Data Constraints Applied to Models of the Ocean General Circulation. Part II: The Transient, Eddy-Resolving Case

PAOLA MALANOTTE-RIZZOLI

Department of Earth, Atmospheric and Planetary Sciences, Center for Meteorology and Physical Oceanography, Massachusetts Institute of Technology, Cambridge, Massachusetts

WILLIAM R. HOLLAND

National Center for Atmospheric Research, Boulder, Colorado

(Manuscript received 1 September 1987, in final form 28 December 1987)

ABSTRACT

In Part I of the present work we performed assimilation experiments with a multilayer, quasi-geostrophic (QG) eddy-resolving model of the ocean general circulation. In Part I we studied the quasi-linear, steady state and the assimilated data were density measured along hydrographic sections. The major result of this study was that the most effective sections are long, meridional ones located at distance from the western boundary. The model estimates are significantly improved over the entire region extending from the data section to the western boundary itself.

In this second part we extend the study to the more realistic time-dependent, fully eddy-resolving ocean. Again we capitalize upon the two assumptions that the available models are imperfect and that data are measured only locally at meridional sections. The location of the sections are chosen according to (i) distance from the western boundary; (ii) energetics of the region. Also, here we compare assimilation of density alone versus density and velocity.

A crucial problem emerges when assimilating data into a fully nonlinear, time-dependent model, that is the problem of model predictability. The assimilated data can in fact be viewed as "perturbations" introduced into the model at a specific location. The important question is then: is data insertion performed only locally, i.e., along sections, sufficient to "drive" the model to the reference ocean overcoming the model inherent loss of predictability?

Different data sections are compared and the model performance is quantified monitoring two global rms (root mean square) errors, the rms DIFF1 between the model with inserted data and the reference ocean and the rms DIFF2 between the model with inserted data and without.

Two major results emerge from the present study. First, and differently from the quasi-linear steady case, a single data section is very ineffective in driving the model towards the reference ocean over time scales of ~ 100 days, comparable with the time scale of predictability loss. The rms-error DIFF2 is used to quantify the effectiveness of the different section as the "true" rms-error DIFF1 exhibits only random fluctuations around a mean equilibration value. The overall error level depends upon the balance between criteria (i) and (ii) above. Results are rationalized by dynamical considerations showing that the internal boundary forcing provided by the data insertion is equivalent to an additional stress-curl (vorticity source) imposed impulsively along a line in each layer. Also, the assimilation of barotropic and baroclinic information versus baroclinic only (velocity and density versus density only) has no effect on the error levels and error growth rates on the short time scale of mesoscale variability. In general, the error growth rates are not significantly different for any of the considered sections, both for the global rms errors measured over the entire basin and for local rms-errors measured over localized regions. On the short time scale of mesoscale variability, all the considered sections are equally ineffective.

A single section of data is shown instead to be quite effective in driving the model to the reference ocean if the data insertion process is carried out for time durations longer than the model equilibration time. With ten years of data assimilation, the climatological mean of the model becomes extremely similar to the climatological mean of the reference ocean. This result can now be quantified using the "true" rms-error DIFF1, which exhibits an unambiguous decreasing trend during the last years of assimilation, thus improving the estimate of the climatology up to 25%. Thus, single hydrographic sections might still be useful in providing a better model climatology if time series of data were available longer than the model equilibration time.

1. Introduction

In a previous paper (Malanotte-Rizzoli and Holland 1986, hereafter referred to as Part I) a first step was

Corresponding author address: Dr. Paola M. Rizzoli, Dept. of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Room 54-1420, Cambridge, MA 02139.

taken in the process of assimilating data into models of the ocean general circulation. We made use of a multilayer, quasi-geostrophic (QG), fine resolution model capable of examining fully turbulent, eddy resolved ocean circulation on the basin scale. In the study of Part I, however, the model parameters were chosen in such a way that the time-dependent model spins up to a steady-state. This steady state is weakly nonlinear and highly frictional. Strong vertical friction plays the role of eddy fluxes in driving the circulation in the deep layers. In Part I we purposedly chose to treat the steady, quasi-linear case to be able to make use of simple analytical calculations for understanding the dynamics governing the process of data insertion.

Three major questions were addressed and answered in Part I. The chosen data (simulated by the model itself in a "reference" experiment) were density (or temperature) measurements along long hydrographic or tomographic sections. We varied the location of the section as well as its orientation. The major findings were that, in the highly frictional ocean, the most effective sections are long meridional ones located at a distance from the western boundary. The model estimates were then significantly improved over the broad region extending from the data section to the western boundary itself. Also, even though the chosen ocean is quasi-linear (advective effects are minimal), the nonlinearity associated with vortex stretching is crucial in determining the spreading of information from the data insertion region. It manifests itself through a quite important steering effect upon the spread of information for which in the subtropical (subpolar) gyre the region southwest (northwest) of the data section is consistently preferred for the improvement of estimates.

In this second part we extend the results of Part I and study the more complex situation of the time-dependent, fully eddy-resolving ocean. Explicit vertical viscosity is now small and eddy fluxes drive the deep circulation. The model never reaches a steady state. The "equivalent" of the steady circulation of Part I is now given by the "climatological" mean, that is the average over long time durations (of order 5 years) of the individual circulation "snapshots." The climatological mean is evaluated after the model has been spun-up, that is after it has reached stable statistics. In the steady case of Part I, information spreads along the characteristics of the steady state dynamical equations. The role played by the characteristics is now played by the transient Rossby eddies, both barotropic and baroclinic. It can be anticipated that this spreading will be much more effective when carried out by the long Rossby waves, westward propagating, than by the short Rossby waves reflected at the western boundary and quickly damped. Thus, the westward part of the ocean will be again preferred for this spreading.

In the experiments analyzed in this paper we capitalize again upon the two assumptions that (i) the available dynamical models are imperfect and (ii) oceanographic data are measured locally. Thus, we choose again data sections located in regions of the gyre dominated by different dynamics. Our ocean is again formed by the subtropical and subpolar gyres, driven by idealized double-gyre of steady winds, like in Part I. Using the results of Part I, however, we discuss only meridional data sections as the most effective in allowing westward propagation of information.

The approach followed to study the process of data insertion is like in Part I. We first construct the "reference" ocean E1, that is a baseline case with standard parameters (see Table 1) for a fully eddy-resolving double gyre circulation. The "imperfect" model is now constructed by assuming that our knowledge of the model physics is slightly wrong in all the parameterizations. Thus, the "model" run E2 is carried out by perturbing the amplitudes of all the model parameters, that is the biharmonic and bottom friction coefficients, A_4 and R, respectively; the reduced gravities $g'_{3/2}$, $g'_{5/2}$; the wind stress intensity T_0. Notice that the perturbation is applied to the wind stress amplitude; no phase perturbation is introduced. The majority of the experiments discussed in the sections to follow relate to a model ocean E2 which is "wrong" by 2% in all the parameters. One experiment will also be discussed in which the model ocean E2 is more strongly perturbed, with 5% error in all the parameters. Then, we measure data in the reference ocean E1 and we insert them along meridional sections of E2, rerunning the "imperfect" model with the internal forcing induced by the data insertion. These assimilation experiments are called J_n where n refers to different experiments.

The location of the data sections are chosen according to two criteria: (i) distance from the western boundary, thus allowing for a greater (or smaller) area over which long Rossby waves can propagate westward and (ii) energetics of the region, comparing for instance sections across the gyre Sverdrup interior with sections crossing the highly energetic, eddy-rich Gulf Stream System. Criteria (i) and (ii) allow a comparison between spreading of information due to linear processes (Rossby wave propagation) versus nonlinear processes (nonlinear advection by the Gulf Stream jet). Also, the

TABLE 1. Experimental parameters.

Coriolis parameter $f = f_0 + \beta y$	$f_0 = 9.3 \times 10^{-5}$ s^{-1} $\beta = 2 \times 10^{-11}$ m^{-1} s^{-1}
Layer thicknesses	$H_1 = 300$ m (surface layer) $H_2 = 700$ m (intermediate layer) $H_3 = 4000$ m (deep layer)
Reduced gravities	$g'_{3/2} = 0.0357$ (first interface) $g'_{5/2} = 0.0162$ (second interface)
Biharmonic friction	$A_4 = 8 \times 10^9$ m^4 s^{-1}
Bottom friction	$R = 5 \times 10^{-7}$ s^{-1}
Wind stress intensity	$T_0 = 1 \times 10^{-4}$ m^2 s^{-2}
Horizontal extent of the basin	L_x (east–west) = L_y (north–south) = 4000 km
Grid resolution	$\Delta_x = 20$ km

inserted data are of two types. As in Part I, we perform experiments in which only density (or temperature) is measured at the data sections, thus assimilating only the baroclinic modes, or pynocline displacements. Here we further perform experiments in which density and velocity are measured at the same sections, thus assimilating the baroclinic and barotropic mode. Thus, we can directly compare the effectiveness of velocity versus density information.

A crucial problem arises when performing assimilation experiments with a fully nonlinear, time-dependent model: the problem of model predictability. It is well known from meteorology that any error introduced in the initial state, no matter how small, will grow exponentially and change completely the model evolution from the evolution without the initial error. This loss of memory of the initial state, or loss of predictability, is characterized by a predictability time scale of the order of the time scale of mesoscale variability. This predictability time scale can be quantified as the e-folding time for the growth of the root mean square (rms) error. In the insertion experiments that follow the inserted data can be viewed as "perturbations" introduced at a specific location. Let the model be the operator L acting upon the vector $\boldsymbol{\phi}$ to produce the "predicted" values $\mathbf{d}_p$:

$$L\{\boldsymbol{\phi}\} = \mathbf{d}_p. \tag{1a}$$

This run is the model ocean E2. At a location $\mathbf{x} = \mathbf{x}_0$ the predicted values $\mathbf{d}_p$ are replaced by values $\mathbf{d}_m$ measured in the reference ocean E1. They can be thought of as composed by $\mathbf{d}_p$ plus a perturbation $\boldsymbol{\varepsilon}$:

$$\mathbf{d}_m = \mathbf{d}_p + \boldsymbol{\varepsilon}.$$

Thus the assimilation experiment J_n corresponds to:

$$L\{\boldsymbol{\phi}\} = \mathbf{d}_p \quad \text{at} \quad \mathbf{x} \neq \mathbf{x}_0$$
$$L\{\boldsymbol{\phi}\} = \mathbf{d}_m = \mathbf{d}_p + \boldsymbol{\varepsilon} \quad \text{at} \quad \mathbf{x} = \mathbf{x}_0. \tag{1b}$$

The information $\boldsymbol{\varepsilon}$ provided by the measured values acts like a perturbation and will grow (≡ spread) in time and in space, propagating away from the data insertion section. However, if the time frequency of data insertion is sufficiently high and if the region of insertion is sufficiently broad and densely covered, the forcing provided by $\boldsymbol{\varepsilon}$ will drive the model with data insertion J_n towards the reference ocean E1. In fact, in the limit in which data are inserted at all grid points and all times, (1b) becomes:

$$L\{\boldsymbol{\phi}\} \equiv \mathbf{d}_m \quad \text{for all } \mathbf{x} \tag{1c}$$

which is exactly the reference ocean E1. As we insert data at all time steps, the relevant question is: is data insertion performed only locally, i.e., along oceanographic sections, sufficient to drive the model to the reference ocean E1 overcoming the loss of predictability intrinsic in the model itself? Are the different sections chosen according to the above criteria differently effective?

Also, two time scales must be distinguished. The first is the short time scale of evolution of mesoscale Rossby eddies, of the order of a few months. The second is the climatological time scale of the order of many years. Local data insertion can be completely ineffective on the short time scale for improving the model estimates. It may, however, be effective on the climatological time scale when the duration in time of the data insertion process is longer than the model equilibration time and may compensate for the smallness of the insertion region, thus leading to a better estimate of the climatological mean circulation.

The paper is organized as follows. In section 2 we discuss the main body of assimilation experiments, comparing the effectiveness of the different assimilation sections according to the criteria discussed above. We compare pairs of assimilation experiments in which only one of these criteria is changed. Analytical considerations are also given in which the process of data assimilation, which is a forcing at an internal boundary in the model, is shown to be equivalent to a line source of vorticity, equivalent therefore to assigning an additional stress curl along a line in each layer of the model. We discuss the assimilation experiments in the light of this interpretation.

In section 3 we examine the effects on the climatological average of the circulation of data assimilation along a single section. We perform a further experiment with a model E2 which is "more imperfect" (5% perturbation of the parameters of the reference ocean E1). Finally, in section 4 we give the conclusions of the present work and the suggestions for future research.

2. Assimilation of local data in the transient case

a. The reference and perturbed oceans E1 and E2

The QG model formulation with N arbitrary layers is a straightforward extension of the two-layer case described by Holland (1978). A description of the semidiscrete form of the equations (in which the vertical discretization has already been done) has been given in Part I, to which the reader is referred for details. The wind stress pattern used to spin up the model is an idealized double gyre of steady wind-forcing, given by $\tau = -T_0 \cos(2\pi y/L_y)$. Table 1 gives the values of the physical parameters used to construct the reference ocean E1. This is spun up to statistical equilibrium and then run for 10 successive years to obtain climatological averages over long time durations.

Figures 1a, b show, respectively, the climatological averages of the surface layer streamfunction ψ_1 for the first 5 years (Fig. 1a) and the entire duration of 10 years (Fig. 1b). The reference ocean E1 is characterized by a rather broad recirculation region south of the Gulf Stream jet. Also, the meridional sections at which data are measured in E1 have been drawn in Fig. 1a.

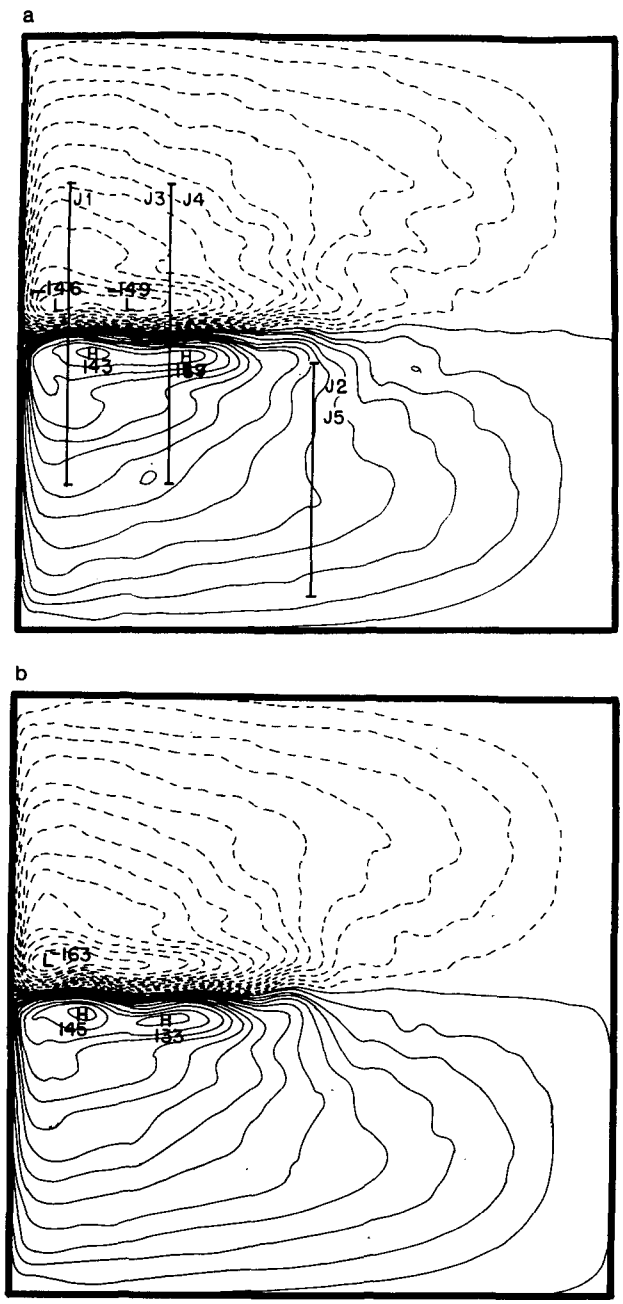

FIG. 1. Climatological means of the surface layer streamfunction ψ_1 of the reference ocean E1. (a) Mean of the first 5 years. Contours from -140×10^2 to 140×10^2 m^2 s^{-1}. Contour interval 10^3 m^2 s^{-1}. (b) Mean of the entire 10-year period. Contours from -160×10^2 to $+140 \times 10^2$ m^2 s^{-1}. Contour interval 10^3 m^2 s^{-1}. The meridional sections at which data are inserted in the assimilation experiments are shown.

The perturbed ocean E2 is constructed by *reducing* by 2% the reduced gravities $g'_{k+1/2}$, the biharmonic friction A_4 and reducing also by 2% the wind-stress intensity T_0. The bottom friction R is instead *increased* by 2%. Thus, the perturbed ocean E2 is less driven by the wind at the surface and more damped by friction at the bottom. The resulting circulation patterns are, however, only slightly different (weaker) as the perturbation is small. Figures 2a, b show the climatological averages of the surface layer streamfunction ψ_1 for E2, respectively for the first 5 years (Fig. 2a) and the entire 10 years (Fig. 2b). As results will be shown mostly for the surface layer streamfunction ψ_1, this will be im-

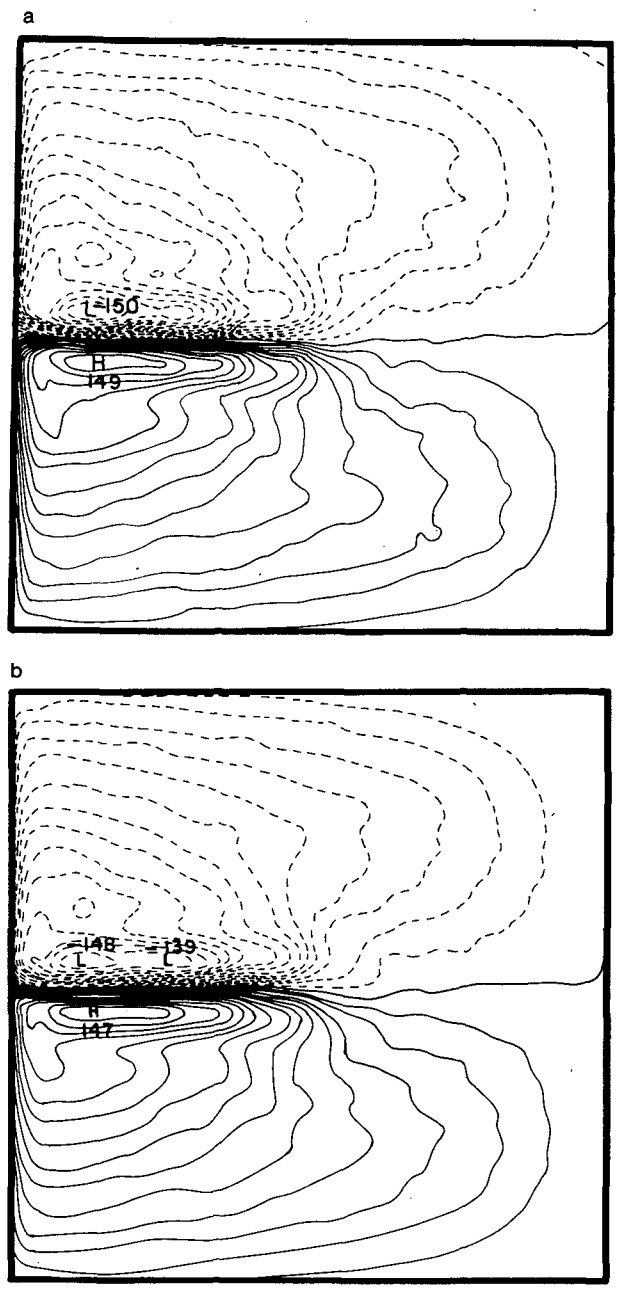

FIG. 2. Climatological means of the surface layer streamfunction E2 for the model ocean perturbed by 2%. (a) Mean of the first 5 years. Contours from -140×10^2 to $+140 \times 10^2$ m^2 s^{-1}. (b) Mean of the entire 10 year period. Contours from -140×10^2 to 140×10^2 m^2 s^{-1}.

plicitly assumed in the figures that follow, unless otherwise specified.

The difference in the intensity of the high maxima between Figs. 1a and 2a; 1b and 2b is small. For Figs. 1a and 2a it is 6×10^2 m^2 s^{-1} compared to the overall maximum of 140×10^2 m^2 s^{-1}. For Fig. 1b and 2b the difference is 2×10^2 m^2 s^{-1} compared to the same overall maximum. However, the total range for the perturbed ocean E2 of Fig. 2b (from -140×10^2 to $+140 \times 10^2$ m^2 s^{-1}) is slightly smaller than for the reference ocean E1 of Fig. 1b (from -160×10^2 to 140×10^2). Perturbed ocean E2 is thus slightly less intense than E1 even though the recirculation region south of the Gulf Stream is tighter and somewhat stronger gradients exist in E2 than in E1 (the contour interval is the same in Figs. 1a, b and 2a, b).

The climatologies of Figs. 1 and 2 filter out the intense mesoscale eddy field associated with the meandering Gulf Stream jet. They are quite evident in any daily snapshot, an example of which is shown in Fig. 3, at day 1800 after spinup for the reference ocean E1.

b. *The assimilation experiments J_n*

We measure data at chosen sections in E1 and insert them in E2, continuously in time (i.e., at every time step) and with the same spatial smoothing around the section discussed in Part I. The perturbed model E2 with the internal forcing provided by the data insertion is then run through the adjustment time and for 10 full years after it. These assimilation experiments are called J_n.

FIG. 3. Snapshot of the surface layer streamfunction at day 1800 after spinup for the reference ocean E1.

Three criteria are used to compare the experiments:

(i) distance of the (meridional) section from the western boundary;
(ii) energetics of the insertion region;
(iii) assimilation of only baroclinic versus barotropic and baroclinic modes.

Table 2 shows the characteristics of the sections shown in Fig. 1a (only meridional sections are discussed here).

As for the steady case of Part I, two diagnostic tools can be constructed, the difference fields (i) between the assimilation experiment J_n and the reference ocean E1 and (ii) between J_n and the perturbed ocean E2. The related rms error can then be evaluated for each of these difference fields. Different error measurements can also be constructed, from the global rms-error over the entire domain to local ones over subportions of the domain and local regions.

The predictability problem of the model is illustrated by Fig. 4, which gives two global rms-errors for the surface layer difference streamfunction of experiment J2 of Table 2 over the first 900 days of the assimilation experiment. Specifically, curve A gives the rms-error for the difference field:

$$\text{DIFF2} = \text{J2} - \text{E2}$$

between the model with and without data insertion and curve B gives the rms-error for:

$$\text{DIFF1} = \text{J2} - \text{E1},$$

that is the "true" error with respect to the reference ocean. Curve B illustrates the typical behavior of the rms-error in a predictability experiment when the model has lost the memory of its initial state. The two oceans, the reference E1 and the model with data insertion J2, are completely uncorrelated; the rms-error exhibits the typical random oscillation around an average statistical value, showing that the two oceans represent two independent realizations of an ensemble.

Curve A shows the rms-error between the perturbed model *with* data insertion J2 and the model *without* data E2. At initial time the error is very small. However, as discussed in the introduction, the inserted data now acts as an initial perturbation given to the model; this perturbation grows exponentially. By $T = 100$ days, curve A cannot be distinguished from B and exhibits the same random oscillation pattern around the same equilibration value. A measure of the time scale for loss of predictability in these experiments is therefore:

$$T_{\text{predictability}} \leq 100 \text{ days}.$$

By 100 days, the reference ocean E1, the perturbed model E2 and the model with data insertion J2 are all statistically independent realizations.

The behavior of the global rms error DIFF1 with respect to the reference ocean, curve B of Fig. 4, may

TABLE 2. Assimilation experiments with meridional sections.

Experiment	Location of the section	Type of data
J1	x = 300 km 1000 km $\leq y \leq$ 3000 km	Baroclinic (density)
J2	x = 2000 km 500 km $\leq y \leq$ 1500 km	Baroclinic (density)
J3	x = 1000 km 1000 km $\leq y \leq$ 3000 km	Baroclinic (density)
J4	x = 1000 km 1000 km $\leq y \leq$ 3000 km	Barotropic and baroclinic (density and velocity)
J5	x = 2000 km 500 km $\leq y \leq$ 1500 km	Barotropic and baroclinic (density and velocity)

seem at first surprising. One would in fact expect that inserting the "right" data into a "wrong" model would tend to drive the "wrong" model closer to the truth and so reduce DIFF1. Curve B of Fig. 4 shows instead that such an improvement, if present at all, is completely obscured by the inherent loss of predictability of the model itself. Over time scales comparable with the predictability time of 100 days there is no apparent reduction at all in DIFF1, which grows instead at initial time and successively oscillates around the equilibration value.

This section focuses upon comparing the effectiveness of the different sections of Table 2 upon the two measures of error DIFF1 and DIFF2. First, we point out that on the short time scale (~100 days) there is *no* improved correction in the global rms-error relative to the baseline ocean E1, that is in DIFF1, for any of the assimilation experiments of Table 2. The behavior of the "true" error DIFF1 is completely insensitive to the change of the section location in more or less energetic regions of the ocean and/or to the assimilation of baroclinic *only* versus baroclinic *and* barotropic data. DIFF1 never decreases, but rather shows the oscillating behavior around a mean value as in curve B of Fig. 4. It is impossible to quantify any improvement/worsening from the random oscillation of DIFF1. A quantified measure can be defined only examining the behavior of DIFF2, that is looking at the inserted data as an initial perturbation, from the point of view of a predictability experiment. DIFF2 grows from approximately zero at initial time (when J_n and E2 are different only along a line) until reaching a statistically "equilibrated" oscillation around the same mean value of DIFF1 (the three oceans E1, E2 and J_n have comparable energetics). Figure 4 shows that such a growth occurs over the predictability time of 100 days and that levels in DIFF2 beyond ~30 000 m^2 s^{-1} are not different from the predictability error. Thus we shall discuss the behavior of DIFF2 during the first 180 days of the assimilation experiments of Table 2. The growth rates of DIFF2 to reach the equilibration value of 30 000 m^2 s^{-1} can be quantified. If any of the sections of Table 2 shows a significant reduction in growth rate of DIFF2 compared with the others, this section will be the most effective. Figure 5 shows the global DIFF2 = J_n − E2 for the surface layer during the first 180 days of each of the experiments of Table 2. Curve (a) gives DIFF2 for section J1; (b) for J2 and so forth. Different comparisons can be made between the various sections.

1) PAIR J1–J3

Same energetic region: the eddy-rich Gulf Stream system.

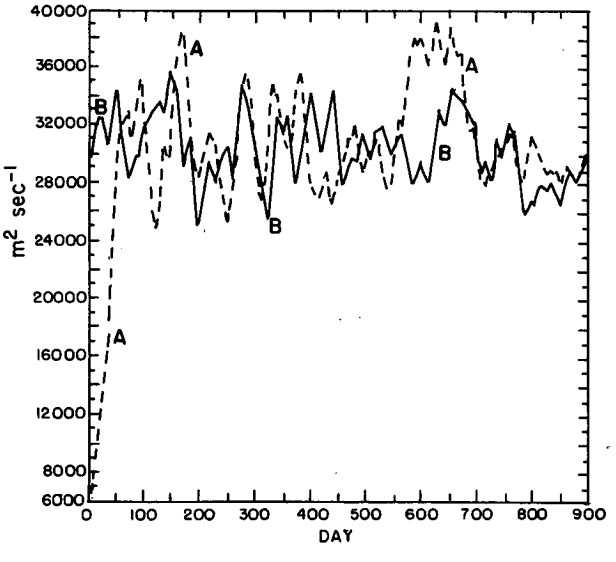

FIG. 4. Global rms-errors in m^2 s^{-1} for the surface layer of experiment J2 of Table 2 as a function of time in days. Curve A: rms-error of DIFF2 = J2 − E2. Curve B: rms-error of DIFF1 = J2 − E1.

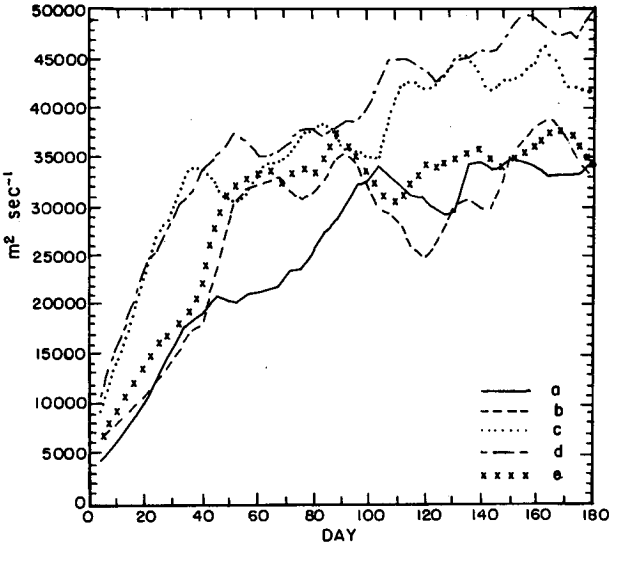

FIG. 5. Global rms-error in m^2 s^{-1} for the surface layer of DIFF2 = J_n − E2 as a function of time in days. (a) J1 − E2; (b) J2 − E2; (c) J3 − E2; (d) J4 − E2; (e) J5 − E2.

Difference: distance from the western boundary.

The maximum error level over 180 days is greater in J3 [Gulf Stream system, 47 000 $m^2 s^{-1}$, curve (c)], than in J1 (Sverdrup interior, 35 000 $m^2 s^{-1}$, curve (a)) by about 25%. At day 100, however, DIFF2 is 33 000 $m^2 s^{-1}$ in J1 and 35 000 $m^2 s^{-1}$ in J3, that is of the same order.

2) PAIRS J3–J2 AND J4–J5

The main difference is in the energetics of the section location. The maximum error is of 47 000 $m^2 s^{-1}$ for J3 (Gulf Stream system, curve (c)) versus 39 000 $m^2 s^{-1}$ for J2 (Sverdrup interior, curve (b)), the former one being greater by 17%. The maximum error is 50 000 $m^2 s^{-1}$ for J4 [Gulf Stream system, curve (d)] versus 38 000 $m^2 s^{-1}$ for J5 (Sverdrup interior, curve (e)), the former being greater by ~25%. At day 100 DIFF2 is 30 000 $m^2 s^{-1}$ in J2; 35 000 $m^2 s^{-1}$ in J3; 39 000 $m^2 s^{-1}$ in J4; 33 000 $m^2 s^{-1}$ in J5.

3) PAIR J1–J2

Different distances from western boundary.

Different energetics of the section location. The maximum error is of 35 000 $m^2 s^{-1}$ in J1 [Gulf Stream system, curve (a)] and of 38 000 $m^2 s^{-1}$ in J2 [Sverdrup interior, curve (b)], the two being of the same magnitude. At day 100 DIFF2 is 33 000 $m^2 s^{-1}$ in J1 and 30 000 $m^2 s^{-1}$ in J2.

4) PAIRS J2–J5 AND J3–J4

Assimilation of the baroclinic mode *only* (density) versus the baroclinic *and* barotropic mode (density and velocity).

The maximum error is 38 000 $m^2 s^{-1}$ in J2 [curve (b)] and 38 000 $m^2 s^{-1}$ in J5 [curve (e)], both located in the Sverdrup interior. In the Gulf Stream system, the maximum error is 47 000 $m^2 s^{-1}$ in J3 [curve (c)] and 50 000 $m^2 s^{-1}$ in J4 (curve (d)). At day 100, DIFF2 is 30 000 $m^2 s^{-1}$ in J2 and 33 000 $m^2 s^{-1}$ in J5; 35 000 $m^2 s^{-1}$ in J3 and 39 000 $m^2 s^{-1}$ in J6.

A final comparison concerns the growth rates of DIFF2 in the various experiments. The equilibration value is assumed to be 30 000 $m^2 s^{-1}$ and the growth rate is defined as

$$\frac{30\,000 - \text{DIFF2}|_{t=0}}{\text{time interval to reach equilibration}}.$$

Thus, the growth rates are, in $m^2 s^{-1}$/day:

J1	J2	J3	J4	J5
300	480	500	500	480

Again notice the identical growth rates in the pairs J2–J5 and J3–J4. The smallest growth rate is in J1, the section nearest the western boundary.

Some (qualitative) conclusions can be inferred from the above measures of DIFF2. The overall error level depends upon a balance between the energetics of the region and the distance of the section from the western boundary. This second factor defines the area, which can be filled out by perturbation mesoscale eddies in their westward propagation. For sections at the same distance, with an equal area between the section and the western boundary, the error level depends on the overall energetics of the inserted data. For regions with similar energetics, the more distant the section the greater the magnitude of DIFF2 (compare J1 with J3–J4). The two factors can compensate. Sections J2–J3 are in the relatively quiescent Sverdrup interior but the area filled out by westward propagating eddies is broad. Sections J3–J4 have a smaller area of influence but are located in the eddy-rich Gulf Stream region. The corresponding overall levels in DIFF2 and growth rates are comparable.

Two important points must be finally discussed. First, the error levels and growth rates of DIFF2 discussed above are not significantly different. This confirms the results obtained by visual inspection of DIFF1, which was impossible to quantify. On the short times of mesoscale variability all the sections are *equally ineffective* in improving the model error on the global space scale.

The second point concerns why the barotropic information has so little effect and the growth rates obtained with only baroclinic versus baroclinic and barotropic data are identical. First, we point out that it may be very difficult, if not impossible, to distinguish and quantify the effectiveness of different types of data with only a single section of them. This is in effect what the above results tell us. Second, even though the data coverage were sufficiently broad to reduce the rms-error and allow us to distinguish between data types, quasi-geostrophic dynamics is characterized by flow of energy, and information, from the small (baroclinic) scales to the large (barotropic) ones. Baroclinic data will quickly excite, through nonlinear interactions, the fast barotropic component. Examination of the individual daily maps show that by day 4 the barotropic mode (initially zero when assimilating only density) is fully developed. Thus, it does not make much difference in the overall growth rates to assign the barotropic mode at initial time or to have it fully developed after four days. The interesting speculation is then about the possible difference in the rms-error behavior when assimilating only baroclinic versus only barotropic information. The above results, however, show the equal inefficacy of individual sections for which the differences in the measured error growths are not significant. Such a comparison should then be carried out with a much greater spatial coverage to provide any conclusive and quantifiable result.

We summarize the overall result of the previous discussion: a single section of assimilated data is com-

pletely ineffective in driving the model toward reality over time scales comparable with the mesoscale variability. As it will be shown in section 2c, the internal forcing provided by the single section of inserted data is too localized in physical space to induce any immediate correlation between the model and the reference ocean E1, at least on the short time scale $T \approx 100$ days characteristic of predictability loss. A single section of data might, however, reduce the level of rms-error (i) in localized regions in space; (ii) over the very long time scales characteristic of the mean circulation (the "climatology"). Case (i) will be discussed in section 2c and case (ii) in section 3.

c. Dynamical considerations

It is well known that the information provided by a source located in the interior of an infinite β-plane propagates westward, as shown by the evaluation of the radially symmetric Green's function of the steady problem (Rhines 1983). Here we want to show that the internal boundary forcing provided by the data insertion at a section is equivalent to a vorticity source along a line, equivalent therefore to an additional stress-curl imposed locally. We shall consider for simplicity the insertion problem in which the streamfunction is directly assigned at the data section. Also, we shall consider a linearized equivalent barotropic model as embodying in its simplified dynamics the relevant physical behavior observed in the fully nonlinear 3-layer QG model.

Consider first as the canonical example the homogeneous Laplace's problem with inhomogeneous boundary conditions:

$$\nabla^2\psi = 0, \quad 0 \leqslant x \leqslant a; \quad 0 \leqslant y \leqslant b \tag{2a}$$

with $\psi = 0$ over three sides and

$$\psi = \psi_b(y) \quad \text{at} \quad x = a.$$

The solution of (2a) is (Morse and Feshback 1953, p. 801)

$$\psi(x, y) = \int_0^b \psi_b(\eta) G_b(x, y/\eta) d\eta \tag{2b}$$

where the Green's function of the problem is

$$G_b(x, y/\eta) = \frac{2}{b} \sum_n \frac{\sinh(n\pi x/b)}{\sinh(n\pi a/b)} \sin(n\pi y/b) \sin(n\pi\eta/b) \tag{2c}$$

as a direct separation of variables and expansion of $\psi_b(y)$ in Fourier series can prove. Consider also the Poisson's problem with homogeneous boundary conditions:

$$\nabla^2\psi = -4\pi\rho(x, y) \tag{3a}$$

and $\psi = 0$ over all four sides of the domain.

The solution to (3a) is

$$\psi(x, y) = \int_0^a d\xi \int_0^b d\eta G(x, y/\xi, \eta)\rho(\xi, \eta) \tag{3b}$$

and the Green's function of the problem is

$$G(x, y/\xi, \eta) = \sum_n \frac{8 \sin(n\pi y/b) \sin(n\pi\eta/b)}{n \sinh(n\pi a/b)} \times \sinh(n\pi x/b) \sinh\left[\frac{n\pi}{b}(a - \xi)\right] \quad \text{for} \quad x < \xi. \tag{3c}$$

It can be shown that (Morse and Feshback 1953, same pages)

$$\left[\frac{\partial}{\partial\xi} G(x, y/\xi, \eta)\right]_{x<\xi, \xi\to a} = -4\pi G_b(x, y/\eta) \tag{4}$$

and the solution ψ is the same for both problems. In the above problems ψ is the potential arising from an interior distributed source $\rho(x, y)$ or from an assigned boundary distribution $\psi_b(y)$. These latter two are thus equivalent, and from the knowledge of the boundary function one can find the equivalent source distribution.

The above can be generalized to the linear equivalent barotropic model in a square domain

$$0 \leqslant x \leqslant a; \quad 0 \leqslant y \leqslant b.$$

Consider first the homogeneous time-dependent problem with inhomogeneous boundary conditions:

$$\nabla^2\psi_t - \frac{1}{R^2}\psi_t + \beta\psi_x = 0 \tag{5a}$$

with $\psi = 0$ over three sides and $\psi = \psi_b(y)$ at $x = a$. In (5) R is the baroclinic Rossby radius. We look for the solution to (5a) in the form of steadily propagating waves. To do so we examine the system moving with the wave speed c_n (the general solution is a linear superposition of Rossby waves), $s = x - c_n t$. In this system (5a) becomes

$$\nabla^2\psi_s - \left(\frac{1}{R^2} + \frac{\beta}{c_n}\right)\psi_s = 0 \tag{5b}$$

or

$$\nabla^2\psi - \left(\frac{1}{R^2} + \frac{\beta}{c_n}\right)\psi = 0$$

as the integration constant can be assumed to be zero by taking $\psi = 0$ and $\nabla^2\psi = 0$ on three sides and: $\psi = \psi_b(y)$ and $\nabla^2\psi = [(1/R^2) + (\beta/c_n)]\psi_b$ at $s = s_a = a - ct$, for all times. The solution to (5b) is easily found to be

$$\psi(s, y) = \int_0^b \psi_b(\eta) G_b(s, y/\eta) d\eta \tag{6a}$$

with the Green's function given by

$$G_b(s, y/\eta) = \frac{2}{b}\sum_n \sin\left(\frac{n\pi y}{b}\right)\sin\left(\frac{n\pi\eta}{b}\right) \times \frac{\sinh\left\{\left[\frac{1}{R^2}+\frac{\beta}{c_n}+\left(\frac{n\pi}{b}\right)^2\right]^{1/2} s\right\}}{\sinh\left\{\left[\frac{1}{R^2}+\frac{\beta}{c_n}+\left(\frac{n\pi}{b}\right)^2\right]^{1/2} s_a\right\}} = \frac{2}{b}\sum_n \sin\left(\frac{n\pi y}{b}\right)\sin\left(\frac{n\pi\eta}{b}\right)\frac{\sinh(\alpha s)}{\sinh(\alpha s_a)} \quad (6b)$$

with

$$\alpha = \left[\frac{1}{R^2}+\frac{\beta}{c_n}+\left(\frac{n\pi}{b}\right)^2\right]^{1/2}.$$

Consider next the inhomogeneous problem with homogeneous boundary conditions:

$$\nabla^2\psi_t - \frac{1}{R^2}\psi_t + \beta\psi_x = \mathrm{curl}_z\tau \equiv 4\pi\rho_x(x, y) \quad (7a)$$

with $\psi = 0$ on all four sides. Again seek a solution in the form of steadily propagating waves and move to the moving frame $s = x - c_n t$. In it (7a) is

$$\nabla^2\psi - \left(\frac{1}{R^2}+\frac{\beta}{c_n}\right)\psi = -\frac{4\pi}{c_n}\rho(s, y) \quad (7b)$$

where, again, an integration constant has been put to zero. Equation (7b) is the equivalent to (3a) for the source $\rho(x, y)$ with $\mathrm{curl}_z\tau \equiv 4\pi\rho_x(x, y)$. The solution to (7b) is

$$\psi(x, y) = \int_0^{s_a} d\xi \int_0^b d\eta G(s, y/\xi, \eta)\rho(\xi, \eta) \quad (8a)$$

with the Green's function given by

$$G(s, y/\xi, \eta) = \sum_n \frac{8\pi}{\alpha b}\frac{\sin(n\pi y/b)\sin(n\pi\eta/b)}{\sinh(\alpha s_a)} \times \sinh(\alpha s)\sinh[\alpha(s_a - \xi)] \quad (8b)$$

for $s < \xi$. Again it is easy to show that

$$\left[\frac{\partial}{\partial\xi}G(s, y/\xi, \eta)\right]_{s<\xi,\xi\to s_a} = -4\pi G_b(s, y/\eta). \quad (9)$$

Then the stress distribution equivalent to the internal boundary forcing can be found from

$$\int_0^{s_a}\int_0^b G(s, y/\xi, \eta)\rho(\xi, \eta)d\xi d\eta = \int_0^b \psi_b(\eta)G_b(s, y/\eta)d\eta. \quad (10)$$

In the insertion experiments like J2, the $\psi_b(y)$ dependence at the data section is fixed by the assigned wind stress curl. Specifically, a section like J2 is equivalent to assigning:

$$\psi_b(y) = \sin\left(\frac{\pi y}{b}\right) \quad (11)$$

to obtain a single gyre-solution. One then gets

$$\rho(\xi, \eta) \equiv A(\xi)\sin\left(\frac{\pi\eta}{b}\right) = \frac{\alpha}{4\pi}\frac{\delta(s_a - \xi)}{\sinh[\alpha(s_a - \xi)]}\sin\left(\frac{\pi\eta}{b}\right) \quad (12a)$$

as direct substitution into (10) shows. The equivalent stress curl distribution is then:

$$\mathrm{curl}_z\tau = 4\pi\rho_s(s, y). \quad (12b)$$

Thus, the problem of continuous data insertion at a single meridional section is equivalent to assigning an additional stress curl as a line source in each of the three layers of the QG model, in addition to the wind stress distributed all over the surface layer. The above solution is for the steady equilibrium response which is found in terms of steadily propagating waves. If, moreover, in (5b) or (7b) ψ_{sss} ($\equiv\psi_{xxt}$) $\ll \psi_s$ ($\equiv\psi_t$), it can be shown that the speed is given by

$$c_n = \frac{-\beta}{(1/R^2) + (n\pi/b)^2} \quad (13)$$

that is the speed of long, nondispersive Rossby waves in the square basin. In the data assimilation experiments along an internal section the data insertion is started impulsively. It is then equivalent to the problem treated by Anderson and Gill (1975) of an ocean, initially motionless, adjusting to a suddenly imposed local wind stress (see also Hendershott 1987). In particular, the problem fully equivalent to that treated by Anderson and Gill is for the difference field:

$$\mathrm{DIFF2} = J_n - \mathrm{E2} \quad (14)$$

in which the data are inserted at $t = 0$ at the data section which constitutes an eastern boundary. DIFF2 $\equiv 0$ at $t = 0$. The transient waves excited at $t = 0$ move westward at the speed c_n, given by (13) for the fastest waves. The Rossby eddies thus fill out the region immediately westward of the data section, while the region eastward of it remains relatively void of eddies.

The above dynamical considerations allow us to rationalize the qualitative results of section 2b. For an insertion section far away from the western boundary the perturbation (≡ information) assigned at the section is rapidly transmitted westward by the longest waves, which can fill out a vast area. If the section is near the western boundary, however, these long waves are immediately reflected as short waves which can propagate the perturbation eastwards at only ⅛ of the long-wave speed (Anderson and Gill 1975). Short eastward moving Rossby waves, being so slow, are damped much

more effectively and they remain confined in a thin western boundary layer. In fact, this is the behavior of the eddy field actually observed in the region eastward of the data section, which grows much more slowly and remains confined near the insertion section (not shown).

If the data section crosses a region where inertial effects are important, like the energetic Gulf Stream jet, then a further mechanism becomes effective. The propagation of perturbation eastward due to nonlinear advection by the Gulf Stream jet becomes competitive with the eastward propagation induced by the linear short Rossby waves. However, the nonlinear advective speed is also smaller than the westward speed of the long Rossby waves, in the rough ratio of 1 to 15 m s^{-1} (Hendershott 1987). Thus, the growth of error of the experiments of Table 2 can be rationalized mainly in terms of Rossby wave dynamics. Obviously, the error growth will be bigger the more energetic are the Rossby eddies.

To conclude this section we focus on localized measures of rms-error versus the global rms previously discussed. Local rms errors were monitored in time for the following regions when the section is J2 of Table 2: (i) the Southern subtropical gyre: $0 \leq x \leq 4000$ km and $0 \leq y \leq 2000$ km; (ii) the Northern subpolar gyre: $0 \leq x \leq 4000$ km and $2000 \leq y \leq 4000$ km; (iii) the region west of the insertion section: $0 \leq x \leq 2000$ km and $0 \leq y \leq 2000$ km; (iv) the region east of the insertion section: $2000 \leq x \leq 4000$ km and $0 \leq y \leq 2000$ km; (v) the region immediately surrounding the section: $1500 \leq x \leq 2500$ km and $500 \leq y \leq 1500$ km.

In the light of the previous considerations based on Rossby wave dynamics, the following results for DIFF2 = J2 − E2 become obvious. For a section like that of experiment J2 in the southern gyre, the rms-ii (northern gyre) and rms-i (southern gyre) are the same, ~35 000 m^2 s^{-1}. rms-iv (eastward sector, ~20 000 m^2 s^{-1}) is much smaller than rms-iii (westward sector, ~40 000 m^2 s^{-1}); rms-v (surrounding region) is equally small as rms-iv, ~20 000 m^2 s^{-1}, not surprisingly as the area is the smallest and around the region of direct insertion. The above figures are for the overall error level after loss of predictability. The actual rates of error growth are of the same order of the growth rate for the global rms. The above confirms the inefficacy of single data sections in improving the model estimates, with the error level being reduced only locally in the region around the section, but the error growth rates remaining the same. The vorticity input along a single line provided by the data is too localized in space to drive the model toward the reference ocean. The unpredictability of the model overcomes the tendency of the assimilation process to improve the model estimates, at least on the transient time scales comparable with the predictability time. The effect of local data insertion, if any, must then be sought not in providing a different effectiveness for different regions in the physical domain of the model but rather in providing a different effectiveness upon the different time scales intrinsic in the model, the climatological trend versus mesoscale variability.

3. The effect of data assimilation on the climatological trend

In this section we study the effect of continuous data insertion on the climatological trend of the model. Figures 1a, b show the first 5-year and 10-year climatological means of the reference ocean E1. Figures 2a, b show the same means for the ocean E2 perturbed by 2% with respect to E1. As discussed in section 2a, the circulation patterns of E1 and E2 are slightly different, with E2 being overall slightly weaker (less wind driven and more bottom damped). However, in the climatological means of Figs. 1 and 2 the recirculation region south of the Gulf Stream jet is higher in E2 than in E1 and with stronger gradients, with a more compact region of high values compared to a somewhat broader area in E1.

Figures 6a, b show the first 5 years and 10 years climatological means for experiment J2 of Table 2, in which data from E1 are inserted into E2 continuously in time for the entire period of 10 years. The data section is also shown. The general, qualitative appearance of Figs. 6a, b is more similar to the reference ocean E1 of Figs. 1a, b than to the perturbed ocean without data insertion E2 of Figs. 2a, b. Specifically, in the climatological means of J2, the location of the highs and lows on either side of the Gulf Stream relative to the western boundary is extremely similar to that in E1. We need to quantify this similarity.

In section 2 we examined the behavior of the rms error DIFF2 = J_n − E2 as it was impossible to quantify an improvement/worsening induced by the data process when examining the "true" error DIFF1 = J_n − E1. Over a time scale of the order of the predictability time ~ 100 days, DIFF1 exhibited only statistically random fluctuations around the average predictability error. When, however, inserting data continuously for very long durations in time, over 10 years, a definite trend appears in DIFF1, as shown in the following. Random oscillations of DIFF1 are still superimposed to this general trend, but the latter one can be quantified. Thus, in this section we can use the more appropriate DIFF1 as rms-error.

Figure 7 shows the global rms error (J2 − E1) for the surface layer difference streamfunction and for the last 1100 days of the assimilation experiment, from day 2500 to 3600. The continuous insertion of data, even though only at a section, manifests itself through the decreasing trend of the error, of ~13% over 10 years with respect to the equilibration value. Transient fluctuations are superimposed on the decreasing trend.

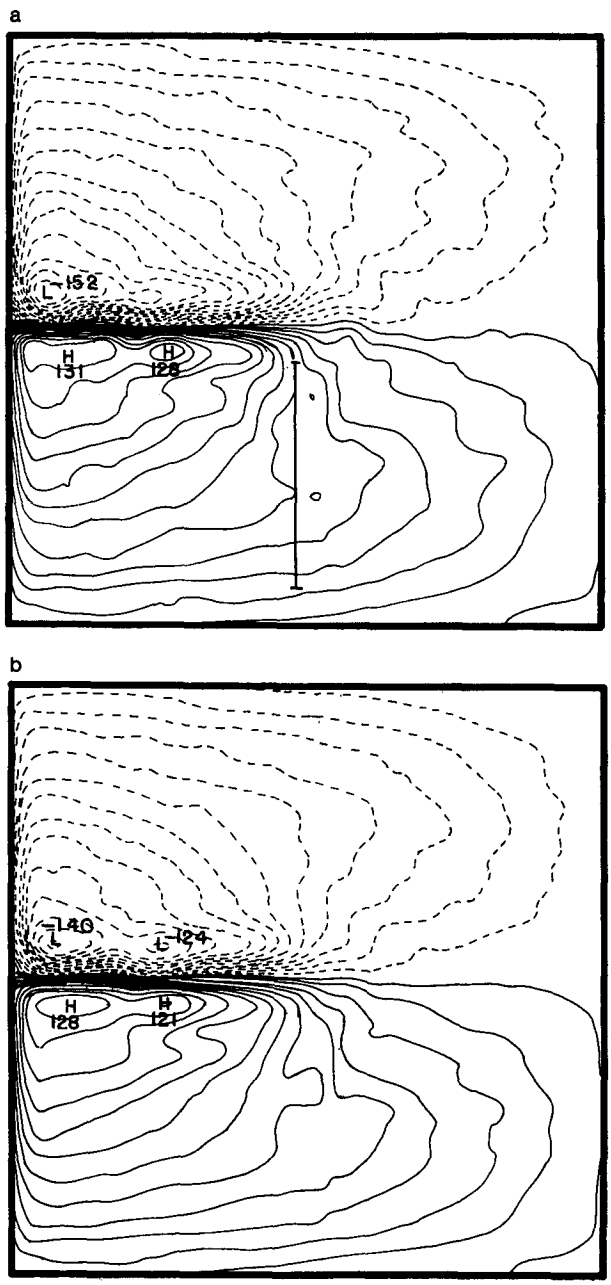

FIG. 6. Climatological means of the surface layer streamfunction ψ_1 for the insertion experiment J2 of Table 2. (a) Mean of the first 5 years. Contours from -150×10^2 to 130×10^2 m^2 s^{-1}. Contour interval 10^3 m^2 s^{-1}. (b) Mean of the entire 10-year period. Contours from -140×10^2 to $+120 \times 10^2$ m^2 s^{-1}. Contour interval 10^3 m^2 s^{-1}.

This trend appears unambiguously during the second half of the experiment. It would not change taking shorter or longer intervals in time, and the trend is much smaller during the first years of the insertion process.

The perturbed ocean E2 is however too weakly perturbed with respect to E1 for this experiment to be conclusive and the resulting percentage decrease in DIFF1 may not be statistically significant. A more perturbed ocean was thus constructed by perturbing E1 with a 5% increase in the reduced gravities $g'_{k+1/2}$ and the biharmonic friction A_4 and by increasing by 5% the wind stress amplitude T_0. The bottom friction R was instead decreased by 5%. This new perturbed ocean E'_2 is more intensely driven and less bottom damped than the reference ocean E1. Figures 8a–c show the first 5-year, second 5-year and overall 10 years means of E'_2. The first 5-year mean of Fig. 8a is already more intense, even though slightly, than the corresponding 5-year mean of E1 shown in Fig. 1a. The recirculation region is tight and gradients around it are stronger. The E'_2 becomes even stronger during the second 5-years, the mean of which is shown in Fig. 8b. The recirculation high has increased by about 10% from the first 5-years of Fig. 8a and is about 20% higher than the means of E1 shown in Fig. 1a, b. The overall 10-year mean of E'_2 shown in Fig. 8c reflects the stronger pattern of the second 5 years of simulation, even though weakened by the first 5 years. Obviously the model has very long time scales in its response and data insertions longer than 5 years are necessary to affect its equilibration.

The assimilation experiment with the meridional data section in the southern subtropical gyre, equivalent to J2 of Table 2, is then repeated for the new perturbed ocean E'_2 and is called I2. Figures 9a–c show respectively the first 5-year, second 5-year and overall 10-year

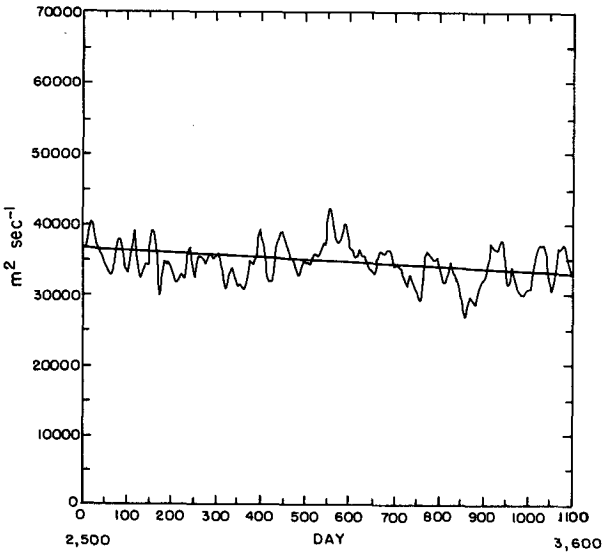

FIG. 7. Global rms-error for the surface layer difference streamfunction. DIFF1 = J2 − E1 in m^2 s^{-1} as a function of time from day 2500 to day 3600.

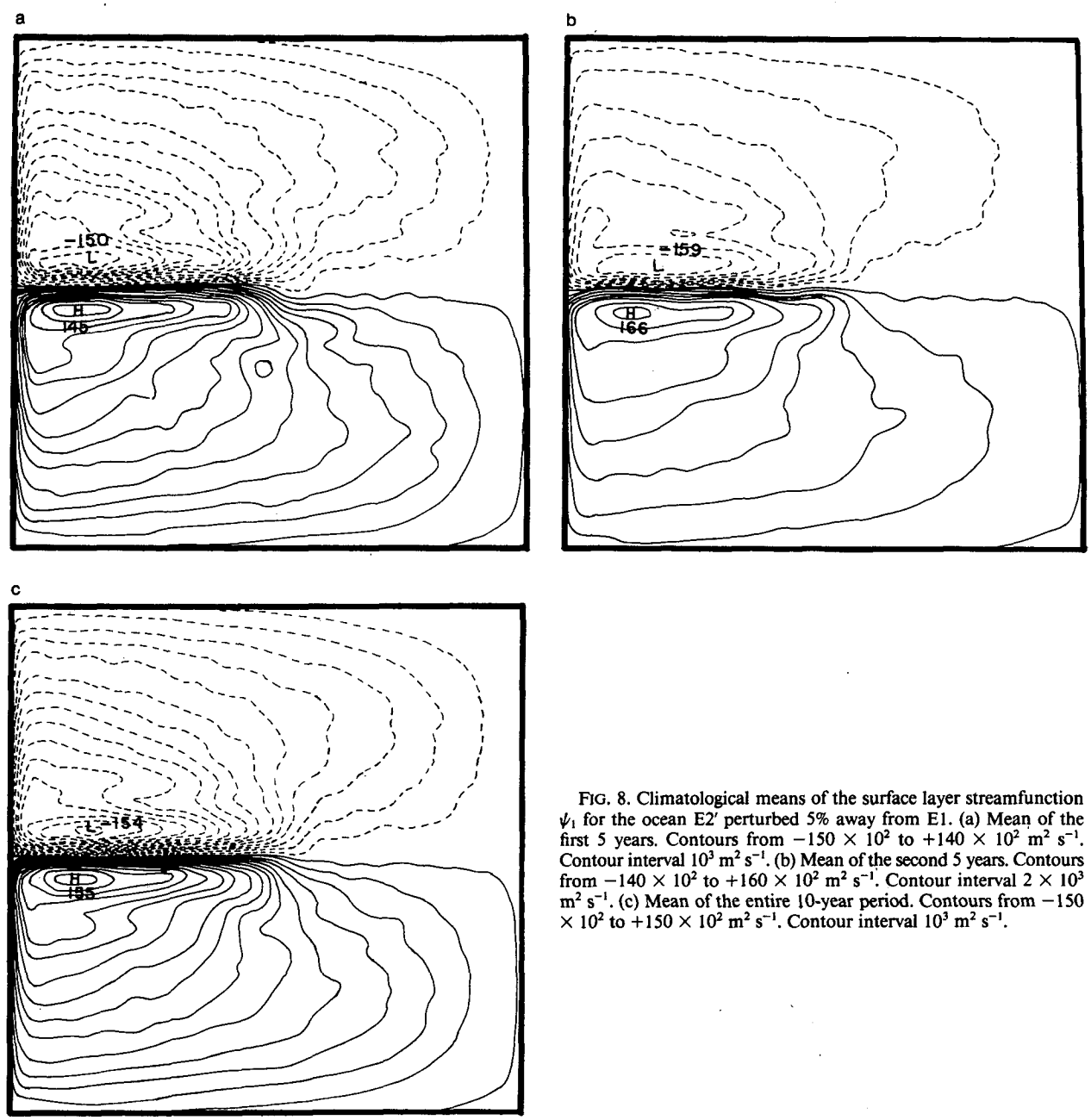

FIG. 8. Climatological means of the surface layer streamfunction ψ_1 for the ocean E2′ perturbed 5% away from E1. (a) Mean of the first 5 years. Contours from -150×10^2 to $+140 \times 10^2$ m^2 s^{-1}. Contour interval 10^3 m^2 s^{-1}. (b) Mean of the second 5 years. Contours from -140×10^2 to $+160 \times 10^2$ m^2 s^{-1}. Contour interval 2×10^3 m^2 s^{-1}. (c) Mean of the entire 10-year period. Contours from -150×10^2 to $+150 \times 10^2$ m^2 s^{-1}. Contour interval 10^3 m^2 s^{-1}.

means for the ocean with data insertion I2. The data section is also shown. The first 5-year means of I2 in Fig. 9a is still rather stronger than the reference ocean E1, by about 10%, and much more similar to the perturbed ocean E_2' in the compact and strong-gradients recirculation region south of the stream. We are still below the time scales necessary for the model to reach a new equilibration affected by the data insertion. The second 5-year mean of I2 shown in Fig. 9b has been "driven" towards the reference ocean E1 with the weakening and broadening of the recirculation pattern south of the stream, which shows now a diffuse area of high values. The location of highs and lows in the recirculation region is now extremely similar to E1. The 10-year mean of Fig. 9c thus very strongly resembles the 10-year mean of E1, shown in Fig. 1b, in shape and intensity.

To quantify the above visual similarities, Fig. 10 shows the global rms-"true" error for the surface layer DIFF1 = I2 − E1 again for the last time interval, day

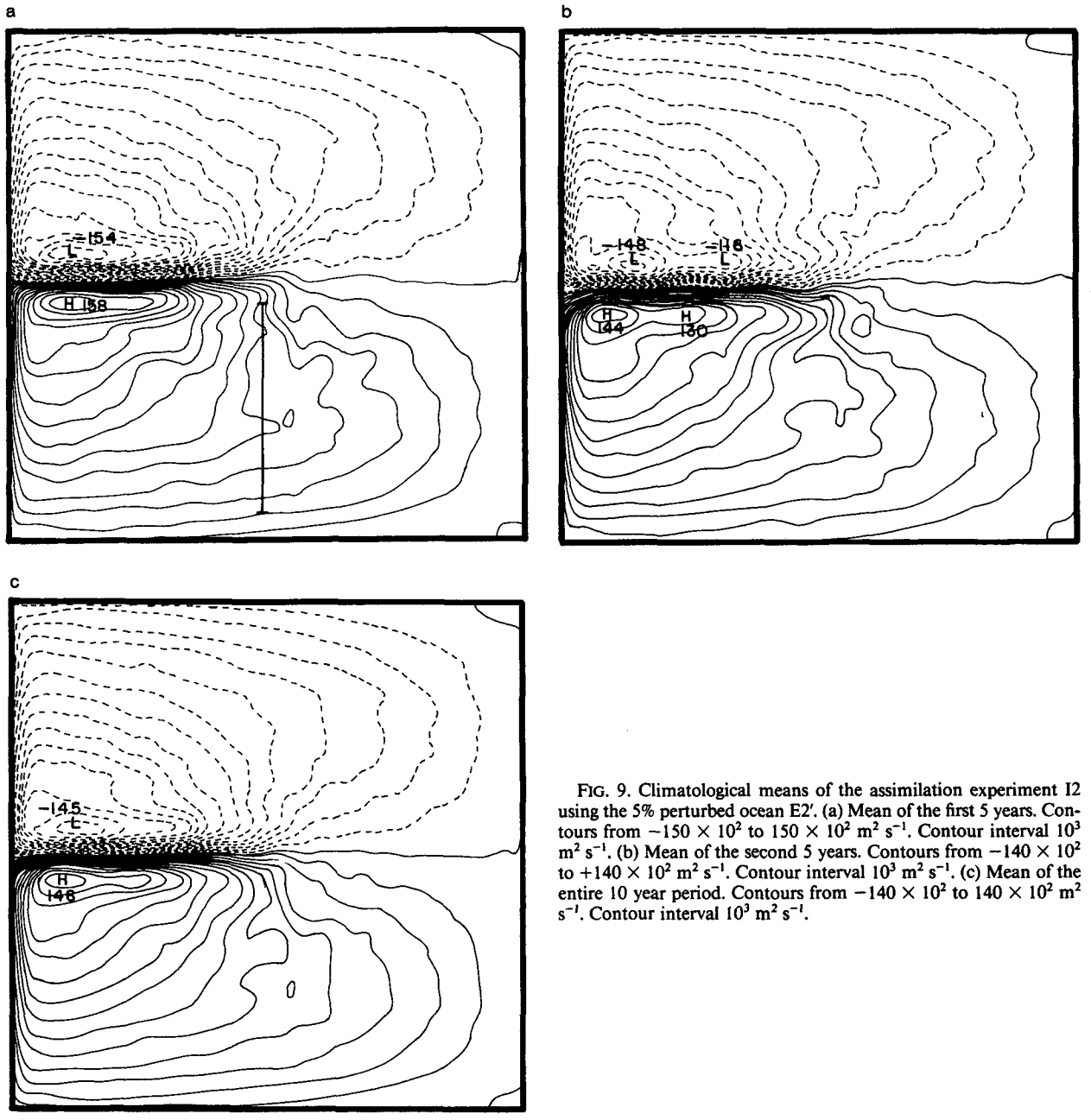

FIG. 9. Climatological means of the assimilation experiment I2 using the 5% perturbed ocean E2'. (a) Mean of the first 5 years. Contours from -150×10^2 to 150×10^2 m^2 s^{-1}. Contour interval 10^3 m^2 s^{-1}. (b) Mean of the second 5 years. Contours from -140×10^2 to $+140 \times 10^2$ m^2 s^{-1}. Contour interval 10^3 m^2 s^{-1}. (c) Mean of the entire 10 year period. Contours from -140×10^2 to 140×10^2 m^2 s^{-1}. Contour interval 10^3 m^2 s^{-1}.

2500–3600, when a trend appears unambiguously. The decreasing trend of the error is now rather more marked than in Fig. 7 with an overall decrease of ~25% over 10 years from the initial equilibration value which is ~40 000 m^2 s^{-1}. The above results show that insertion of local data, if carried on continuously for time durations rather longer than the model equilibration time, can be quite effective in reducing the error of the long time scales of the circulation and in significantly improving the climatological trend.

A comparison might have been carried out to investigate which of the different sections of Table 2 is most effective for the climatology. All sections, in fact, might be equally effective over 10 years of assimilation, as all were equally ineffective on the short time scale. We did not attempt such a comparison, as the requirement that the data assimilation process must be longer than the model equilibration time to be effective, makes the result very improbable to be achieved with the present datasets.

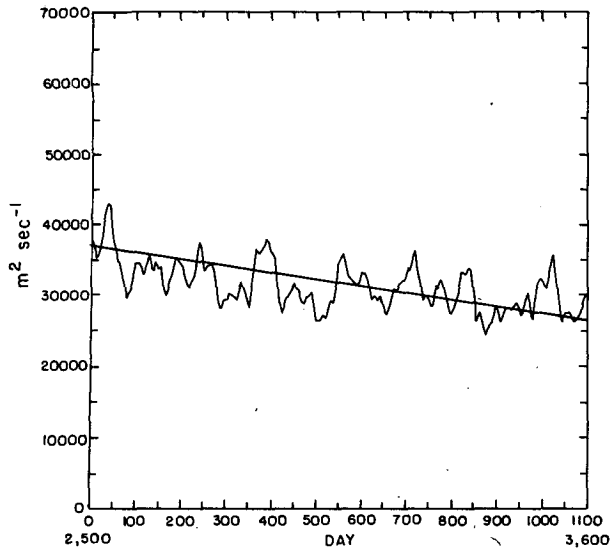

FIG. 10. Global rms-error for the surface layer difference streamfunction DIFF1 = I2 − E1 in $m^2 s^{-1}$ as a function of time from day 2500 to day 3600.

4. Conclusions

A series of assimilation experiments has been carried out and discussed in which data are inserted along single oceanographic sections into a fully eddy-resolving, time-dependent model. Following the results of Part I of the present work, these sections have been chosen to be meridional as the most effective in allowing the information provided by the data to spread towards the western boundary. The role played by the characteristics of the equations of the steady problem is here shown to be played by the long, westward propagating Rossby waves excited by the data insertion at the section.

In the time-dependent, fully nonlinear case, however, the insertion of local data is equivalent to imposing a perturbation at initial time at chosen gridpoints in the model. As is well known from the predictability problem of any general circulation model, this perturbation grows exponentially until the model evolution becomes statistically independent from the evolution without initial perturbations. A single data section in physical space is too localized a forcing to "drive" the "wrong" model towards the "reference" ocean. Global (≡ over the entire basin) versus local rms error estimates show that a single section is very ineffective in improving the model performance on the time scales of mesoscale variability, 100 days, comparable to the time scale of predictability loss. This inefficacy holds for whatever location is chosen for the data section inside the gyre and for whichever subregion of the gyre is monitored. The level of the rms-error DIFF2 is smaller east of the section than west, as Rossby eddies propagating eastward are short, slow and quickly damped by viscosity. The level of rms-error is smaller when the section is near the western boundary, even though crossing the highly energetic Gulf Stream system. In this case, in fact, even though the excited perturbations are energetic they are slowly carried eastward by the inertial jet. Obviously, the level of rms-error is the smallest when evaluated locally just around the data section itself. However, the error growth rates are roughly the same, with no significant differences. Random fluctuations are superimposed to the mean predictability error. The "wrong" ocean with data insertion becomes a realization of the ocean circulation completely uncorrelated both with the "wrong" ocean itself and the "reference" ocean.

If local oceanographic sections will thus be inefficient in improving model estimates over short time scales, they may be quite effective in improving the climatological trend of the model if the continuous data insertion is performed for time durations longer than the model equilibration time. The present results indicate a significant improvement over a time scale of 10 years, with the climatological mean circulation being much more similar to the climatology of the reference ocean than to the climatology of the wrong model without data. This effect is not simply qualitative and visual. It is quantified by the time evolution of the global rms-error which in the last ~3 years of the assimilation experiment shows a well-defined decreasing trend. This trend is not a result of the specific time interval chosen, but it is rather smaller at the beginning and manifests itself unambiguously only during the last period of the 10-year experiment. The transient scales are still randomly fluctuating but the climatological mean is significantly improving. Thus, individual oceanographic sections might still be useful if very long time series of data were available.

The result that a single data section is quite ineffective is not surprising. More surprising is instead the possibility of obtaining better climatologies from the continuous assimilation of local data. The requirement, however, that the data assimilation process must be sufficiently long in time to be effective (longer than the model statistical equilibration time) makes the result rather improbable to be achieved with actual datasets.

Thus a series of important questions arises. What is the minimum coverage in space to be monitored for the data assimilation process to reduce the error and overcome the error growth given by loss of predictability? If the data region is wide enough, can the assimilation process be consequently reduced in time and still provide significant amelioration of model estimates? Which is the optimal combination of areal extension of the data and duration of time series for the assimilation to provide significant decreases in the model errors? How do they depend on the location of the measurements, for instance if they are in the gyre

interior only but over all the ocean depth or over all the gyre but only in the surface layer? When and how are the model climatological mean and trend improved by these different and possible combinations? These are only a few of the questions posed by the present results that challenge the ocean modeling community as far as data assimilation is concerned.

Acknowledgments. This research was carried out with the support of the National Science Foundation, Grant OCE-8614369 and of the National Aeronautics and Space Administration, NASA project UPN 161-25-63-03. Acknowledgment is made to the National Center for Atmospheric Research, which is sponsored by the National Science Foundation, for the computing time used in this research. Ms. Dorothy Frank carefully typed the manuscript and Ms. Isabelle Kole drafted the figures.

REFERENCES

Anderson, D. L. T., and A. E. Gill, 1975: Spin-up of a stratified ocean, with applications to upwelling. *Deep-Sea Res.*, **22**, 583–596.

Hendershott, M. C., 1987: Single layers models of the general circulation, *General Circulation of the Ocean*, D. I. Abarbanel and W. R. Young, Eds., Springer-Verlag, 202–269.

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

Malanotte-Rizzoli, P., and W. R. Holland, 1986: Data constraints applied to models of the ocean general circulation. Part I: The steady case. *J. Phys. Oceanogr.*, **16**, 1665–1682.

Morse, M. M., and H. Feshback, 1953: *Methods of Theoretical Physics, Vol. I*, McGraw-Hill, 997 pp.

Rhines, P. B., 1983: *Lectures in Geophysical Fluid Dynamics. Lecture in Applied Mathematics*, Vol. 20, 3–58.