relative prediction error for* 500 *runs with different initial configurations and random initial velocities:* (1) *perfect configuration with the biased predictand* (*series* 1), (2) *perfect configuration with the predictand at the center* (*series* 2), (3) *uniformly distributed predictors with the biased predictand* (*series* 3), (4) *uniformly distributed predictors with the predictand at the center* (*series* 4). (b) *Same as in* (a) *with the initial velocities proportional to the positions:* $\mathbf{v}_j(0) = k\mathbf{r}_j(0)$.

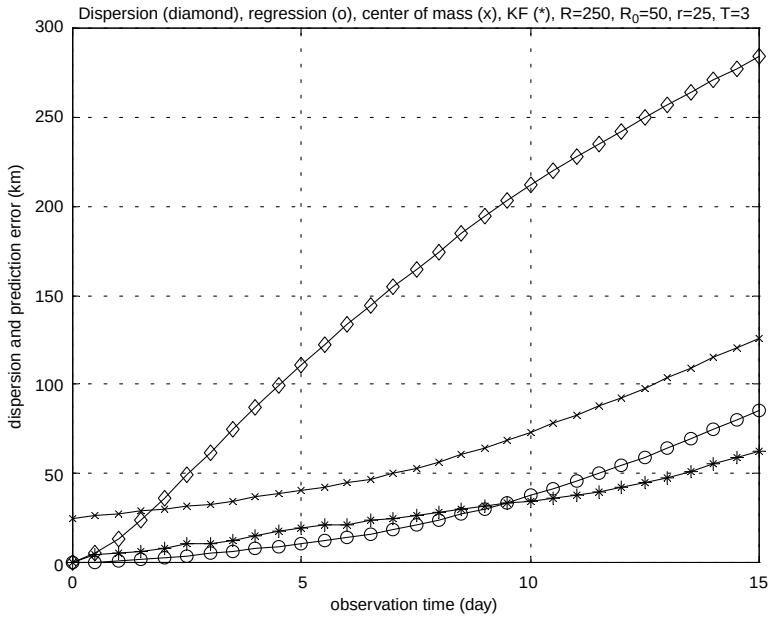
In general, there is not any visible difference in the prediction skill for the different velocity initializations. However, this conclusion is relevant only to 15 days' prediction. In practice we are more interested in a 3- to 5-day forecast, and in such time scales, the difference could be essential.

6. Comparison with KF and CM algorithms (simulations). The goal of the experiments discussed in this section is to compare the performance of RA and KF in both cases: the zero mean flow and a linear shear mean flow. The KF algorithm is based on the system of stochastic differential equations (19) for the M-particle motion. Since the diffusion matrix depends on the state variable $\mathbf{z}$, the classical Kalman filter cannot be applied to this system. What was proposed and studied in Özgökmen et al. (2000), (2001) and Piterbarg (2001b) is as follows. Pretend that the diffusion matrix is constant and write the KF equations for the optimal prediction of the unobserved particle velocity and position. Of course, these equations include the diffusion matrix. Then recall that it depends on the positions of all the particles and simply plug the observed positions for predictors and predictand forecast at the previous time step. We call this procedure a Kalman filter–type algorithm or, for short, the KF algorithm. An exact theoretical error analysis for the KF is very difficult. A Monte Carlo study showed that it gives a reasonable prediction if the model parameters are known (Piterbarg (2001b)). We follow the same approach here: the Lagrangian correlation time τ and the forcing covariance tensor $\mathbf{B}(\mathbf{r}_1, \mathbf{r}_2)$ are the same for generating Lagrangian trajectories and prediction formulas. However, the time steps are different: 1 hour for simulations and 12 hours for prediction. Thus, the KF has a big advantage over the RA, which does not use any information on the flow statistics.

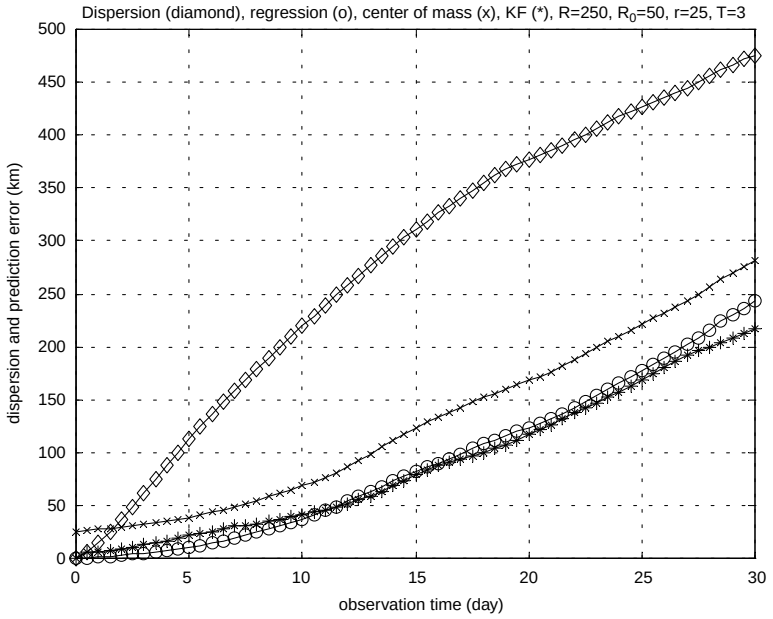
In the first series of experiments we considered the zero mean flow and fixed $\tau = 3$ days, $R = 250$ km, $R_0 = 50$ km, $M = 7$. Initially, the predictors are located in the vortices of the right hexagon, and the velocities are proportional to the position vectors; that is, the initialization (33) and the perfect configuration were used.

If the predictand is placed some distance from the center ($r_0 = 25$ km), then the mean square error of the regression algorithm is slightly lower than that of the KF (Figure 7(a)). Both algorithms are doing quite well compared with the dispersion (diamonds). This is because the initial cluster radius is 5 times less than the spatial correlation radius. The center of mass prediction (crosses) gives clearly worse prediction. After 22 days the performance of KF and regression is pretty much the same (Figure 7(b)).

If the predictand is placed at the center under the same experimental conditions, then the KF prediction turns out to be better (Figure 8(a)). As we mentioned before and which follows from the analytical formulas, the regression and center of mass methods give the same result in this case. For the midterm prediction (up to 30 days), this trend is confirmed (Figure 8(b)). For observation time $T = 30$ days the KF error is about 165 km, while the regression error is around 270 km under the dispersion 450 km. After changing τ to 2.5 days the general picture almost did not change (Figure 9(a)) and the conclusion is the same: KF performs better. However, if we introduce a mean flow, not very strong, the picture changes drastically. For a gyre given by $\omega = 0.1$ and $\gamma = 0$ the performance of KF is very poor (Figure 9(b)). The error reaches 160 km for a 15-day forecast, almost 80% of the dispersion (230 km), while the error of the regression algorithm is acceptable (60 km). Consider a different shear with no rotation: $\omega = 0$ and $\gamma = 0.05$. The conclusion is the same: the performance of the regression is clearly better (Figure 9(c)). The error of KF is

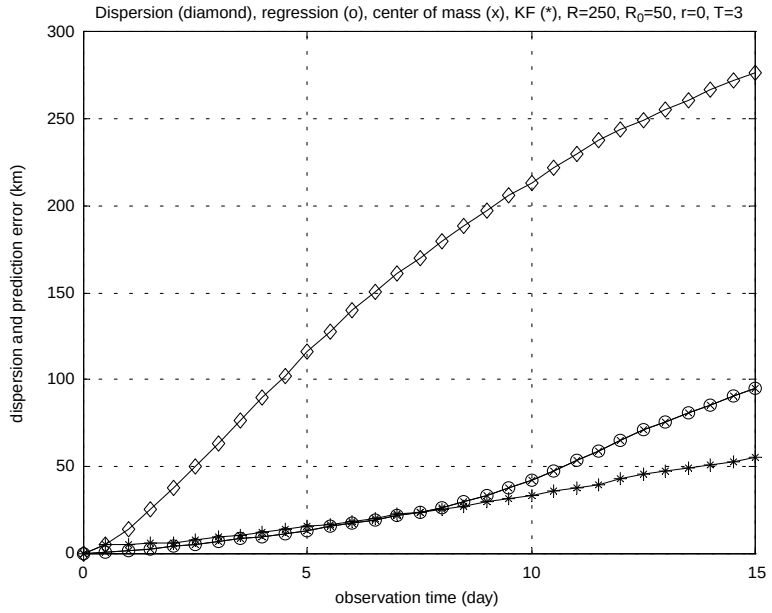


(a)

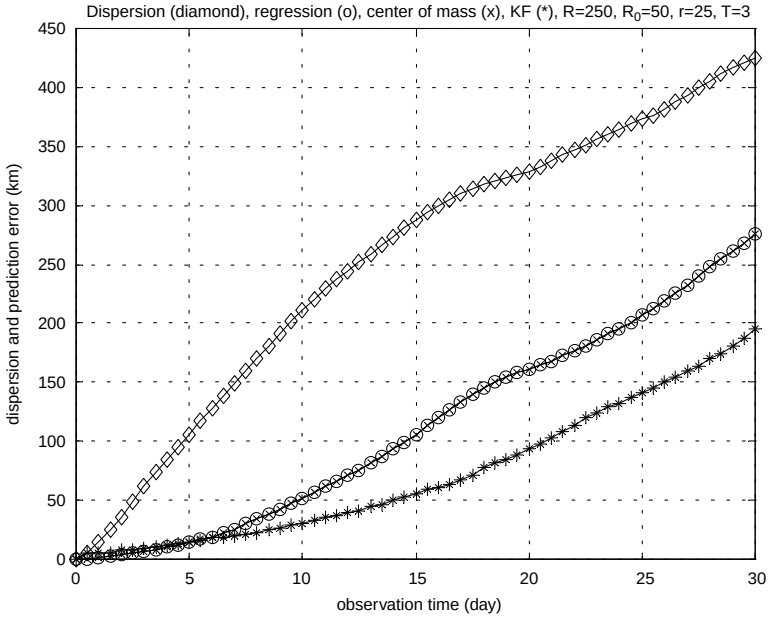


(b)

FIG. 7. (a) *Comparison of the dispersion (diamonds) and prediction error for the RA (circles), the CM (x), and the KF method for the maximum observation time* $T = 15$ *days and zero mean flow. The number of predictors* $p = 6$. *The predictors are located in vertices of a right hexagon. Lagrangian correlation time* $\tau = 3$ *days, Eulerian velocity correlation radius* $R = 250$ *km, initial hexagon radius* $R_0 = 50$ *km, distance of the predictand from the hexagon center* $r_0 = 25$ *km.* (b) *Same as in* (a) *for the maximum observation time* $T = 30$ *days. A different experiment.*

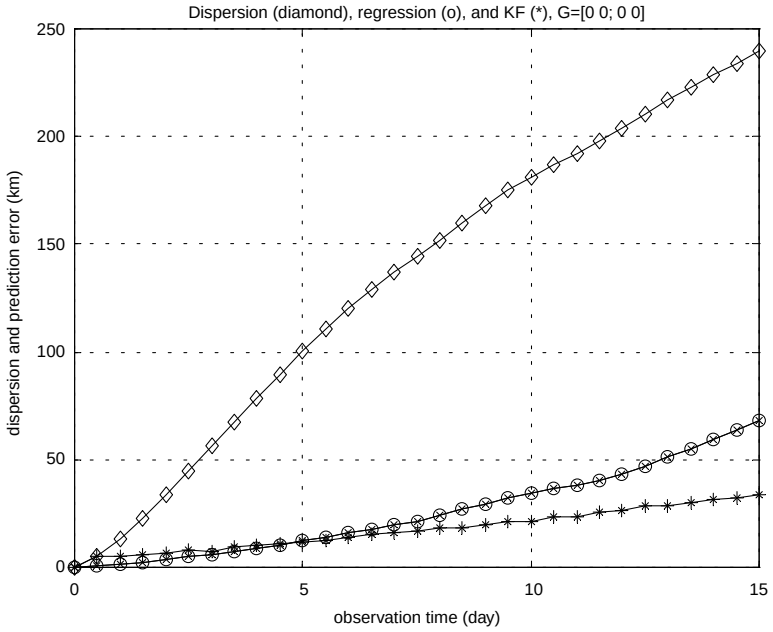


(a)

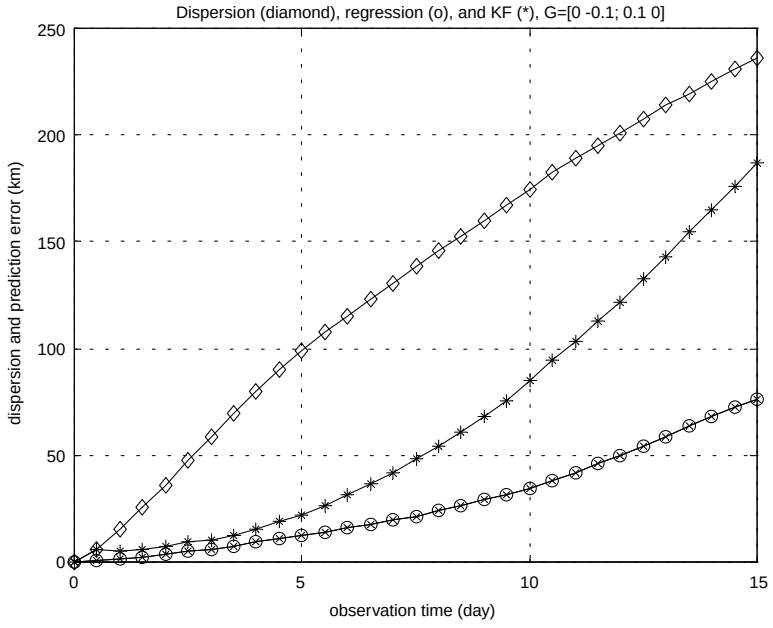


(b)

FIG. 8. (a) *Same as in Figure* 7(a) *with the predictor initially located at the center* ($r_0 = 0$). (b) *Same as in Figure* 8(a) *for the maximum observation time* $T = 30$ *days. A different experiment.*

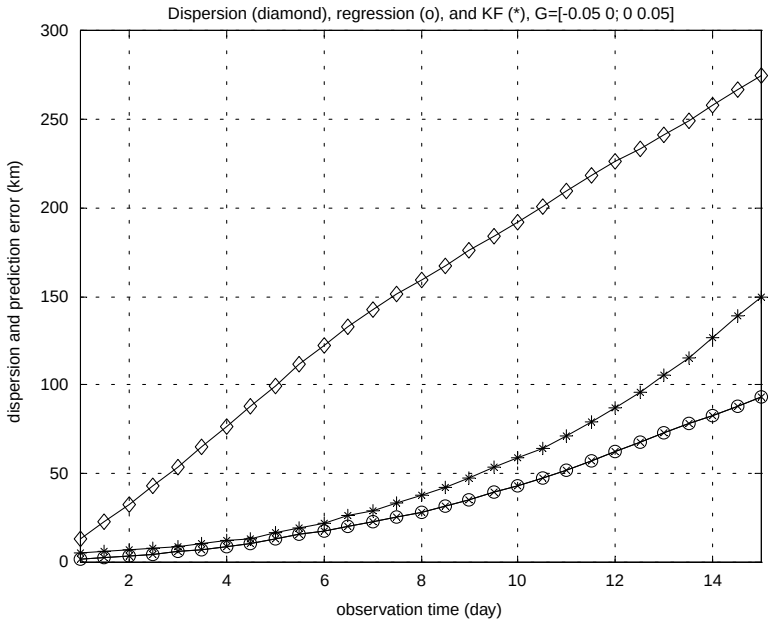


(a)

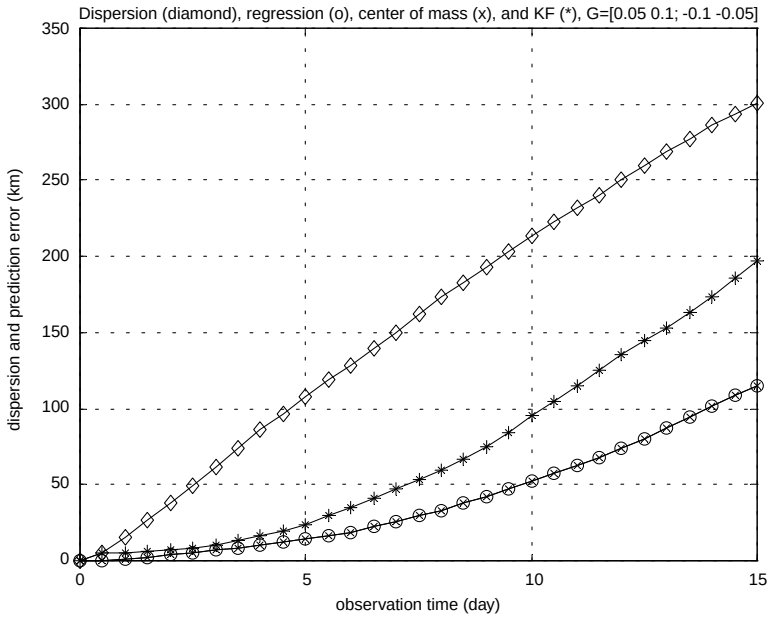


(b)

FIG. 9. (a) *Same as in Figure* 8(a) *with slightly different Lagrangian correlation time* 2.5 *days.* (b) *Same as in Figure* 9(a) *for a nonzero shear:* $\gamma = 0.05\,day^{-1}$, $\omega = 0$.



(c)



(d)

FIG. 9. (c) *Same as in Figure* 9(a) *for a nonzero shear:* $\gamma = 0$, $\omega = 0.1\,day^{-1}$. (d) *Same as in Figure* 9(a) *for a nonzero shear:* $\gamma = 0.05$, $\omega = 0.1\,day^{-1}$.

about 140 km, while for the regression it is the same: 60 km. Finally, combining both cases ($\omega = 0.1$ and $\gamma = 0.05$) we observe that the KF error is almost twice as much as the RA error (Figure 9(d)). Thus, with no doubt the regression algorithm performs better in the presence of a deterministic linear shear flow. This is because it is based on the assumption of linear dependence of the particle current position on the initial position. In fact, for purely linear flow the RA gives the exact prediction. Thus, the presence of the mean flow implies the better performance of the regression algorithm.

7. Comparison with KF and CM algorithms (real data). The final stage in this study is to apply RA to predict the motion of oceanic drifters released in a cluster and compare its performance with that of the simulations. It was found in Özgökmen et al. (2001) that during the period in which drifters remain close to one another as a tight cluster (quantified by the number of drifters within the velocity space correlation scale R), the CM method is a simple yet effective means of predicting the drifter location. However, how the prediction accuracy of RA compares to that of CM and the far more complicated technique KF for oceanic drifters needs to be investigated.

The drifter data are obtained from the NOAA Atlantic Oceanographic and Meteorological Laboratory, Global Drifter Center, by searching the entire 1988–1996 data set for a group of 5 or more drifters released within the velocity space correlation scale R. The drifter data are used as provided by the Global Drifter Center, which lists the drifter positions in six-hour intervals after standard quality control procedures (e.g., Hansen and Poulain (1996)) and no further processing has been applied. A total of 7 clusters, each consisting of 5–10 drifters, has been analyzed. In the following, we concisely present results from 3 of these clusters, since the main conclusions remain the same for others. These 3 drifter clusters have been released in the tropical Pacific Ocean, which is a region characterized by strong currents and shears and lacking the effect of coastlines or boundaries. The mean currents (Figure 10) are calculated using the technique described in Bauer et al. (1999) from the entire drifter data set collected under the World Ocean Circulation Experiment (WOCE) during 1988–1996. This figure depicts the general circulation pattern in this region, which is governed by the westward North Equatorial Current north of $10°N$, the eastward North Equatorial

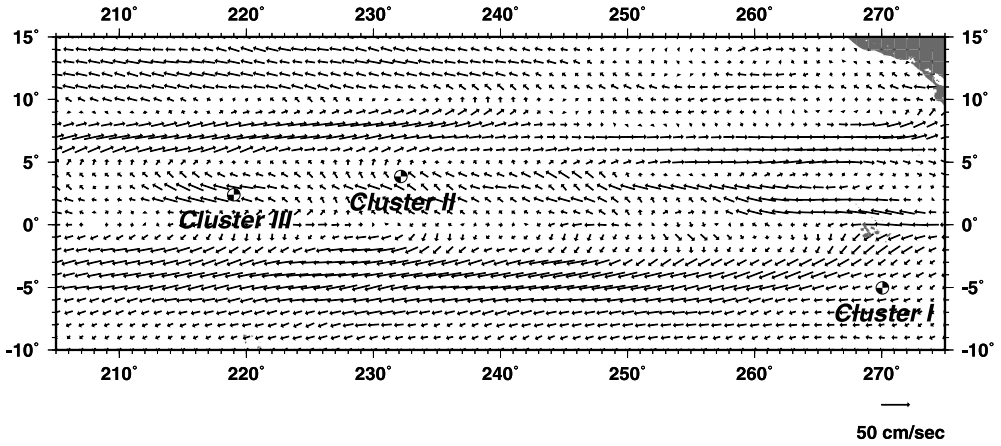


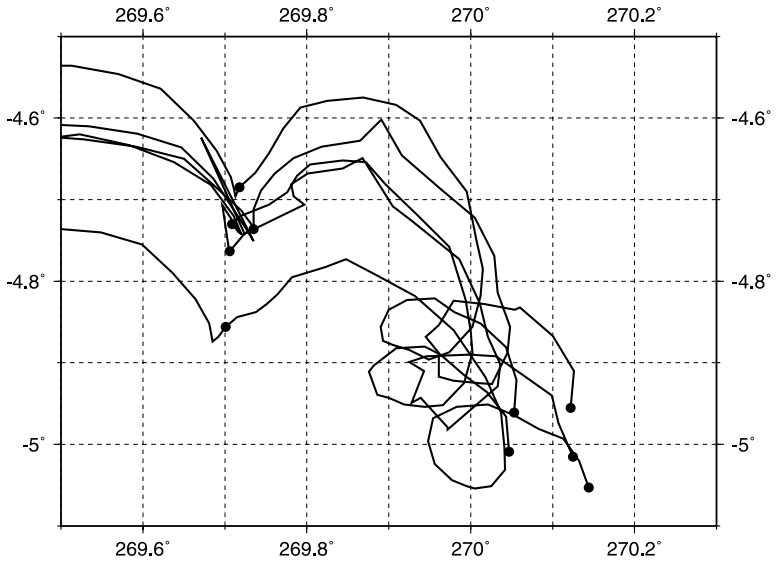
FIG. 10. *The climatological mean flow field depicting the major currents in the tropical Pacific Ocean and the initial release locations of clusters* I, II, *and* III.

Countercurrent between $4°N$ and $9°N$, and the westward South Equatorial Current, which extends across the equator to $10°S$. Drifters in the first cluster (cluster I) have been released in the South Equatorial Current, whereas the others in clusters II and III have been launched just south of the North Equatorial Countercurrent. The mean currents, however, are not a good indicator of drifter motion (Özgökmen et al. (2000), (2001)) and are discussed here only to provide the general surface flow characteristics of this oceanic region. It is also important to point out that in order to be able to deploy these drifters on tight grids, a real-time analysis of a variety of data sets, including current meter profilers and satellite data images, has been necessary for a detailed dynamical analysis due to strong currents in this region (e.g., Flament et al. (1996)). Finally, the space correlation radius of the velocity field in the tropical Pacific Ocean is taken as the Rossby deformation radius $R = 250$ km (Cushman-Roisin (1994)), and the Lagrangian correlation time is taken as $\tau = 3$ days, based on the analysis of drifter motion in the WOCE data set (Bauer et al. (1999)).

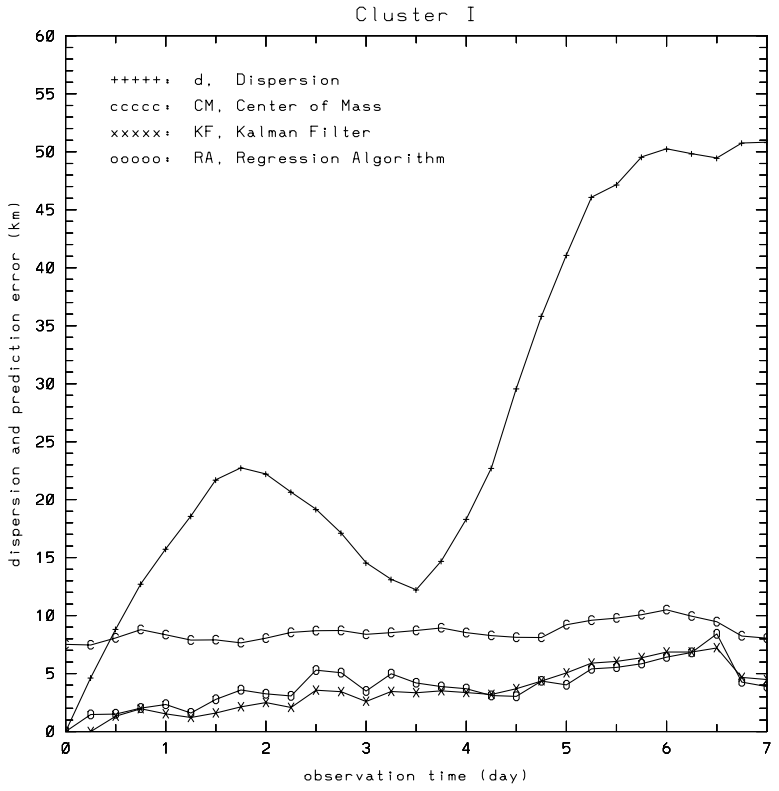
Clusters I–III are sorted according to the difficulty of prediction, quantified by the initial scale of the cluster, the velocity variance and prediction period. Cluster I consists of 5 drifters launched within a scale ($\sim R_0$) of approximately 10 km from each other (Figure 11(a)), and the prediction algorithms are applied for 7 days of particle motion, during which the velocity variance is approximately $\sigma_v^2 = 430$ km^2/day^2. During 7 days of motion, these drifters do not spread apart significantly. Given $R_0 \ll R$ and the low velocity variance, one can anticipate very good performance by RA based on the results from theory and stochastic simulations. Dispersion $d(t)$ and prediction errors $s(t)$ from RA, CM, and KM are calculated by sequentially selecting each drifter as predictand and the remaining others in the cluster as predictors, corresponding to the root mean square of that of all cluster particles. The results are shown in Figure 11(b) for cluster I for an observation time of 7 days. This figure shows that prediction errors of both KF and RA are less than that of CM during the observation period and that RA is as accurate as KF. More quantitatively, dispersion reaches approximately 51 km at $T = 7$ days, error from CM is 8 km ($s_r = 0.16$), and error from KF and RA is about 5 km ($s_r = 0.1$).

Cluster II consists of 7 drifters that are also released with a mean diameter of approximately 10 km, but disperses much faster than Cluster I due to a higher velocity variance of $\sigma_v^2 = 720$ km^2/day^2, and the mean cluster diameter reaches 25 km and 50 km after 7 and 14 days of observation time, respectively (Figure 12(a)). Dispersion and prediction errors for cluster II over an observation period of 14 days are shown in Figure 12(b), and the conclusion remains the same as for cluster I; prediction errors of both KF and RA are less than that of CM during the observation period, and RA is as accurate as KF. Dispersion reaches approximately 136 km at $T = 14$ days, error from CM is 44 km ($s_r = 0.32$), and errors from KF and RA are about 26 km ($s_r = 0.19$). The sensitivity of the prediction accuracy of RA to the number of predictors p is investigated by randomly eliminating drifters from cluster II. Figure 12(c) shows the dispersion curve based on the entire cluster and prediction errors calculated for $p = 6$ (same as in Figure 12(b)), $p = 5$, $p = 4$, and $p = 3$. When $p = 3$, a drastic reduction of prediction accuracy takes place, which is found to be independent of the combination of chosen predictors in this cluster. Otherwise, the prediction accuracy gradually decreases as the number of predictors is decreased from 6 to 4, but the accuracy of the method using 4 to 6 predictors remains essentially constant for $T \leq \tau$ or for $T \leq 3$ days.

The motion of cluster III, consisting of 10 drifters, is investigated for 21 days

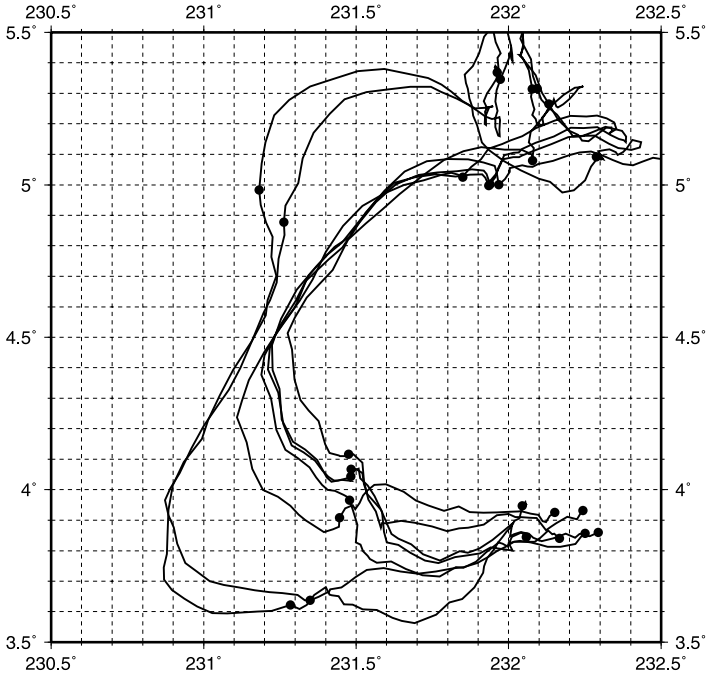


(a)

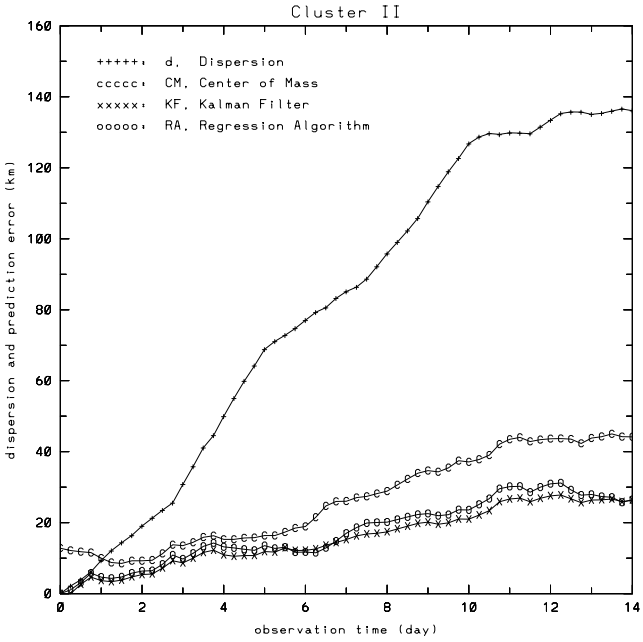


(b)

FIG. 11. (a) *Drifters trajectories in cluster* I. *The circles mark* 7*-day intervals.* (b) *Comparison of the dispersion,* $d(t)$*, and prediction errors,* $s(t)$*, of RA, CM, and KM for an observation time of* 7 *days for cluster* I.

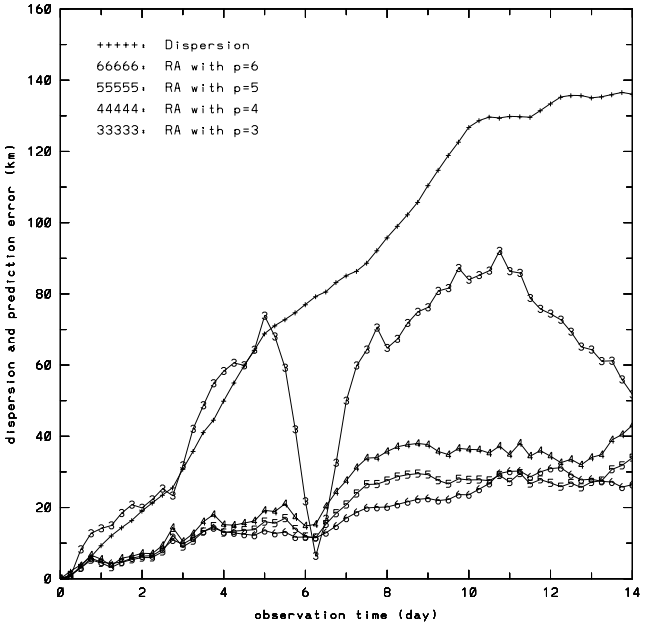


(a)



(b)

FIG. 12. (a) *Drifters trajectories in cluster* II. *The circles mark* 7*-day intervals.* (b) *Comparison of the dispersion,* $d(t)$*, and prediction errors,* $s(t)$*, of RA, CM, KM for an observation time of* 14 *days for cluster* II.



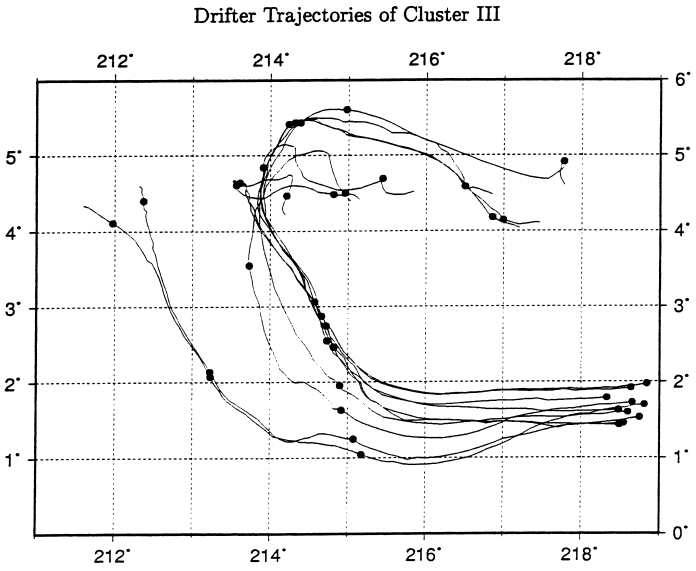
(c)

FIG. 12. (c) *Sensitivity of the prediction error of RA to the number of predictors in cluster* II.

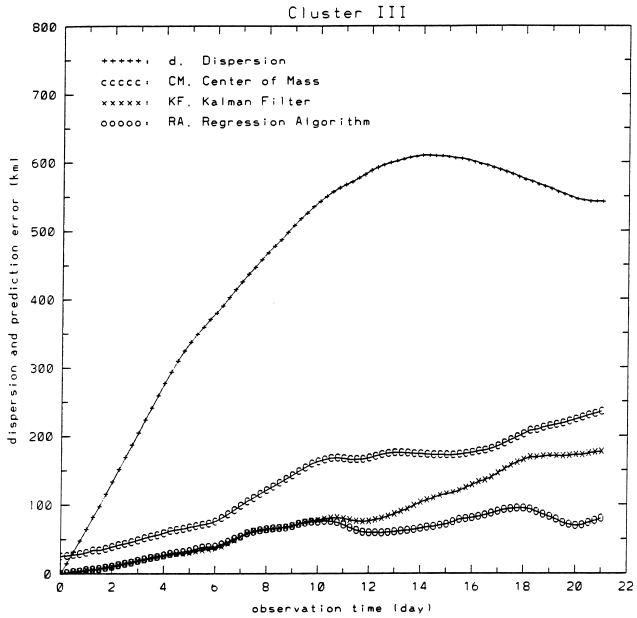
during which velocity variance is $\sigma_v^2 = 2240$ km^2/day^2, or the highest of the three clusters. These drifters were released over an area with an approximate diameter of 30 km, but this scale increases to approximately 100, 180, and 250 km after 7, 14, and 21 days, respectively (Figure 13(a)). Dispersion and prediction errors for cluster III are shown in Figure 13(b). During the first 10 days, prediction errors of both KF and RA are approximately the same and less than that of CM, but during the second half of the observation period, the error of KF increases faster than that of RA. This increase appears to be related to the inability of the KF algorithm to follow the bifurcation of some drifters in a larger group as effectively as the RA technique. Dispersion is 426 km (543 km), error from CM is 99 km (235 km) or $s_r = 0.23$ ($s_r = 0.43$), error from KF is 55 km (176 km) or $s_r = 0.13$ ($s_r = 0.32$), and error from RA is 54 km (80 km) or $s_r = 0.13$ ($s_r = 0.15$) at $T = 7$ days ($T = 21$ days).

All in all, the real data comparison of different prediction algorithms is in good qualitative agreement with the simulation results. Even the prediction error values are of the same order, as our simple error theory concludes. Deviations are related to oversimplifications accepted in the considered stochastic model such as the shear flow linearity and fluctuations isotropy.

In summary, the algorithm described in this study presents several important simplifications with respect to the KF method developed and investigated by Piterbarg (2001b) and Özgökmen et al. (2000), (2001): (i) This algorithm does not require any parameters, such as the Lagrangian parameters describing the characteristics of the underlying flow, the velocity correlation space scale R, and the Lagrangian correlation time scale τ. (ii) RA does not utilize the mean flow field, the calculation of which requires large data sets and the associated subgrid scale interpolation introduces further errors. (iii) RA does not need to be initialized with turbulent velocity fluctuations at



(a)



(b)

FIG. 13. (a) *Drifter trajectorie in cluster* III. *The circles mark* 7*-day intervals.* (b) *Comparison of the dispersion,* $d(t)$*, and prediction errors,* $s(t)$*, of RA, CM, KM for an observation time of* 21 *days for cluster* III.

the launch location. (iv) RA is not based on the integration of velocity field to estimate the particle position, which necessarily leads to accumulation of velocity errors as errors of drifter location. (v) Consequently RA is computationally far simpler than KF. Despite these simplifications, it is found on the basis of several oceanic clusters that RA outperforms CM and that RA is as accurate as KF. Also, predictions from RA appear to remain applicable over a time scale of $T \gg \tau$, or much longer than one would anticipate. In future studies, it will be investigated theoretically and numerically how this method performs when $R_0 \approx R$, which is likely to be the case in mid- and high-latitude oceans and for bifurcating clusters.

Appendix.

A.1. Prediction error in terms of separation. Using the definition (4) and expression (6) for b_k, obtain

$$\begin{aligned} s^2 &= E|\hat{\mathbf{r}}_M(t) - \mathbf{r}_M(t)|^2 \\ &= E\{(\mathbf{r}_c(t) + \mathbf{S}(t)\mathbf{S}(0)^{-1}(\mathbf{r}_M(0) - \mathbf{r}_c(0)) - \mathbf{r}_M(t))^T(\mathbf{r}_c(t) \\ &\quad + \mathbf{S}(t)\mathbf{S}(0)^{-1}(\mathbf{r}_M(0) - \mathbf{r}_c(0)) - \mathbf{r}_M(t))\} \\ &= E\{(\mathbf{r}_c(t) - \mathbf{r}_M(t))^T(\mathbf{r}_c(t) - \mathbf{r}_M(t))\} \\ &\quad + 2E\{(\mathbf{r}_c(t) - \mathbf{r}_M(t))^T\mathbf{S}(t)\}\mathbf{S}(0)^{-1}(\mathbf{r}_M(0) - \mathbf{r}_c(0)) \\ &\quad + (\mathbf{r}_M(0) - \mathbf{r}_c(0))^T\mathbf{S}(0)^{-1}E\{\mathbf{S}(t)^T\mathbf{S}(t)\}\mathbf{S}(0)^{-1}(\mathbf{r}_M(0) - \mathbf{r}_c(0)) \\ &= E\{(\mathbf{r}_c(t) - \mathbf{r}_M(t))^T(\mathbf{r}_c(t) - \mathbf{r}_M(t))\} \\ &\quad + 2\textstyle\sum_{k=1}^{p} b_k E\{(\mathbf{r}_c(t) - \mathbf{r}_M(t))^T(\mathbf{r}_k(t) - \mathbf{r}_c(t))\} \\ &\quad + \textstyle\sum_{k,l=1}^{p} b_k b_l E\{(\mathbf{r}_k(t) - \mathbf{r}_c(t))^T(\mathbf{r}_l(t) - \mathbf{r}_c(t))\}. \end{aligned} \tag{A1}$$

Then, dropping the time argument for brevity,

$$\begin{aligned} E\{(\mathbf{r}_k - \mathbf{r}_c)^T(\mathbf{r}_l - \mathbf{r}_c)\} &= \tfrac{1}{p^2}\textstyle\sum_{i,j=1}^{p} E\{(\mathbf{r}_k - \mathbf{r}_i)^T(\mathbf{r}_l - \mathbf{r}_j)\} \\ &= -\tfrac{1}{2p^2}\textstyle\sum_{i,j=1}^{p}(\rho_{kl} + \rho_{ij} - \rho_{ki} - \rho_{lj}) \\ &= -\tfrac{1}{2}\rho_{kl} + \tfrac{1}{2p}\textstyle\sum_{i=1}^{p}(\rho_{ki} + \rho_{li}) - \tfrac{1}{2p^2}\textstyle\sum_{i,j=1}^{p}\rho_{ij}. \end{aligned} \tag{A2}$$

In the latter we used the relation

$$E\{\mathbf{r}_k{}^T\mathbf{r}_l\} = \frac{1}{2}(-\rho_{kl} + E\{\mathbf{r}_k{}^T\mathbf{r}_k\} + E\{\mathbf{r}_l{}^T\mathbf{r}_l\}).$$

By substituting (A2) into (A1), using the obvious relation

$$\sum_{k=1}^{p} b_k = 0,$$

and changing the summation indexes, we obtain (5).

A.2. Separation for Brownian flow (close initial positions). The function

$$\rho(t, r) = E\{(\mathbf{r}_1(t) - \mathbf{r}_2(t))^2\},$$

where $|\mathbf{r}_1(0) - \mathbf{r}_2(0)| = r$, satisfies

$$\frac{\partial \rho}{\partial t} = L_s\rho\ , \qquad \rho(0, r) = r^2, \tag{A3}$$

where in the isotropic case the generator for the separation process is

$$L_s = \frac{b_0 - b_N(r)}{r}\frac{\partial}{\partial r} + (b_0 - b_L(r))\frac{\partial^2}{\partial r^2}\ .$$

Substitute expansions (13), (15) into (A3). The result is

$$\frac{d\rho_1}{dt} = \overline{\beta}\rho_1, \quad \rho_1(0) = 1, \qquad \frac{d\rho_0}{dt} = \beta_0\rho_0 + \overline{\gamma}\rho_1, \quad \rho_0(0) = 0,$$

where

$$\overline{\beta} = \beta_N + \beta_L, \quad \beta_0 = 2\beta_N + 6\beta_L, \qquad \overline{\gamma} = \gamma_N + \gamma_L.$$

Solving the latter equations we obtain (16).

A.3. Dispersion for the flow with memory in presence of linear shear flow. For simplicity assume $\mathbf{U} = 0$. From (20) it follows that

$$\mathbf{r}(t) = \bar{\mathbf{r}}(t) + \int_0^t \exp(\mathbf{G}(t-s))\mathbf{v}(s)ds,$$

where $\mathbf{v}(t)$ is a two-dimensional Ornstein–Uhlenbeck process with the covariance $\sigma_v^2\exp(-t/\tau)\mathbf{I}$ and

$$\bar{\mathbf{r}} = \exp(\mathbf{G}t)\mathbf{r}_0 + (\exp(\mathbf{G}t) - \mathbf{I})\mathbf{G}\mathbf{U}.$$

Set

$$d_0^2 = (\bar{\mathbf{r}} - \mathbf{r}_0)^2;$$

then

$$d^2 = d_0^2 + \sigma_v^2\int_0^t\int_0^t Sp(\exp(\mathbf{G}(t-s_1))\exp(\mathbf{G}^T(t-s_2)))\exp(-|s_1-s_2|/\tau)ds_1ds_2,$$

where $Sp(\mathbf{A})$ means the race of matrix A. Then we use the following relations:

$$Sp(\mathbf{A}) = \lambda_1(\mathbf{A}) + \lambda_2(\mathbf{A}), \quad \lambda(\exp(\mathbf{A})) = \exp\lambda((\mathbf{A})),$$

where $\lambda(\mathbf{A})$ is an eigenvalue of $\mathbf{A}$ and arrive at (21).

A.4. Separation for the flow with memory (close initial positions and velocities). In the case considered, the separation satisfies

$$\frac{\partial\rho}{\partial t} = L_s\rho, \quad \rho|_{t=0} = x^2 + y^2, \tag{A4}$$

with the generator (24) written in the coordinatewise form

$$\begin{aligned} L_s = {} & (\gamma y + \omega x + u)\frac{\partial}{\partial x} - (\gamma y + \omega x + v)\frac{\partial}{\partial y} - \frac{1}{\tau}u\frac{\partial}{\partial u} - \frac{1}{\tau}v\frac{\partial}{\partial v} \\ & + (b_0 - b_N(r) - \frac{x^2}{r^2}(b_L(r) - b_N(r))\frac{\partial^2}{\partial u^2} + (b_0 - b_N(r) - \frac{y^2}{r^2}(b_L(r) - b_N(r))\frac{\partial^2}{\partial v^2} \\ & - \frac{2xy}{r^2}(b_L(r) - b_N(r))\frac{\partial^2}{\partial u\partial v}. \end{aligned}$$

Substituting expansions (25), (27) into (A4) we obtain (27), where

$$\mathbf{A} = \begin{pmatrix} 2\gamma & -2\omega & 0 & 2\beta_L & 0 & 2\beta_N & 0 & 0 & 0 & 0 \\ \omega & 0 & -\omega & 0 & -2\beta & 0 & 0 & 0 & 0 & 0 \\ 0 & 2\omega & -2\gamma & 2\beta_N & 0 & 2\beta_L & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2/\tau & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2/\tau & 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 & -2/\tau & 0 & 0 & 0 & 1 \\ 2 & 0 & 0 & 0 & 0 & 0 & \gamma - 1/\tau & 0 & -\omega & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 & \gamma - 1/\tau & 0 & -\omega \\ 0 & 2 & 0 & 0 & 0 & 0 & \omega & 0 & -\gamma - 1/\tau & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & \omega & 0 & -\gamma - 1/\tau \end{pmatrix},$$

where $\beta = \beta_N - \beta_L$,

$$\mathbf{a}_0 = (1\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0)^T.$$

For the zero mean flow $\gamma = 0$, $\omega = 0$, and (27) reduces to (29) for $\rho_0 = a_4 + a_6$, $\rho_1 = a_1 + a_3$, and $\rho_{01} = a_7 + a_{10}$.

REFERENCES

S. Bauer, M. S. Swenson, A. Griffa, A. J. Mariano, and K. Owens (1999), *Eddy-mean flow decomposition and eddy-diffusivity estimates in the tropical Pacific Ocean*, J. Geophys. Res., 103, pp. 30855–30871.

P. Baxendale and T. Harris (1986), *Isotropic stochastic flows*, Ann. Probab., 14, pp. 1155–1179.

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

B. Cushman-Roisin (1994), *Introduction to Geophysical Fluid Dynamics*, Prentice-Hall, Englewood Cliffs, NJ, p. 285.

R. E. Davis (1991a), *Lagrangian ocean studies*, Ann. Rev. Fluid Mech., 23, pp. 43–64.

R. E. Davis (1991b), *Observing the general circulation with floats*, Deep-Sea Research, 38, pp. 5531–5571.

P. J. Flament, 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, pp. 610–613.

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

A. Griffa, K. Owens, L. Piterbarg, and B. Rozovskii (1995), *Estimates of turbulence parameters from Lagrangian data using a stochastic particle model*, J. Marine Res., 53, pp. 371–401.

D. V. Hansen and P.-M. Poulain (1996), *Quality control and interpolations of WOCE/TOGA drifter data*, J. Atmos. Ocean Techn., 13, pp. 900–909.

H. Kunita (1990), *Stochastic Flows and Stochastic Differential Equations*, Cambridge University Press, Cambridge, UK.

R. S. Liptser and A. N. Shiryaev (1978), *Statistics of Random Processes*, Springer-Verlag, Berlin.

A. S. Monin and A. M. Yaglom (1975), *Statistical Fluid Mechanics: Mechanics of Turbulence*, MIT Press, Cambridge, MA.

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

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

L. I. Piterbarg (1998), *Drift estimation for Brownian flows*, Stochastic Process. Appl., 79, pp. 132–149.

L. I. Piterbarg (2001a), *The top Lyapunov exponent for a stochastic flow modeling the upper ocean turbulence*, SIAM J. Appl. Math., 62, pp. 777–800.

L. I. Piterbarg (2001b), *Short-term prediction of Lagrangian trajectories*, J. Atmos. Ocean Techn., 18, pp. 1398–1410.

T. Schneider (1998), *Lagrangian Drifter Models as Search and Rescue Tools*, M.S. thesis, Dept. of Meteorology and Physical Oceanography, University of Miami, Miami, FL.

D. J. Thomson (1986), *A random walk model of dispersion in turbulent flows and its application to dispersion in a valley*, Quart. J. R. Met. Soc., 112, pp. 511–529.

E. Zambianchi and A. Griffa (1994), *Effects of finite scales of turbulence on dispersion estimates*, J. Marine Res., 52, pp. 129–148.

C. L. Zirbel and E. Cinlar (1996), *Dispersion of particle systems in Brownian flows*, Adv. in Appl. Probab., 28, pp. 53–74.