2001

Reconstructing Basin-Scale Eulerian Velocity Fields From Simulated Drifter Data

M. Toner

A. C. Poje

A. D. Kirwan
Jr.

C. K. R. T. Jones

B. L. Lipphardt

See next page for additional authors

Original Publication Citation

Toner, M., Poje, A.C., Kirwan, A.D., Jones, C., Lipphardt, B.L., & Grosch, C.E. (2001). Reconstructing basin-scale eulerian velocity fields from simulated drifter data. *Journal of Physical Oceanography, 31*(5), 1361-1376. doi: 10.1175/1520-0485(2001)0312.0.co;2

Authors
M. Toner, A. C. Poje, A. D. Kirwan, C. K. R. T. Jones, B. L. Lipphardt, and C. E. Grosch

Reconstructing Basin-Scale Eulerian Velocity Fields from Simulated Drifter Data

M. TONER,* A. C. POJE,+ A. D. KIRWAN JR.,* C. K. R. T. JONES,+ B. L. LIPPHARDT,* AND C. E. GROSCH#

* College of Marine Studies, University of Delaware, Newark, Delaware
+ Lefschetz Center for Dynamical Systems, Brown University, Providence, Rhode Island
# Center for Coastal and Physical Oceanography, Old Dominion University, Norfolk, Virginia

(Manuscript received 6 October 1999, in final form 30 August 2000)

ABSTRACT

A single-layer, reduced-gravity, double-gyre primitive equation model in a 2000 km × 2000 km square domain is used to test the accuracy and sensitivity of time-dependent Eulerian velocity fields reconstructed from numerically generated drifter trajectories and climatology. The goal is to determine how much Lagrangian data is needed to capture the Eulerian velocity field within a specified accuracy. The Eulerian fields are found by projecting, on an analytic set of divergence-free basis functions, drifter data launched in the active western half of the basin supplemented by climatology in the eastern domain. The time-dependent coefficients are evaluated by least squares minimization and the reconstructed fields are compared to the original model output. The authors find that the accuracy of the reconstructed fields depends critically on the spatial coverage of the drifter observations. With good spatial coverage, the technique allows accurate Eulerian reconstructions with under 200 drifters deployed in the 1000 km × 1400 km energetic western region. The base reconstruction error, achieved with full observation of the velocity field, ranges from 5% (with 191 basis functions) to 30% (with 65 basis functions). Specific analysis of the relation between spatial coverage and reconstruction error is presented using 180 drifters deployed in 100 different initial configurations that maximize coverage extremes. The simulated drifter data is projected on 107 basis functions for a 50-day period. The base reconstruction error of 15% is achieved when drifters occupy approximately 110 (out of 285) 70-km cells in the western region. Reconstructions from simulated mooring data located at the initial positions of representative good and poor coverage drifter deployments show the effect drifter dispersion has on data voids. The authors conclude that with appropriate coverage, drifter data could provide accurate basin-scale reconstruction of Eulerian velocity fields.

1. Introduction

There has been an explosion in Lagrangian information about the ocean during the past two decades (Sombardier 1992; Sombardier and Niiler 1994). In the Gulf of Mexico, for example, over 300 satellite-tracked drifters were released between 1994 and 1995 in the Louisiana–Texas shelf. In the West Florida Shelf study even more drifters were released during 1996–97. A comparable number of drifters were released as part of consortium of oil companies during the 1990s in the abyssal Gulf of Mexico and a number of analyses of these observations have now appeared in the literature (Lewis et al. 1989; Forristall et al. 1992; Vidal et al. 1992; Glenn and Ebbsmeyer 1993).

There are at least four reasons for the emphasis on Lagrangian data. First, velocity data obtained by satellite tracking of drifters are inexpensive relative to the same amount of data obtained from moorings and ships. Second, drifter data are available to operational users in real time whereas there can be considerable delay in transmitting data from conventional sources. Third, drifters are the only practical method for obtaining synoptic surface current data over regional to basin scales. Finally, in many applications the Lagrangian information (particle pathways, dispersion) is itself of prime importance. For these reasons it is likely that drifter deployments of similar and even greater numbers will continue into this millennium in other parts of the World Ocean.

Despite its preponderance, Lagrangian data tends to be underutilized in nowcasting the ocean state and in data assimilation schemes for predictive modeling. There are a number of technical reasons for this. Unlike data from moorings, Lagrangian velocity data are irregularly spaced and tend either to bunch together or advect out of regions of interest. In contrast, Eulerian data are easily reconciled with numerical models.

As succinctly summarized by Davis (1985b), the problem of mapping observations is achieved either by least squares projection of the data onto a set of or-

Corresponding author address: Dr. Michael Toner, Graduate School of Marine Studies, University of Delaware, Robinson Hall, Newark, DE 19716-3501.
E-mail: toner@udel.edu

thogonal basis functions, by optimal linear interpolation, or, as proposed by Davis, by some combination of both. Optimal interpolation requires statistical information about the uncertainties in both the observations and the true field. The latter is usually not available in oceanography. Owens (1991) addressed many of the issues in developing a statistical description of the mean circulation in the Atlantic from SOFAR floats.

The focus here is reconstructing daily maps of the Eulerian velocity. So as not to prejudice the reconstruction with a priori statistical knowledge of the velocity field, we use a least squares approach to fit geometrical orthogonal basis functions to the simulated drifter data. The basis functions are chosen using general characteristic flow scales and knowledge of the climatological energetics of the velocity field. Details and attributes of this approach are discussed in the next section; here it is noted that this spectral approach is global in that it can be applied to the entire model domain or any subset thereof provided some information exterior to the subdomain is available, such as a climatological mean. Moreover, the basis functions have simple physical interpretations that describe energetics of the flow.

The geometrical orthogonal function technique does not provide the maximum data compression as does the empirical orthogonal function (EOF) technique. Using the ideas of Davis (1985a) it is possible to combine the EOF approach with optimal estimation and provide an objective analysis without full statistical knowledge of the signal. Such an approach, however, is not consistent with our approach of excluding a priori knowledge of the flow specifics. For the purposes of this study, physical attributes are more important than data compression.

Of fundamental importance in utilizing Lagrangian data is the sensitivity of reconstructed Eulerian velocities to the attendant characteristics of the input data. To address this question, we conduct controlled experiments with velocity fields produced by a reduced-gravity, primitive equation basin model. This standard double-gyre model on a 2000-km square domain, previously used by Poje and Haller (1999) to study cross-stream, Lagrangian mixing, necessarily eliminates stratification, physical side and bottom conditions, and realistic wind forcing, all certainly important in real oceanographic settings. These simplifications, however, allow us to address scale and sampling issues unambiguously. The geometry of the double-gyre model also provides a flow with wide variations in eddy kinetic energy density across the domain.

In the spirit of previous work by Rao and Schwab (1981), Eremeev et al. (1992a), and Cho et al. (1998) we objectively reconstruct the Eulerian current fields of this model from Lagrangian data and assess how the accuracy of the reconstructed velocity field depends on the spatial coverage of the drifter paths. Launch patterns that accentuate coverage extremes are used for this purpose. Only the velocity data from the drifters are used; no effort is made to utilize trajectory information. A comparison of reconstructed velocity fields generated with drifter data is made with reconstructed velocity fields generated with mooring data. The locations of the moorings are the initial launch locations of the drifters, so the same coverage extremes are realized.

The focus here is on the practical aspects of reconstructing the Eulerian fields. The accuracy metric is the kinetic energy of the vector difference between the original model and the reconstructed velocity field. In this analysis, the model results are considered ''truth'' and the trajectory data are considered error free. It is hoped this is the first step in moving toward objective drifter deployment strategies.

The rest of the paper is organized as follows. The next two sections summarize, respectively, the analysis methodology and the numerical model. In section 4, the mode selection criteria and model energetics are presented. This section introduces the measurement of error in the reconstructed field. Section 5 examines topics such as the number of drifters required, climatological data sampling, deployment issues, and mooring comparisons are addressed. In the last section, the conclusions of the study are summarized and implications are discussed.

2. Normal mode analysis

One of the difficulties in comparing Eulerian and Lagrangian velocities is that the locations of the former are fixed for the observation period, while the latter provide data from constantly changing locations. One remedy might be to interpolate the Lagrangian data onto the Eulerian grid. An enormous variety of interpolation schemes are available and many rely on empirical estimates of correlation scales. Optimal methods would require both Eulerian and Lagrangian correlation scales, which are not available for the World Ocean. Moreover this approach forces all interpolator uncertainty into the Lagrangian data. Spectral approaches are another alternative. As with any method they also have deficiencies; however, for the goals of this paper they are appropriate. We have elected to follow an approach used by Rao and Schwab (1981) in an analysis of a current meter array from Lake Ontario. This approach was generalized by Eremeev et al. (1992a) to apply to a three-dimensional incompressible velocity field. Since then it has been used by Eremeev et al. (1992b, 1995a,b) and more recently by Cho et al. (1998), Lipphardt et al. (1998, 2000), and Schulz (1999). Lipphardt et al. (1998) give a succinct summary of the general approach applicable to an arbitrarily shaped domain that may include both open and closed boundaries.

The three-dimensional incompressible velocity field is expressed in terms of two scalar potentials as

$$\mathbf{u} = \nabla \times [\mathbf{k}(-\Psi) + \nabla \times (\mathbf{k}\Phi)]. \qquad (1)$$

Here, $\mathbf{k}$ is the unit vector in the vertical direction. Equation (1) generalizes the model used by Rao and Schwab

(1981) and Cho et al. (1998) to account for a nonquasigeostrophic component of flow that could be present whenever there are intense gradients such as eddy shedding. This form ensures the velocity field is exactly incompressible in three dimensions. However, this study does not make use of vertical velocities.

Projection of the vertical component of relative vorticity from (1) gives a Helmholtz equation for Ψ with homogeneous Dirichlet boundary conditions; Ψ is expanded using eigenfunctions, which are called *Dirichlet* modes (ψ_n). These modes represent the quasigeostrophic component of the flow. They are solutions to

$$\nabla^2\psi_n + \lambda_n\psi_n = 0, \qquad \psi_n|_{\text{boundary}} = 0. \tag{2}$$

From (1), the gradients of ψ_n are expressed as

$$(u_n^D, v_n^D) = \left(\frac{-\partial\psi_n}{\partial y}, \frac{\partial\psi_n}{\partial x}\right). \tag{3}$$

Projection of the vertical velocity component gives a Helmholtz equation for Φ with homogeneous Neumann boundary conditions. As with the Dirichlet eigenfunctions Φ is expanded using eigenfunctions, which are called *Neumann* modes (ϕ_m). They are solutions to

$$\nabla^2\phi_m + \mu_m\phi_m = 0, \qquad (\mathbf{p}\cdot\nabla\phi_m)|_{\text{boundary}} = 0, \tag{4}$$

where $\mathbf{p}$ is the outward pointing boundary normal. The ϕ_m may be thought of as velocity potential or divergence modes, with zero relative vorticity. From (1), the gradients of ϕ_m are expressed as

$$(u_m^N, v_m^N) = \left(\frac{\partial\phi_m}{\partial x}, \frac{\partial\phi_m}{\partial y}\right). \tag{5}$$

With arbitrarily shaped coastlines (2) and (4) are readily solved numerically. Note that the homogeneous boundary conditions ensure no flow across the boundaries. If the domain has open boundaries, then these solutions must be augmented with a boundary flow solution that uses known flow across the boundary. See Lipphardt et al. (2000) for details and an application to Monterey Bay. This latter factor is important when studying an open subdomain. In the present case, the domain geometry is so simple we use the analytic eigenfunctions (sines and cosines) rather than numerically generated values.

Thus the Eulerian velocity fields can be represented as

$$u_{\text{proj}}(x, y, t) = \sum_{n=1}^{N} A_n(t)u_n^D(x, y) + \sum_{m=1}^{M} B_m(t)u_m^N(x, y) \tag{6}$$

and

$$v_{\text{proj}}(x, y, t) = \sum_{n=1}^{N} A_n(t)v_n^D(x, y) + \sum_{m=1}^{M} B_m(t)v_m^N(x, y). \tag{7}$$

The u_n^D, u_m^N, v_n^D, and v_m^N are known (sines and cosines in our case), while the amplitudes $A_n(t)$ and $B_m(t)$ are determined from observations. Specifically, if $\mathbf{u}_{\text{ob}}(x_i(t), y_i(t), t)$ are I observations at time t, then the observation vector of length $2I$ is

$$\mathbf{o} = \begin{bmatrix} u_{\text{ob}}(x_1(t), y_1(t), t) \\ v_{\text{ob}}(x_1(t), y_1(t), t) \\ u_{\text{ob}}(x_2(t), y_2(t), t) \\ v_{\text{ob}}(x_2(t), y_2(t), t) \\ \vdots \\ u_{\text{ob}}(x_I(t), y_I(t), t) \\ v_{\text{ob}}(x_I(t), y_I(t), t) \end{bmatrix}. \tag{8}$$

The unknown modal amplitudes

$$\mathbf{c} = [A_1(t), A_2(t), \cdots, A_N(t), B_1(t), B_2(t), \cdots, B_M(t)]^{\mathrm{T}} \tag{9}$$

are determined as a least squares minimization of

$$\min_{\mathbf{c}} \|\mathbf{M}\mathbf{c} - \mathbf{o}\|_2, \tag{10}$$

where $M_{l,j}$ is a $2I \times (N + M)$ matrix, $[\ ]^{\mathrm{T}}$ is the transpose operation, and $\|\ \|_2$ denotes the standard quadratic scalar norm. Details on the theory of least squares minimization can be found in any standard matrix algebra text such as Golub and van Loan (1983). The jth index of $M_{l,j}$ refers the u ($l = 1, 3, 5, \cdots, 2I - 1$) or v ($l = 2, 4, 6, \cdots, 2I$) component of the basis function velocity derived from the appropriate streamfunction ($j = 1, \cdots, N$) or velocity potential ($j = N + 1, \cdots, N + M$) evaluated at the location of the ith observation in (8). The reconstructed Eulerian velocity at any time is given by the right-hand sides of (6) and (7).

3. Numerical model

Here we use a standard, reduced-gravity primitive equation model to provide both the base Eulerian velocity field we seek to reconstruct and the numerically generated drifter observations. Reduced-gravity, 1.5-layer primitive equation and quasigeostrophic models have been in use by the community for at least the last two decades (Holland 1978; Marshall 1984), and more recently, such models have been used to examine Lagrangian and Eulerian transport (Figueroa 1994; Figueroa and Olsen 1994), multiple equilibria, and low-dimensional reconstructions (Jiang et al. 1995; Berloff and Meacham 1998) as well as statistical questions involving low-frequency climate variability (McCalpin 1995; McCalpin and Haidvogel 1996).

In advection form, the governing equations are

$$\begin{aligned} &\frac{\partial h}{\partial t} + \frac{\partial(uh)}{\partial x} + \frac{\partial(vh)}{\partial y} = 0 \\ &\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} = -g'\frac{\partial h}{\partial x} + f_0(1 + \beta y)v \\ &\qquad\qquad + \nu\nabla^2 u + F^u \\ &\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} = -g'\frac{\partial h}{\partial y} - f_0(1 + \beta y)u \\ &\qquad\qquad + \nu\nabla^2 v + F^v, \end{aligned} \tag{11}$$

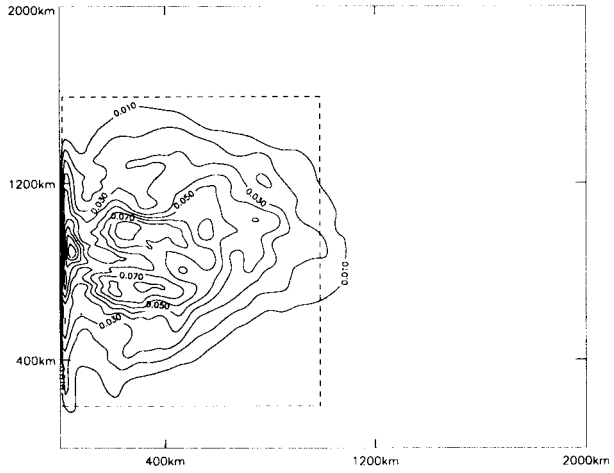


FIG. 1. Contours of the mean eddy kinetic energy, $\frac{1}{2}[\overline{(u-\overline{u})^2} + \overline{(v-\overline{v})^2}]$, for the 10-yr history of the double-gyre flow. The kinetic energy ranges from 0 to 0.099 $m^2\ s^{-2}$ and contour lines are shown in 0.01 $m^2\ s^{-2}$ increments. The active subdomain, *D*, is marked by the dash-lined box.

where the vector **F** represents the imposed wind stress given by

$$F^u = \frac{\tau_0}{\rho H_0 \sin}\left(\frac{2\pi y}{L_y}\right)$$

and $F^v = 0$. The inclusion of a Laplacian diffusion (a simple model of subgrid-scale turbulent transport and dissipation) allows the imposition of Dirichlet boundary conditions on the velocity field. The height field varies freely on the boundaries.

Equations (11) are solved by standard finite differences using the MP-DATA advection algorithm of Smolarkiewicz and Margolin (1993) on a C grid. The domain is a 2000-km square; spatial resolution is 10 km and the computation time step is 15 minutes. Full details of the numerical procedure are given in Jones et al. (1997). The flow parameters are chosen to coincide with those in Poje and Haller (1999).

Figure 1 shows the average eddy kinetic energy density in the domain during the 10 years prior to the reconstruction period. The western half of the domain is dominated by a strong free jet formed by two legs of the western boundary current along with a number of large-scale (100-km radius) cyclonic and anticyclonic eddies shed intermittently by jet rollup. Outside this active region, the flow is quiescent with fluctuations from the short time mean dominated by small amplitude Rossby waves.

The Eulerian fields were archived every 24 hours for a 50-day reconstruction period beginning in the 15th model year after spinup from rest. As detailed in Poje and Haller (1999), during these 50 days a hyperbolic point in the center of the domain strengthens, peaks, and then weakens as a cold core ring detaches from the central jet. This provides an interesting period to study the reconstruction error because it represents the typical topology change of the flow field during an eddy-shedding event.

Our goal is to reconstruct the Eulerian velocity field in the active, eddy-shedding region *D* by tracking drifters launched there. Figure 2 shows the velocity field of the model at day 1 and day 49 of the reconstruction period. During this period, one eddy is shed from the central jet and a previously shed eddy impacts the western boundary.

4. Mode selection

The rectangular geometry of the domain affords an analytic representation of the basis functions. The solutions of (2) and (4) take the form

$$\psi_{\mathbf{n}}(x, y) = \frac{2}{\pi}(\mathbf{n}\cdot\mathbf{n})^{-1/2} \sin\left(\frac{n_x}{l_x}\pi x\right) \sin\left(\frac{n_y}{l_y}\pi y\right) \quad (12)$$

and

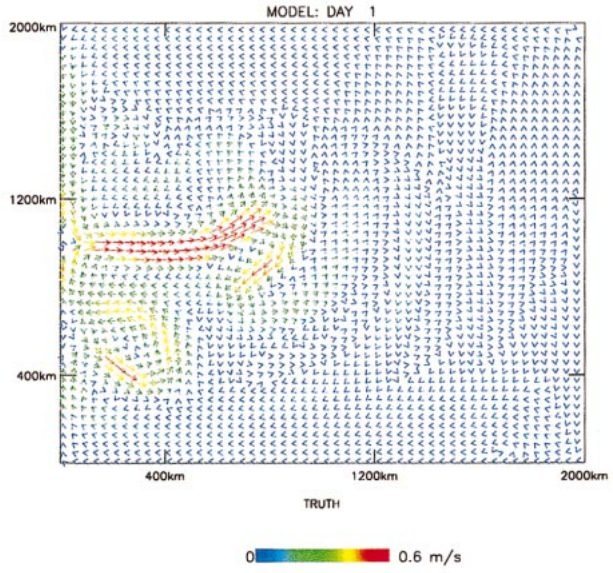
$$\phi_{\mathbf{n}}(x, y) = \frac{2}{\pi}(\mathbf{n}\cdot\mathbf{n})^{-1/2} \cos\left(\frac{n_x}{l_x}\pi x\right) \cos\left(\frac{n_y}{l_y}\pi y\right), \quad (13)$$

where $\mathbf{n} = (n_x/l_x, n_y/l_y)$. The mode-number-dependent scaling ensures that the square of the amplitude corresponds to the modal energy content.

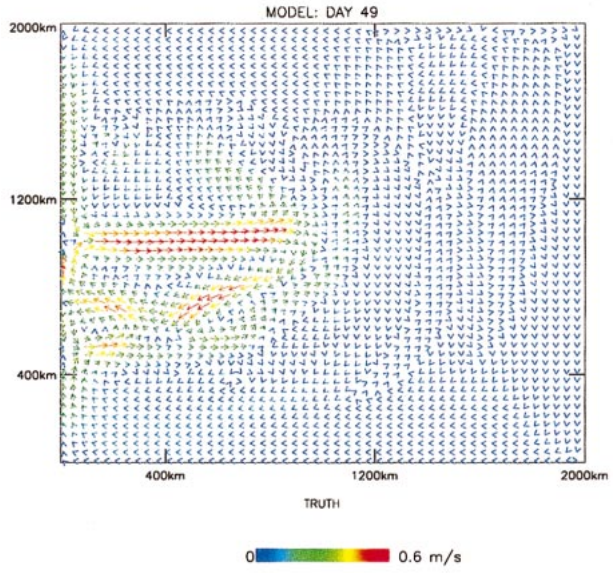
The methodology of the normal mode analysis allows a 2D divergence-free approximation to the flow by selecting only streamfunction modes. Energy in the velocity potential modes gives the spatial structure of the 2D divergent flow, which in the present model estimates the nonquasigeostrophic component. Of course 2D divergence-free flow can have nonquasigeostrophic components as well. Viscous effects such as wind stress and turbulent stresses will produce a nonquasigeostrophic response.

The energetic scales of the double-gyre flow are bounded below by either the Rossby radius of deformation, $l_d = (1 + \beta y)\sqrt{gh}/f$, or the viscous Munk length, $l_m = \nu/U$. Since the basis functions do not satisfy the no-slip boundary conditions of the flow, the projection is restricted to interior grid points only. Estimates of the small-scale features are therefore based only on the Rossby radius, 38–52 km. Such a scaling implies that the highest mode number, n_x or n_y, of the basis functions is $l_x/(2L_d) = l_y/(2L_d) = 28$. Thus, flow structures with scales as small as a Rossby diameter are resolved.

Based on the spatial scale only, a divergence-free approximation will entail projection on 784 modes. The inclusion of the divergent, potential modes doubles this number. Such a large number of modes presents two problems. First, we do not expect to have a sufficient number of observations (drifters) to adequately resolve over 1500 degrees of freedom at any time. Second, the computational effort required to minimize large least squares problems is prohibitive.



(a) Day one



(b) Day 49

FIG. 2. Model velocity fields at the beginning and end of the test reconstruction period.

Selecting an appropriate subset of modes for the projection initially requires some knowledge of the Eulerian field. As a first step toward identifying the most energetic basis functions, we project the Eulerian data obtained from 40 statistically independent time slices of the model flow taken once every 90 days for 10 years prior to the reconstruction period. This approach mimics the situation where the specifics of the Eulerian flow field are unknown during the reconstruction period, but some historical record of general flow characteristics may be obtained. The imperfect knowledge of the Eulerian field is further enhanced by using only one-fourth of the model grid points (i.e., $\Delta x = 20$ km) in the projection.

Figure 3 shows the energy distribution of the historical velocity fields projected on the full set of 784 + 784 basis functions. An attribute of this spectral technique is that the geometric properties of the basis functions are associated with phenomenological scales. For example, the most energetic mode, $\psi_{(1,2)}$, is a streamfunction with one-half sine wave in the x direction and a full sine wave in the y direction. This is exactly the geometry of the the basinwide double gyre. The color scale in Fig. 3 identifies modes that contain at least 1% of this energy. As expected, in this parameter range, where the dynamics are largely quasigeostrophic, the 2D divergence-free streamfunction modes contain the vast majority of the energy.

In addition to the dominant double-gyre mode, large streamfunction energies exist in modes $\psi_{(2,2)}$–$\psi_{(8,2)}$ and $\psi_{(6,1)}$–$\psi_{(8,1)}$. The geometric structure of these modes indicate that they represent the westward intensification of the gyre structure.

Similarly, the most energetic velocity potential modes identify 2D divergent effects associated with the double-gyre Ekman pumping, $\phi_{(1,2)}$; the ageostrophic components of the jet, $\phi_{(2,6)}$ and $\phi_{(2,8)}$; and the westward intensification, $\phi_{(4,2)}$–$\phi_{(9,2)}$ and $\phi_{(6,1)}$–$\phi_{(8,1)}$.

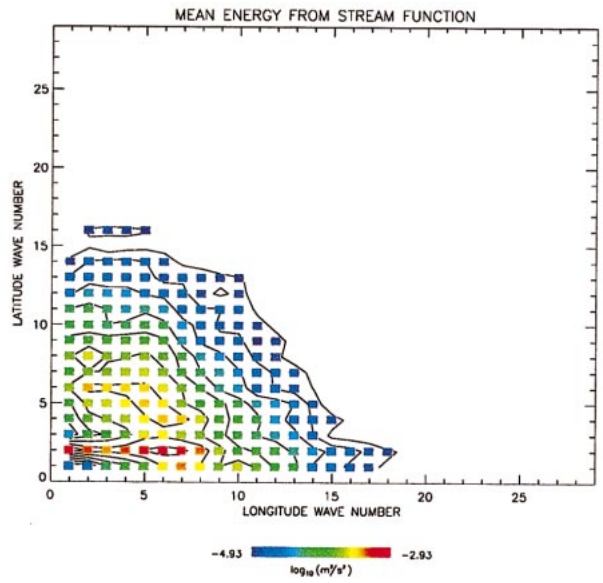
Note the different pattern of energy distribution in the wavenumber plane between the streamfunction and velocity potential modes. The velocity potential modes contain negligible energy in symmetric flow structures, such as eddies.

Reconstruction error relative to the model truth on an interior subdomain D arises when fewer degrees of freedom exist in the basis function set than in the model field. This will always be the case. To quantify the reconstruction error for the test problem, we choose the kinetic energy of the difference vector normalized by the energy of the model field:

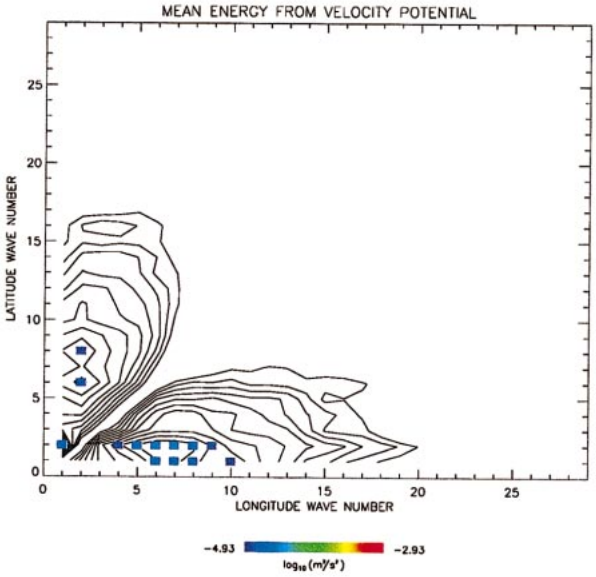
$$E(D, t) = \frac{\int_D (\mathbf{u}(\mathbf{t})_{\text{model}} - \mathbf{u}(\mathbf{t})_{\text{proj}})^2 \, dA}{\int_D (\mathbf{u}(\mathbf{t})_{\text{model}})^2 \, dA}, \qquad (14)$$

where D is the region (10 km $\leq x \leq$ 1000 km, 200 km $\leq y \leq$ 1600 km). The temporal mean of the error over the 50-day test period will be denoted $\overline{E(D)}$.

Figure 4 shows the falloff of $\overline{E(D)}$ with increasing numbers of modes. All 198 $\times$ 198 interior grid points



(a) Streamfunction modes



(b) Velocity potential modes

FIG. 3. Mean energy distribution in wavenumber space for the streamfunction and velocity potential modes computed by projecting a 10-yr model history on 1568 basis functions. Interior grid points were uniformly observed at a spatial resolution of 20 km, sampled every 90 model days. Modes containing 1% or more of the energy of the most energetic mode are shown in color on a log-energy scale.

were used in each projection. Selection criteria of the mode subsets is based on the 10-yr historical energy content of the basis functions. Table 1 lists the specifics of each mode subset.

The largest subset includes all modes from Fig. 3 containing energy within 1% of the double-gyre mode $\psi_{(1,2)}$. Of the 191 modes in this subset, 13 are velocity potential. The smallest subset has 65 streamfunction and no velocity potential modes.

Note the change in slope of the reconstruction error curve in Fig. 4 at the 107 mode subset. After this subset, velocity potential modes are included in the reconstruction. Between 65 and 107 modes the error is reduced by 0.33% per mode, while between 107 and 191 modes, error is decreased by 0.12% per mode. The decreasing rate in error reduction indicates the velocity field may be represented with relatively few modes.

The global energy captured for each reconstruction listed in Table 1 is computed by summing the square of the eigenfunction amplitudes (9) during the reconstruction period and normalizing by the kinetic energy of the model. The reconstruction error $\overline{E(D)}$ is very close to the uncaptured global energy of the reconstructions.

Mode selection for this model may be summarized as follows: Analytic basis functions provide arbitrary spatial resolution, which in theory could reproduce the exact model velocity on the computational grid. Physical scales, such as the Rossby radius, provide a practical minimum spatial resolution scale. Energetics from sampling the 10-yr historical model flow reduces the number of basis functions further. Consistent with the physics of the model, the streamfunction modes contain most of the energy.

5. Basin Eulerian reconstructions from drifter data

Our goal is to reconstruct the model Eulerian velocity field from a limited amount of simulated Lagrangian drifter data. Given the topology of the model double gyre, we focus on accurately representing the energetic western region identified in Fig. 1. Large velocities in this region exacerbate the difference between Lagrangian and Eulerian data.

The least squares reconstruction requires that the matrix problem is overdetermined; the number of observations must be greater than the number of unknown modal amplitudes. Since we wish to investigate the reconstruction error associated with minimal numbers of drifter observations, the choice of a particular normal mode analysis (NMA) basis set for the drifter experiments is dictated by the conflicting requirements of a low base reconstruction error and a limited number of modes (see Fig. 4). As a compromise, we chose the 107 mode basis set consisting only of 2D divergence-free streamfunction modes. This basis set provides quasi-geostrophic approximation to the primitive equation

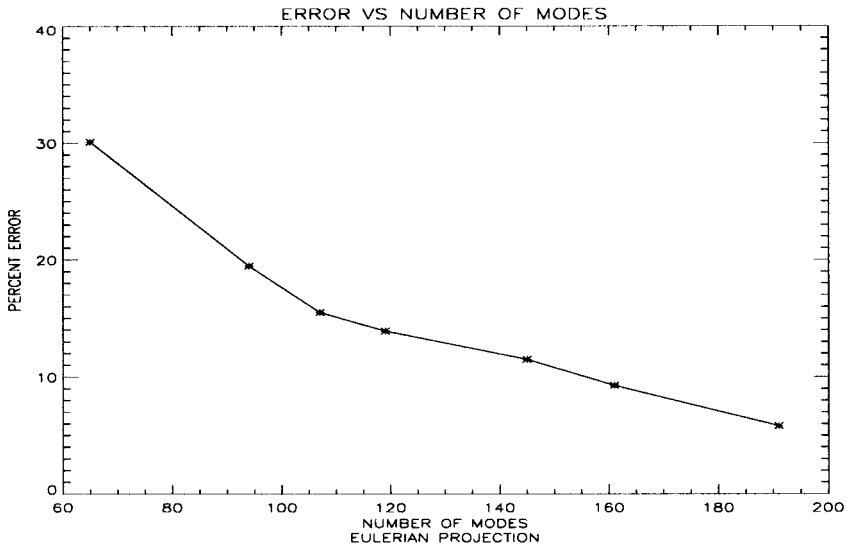


FIG. 4. Mean reconstruction error $\overline{E(D)}$ as a function of the number of modes in the projection. All interior grid points were used in each projection.

model with a base error, using the entire Eulerian model field as the observation vector, of 15%.

Figure 5 illustrates the distribution of energy in wavenumber space for the 50-day reconstruction period. The flow is more anisotropic than the 10-yr historical sample indicates in Fig. 3. For example, Fig. 3 shows a smooth distribution of energy between $\psi_{(1,2)}$ and $\psi_{(8,2)}$ while Fig. 5 exhibits a local energy peak at $\psi_{(9,2)}$. Two energy bands centered at $\psi_{(2,8)}$ and $\psi_{(2,6)}$ appear in Fig. 5 but not in Fig. 3.

The anisotropic features exhibited during the 50-day reconstruction period are additional indicators to assess the quality of reconstruction from drifter data. The energy distribution in Fig. 5 will serve as a reference to compare specific reconstructions.

Drifter data were obtained by initializing fluid tracers in the region D and computing Lagrangian trajectories using third-order polynomial spatial and linear temporal interpolation of the model velocity fields with standard fourth-order Runge–Kutta integration at 2.4-h time steps. As such, these simulated drifters approximate fluid particles up to the accuracy of the numerical scheme.

TABLE 1. Details of the mode subsets with full observation of the Eulerian field.

Energy of $\psi_{(1,2)}$ in the least energetic mode (%)	Streamfunction/ velocity potential modes	Global energy captured (%)	Reconstruction error $\overline{E(D)}$ (%)
1.0	178/13	95	5.8
1.5	153/8	92	9.3
2.0	140/5	89	12
3.0	117/2	87	14
4.0	107/0	85	16
5.0	94/0	81	20
10	65/0	70	30

Drifter velocities and position are stored once per simulation day.

The drifter data do not adequately sample the quiescent region outside D. Although the reconstruction error does not measure the error outside D, the basis functions require observations there to avoid unrealistic global behavior. Thus, the drifter observations are augmented with a limited number of uniformly spaced Eulerian observations outside the region D taken from the 10-yr historic mean model field. Spacing of the uniform climatological sample is chosen to prevent data voids

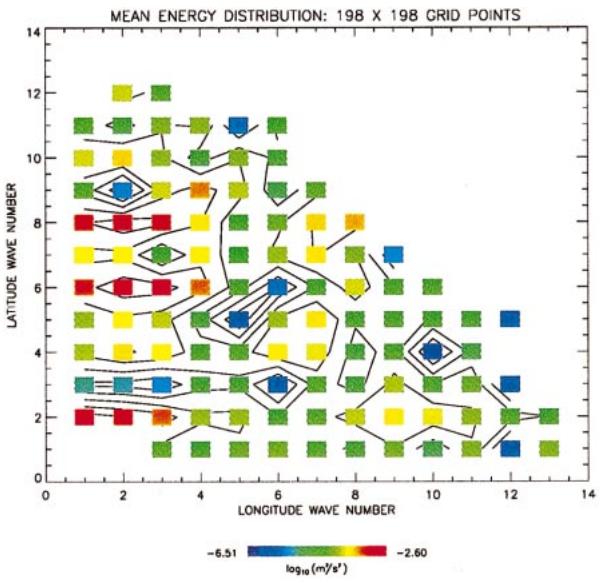


FIG. 5. Mean energy distribution in wavenumber space for the 107 divergence-free basis functions during the reconstruction period using all 198 × 198 interior grid points as an observation vector.

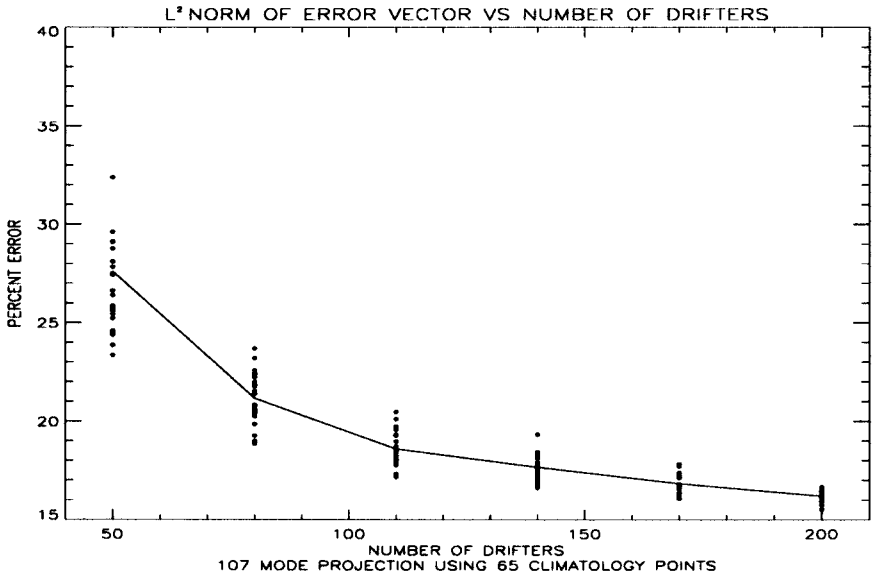


FIG. 6. Reconstruction error vs number of Lagrangian observations, N_{drift}. For each value of N_{drift}, 25 different random initial deployments are shown. The average error is given by the solid line.

larger than the smallest scale basis function. Such data voids cause undersampling that produces spurious flow structures in the reconstruction, a phenomenon referred to as *noodling*. The smaller structures present in the 107 modes chosen have spatial scales ranging from 250 km in the x and y directions ($\psi_{(8,8)}$) to 154 km in the x direction ($\psi_{(13,2)}$). A uniform sample of 200 km in the quiescent region gives 65 sample points and is adequate to prevent unrealistic behavior. The reconstruction error is not sensitive to the climatological sampling strategy. The drifter and climatological velocities together make the observation vector (8).

The relation between $\overline{E(D)}$ and the number of drifters deployed is shown in Fig. 6 using random launch locations of 50 to 200 drifters. For each fixed number of drifters, 25 experiments are conducted and the ensemble average reconstruction error is shown by the solid line. The average error is inversely proportional to the number of observations, N_{drift}, and asymptotes to the base error computed using the full Eulerian dataset as input. Notably, velocity observations from 200 randomly launched drifters can provide reconstructions within 1% of the base error.

The asymptotic approach of reconstruction error to the base error as N_{drift} increases under random deployments is verified for all mode subsets listed in Table 1. The saturation level depends on the number of modes used, but does not exceed 2% of the model grid points in the subdomain D.

Two important issues are unresolved by Fig. 6. Since the drifter deployment for each reconstruction in an ensemble is random, data voids were minimized as N_{drift} increased. Does launching a significant number of drifters guarantee an accurate reconstruction regardless of the deployment strategy? Does the error increase or decrease as the drifters leave the random launch locations? The next section attempts to quantify the first issue, while the following section examines the second.

a. Drifter coverage

There exists a considerable amount of variability in the error in experiments using a fixed number of drifters. This variability increases dramatically as fewer drifters are deployed. For example, as seen in Fig. 6, the error spread in the reconstruction for experiments using 80 drifters overlaps that of experiments that use 140 drifters. The reconstructed fields are sensitive not only to the number of Lagrangian observations but also to the quality of these observations.

Poje and Haller (1999) show how the model flow is governed by geometrical aspects of the invariant manifold structure. In particular, the initial launch location will to a large extent determine the dispersion of the fluid particles. Without detailed knowledge of the flow specifics, general physical principles may be employed to estimate the dynamic scales. Drifters located within a Rossby diameter of each other should provide redundant information for the 107 mode reconstruction, effectively reducing the number of independent observations.

To quantify coverage of the domain, we introduce the concept of an active cell. Let C_i with $i = 1, \cdots, N_{\text{cell}}$ be a partition of the subdomain D such that $D = \cup_{i=1}^{N_{\text{cell}}} C_i$ and $C_i \cap C_j = \varnothing$. A cell C_i is said to be active if at least one drifter is located in the cell. Denote a set of drifters by $\mathbf{x}_j(t)$ with $j = 1, \cdots, N_{\text{drift}}$ and define

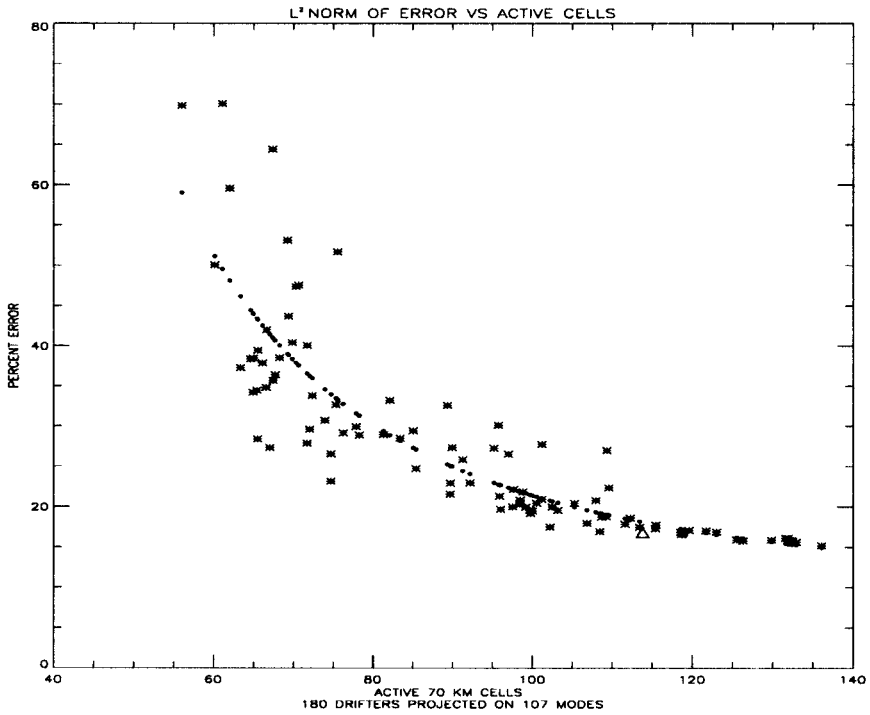


FIG. 7. Reconstruction error $\overline{E(D)}$ vs coverage $\overline{\alpha}$. The linear correlation between $\overline{E(D)}$ and $1/\overline{\alpha}^{2.6}$ (dotted curve) is $r = 0.87$. One hundred experiments with initial coverage ranging from 15 to 180 randomly chosen cells generate the pattern. A uniform initial distribution experiment, which also occupies 180 cells, is shown (triangle) for comparison.

$$A\left(C_i, \bigcup_{j=1}^{N_{\text{drift}}} \mathbf{x}_j(t)\right) = \begin{cases} 1, & \text{if } C_i \cap \bigcup_{j=1}^{N_{\text{drift}}} \mathbf{x}_j(t) \neq \varnothing \\ 0, & \text{otherwise.} \end{cases} \tag{15}$$

The time-dependent coverage function is then

$$\alpha(t) = \sum_{i=1}^{N_{\text{cell}}} A\left(C_i, \bigcup_{j=1}^{N_{\text{drift}}} \mathbf{x}_j(t)\right). \tag{16}$$

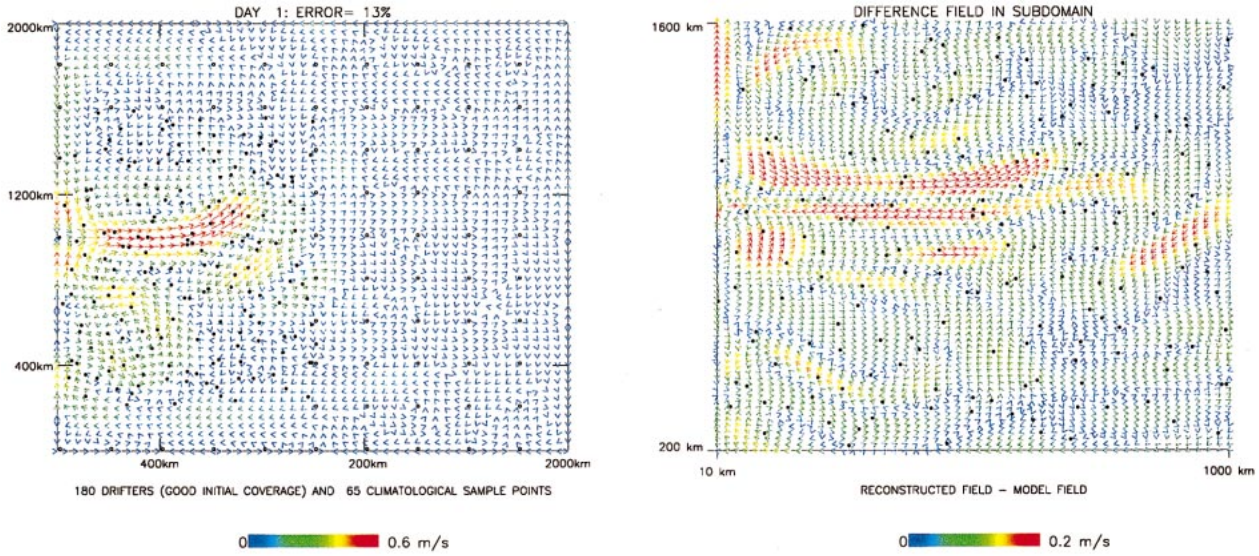
The size and shape of the cell delineates what is considered a redundant observation; that is, if two drifters are located within a cell, their observations are considered redundant. Deployments maintaining a large number of active cells minimize redundant observations.

Specifying the optimal decomposition C_i is not a trivial task. Ideally, a full knowledge of the spatial decorrelation function is needed. However, even with perfect knowledge of the flow field, this is a computationally intensive task. Since we know the physics of the flow more precisely than the spatial decorrelation function, a simple decomposition of D into 70-km square cells is used. This spatial scale represents twice the smallest Rossby radius of the flow.

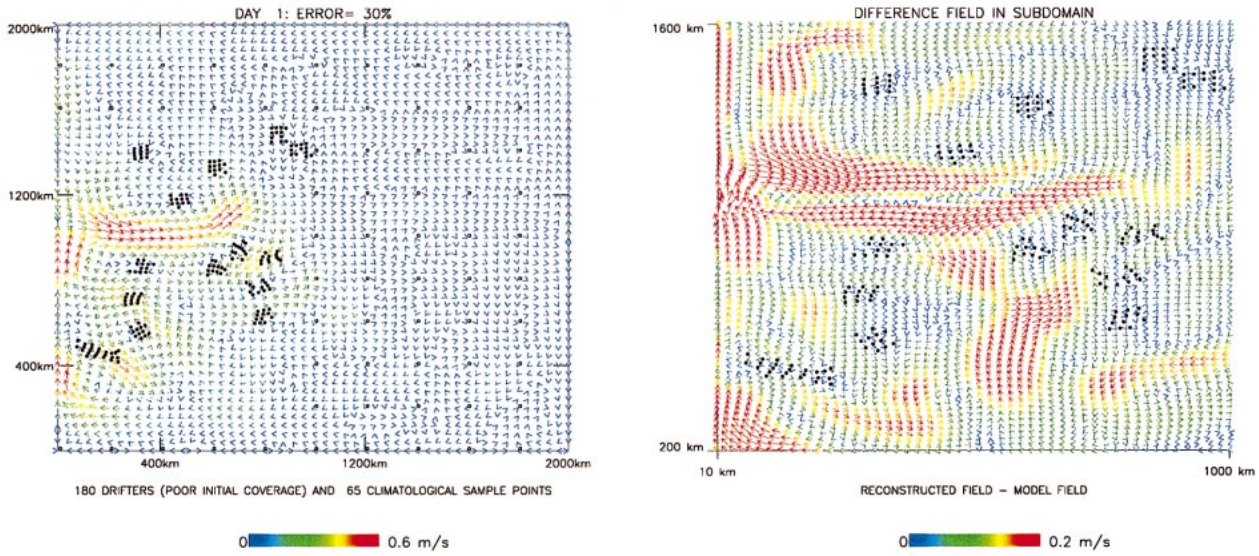
The relationship between coverage and reconstruction error is studied using 180 drifters. Random launch locations are chosen for 100 experiments with initial coverage ranging from 180 to 15 active cells out of 285 possible cells in the domain. Figure 7 shows the relationship between $\overline{E(D)}$ and the mean coverage $\overline{\alpha}$. For reference to the results obtained from the random distribution of initial cells, an experiment with a uniform distribution of 180 drifters throughout the domain is performed. One drifter is deployed approximately every 83 km in the x direction and 93 km in the y, so the initial coverage is 180 cells. The mean error and coverage properties of the uniform deployment fit the general pattern of the other, random deployments. Note that uniform initial distribution does not give the highest mean coverage.

A standard least squares approximation to a geometric model gives $\overline{E(D)} \propto 1/\overline{\alpha}^{2.6}$. The linear correlation coefficient between $\overline{E(D)}$ and $1/\overline{\alpha}^{2.6}$ is $r = 0.87$, indicating that the simple domain decomposition provides a reasonable indication of coverage. Figure 7 indicates that high coverage implies a low reconstruction error. However, for low coverage experiments, a wide range of reconstruction error exists for the same mean coverage. A more detailed, spatially dependent decomposition would improve the correlation.

The asymptotic relation between error and coverage shown in Fig. 7 is robust in that modifying the cell size, the number of drifters, or the number of modes does not alter the results appreciably. For all mode subsets listed in Table 1, the reconstruction error $\overline{E(D)}$ asymptotes toward the base error that is achieved with full observation of the Eulerian velocity field as coverage increases.



(a) Good coverage

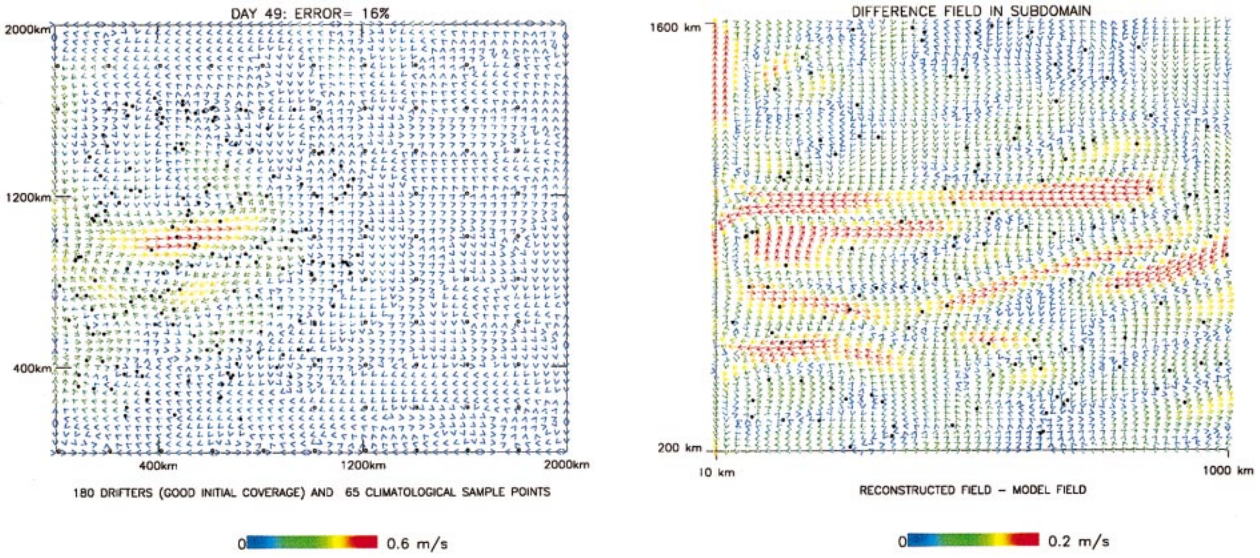


(b) Poor coverage

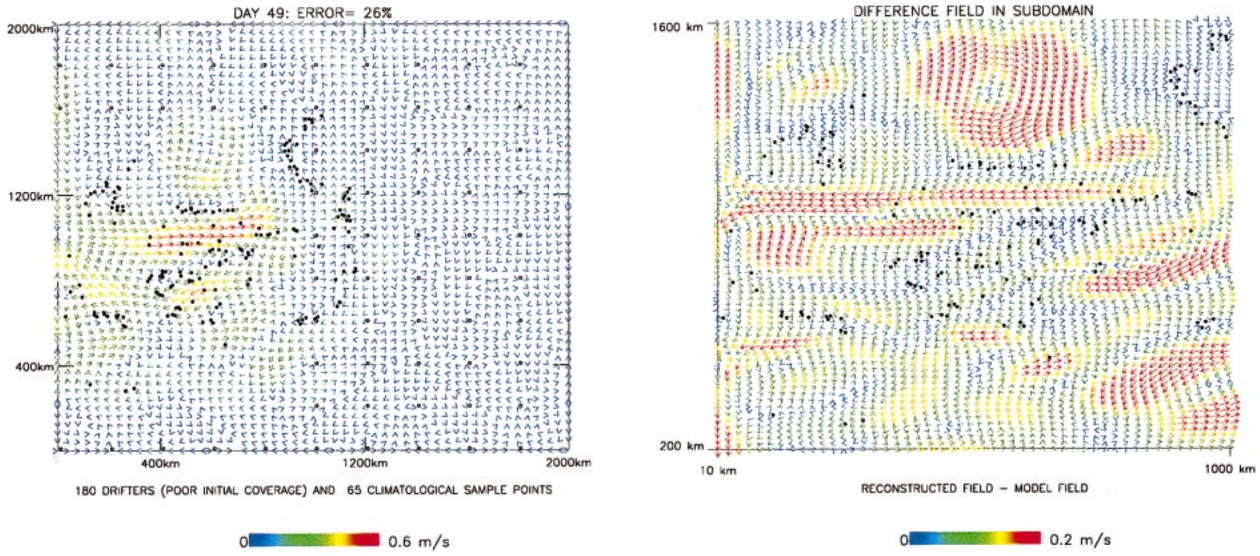
FIG. 8. A comparison between reconstructions from good and poor drifter coverage at day 1.

Figures 8 and 9 compare two reconstructions from the 100 experiments at day 1 and day 49. The *good coverage* experiment launches drifters in 180 cells; the *poor coverage* experiment launches the drifters in 15 active cells.

Day 1 of the reconstruction in Fig. 8 shows noodling in the poor coverage reconstruction. A comparison with Fig. 2 reveals incorrect velocity peaks on the western boundary at 400 and 1000 km and a false eddy in the western domain south of 400 km. The good coverage reconstruction does not exhibit this structure. Details of the difference field in the subdomain D is also shown in Fig. 8. Note that the color scale on the difference field has a maximum of 0.2 m s^{-1} while the color scale



(a) Good coverage



(b) Poor coverage

FIG. 9. A comparison between reconstructions from good and poor drifter coverage at day 49.

on the model and reconstructed fields has a maximum of 0.6 m s^{-1}. Similar patterns exist in the difference field of the good and poor coverage reconstructions, although the strength of the patterns differ. For example, the southern part of the central jet in both fields is underestimated, while the northern part is overestimated. On the northern part of the western boundary of *D*, the western boundary current is underestimated in both fields. The poor coverage reconstruction, however, exhibits structures with significant energy on the central and southern part of the western boundary and in the southern region of *D*.

Day 49 of the reconstruction in Fig. 9 shows the noodling that occurs in the poor coverage reconstruction has moved to the regions north and southwest of the central jet with artificial eddies in the velocity field. Error near the central and southern part of the western boundary in this reconstruction, however, is reduced from that of day 1. Both good and poor coverage reconstructions underestimate the strength and length of the central jet. The difference fields do not exhibit the overestimation in strength of the northern section of the jet as occurred on day 1. Both difference fields have similar patterns in the region south of the central jet, on the northern part of the western boundary of D, and in the region extending from the central portion of the eastern boundary of D.

b. Moorings versus drifters

The present implementation of the NMA technique treats temporally evolving Lagrangian trajectory information as variable position Eulerian measuring stations. Aspects of the flow are manifested through dispersion and the resulting change in coverage. The strong relation between coverage and reconstruction error indicate that the temporal evolution of $E(D, t)$ should be very different for drifters and moorings.

To examine these differences we chose two different initial sets of launch locations corresponding to good and poor spatial coverage. Velocity observations are then taken at subsequent times from these fixed positions (moorings) and along the Lagrangian trajectories (drifters) emanating from these initial conditions. The reconstructed velocity fields are thus identical at the initial time.

Figure 10 indicates the time dependence of $E(D, t)$ computed for each set of 180 moorings or drifters for both good and poor coverage. For reference, $E(D, t)$ is computed from the reconstruction using all 198×198 interior grid points as mooring observations. The time interval is extended to 120 days so the effects of drifter dispersion can be shown.

Note that the reconstruction error is time dependent when fixed mooring data or the entire Eulerian field is used as observations. In contrast to linear flow, where the transient solution decays, the nonlinear double-gyre flow does not approach a steady solution. Instead, there is a time-dependent flux of energy across spectral components as the jet meanders and when eddies are shed, coalesce, and decay.

The good initial coverage experiment has 180 active cells at launch. This coverage can only go down as the drifters advect. The moorings, however, maintain the 180 active cells throughout the experiment. As expected, $E(D, t)$ increases in the drifter experiment and remains low in the mooring experiment. It is interesting that the well-placed moorings have a lower reconstruction error than the reference reconstruction using all 198×198 grid points. Since the chosen subdomain D is well covered in the mooring experiment and only 65 climatology samples are taken outside D, the least squares minimization weighs the drifters almost 3 to 1. The 198×198 observations use no climatology, but the quiescent region is almost three times larger than the subdomain D. As a result, the reference reconstruction has a slightly higher local error $E(D, t)$ due to the high number of observations in the quiescent region.

In the case of poor initial coverage, the opposite effect is observed. The dispersion of Lagrangian particles increases the spatial coverage of the drifter observations in time while the coverage of the moorings remains poor. After 120 days, the errors from both the good and poor initial coverage drifter experiments are near 20%. Unlike the temporal evolution in reconstruction error exhibited by the mooring experiments, data voids generated and removed by drifter dispersion dominate the reconstruction error for both good and poor initial coverage drifter experiments.

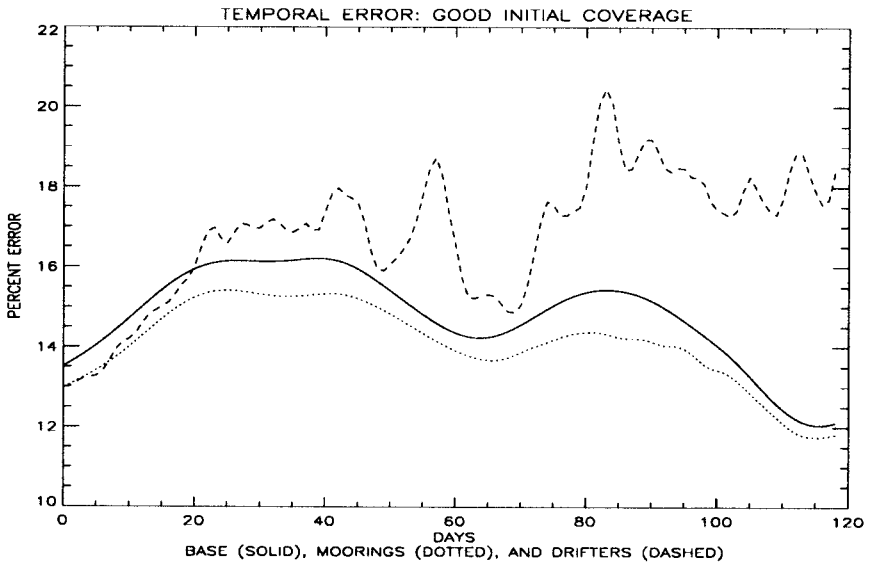
Figure 11 compares the energy distribution of the reconstructions for the first 50 days. The good coverage experiments distribute energy similar to the base projection in Fig. 5. Differences occur in some of the higher wavenumbers. For example, both the drifter and mooring reconstruction show an energy peak at $\psi_{(2,12)}$ not present in Fig. 5.

Noodling in the poor coverage experiments significantly alters the energy distribution in the wavenumber plane. The distinct energy minima at $\psi_{(2,9)}$ and $\psi_{(10,4)}$ captured in the good coverage experiments do not appear with poor coverage.

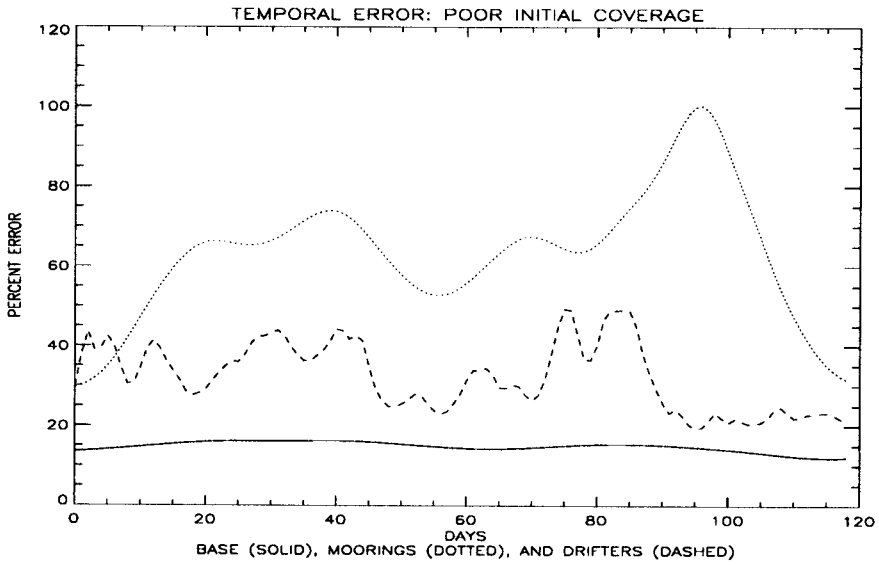
The poor drifter and mooring reconstructions differ appreciably in the energy distribution because the coverage and noodling varies temporally with the drifter trajectories. Conversely, the good drifter and mooring reconstructions are quite similar since noodling is minimal.

6. Conclusions

This study has addressed some of the issues involved in utilizing Lagrangian data. The approach relied on the orthogonal reduction technique of normal mode analysis. The analysis quantified the accuracy of the reconstruction in relation to the number and coverage of observations, and number of spectral modes. The coverage results are generic in that the same pattern of error reduction with improved coverage exists for many different drifter–mode combinations. In this study, we found the Eulerian flow field can be accurately reconstructed with data representing approximately 1.5% of the 12 600 model grid points in the subdomain of interest. The error of the reconstructed velocity fields asymptotes to the base error (achieved with full observation of the Eulerian field) in two independent variables: coverage (with an adequate number of drifters) and number of drifters (with adequate coverage). As a result, with adequate coverage and data, the reconstruc-



(a) Good coverage



(b) Poor coverage

FIG. 10. The time dependence of the error $E(D, t)$ from good or poor coverage reconstructions using 180 drifters or moorings and 107 modes. Moorings are located at the initial position of the drifters. The base reconstruction error is included for comparison.

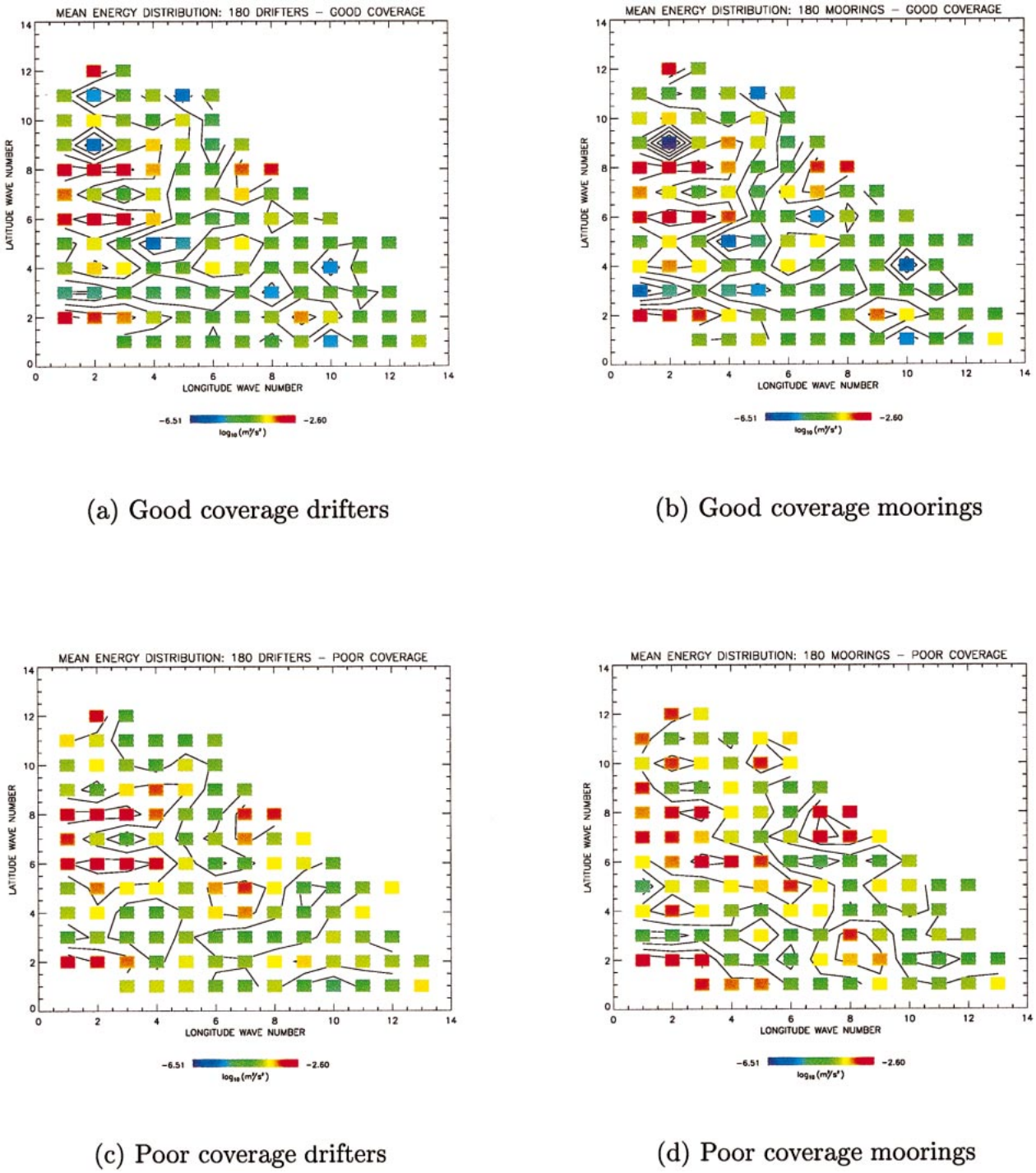


FIG. 11. Mean energy distribution in wavenumber space of the 107 divergence-free basis functions from drifter or mooring reconstructions.

tion error can be improved by increasing the number of modes. The velocity potential modes are crucial to obtaining reconstruction errors under 10% for the double-gyre flow. Clearly, different real or simulated flow fields would require more or less observations to accurately reconstruct the varying importance of the velocity potential modes.

The number of modes used in a particular reconstruction is dictated by several factors. Irregular domains require numerical solutions to the Helmholtz equations,

thus imposing a grid resolution on the basis functions. Additionally, there is a limit on the spatial resolution due to the maximum number of eigenfunctions that may be resolved numerically. Energy distribution of a projection may still be analyzed with numerical basis functions. However, the two-dimensional wave number space is reduced to a one-dimensional eigenvalue space and the geometric associations with flow structures is less obvious.

Knowledge of the 10-yr history of the flow was used to reduce the number of basis functions. The controlled nature of this study, involving analytic basis functions and access to the ''true'' flow field, permitted such an analysis. For basin-scale ocean flows where such exact information is not available, remotely sensed altimeter data and general circulation models could be used to study the mode energetics and, if needed, reduce the number of basis functions.

One limitation of the technique is noodling. This arises when data voids are of comparable size to the spatial scale of the highest wave-numbered basis function. This was prevented from occurring outside the active subdomain, approximately two-thirds of the basin, by a uniform 200-km sample of the static climatology. We chose not to fill voids in the active flow region to identify the consequent error. In practice, appropriate data filling techniques should be employed when data voids of this type occur.

Deployment of drifters with the best possible initial coverage on a basin-scale domain is not practical. Poor initial coverage experiments, while much more practical, vary widely in the reconstruction accuracy due to the heterogeneous dispersive characteristics of the flow. Launch strategies based on some knowledge of the flow should improve the coverage, given a fixed number of deployment locations.

The effect drifter dispersion has on reconstruction error is an important consideration. As illustrated by the drifter versus mooring analysis, poor initial coverage can, after adequate dispersion occurs, still give reasonable reconstruction errors. Conversely, good initial coverage is degraded by dispersion. These results indicate a need for time-dependent launch strategies. Combined with other data sources, this would alleviate data voids.

Treshnikov et al. (1986) proposed a drifter launch scheme based on the reconstruction of the surface pressure field and the energy of the currents. An alternative approach would be to use dynamical systems techniques, such as described for this flow by Poje and Haller (1999), to identify potential data voids.

This study did not make use of the a priori knowledge of model flow features. Approximating the location of hyperbolic points, for example, might give launch sites that disperse drifters quickly. Future work should address approximating such dynamical features in the ocean.

Assessing the quality of reconstruction using the difference vector field energy, an Eulerian metric provides no assessment of Lagrangian features of the reconstructed flow. How the Lagrangian dynamics of the reconstructed field compare to the original remains unanswered.

The reconstructed flow fields have a wide range of application. Nowcasts, model improvement, and rapid environmental assessment are obvious candidates. More subtle applications include drifter deployment strategies and manifold approximation.

Acknowledgments. This research was supported by the Office of Naval Research Grants N00014-99-1-0054, N00014-99-1-0052, and N00014-92-J-1481.

REFERENCES

Berloff, P., and S. P. Meacham, 1998: On the stability of the wind-driven circulation. *J. Mar. Res.*, **56,** 937–993.

Cho, K., R. O. Reid, and W. D. Nowlin Jr., 1998: Objectively mapped stream function fields on the Texas–Louisiana shelf based on 32 months of moored current meter data. *J. Geophys. Res.,* **103,** 10 377–10 390.

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

——, 1985b: Objective mapping by least squares fitting. *J. Geophys. Res.,* **90,** 4773–4777.

Eremeev, V. N., L. M. Ivanov, and A. D. Kirwan Jr., 1992a: Reconstruction of oceanic flow characteristics from quasi-Lagrangian data: 1. Approach and mathematical-methods. *J. Geophys. Res., Oceans,* **97,** 9733–9742.

——, ——, and ——, 1992b: Reconstruction of oceanic flow characteristics from quasi-Lagrangian data: 2 Characteristics of the large-scale circulation in the Black Sea. *J. Geophys. Res., Oceans,* **97,** 9743–9753.

——, ——, ——, and T. M. Margolina, 1995a: Amount of ^{137}Cs and ^{134}Cs radionuclides in the Black Sea produced by the Cherboyl accident. *J. Environ. Radioact.,* **27,** 49–63.

——, ——, ——, and ——, 1995b: Analysis of caesium pollution in the Black Sea by regularization methods. *Mar. Pollut. Bull.,* **7,** 460–462.

Figueroa, H. A., 1994: Eddy resolution versus eddy diffusion in a double gyre GCM. Part II: Mixing of passive tracers. *J. Phys. Oceanogr.,* **24,** 387–402.

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

Forristall, G. Z., K. J. Schaudt, and C. K. Cooper, 1992: Evolution and kinematics of a loop current eddy in the Gulf of Mexico. *J. Geophys. Res.,* **97,** 2173–2184.

Glenn, S. M., and C. C. Ebbesmayer, 1993: Drifting buoy observations of a loop current anticyclonic eddy. *J. Geophys. Res.,* **98,** 20 105–20 119.

Golub, G. H., and C. F. van Loan, 1983: *Matrix Computations.* The Johns Hopkins University Press, 476 pp.

Holland, W. R., 1978: The role of mesoscale eddies in the general circulation of the ocean—Numerical experiments using a wind-driven quasi-geostrophic model. *J. Phys. Oceanogr.,* **8,** 363–392.

Jiang, S., F.-F. Jin, and M. Ghil, 1995: Multiple equilibria, periodic, and aperiodic solutions in a wind-driven, double-gyre, shallow-water model. *J. Phys. Oceanogr.,* **25,** 38–60.

Jones, D. A., A. C. Poje, and L. G. Margolin, 1997: Resolution effects and enslaved finite-dfference schemes for a double-gyre, shallow water model. *Theor. Comput. Fluid Dyn.,* **9,** 269–280.

Lewis, J. K., A. D. Kirwan Jr., and G. Z. Forristall, 1989: Evolution of a warm-core ring in the Gulf of Mexico—Lagrangian observations. *J. Geophys. Res.,* **94,** 8163–8178.

Lipphardt, B. L., A. D. Kirwan Jr., C. E. Grosch, L. M. Ivanov, and J. K. Lewis, 1998: Merging disparate oceanographic data. *Rapid Environmental Assessment,* NATO SACLANT Undersea Research Centre, 211–218.

——, ——, ——, J. K. Lewis, and J. D. Paduan, 2000: Blending HF radar and model velocities in Monterey Bay through normal mode analysis. *J. Geophys. Res.,* **105** (C2), 3425–3450.

Marshall, J., 1984: Eddy-mean flow interatcion in a barotropic ocean model. *Quart. J. Roy. Meteor. Soc.,* **110,** 573–590.

McCalpin, J. D., 1995: The statistics and sensitivity of a double-gyre model: The reduced gravity, quasigeostrophic case. *J. Phys. Oceanogr.,* **25,** 806–824.

——, and S. C. Haidvogel, 1996: Phenomenology of the low-frequency variability in a reduced-gravity, quasigeostrophic double-gyre model. *J. Phys. Oceanogr.,* **26,** 739–752.

Owens, W. B., 1991: A statistical description of the mean circulation and eddy variability in the nowthwestern atlantic using sofar floats. *Progress in Oceanography,* Vol. 28, Pergamon, 257–303.

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

Rao, D. B., and D. J. Schwab, 1981: A method of objective analysis for currents in a lake with applications to Lake Ontario. *J. Phys. Oceanogr.,* **11,** 739–750.

Schulz, W. J., 1999: Ocean surface maps from blending disparate data through normal mode analysis. Ph.D. thesis, Center for Coastal Physical Oceanography Tech. Rep. No. 99-02, Old Dominion University, 102 pp.

Smolarkiewicz, P. K., and L. G. Margolin, 1993: On forward in time differencing for fluids. *Mon. Wea. Rev.,* **121,** 1849–1859.

Sombardier, M. L., 1992: Surface current trackers of the world's oceans—WOCE/TOGA. *Oceanus,* **35,** 6–8.

——, and P. P. Niiler, 1994: Global surface circulation measured by Lagrangian drifters. *Sea Technol.,* **35,** 21–24.

Treshnikov, A. F., V. V. Guretskiy, A. I. Danilov, N. V. Yefremeyev, L. M. Ivanov, and V. M. Smelyanskiy, 1986: Optimal array of satellite-tracked drifting buoys in the Southern Alantic. *Dokl. Akad. Nauk SSSR,* **287,** 430–434.

Vidal, V. M. V., F. V. Vidal, and J. M. Perezmolero, 1992: Collision of a loop current anticyclonic ring against the continental-shelf slope of the western Gulf of Mexico. *J. Geophys. Res.,* **97,** 2155–2172.