Data Constraints Applied to Models of the Ocean General Circulation. Part I: The Steady Case

PAOLA MALANOTTE-RIZZOLI

Department of Earth, Atmospheric, and Planetary Sciences, Center for Meteorology and Physical Oceanography, Massachusetts Institute of Technology, Cambridge, MA 02139

WILLIAM R. HOLLAND

National Center for Atmospheric Research, Boulder, CO 80307*

(Manuscript received 18 October 1985, in final form 27 March 1986)

ABSTRACT

In this work we take a first step in the process of assimilating data into models of the ocean general circulation. The goal is not prediction but rather understanding how the data insertion process affects, and is affected by, the dynamics governing the model. The chosen model ocean is steady, weakly nonlinear and highly frictional. Strong vertical friction plays the role of eddy fluxes in driving the circulation in the deep layers.

In the data insertion process we capitalize upon the two principles that (i) the available dynamical models are imperfect; (ii) oceanographic data are measured locally. Three major questions are addressed; 1) what is the influence of local data insertion in terms of improving estimates of the model general circulation? 2) how does the model dynamics affect the spreading of information from the data insertion region? 3) what can we learn about the model physics from the effects of data insertion?

Density (or temperature) measurements along long hydrographic or tomographic sections or arrays are chosen as data. We vary the location of the section as well as its orientation. In our highly frictional ocean, the most effective sections are meridional, long and located at a distance from the western boundary. Model estimates are then significantly improved over the broad region extending from the data section to the western boundary itself.

Advective effects are minimal and influence the spreading of information only in the intense western boundary current. Rather, the structure of the gyre interior manifests itself through a quite important steering effect exerted by the motion in the intermediate layer upon the spread of information in the surface layer. Due to this effect the region southwest of the data section is consistently preferred for the improvement of the estimates. Simple analytical computations are carried out to rationalize the numerical results. This effect is likely to persist in more realistic, fully eddy-resolving simulations in which the interfacial eddy stresses would play the role here given to vertical friction.

The dependence of spreading of information upon the internal physics and/or external forcing is used to examine what is imperfect in the model parameterizations. In a simple analytical example we scan the two-dimensional parameter space defined by internal friction and wind stress amplitude. The "correct" values of the above parameters cannot be inferred by this simple scanning due to the non-uniqueness of the solution.

1. Introduction

Over the course of the next decade a number of new observational techniques, based upon the availability of satellites (altimetry, scatterometry) and the development of new techniques for observing the interior of the ocean (such as tomography) will begin to give the oceanographer a synoptic, "real time" picture of the oceanic circulation and how it changes with time. This information, when it becomes available, will need rationalization in terms of dynamical descriptions of the relevant phenomena.

Numerical models will play a key role in this process. In contrast to the meteorological community, oceanographers have had little experience in developing the procedures and analysis techniques to effectively assimilate real data into ocean models. Moreover, the essential nature of ocean dynamical processes, particularly those associated with meridional barriers to flow, the gyre structure, and intense western boundary currents, are quite distinct from the atmospheric case. It is incumbent upon the ocean modeling community to begin to develop this experience for itself. That is the motivation for a sequence of studies, of which this is the first. Future papers will examine more complex situations, including transient and highly nonlinear situations that involve fully eddy resolving ocean cir-

* The National Center for Atmospheric Research is sponsored by the National Science Foundation.

culation patterns requiring sophisticated assimilation techniques.

In this exploratory study a number of different kinds of simulated data could be chosen to constrain our models. In fact an important question to be answered in future work is what kinds of observations are most effective in the sense of leading to large and widespread improvement in ocean model simulations. Note that our focus here is not upon boundary conditions for the ocean, such as wind stress and heat flux at the sea surface, but rather on internal observations which are known only locally. We stress the local nature of inserted data as an important feature of many oceanographic measurements, although future satellite observations will provide a more global coverage of some important variables.

The simulated oceanographic data chosen to constrain our models in this first study are temperature data along sections. The reason is simple. Such data are already available synoptically in some places from XBT and CTD shipborne observations. Even more important, acoustic techniques (tomography arrays, inverted echo sounders) are under rapid development and should, in the near future, provide the detailed synoptic evolution of the temperature field in ocean gyres, at least along certain sections. In fact, tomographic experiments have already been carried out in the ocean (Ocean Acoustic Tomography Group, 1982; Cornuelle et al., 1985) and have been simulated on the scale of the ocean gyre using an eddy resolving general circulation model (EGCM; see Holland, 1978) to determine which quantities of importance for ocean general circulation and ocean climate could be measured (Malanotte-Rizzoli and Holland, 1985; Malanotte-Rizzoli, 1985).

The goal of this first set of data insertion experiments is very different from the goal and motivations of the meteorological community as we do not aim at prediction. The technique we use to assimilate data into the model is, however, similar to what is done in meteorology (see, for instance, Ghil et al., 1979). Quite an extensive body of literature exists on methods for the initialization of and data assimilation into numerical models. They have very much in common with other methods variously called optimal estimation, inverse procedures, objective mapping, etc., to such an extent that sometimes the terminology is only a matter of taste. We shall borrow from the existing literature the insertion techniques which have already been developed.

Also, in the present study, we assume the data to be perfect and the model imperfect. In a parallel but independent study, Schröter and Wunsch (1986) investigate the case of imperfect data/perfect model, even though they devote some attention to an imperfect model parameterization of internal friction. They focus on the data errors and on finding an optimization technique capable of dealing with data uncertainty. Our emphasis is on understanding how data insertion in local, limited regions affects the dynamics of the model and how, in its turn, the dynamics governing the model affects the spreading of information from the data insertion section. The case of imperfect model/data with errors is not a trivial extension of what is done here and will be the object of future research.

We make use of a multilayer, quasi-geostrophic (hereafter QG), fine resolution model capable of examining fully turbulent, eddy resolved ocean circulation on the basin scale. In this first study, however, we run the model with parameters chosen in such a way that the time-dependent model spins up to a steady state. In addition this steady state is only slightly nonlinear (explicit vertical viscosity is relatively large to replace the eddy effects that drive the deep ocean) but, as will become clear, the nonlinearity associated with vortex stretching has a remarkable and interesting influence upon the data insertion problem. Moreover, treating the steady, nearly linear case allows us to make use of simple analytic models for understanding what is going on in this first stage of our work.

The paper is organized as follows. Section 2 gives a short description of the QG EGCM and of the model physics used in this set of experiments. Also, the technique of data insertion is described. Section 3 discusses the bulk of the experiments, their motivation and intercomparison. Three major questions are put forward: 1) what is the influence of data insertion in terms of improving estimates of the model ocean circulation? 2) how does the model dynamics affect the data insertion? 3) what can we learn about the model physics from the effects of data insertion? Section 4 analyzes and interprets the experimental results, addressing the first two proposed questions through simple dynamical considerations and further numerical evidence. In section 5, we explore how we can approach the third question and, through a simple analytical example, we discuss the model sensitivity and non-uniqueness of the solution when the imperfectly known internal and/or external parameters of the model are varied. Finally, section 4 gives the conclusions of the present work and the suggestions for future research.

2. The quasi-geostrophic general circulation model

The QG model formulation with N arbitrary layers is a straightforward extension of the two-layer case described by Holland (1978). Here we shall use the semi-discrete form of the equations (in which the vertical discretization has already been done). The horizontal discretization and the form of the finite difference equations will not be discussed here—the interested reader is referred to Chow and Holland (1986) who have recently described the physics and an efficient numerical code for solving this time-dependent boundary value problem.

The governing equations are the vorticity and interface height perturbation equations, and the thermal wind relation:

$$\frac{\partial}{\partial t}\nabla^2\psi_k = J(f+\nabla^2\psi_k, \psi_k) + \frac{f_0}{H_k}(w_{k-1/2} - w_{k+1/2}) + F_k: k = 1 \text{ to } N \quad (1)$$

$$\frac{\partial}{\partial t}h_{k+1/2} = J(h_{k+1/2}, \psi_{k+1/2}) + w_{k+1/2}: k = 1 \text{ to } N-1 \quad (2)$$

$$h_{k+1/2} = \frac{f_0}{g'_{k+1/2}}(\psi_{k+1} - \psi_k). \quad (3)$$

Here the integer subscripts (k) denote the vertical layers (k increasing downward) in which the QG streamfunction is defined (nominally at the center of each of the layers) while fractional subscripts ($k + \frac{1}{2}$) denote the interfaces between layers where vertical velocity and interface height perturbation are defined (Fig. 1a). The variables are the streamfunction (ψ_k) with horizontal velocity components ($u = -\psi_y$, $v = \psi_x$), the interface height perturbation ($h_{k+1/2}$), positive upward, and the vertical velocity ($w_{k+1/2}$), also positive upward. The horizontal coordinates are x (eastward) and y (northward), the Coriolis parameter is $f = f_0 + \beta y$, and the mean layer thicknesses are H_k. The values of f_0 and β are chosen to represent typical midlatitude gyre values (see Table 1 for a listing of all the parameters governing the numerical experiments). The basic background vertical stratification is written in terms of the reduced gravity $g'_{k+1/2} = g\Delta\rho_{k+1/2}/\rho_0$, where $\Delta\rho_{k+1/2}$ is the (positive) density difference between layers $k+1$ and k. Frictional and forcing effects, written symbolically in Eq. (1) as F_k, have been parameterized in different ways in various earlier calculations. Here we include lateral friction of the biharmonic kind (Holland, 1978) and a vertical internal friction between layers, wind forcing at the upper interface and bottom friction. Thus

$$F_k = -A_4\nabla^6\psi_k + F_k^t - F_k^b \quad (4)$$

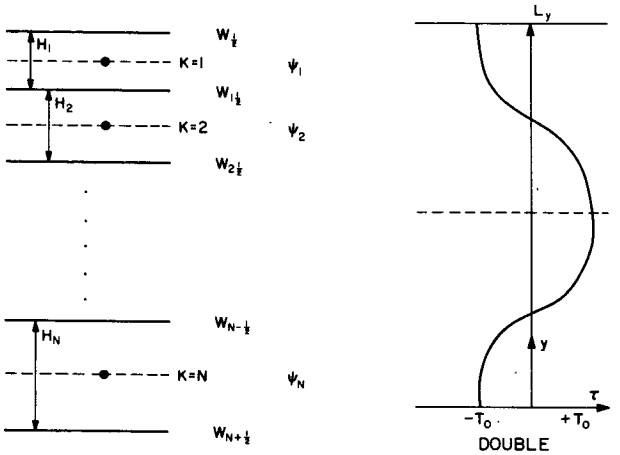

FIG. 1. (a) Schematic diagram for the vertical structure of the three-layer eddy resolving model. (b) The double-gyre of wind-forcing used as wind-stress pattern.

TABLE 1. Experimental parameters.

Coriolis parameter	$f_0 = 9.3 \times 10^{-5}$ s^{-1}
$f = f_0 + \beta y$	$\beta = 2 \times 10^{-11}$ m^{-1} s^{-1}
Layer thicknesses H_k	$H_1 = 300$ m; $H_2 = 700$ m; $H_3 = 4000$ m
Reduced gravities	$g'3/2 = 0.0357$; $g'5/2 = 0.0162$
Biharmonic friction	$A_4 = 8 \times 10^{10}$ m^4 s^{-1}
Vertical internal friction	$\epsilon = 2 \times 10^{-3}$ m s^{-1} "reference" ocean 1×10^{-3} m s^{-1} "model" ocean
Bottom friction	$R = 5 \times 10^{-7}$ s^{-1}
Wind stress intensity	$T_0 = 1 \times 10^{-4}$ m^2 s^{-2}
North–south extent of the basin	$L_y = 4000$ km

where $F_k^t = -\epsilon(\nabla^2\psi_k - \nabla^2\psi_{k-1})/H_k$, except for the uppermost level where $F_1^t = \text{curl } \tau/H_1$, and $F_k^b = +\epsilon(\nabla^2\psi_k - \nabla^2\psi_{k+1})/H_k$, except for the bottom-most level where $F_N^b = R\nabla^2\psi_N$. Note that the effect of the wind forcing, equal to curl τ/H_1, produces an Ekman pumping stretching tendency in the upper layer that is equivalent to the curl of a body force acting on the upper layer. Also note that a gentle bottom slope can be consistently included in the QG framework so that the influence of variable depth can be examined. Then the bottom boundary condition on w is $w_{N+1/2} = -J(H_B, \psi_N)$, with H_B the (positive upward) perturbation bottom topography. At the sea surface, $w_{1/2} = 0$. The advective velocities at the interfaces, needed in Eq. (2), are calculated from a weighted average of the velocities in the layers, i.e., $\psi_{k+1/2} = \alpha_{k+1/2}\psi_{k+1} + (1 - \alpha_{k+1/2})\psi_k$, where $\alpha_{k+1/2} = H_k/(H_k + H_{k+1})$.

These equations can be written in potential vorticity form

$$\frac{DQ_k}{Dt} = F_k \quad (5)$$

where

$$Q_k = \nabla^2\psi_k + f + \frac{f_0}{H_k}(h_{k+1/2} - h_{k-1/2}). \quad (6)$$

For consistency $H_{1/2} \equiv 0$ and $h_{N+1/2} = H_B$. In all the numerical experiments discussed here, $H_B \equiv 0$.

The wind stress patterns used in this study are idealized double gyres of wind-forcing (Fig. 1b), given by $\tau = -T_0 \cos(2\pi y/L_y)$; L_y is the north–south extent of the basin. These cases thus have two equal but opposing regions of vorticity input.

In the experiments we make use of a simple technique for insertion of simulated "data". Although we will be examining model experiments that reach a steady state in this paper, we choose the insertion procedure in anticipation of future work that will involve time dependent observations as well as time dependent model oceans.

As discussed in the next section, we obtain data from a "reference" ocean which is a model run that has reached a steady state. The observations are the steady state, vertical distribution of the "temperature" along

certain sections; in the QG context these are observations of $d\psi/dz$. Malanotte-Rizzoli (1985) discusses the relationship of sound speed perturbations measured with a tomography array to density perturbations, and thus (quasi-geostrophically) to $d\psi/dz$. At each time step of additional "imperfect" model ocean calculations, we insert these "observations" along the section with a degree of spatial smoothing. This spatial smoothing is introduced so as not to introduce grid scale discontinuities by replacing the evolving model solution by observations at only a single row of grid points. On the other hand, smoothing procedures may introduce local, spurious effects near the insertion lines (Bennett, 1985).

As an example, suppose our observations of $d\psi/dz$ are distributed along a line $x = x_0$, $y_1 < y < y_2$, and they are $d\psi/dz = F(y)$. Then at each time step of the evolving solution, we replace the predicted values in the vicinity of this line according to $d\psi/dz^{\text{new}} = d\psi/dz^{\text{old}} + \exp(-r^2/L^2)*(d\psi/dz^{\text{obs}} - d\psi/dz^{\text{old}})$. Here, for $y_1 < y < y_2$, $r^2 = (x - x_0)^2$; for $y < y_1$, $r^2 = (x - x_0)^2 + (y - y_1)^2$; for $y > y_2$, $r^2 = (x - x_0)^2 + (y - y_2)^2$. In the majority of cases to be discussed below, $L = 60$ km (3 grid intervals for the calculations here), of the order of the first baroclinic Rossby radius (44.9 km). Cases with no spatial smoothing, i.e., $d\psi/dz^{\text{new}} = d\psi/dz^{\text{obs}}$ along the data insertion line were also run. The barotropic mode is unaffected by data constraints upon $d\psi/dz$ (temperature). Thus, in the numerical experiments described in the next section, the data insertion constrains the two baroclinic modes of our three layer model, equivalent to constraining the two interfaces, $h_{3/2}$ and $h_{5/2}$, defined by (3). This technique causes the evolving model solution to be modified in the vicinity of the observations in a smooth fashion. The observed and predicted values are blended such that right on the section the values of $d\psi/dz$ are exactly the observed ones while nearby the values are a mix of observed and predicted values.

3. Numerical experiments

We first construct the "reference" ocean, from which the data will be obtained to be used to constrain the "imperfect" model. The "reference" ocean is determined by a numerical experiment in which the model is spun-up under a constant (in time) wind stress allowing for the formation of a double-gyre three-dimensional circulation, until a steady state is reached. Circulation in the lower layers (second and third) occurs because of interfacial friction, with a friction coefficient $\epsilon = 2 \times 10^{-3}$ m s^{-1}. The model final state with this amount of friction is quasi-linear. Figure 2 shows (a) the streamfunction ψ_1 in the upper layer; (b) the streamfunction ψ_2 in the intermediate layer; (c) the streamfunction ψ_3 in the bottom layer. We call E_1 the "reference" ocean run.

The second simulation is considered to be a "model" of the first, in which the physics or the forcing functions are only approximately known and therefore different from the reference ocean. This is the "model" ocean, called E_2, to be corrected with the data insertion. In the present case, E_2 has slightly wrong physics with respect to the reference E_1, represented by a smaller value of the interfacial friction $\epsilon = 10^{-3}$ m s^{-1}. The difference $(E_2 - E_1)$ is the error field of the model E_2 compared to the reference ocean E_1.

A series of experiments is then carried out in which data from E_1 are inserted in the model E_2; the resulting numerical experiments are again run to a final steady state. The data, considered to be synoptic obervations of the temperature along certain sections, i.e., the baroclinic modes, are inserted as discussed in Section 2. Each assimilation experiment is called E_n ($n \geq 3$), and the error field is given by $(E_n - E_1)$. Another difference field, important for the diagnostics, is the INFL-field (influence field) defined as

$$\text{INFL}_n = E_n - E_2$$

in which the effect of the data insertion into the model is elucidated through its "influence" in making the model estimates better or worse in different regions of the gyre. The INFL-field is particularly interesting since it shows visually the propagation of information from the assimilation section or array to other dynamically connected parts of the gyre. It tends to be largest at or near the data insertion section (the greatest improvement is there) and to decay away from it.

Note that the difference field $(E_n - E_1)$, the error with respect to the reference ocean, can be known only if we know the reference circulation pattern over the whole basin through other independent measurements. The influence field $(E_n - E_2)$ on the other hand gives the error of the model constrained by data with respect to the unconstrained model, and is always known. Thus the influence field, resulting from the combination of imperfect models with perfect local observations, is a very useful diagnostic tool to understand how we better (or worsen) the model performance in the different regions of the gyre by the insertion of data. Table 2 summarizes the set of experiments carried out with data insertion.

The following figures show results from some of the above experiments. Figure 3 shows the error field $(E_2 - E_1)$ between the model E_2 and the reference ocean E_1. This is the error without data insertion. As the model E_2 has too small an interfacial friction, the streamfunction fields ψ_1, ψ_2 in the two upper layers have a larger maximum amplitude than in E_1. The lowest layer streamfunction ψ_3, on the other hand, has a smaller maximum amplitude in E_2 than in E_1. In fact, with the same value of the bottom friction but a smaller value of interfacial friction, the motion in the bottom layer is "less driven" by the motion in the upper layers in E_2 than in E_1.

Results from the experiments of Table 2 will be shown primarily for the streamfunction ψ_1 of the upper layer. The following figures should be compared with Fig. 3a, the error field $(E_2 - E_1)$ for ψ_1. As discussed

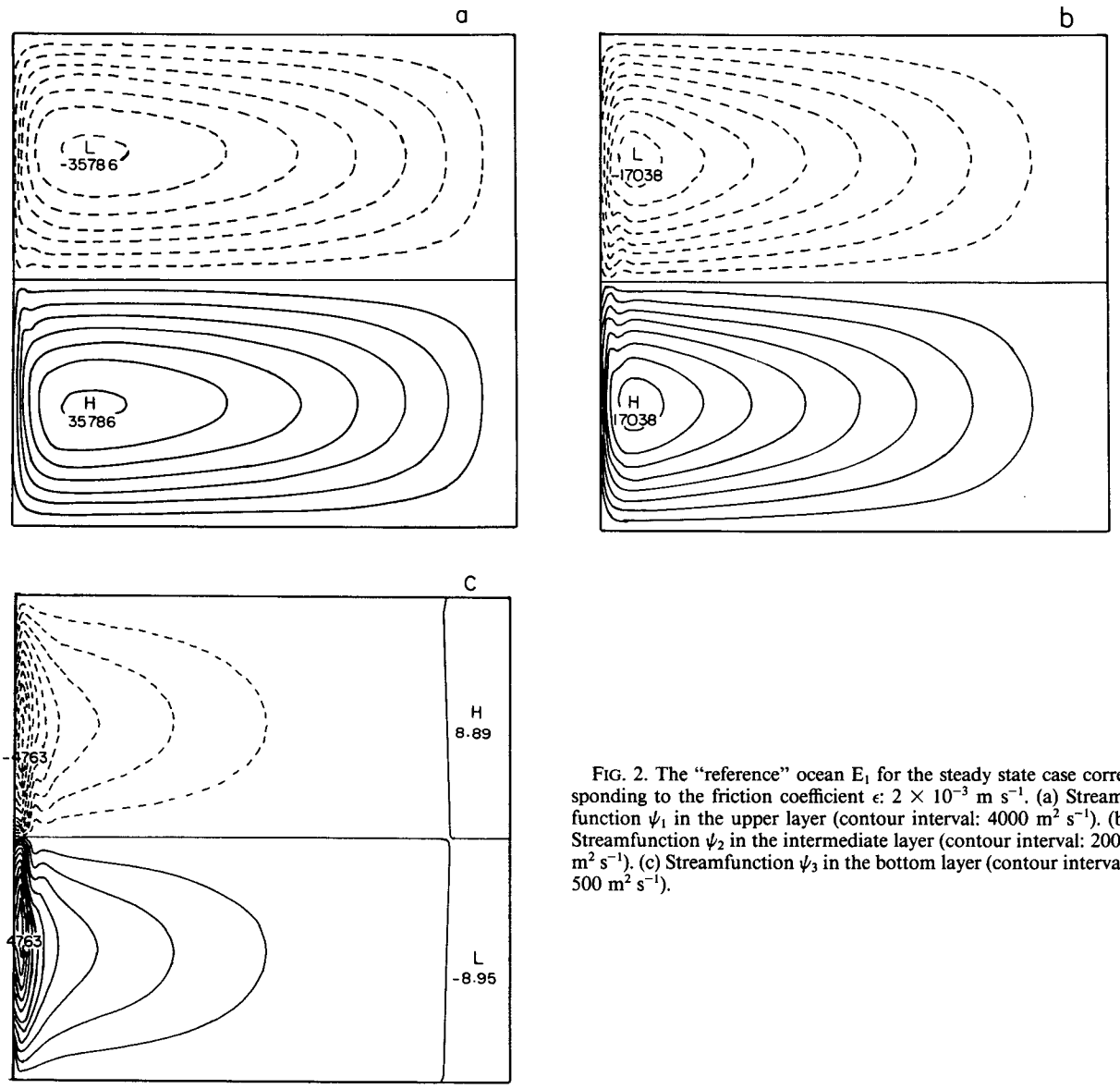

FIG. 2. The "reference" ocean E_1 for the steady state case corresponding to the friction coefficient ϵ: 2×10^{-3} m s^{-1}. (a) Streamfunction ψ_1 in the upper layer (contour interval: 4000 m^2 s^{-1}). (b) Streamfunction ψ_2 in the intermediate layer (contour interval: 2000 m^2 s^{-1}). (c) Streamfunction ψ_3 in the bottom layer (contour interval: 500 m^2 s^{-1}).

above, this error field is positive (overestimate of the streamfunction) in the southern gyre, with a maximum amplitude of 1.6×10^4 m^2 s^{-1}, and a symmetrical negative error in the northern gyre. The figures are organized as follows. For each of them the left panel gives the error field between the model with data inserted and the reference ocean ($E_n - E_1$). Less positive (negative) values compared to $E_2 - E_1$ in the southern (northern) gyre mean that the overestimate (underestimate) is reduced by data insertion. The right panel in each of the following figures gives the INFL-field ($E_n - E_2$), which emphasizes visually the region of "influence" of the data insertion. A negative (positive) INFL-field for the surface layer streamfunction ψ_1 in the southern gyre means that we improve (worsen) the estimate. The reciprocal statement holds for the northern gyre.

Figure 4 shows the error and influence fields for experiment E_3 of Table 2 for the streamfunction ψ_1 of the upper layer. Experiment E_3 is carried out with insertion of the baroclinic modes from E_1 along a meridional section in the middle of the southern gyre. Figure 4a shows the error field ($E_3 - E_1$). The error is obviously drastically reduced along and near the meridional section. In fact, the error in the baroclinic modes is zero right at the section because the model must agree perfectly there with the reference ocean. However the ψ_1 error need not be exactly zero because of any barotropic mode error. To the west side of the assimilation section the maximum error is now 1.1×10^4 m^2 s^{-1}, on the east side 2.6×10^3 m^2 s^{-1}, thus considerably reducing the overestimate provided by the model E_2 without data insertion. The improvement of the estimate in the southern gyre, where the assimila-

tion section is located, is however counterbalanced by a slightly worse estimate in the northern gyre. Here, in fact, the maximum, negative error increases from -1.6×10^4 to -1.9×10^4 m^2 s^{-1}.

This behavior is even clearer in the influence field ($E_3 - E_2$) of Fig. 4b. The negative maximum in the southern gyre shows that we improve the estimates by 8.6×10^3 m^2 s^{-1} at the assimilation section (maximum improvement) over the model (E_2) value of 3×10^4 m^2 s^{-1}, that is an improvement of 30%. Information from the data insertion spreads towards the western boundary, following a pattern which will emerge as typical of the influence fields in all the constrained experiments.

Note that, surprisingly, the region of influence in the southern gyre is not symmetrical about the midlatitude of the gyre, even though the inserted data are perfectly symmetrical about the midpoint. Instead, the influence field stretches toward the southwest with a pronounced asymmetry. The southwestward sector of the basin, moving from the location of data insertion, thus is found to be the region of preferential improvement of the estimates. We shall propose an explanation for this preferential spread of influence in the next section. In the western boundary current the influence field stretches northward in a narrow tonguelike feature confined near the boundary. Even though the high explicit viscosity makes our ocean quasi-linear, nevertheless horizontal advection of relative vorticity is important enough in the boundary current to produce this effect.

Figure 5 shows the streamfunction pattern ψ_1 for the surface layer for experiment E_4 of Table 2. Again, the left panel shows the error between the constrained experiment and E_1 and the right panel shows the INFL field. Experiment E_4 is carried out inserting the data again along a meridional section which now cuts symmetrically both the northern and southern gyres. Due to the symmetry of the section, the improvement in the error field is now symmetrical in the two gyres. Figure 5b shows the influence field ($E_4 - E_2$). Again notice the asymmetry in the pattern of Fig. 5b for which the inserted information exerts its influence spreading towards the southwest (northwest) in the southern (northern) gyre. Also, note the horizontal advective effects by the intense western boundary currents in the two western boundary regions. The northward (southward) spreading tonguelike shapes of the influence field in the boundary region are even closer to the western boundary itself.

Figures 6a, b show respectively, the error and influence fields for experiment E_6 of Table 2. Now the data section is zonal, in the center of the southern gyre. The influence field of Fig. 6b shows again an asymmetrical, southwestward orientation. The two extreme cases of meridional versus zonal data assimilation can be compared by examining the influence fields of Figs. 4b and 6b. First, we again point out that, in these quasi-linear oceans, the assimilated information is not advected away from the data section, and the influence fields do not therefore stretch following the streamlines of the steady circulation. The influence fields are also completely insensitive to the horizontal boundary layers created around the data insertion section by the explicit viscosity A_4. The lateral friction used in these experiments is of the biharmonic type, $A_4\nabla^6\psi$, and for the chosen value for A_4 (Table 1), the boundary layer thickness is $\delta = (A_4/\beta)^{1/5} \approx 21$ km, that is of the order of only one grid unit.

The order one balance in the model is then between the β-effect and vertical friction. The β-effect alone would allow the data insertion influence to spread pseudo-westward all the way to the western boundary if it were not for the rather strong vertical friction that ties the three layers together. Both the westward decay of influences shown in Fig. 4b and the north/south decay of influence found in Fig. 6b are due to this important frictional effect. In fact, if vertical friction were smaller in the basic experiments (both E_1 and E_2), the $INFL_3$-field of Fig. 4b would extend much further westward, following much more closely the planetary vorticity contours in the interior and closing with a thin western boundary layer of Stommel type.

Vertical friction also permits the existence of boundary layers near northern or southern boundaries. When we insert data along a zonal section as in Experiment E_6, boundary layers in the north/south direction will be created in the $INFL_6$-field at the zonal section. A boundary layer analysis leads to the equation

$$\psi_x = \tau^* \sin\left(\frac{2\pi}{L_y} y_i\right) - \psi_{\tilde{y}\tilde{y}}$$

where the (dimensionless) wind stress curl is evaluated at the section latitude y_i. In this boundary layer, planetary advection, wind torque and frictional torque are all of the same order. $\tilde{y} = (\epsilon^*)^{-1/2}(y - y_i)$ is the boundary stretched coordinate, with $\epsilon^* = \epsilon/H_1\beta L_y$. The thickness of this zonal boundary layer is $\epsilon^{*1/2}$, much broader than a meridional, Stommel-type boundary layer whose thickness is ϵ^*. In dimensional quantities, the thickness is $\delta = (\epsilon^*)^{1/2} L_y \approx 816$ km, as seen in the broad north/south spreading of the $INFL_6$-field in Fig. 6b. Had vertical friction been smaller, the $INFL_6$-field would be confined very much closer to the data insertion section, approaching the limit of a straight line from the section to the western boundary when $\epsilon \to 0$.

Since the influence field extends from the data section toward the western boundary, it is obvious that the effectiveness of data insertion in improving field estimates over a larger area depends on the distance of the section from the western boundary itself. This is true because of the quasi-linearity of the experiments and the consequent lack of much horizontal advection, as is clear when examining the error and influence fields of experiment E_7 of Table 2, shown in Fig. 7. In Experiment E_7, the data section is zonal, short and crosses

TABLE 2.

Experiment	Type of data insertion and dynamics	Forcing function
E_3	*Meridional* section in the southern gyre $x = 2000$ km 500 km $\leq y \leq$ 1500 km Fully nonlinear	curl $\tau = -T_0 \cos(2\pi y/L_y)$
E_4	*Meridional* section in both gyres $x = 1000$ km 1000 km $\leq y \leq$ 3000 km Fully nonlinear	as in E_3
E_5	*Square array* in the southern gyre $x = 1000$ km and $x = 2000$ km: 500 km $\leq y \leq$ 1500 km $y = 500$ km and $y = 1500$ km: 1000 km $\leq x \leq$ 2000 km Fully nonlinear	as in E_3
E_6	*Zonal* section in the southern gyre $y = 1000$ km; 1000 km $\leq x \leq$ 3000 km Fully nonlinear	as in E_3
E_7	*Zonal* section across the western boundary current in the southern gyre $y = 1000$ km; $0 \leq x \leq 50$ km Fully nonlinear	as in E_3
E_3-B	*Meridional* section in the southern gyre $x = 2000$ km 500 km $\leq y \leq$ 1000 km Fully nonlinear	as in E_3
E_3-C	*Meridional* section in the southern gyre $x = 2000$ km 1000 km $\leq y \leq$ 1500 km Fully nonlinear	as in E_3
E_6-S	*Zonal* Section as in E_6 Fully linear Smaller $A_4 = 1 \times 10^{10}$ m^4 s^{-2}	as in E_3
E_3-LL	*Meridional* section as in E_3 Fully nonlinear	as in E_3
E_3-L1	*Meridional* section as in E_3 $J(\psi, \nabla^2\psi) = 0$	as in E_3
E_3-L2	*Meridional* section as in E_3 $J(\psi, h) = 0$	as in E_3
E_8	*Meridional* section as in E_3 Fully nonlinear	curl $\tau \equiv 0$
E_8-LL	*Meridional* section as in E_3 Fully linear	as in E_8
E_8-L1	*Meridional* section as in E_3 $J(\psi, \nabla^2\psi) = 0$	as in E_8
E_8-L2	*Meridional* section as in E_3 $J(\psi, h) = 0$	as in E_8

the western boundary current. The dynamical location, very near to the western boundary towards which information spreads, the shortness of the section, and the weakness of horizontal advective effects make the influence region of Fig. 7b rather small and localized. The section of experiment E_7 is the most ineffective in improving the model estimates over a significant region in the present simulations.

4. Explanation of the experiments

In this section we make simple dynamical computations to elucidate the results of the numerical exper-

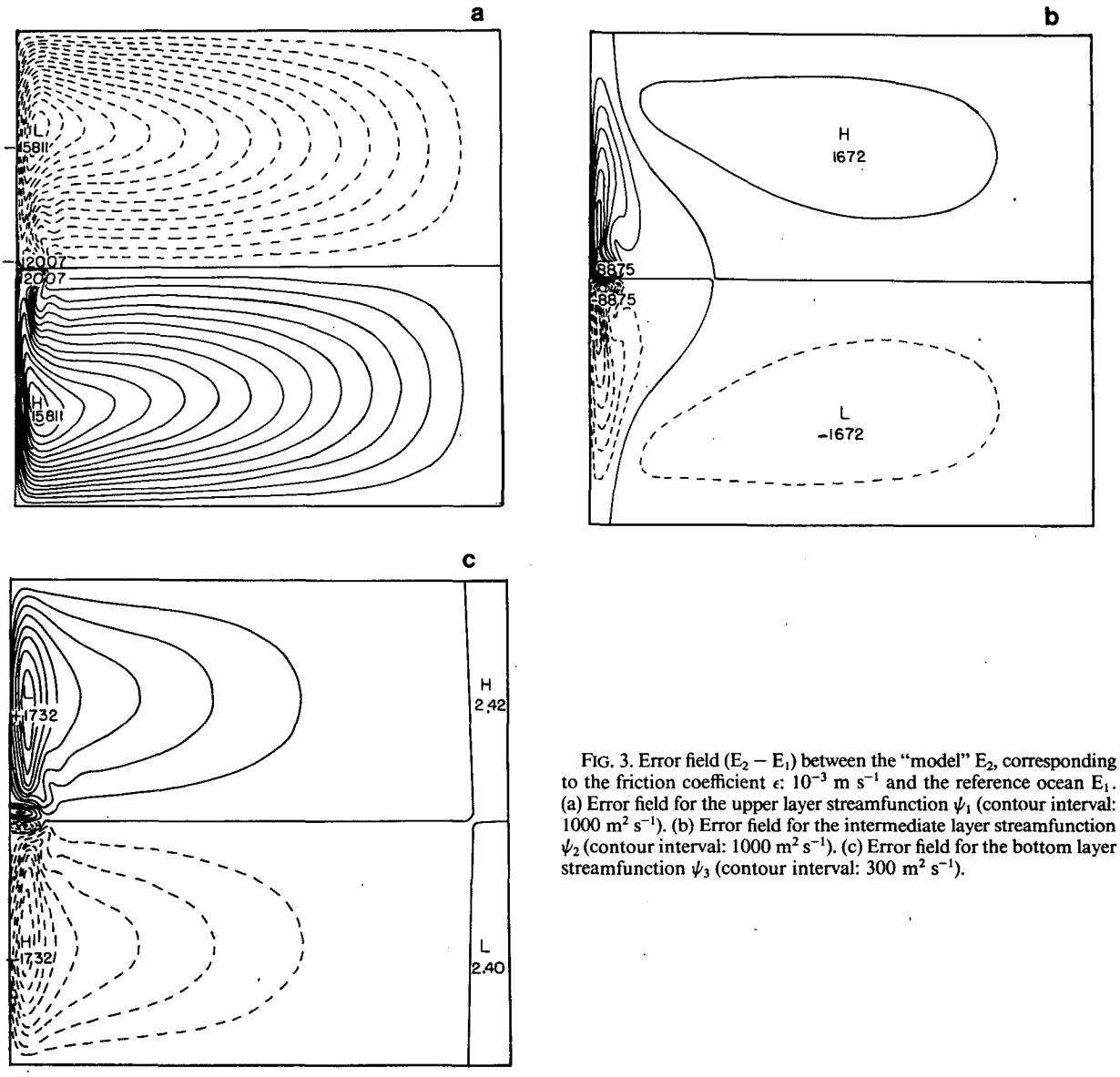

FIG. 3. Error field ($E_2 - E_1$) between the "model" E_2, corresponding to the friction coefficient ϵ: 10^{-3} m s^{-1} and the reference ocean E_1. (a) Error field for the upper layer streamfunction ψ_1 (contour interval: 1000 m^2 s^{-1}). (b) Error field for the intermediate layer streamfunction ψ_2 (contour interval: 1000 m^2 s^{-1}). (c) Error field for the bottom layer streamfunction ψ_3 (contour interval: 300 m^2 s^{-1}).

iments discussed in section 3 and we provide answers to the first two questions addressed in the Introduction. These answers are valid for, and restricted to, the particular kind of steady, highly frictional model physics studied here. Nevertheless, they suggest the direction to be followed in the investigation of more complex simulations.

The first question asked was: how successful is data insertion for improving model estimates of the circulation in different regions of the gyre? The results of the numerical simulations clearly show that the model estimates are significantly improved over the whole region extending from the data insertion section towards the western boundary. The eastern sector, that is the region extending from the data section towards the eastern boundary, is improved only in a narrow, boundary-layer type region near the section itself. This suggests the fundamental role of Sverdrup-type dynamics in which information tends to propagate westward along the characteristics of the interior equations. In the light of the discussion given in section 3, this result means that sections which are long, located some distance from the western boundary, and which tend to be meridional in their orientation are the most effective in improving the model estimates over broad "influence" regions. This type of section would then be preferential in array planning and design. Due to the basic quasi-linearity of the dynamics, nonlinear horizontal advective effects can be recognized only in the western boundary current region, due to the intensity of the western boundary flow. There, horizontal advection produces a typical tonguelike shape in the

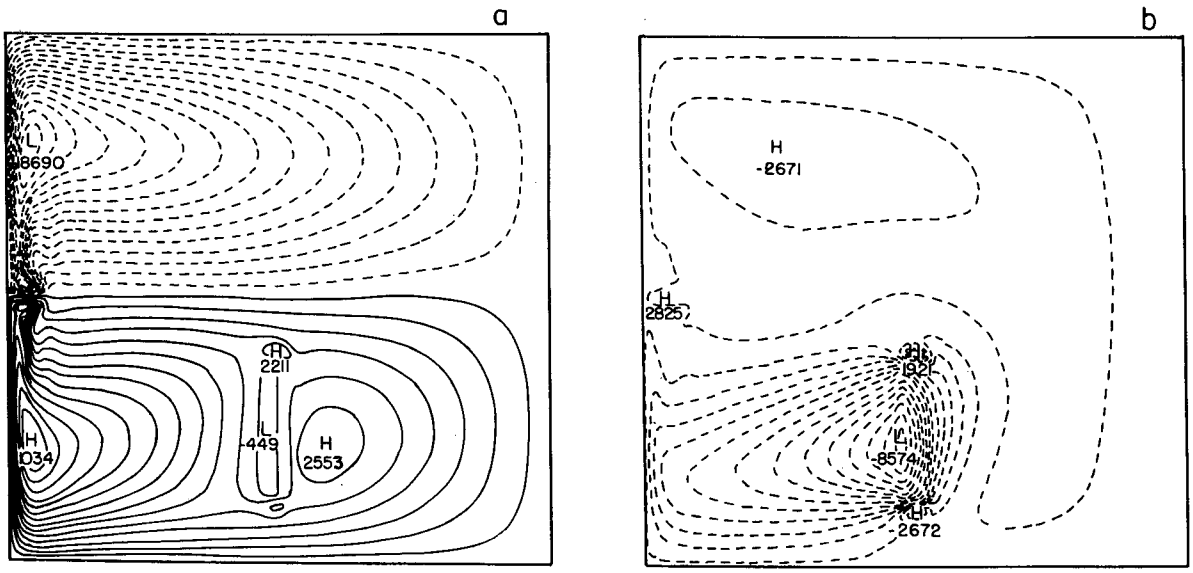

FIG. 4. Diagnostics for the surface layer streamfunction for experiment E_3 of Table 2. (a) Error field ($E_3 - E_1$); contour interval: 1000 $m^2 s^{-1}$. (b) Influence field ($E_3 - E_2$); contour interval: 400 $m^2 s^{-1}$.

influence field, which spreads northward. Advection is however too weak for this narrow tongue to carry the influence of the data insertion back into the gyre interior and sections like that of Experiment E_7 are the most ineffective for our model physics. The important nonlinearity in the present simulations is thus associated with vertical vortex stretching at the two pycnocline interfaces (discussed below).

We can explain the difference in the length scales of the influence field, and consequent effectiveness of the data insertion, between the regions west and east of the data section with the following simple analytical computation. Consider the analogue one-layer problem for the influence field $\psi^{(1)}$, that is steady motion with neglect of horizontal advection of relative vorticity. Neglect also horizontal diffusion of momentum, as the horizontal viscosity induced boundary layers are extremely thin. Then, the equation for a single layer model E_2, without data insertion, is

$$\beta\psi_x^{(2)} = \frac{\text{curl}\tau}{H_1} - \frac{\epsilon}{H_1}\nabla^2\psi^{(2)} \qquad (11a)$$

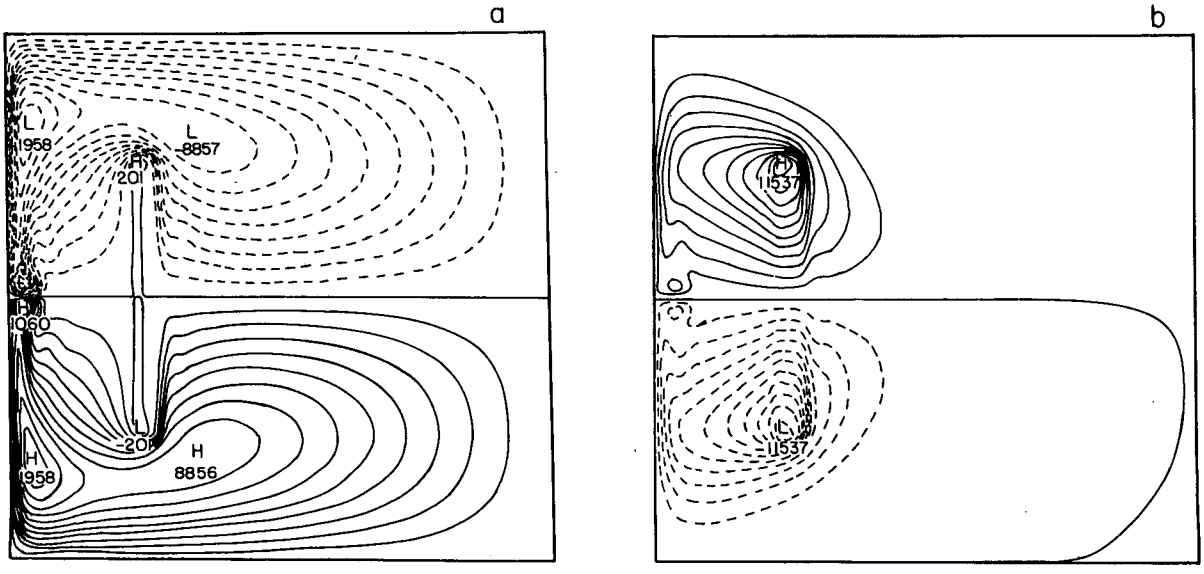

FIG. 5. Diagnostics for the surface layer streamfunction for Experiment E_4 of Table 2. (a) Error field ($E_4 - E_1$); contour interval: 1000 $m^2 s^{-1}$. (b) Influence field ($E_4 - E_2$); contour interval: 1000 $m^2 s^{-1}$.

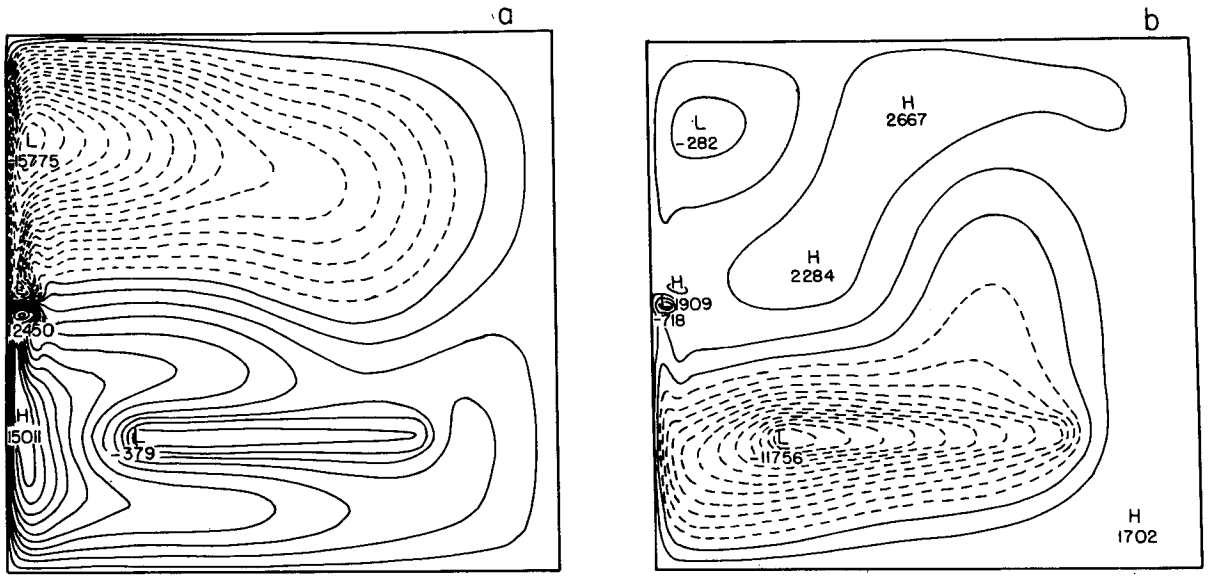

FIG. 6. Diagnostics for the surface layer streamfunction for experiment E_6. (a) Error field ($E_6 - E_1$); contour interval: 1000 $m^2 s^{-1}$. (b) Influence field ($E_6 - E_2$); contour interval: 900 $m^2 s^{-1}$.

where the superscript (2) refers to run E_2 with imperfect physics. $\psi^{(2)} = 0$ along the perimeter of the basin. In an experiment with data insertion E_n the equation will remain identical to (11a), but the streamfunction will have to satisfy the data constraints along the section in the basin interior. Consider specifically problem E_3 in which the data section is meridional in the interior of the subtropical gyre:

$$\beta\psi_x^{(3)} = \frac{\text{curl}\tau}{H_1} - \frac{\epsilon}{H_1}\nabla^2\psi^{(3)} \tag{11b}$$

with $\psi^{(3)} = 0$ along the perimeter of the basin and

$$\psi^{(3)} = \psi^{(1)} \quad \text{at} \quad x = x_0, \quad y_1 \leqslant y \leqslant y_2$$

where $\psi^{(1)}$ are the "data" from the reference ocean E_1, which has a different friction $\epsilon' = 2\epsilon$. The influence field is the difference between the model with data insertion and without:

$$(\text{INFL})_3 = E_3 - E_2.$$

Thus, we define the "influence streamfunction" as

$$\psi^{(I)} = \psi^{(3)} - \psi^{(2)} \tag{12a}$$

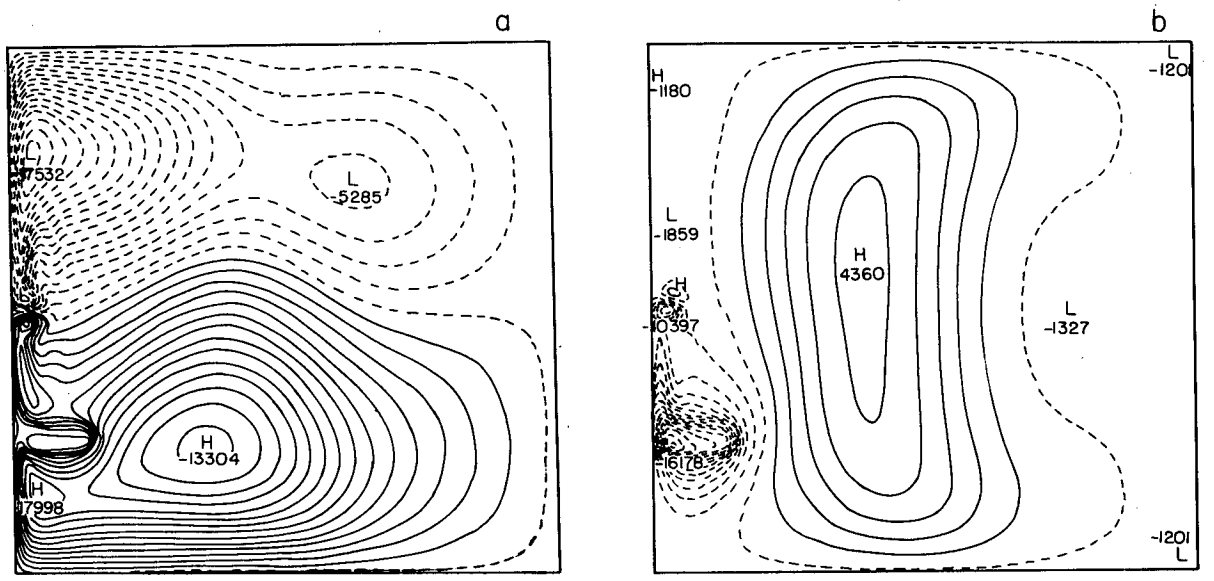

FIG. 7. Diagnostics for the surface layer streamfunction for experiment E_7 of Table 2. (a) Error field ($E_7 - E_1$); contour interval: 1000 $m^2 s^{-1}$. (b) Influence field ($E_7 - E_2$); contour interval: 1000 $m^2 s^{-1}$.

and the influence problem is

$$\beta\psi_x^{(I)} + \frac{\epsilon}{H_1}\nabla^2\psi^{(I)} = 0 \tag{12b}$$

where the boundary conditions

$$\psi^{(I)} = 0 \quad \text{on the perimeter of the basin and}$$
$$\psi^{(I)} = \psi^{(1)} - \psi^{(2)} = F(y) \quad \text{at} \quad x = x_0, \quad y_1 \leq y \leq y_2. \tag{12c}$$

We emphasize that the problem posed by (12a, b, c) is the analogue one-layer problem of the full three-layer model. However, even in the context of the one-layer model, the "circulation" $\psi^{(I)}$ is not zero, but is driven by the internal boundary forcing (12c). Equation (12b) is not a Stommel problem as the interfacial friction $\epsilon = 10^{-3}$ m s^{-1} is important and the two terms have the same order of magnitude even in the basin interior. In fact it is the frictional effect which allows the streamlines for the influence streamfunction $\psi^{(I)}$ to cross the planetary vorticity contours. The data constraints imposed at $x = x_0$ determine the y-shape of $\psi^{(I)}$, as (12b) is separable. Both the reference ocean $\psi^{(1)}$ and the model $\psi^{(2)}$ have a y-structure determined by the shape of

$$\text{curl}\tau = -\frac{\partial\tau}{\partial y} = -T_0\frac{2\pi}{L_y}\sin(2\pi y/L_y)$$
$$\text{with} \quad L_y = 4000 \text{ km}.$$

Thus the interior constraint for the influence field is

$$\psi^{(I)}(x_0) = A(x_0)\sin(2\pi y/L_y) \quad \text{for} \quad y_1 \leq y \leq y_2. \tag{13a}$$

We change the zonal variable to $x' = x - x_0$, so that $-L_1 \leq x' \leq 0$ defines the westward sector and $0 \leq x' \leq +L_2$ the eastward one. From (13a) and variable separation a solution of the type

$$\psi^{(I)} = A(x')\sin(2\pi y/L_y) \tag{13b}$$

is sought and (12b) becomes

$$A_{x'x'} + \frac{\beta H_1}{\epsilon}A_{x'} - (2\pi/L_y)^2 A = 0 \tag{14}$$

with the condition (13a) applied at $x' = 0$. Equation (14) has the solution

$$A(x') = p\exp\left\{-\frac{\beta H_1}{2\epsilon} - [(\beta H_1/2\epsilon)^2 + (2\pi/L_y)^2]^{1/2}\right\}x' + q\exp\left\{-\frac{\beta H_1}{2\epsilon} + [(\beta H_1/2\epsilon)^2 + (2\pi/L_y)^2]^{1/2}\right\}x'. \tag{15}$$

In the westward sector, for the solution to be bounded as $x' \to L_1\ (-\infty)$, $p \approx 0$; in the eastward sector as $x' \to +L_2\ (+\infty)$, $q \approx 0$. Thus

$$A(x')_{\text{westward}} \approx A(0)\exp\left\{-\frac{\beta H_1}{2\epsilon} + [(\beta H_1/2\epsilon)^2 + (2\pi/L_y)^2]^{1/2}\right\}x'$$

$$A(x')_{\text{eastward}} \approx A(0)\exp\left\{-\frac{\beta H_1}{2\epsilon} - [(\beta H_1/2\epsilon)^2 + (2\pi/L_y)^2]^{1/2}\right\}x'. \tag{16}$$

The decay scales (e-folding scales) of the solution in the two sectors are

$$|\Delta x|_{\text{westward}} = \frac{1}{-(\beta H_1/2\epsilon) + [(\beta H_1/2\epsilon)^2 + (2\pi/L_y)^2]^{1/2}} \approx 2600 \text{ km}$$

$$|\Delta x|_{\text{eastward}} = \frac{1}{-(\beta H_1/2\epsilon) - [(\beta H_1/2\epsilon)^2 + (2\pi/L_y)^2]^{1/2}} \approx 160 \text{ km}. \tag{17}$$

Equation (17) shows why the effect of data insertion is much more effective in improving the model estimates from the data location toward the western boundary rather than towards the eastern one. Even though the above computation is linearized, it provides the explanation for the east–west asymmetry of the influence field observed in all the numerical simulations.

There is a further asymmetry, however, in the influence field which is not explained by the above linearized solution, symmetrical in y direction. All the influence fields, in fact, show a considerable north–south asymmetry, for which the preferential spreading of influence streamlines is to the southwestward region in the southern gyre. By explaining the north–south asymmetry we provide also an answer to the second question addressed in the Introduction, that is how the influence field is affected by the dynamics built in the model.

In the above linearized computation we neglected the important nonlinearity of the present simulations, namely vortex stretching. We can attack much more rigorously the influence problem for experiment E_3 considering the steady model in its three-layer version with the vortex stretching term, and neglecting only horizontal friction. The equation for the surface layer streamfunction for the reference ocean E_1 with friction ϵ_1 is then

$$\beta\psi_{1x}^{(1)} = \frac{\text{curl}\tau}{H_1} - \frac{\epsilon_1}{H_1}\nabla^2(\psi_1^{(1)} - \psi_2^{(1)}) + \frac{f_0}{H_1}J(h_{3/2}^{(1)}, \psi_{3/2}^{(1)}) \tag{18}$$

where superscripts indicate the experiment considered and subscripts indicate the level at which the fields are given (Fig. 1a). From the definitions for $h_{3/2}$, $\psi_{3/2}$ (section 2), (18) can be rewritten exactly as

$$\beta\psi_{1_x}^{(1)} = \frac{\text{curl}\tau}{H_1} - \frac{\epsilon_1}{H_1}\nabla^2(\psi_1^{(1)} - \psi_2^{(1)}) - \frac{f_0^2}{g'_{3/2}H_1}J(\psi_1^{(1)}, \psi_2^{(1)}) \tag{18a}$$

which can further be written as

$$\frac{\partial}{\partial y}\left(\beta y+\frac{f_0^2}{g'_{3/2}H_1}\psi_2^{(1)}\right)\psi_{1_x}^{(1)}-\frac{-f_0^2}{g'_{3/2}H_1}\frac{\partial\psi_2^{(1)}}{\partial x}\psi_{1_y}^{(1)}$$

$$=\frac{\text{curl}\tau}{H_1}-\frac{\epsilon_1}{H_1}\nabla^2\psi_1^{(1)}+\frac{\epsilon_1}{H_1}\nabla^2\psi_2^{(1)}. \quad (18b)$$

In (18b) notice that the left-hand side is the quasi-geostrophic approximation of the potential vorticity conservation statement, while the right-hand side gives the sources and sinks of vorticity (wind torque; vertical friction; driving from the lower layer). In the absence of sources and sinks of vorticity, motion in the surface layer would occur along the characteristics for the surface layer equation, given by

$$f_0+\beta y+\frac{f_0^2}{g'_{3/2}H_1}\psi_2^{(1)}=\text{constant}. \quad (18c)$$

The intermediate-layer streamfunction $\psi_2^{(1)}$ thus acts like an effective height and exerts an important steering effect upon the motion of the surface layer in the reference ocean E_1. The characteristics (18c) for the surface layer are shown in Fig. 8.

The equations for the surface layer streamfunction both in the model E_2 and in the model with data insertion E_3 are completely analogous to (18b) and the characteristics again are given by (18c). Thus Fig. 8 gives the shape of the characteristics for the surface layer motion in all the three experiments, E_1, E_2 and E_3.

The influence problem can now be rigorously formulated, using the definition (12a) for the three layers, and the characteristics for the surface layer influence fluid $\psi^{(I)}$ can easily be found to be

$$f_0+\beta y+\frac{f_0^2}{g'_{3/2}H_1}\psi_2^{(3)}=\text{constant}. \quad (19)$$

That is, they are identical to the characteristics in problem E_3. We remind the reader that superscripts indicate the experiment considered E_i, $i=1,2,3$ and subscripts the vertical layer. They too will have the shape shown in Fig. 8. The terms in the rhs of the influence streamfunction equation, analogous to (18b), represent sources and sinks of the influence vorticity, that is the forces which would allow the influence-field circulation to cross the characteristics.

Compare Fig. 8 with the influence field of E_3 shown in Fig. 4b; the latter one shows the same southwestward tendency as the characteristics of Fig. 8. The fact that the shape of the characteristics is so faithfully reproduced by the $\psi_1^{(I)}$-field of Fig. 4b clearly indicates how remarkable and important is the steering effect exerted by the motion in the intermediate layer upon the circulation in the surface one, and the subsequent preferential spreading of information to the southwestward sector. It also shows that the source/sink terms for the influence vorticity are small and do not induce significant crossing of the characteristics.

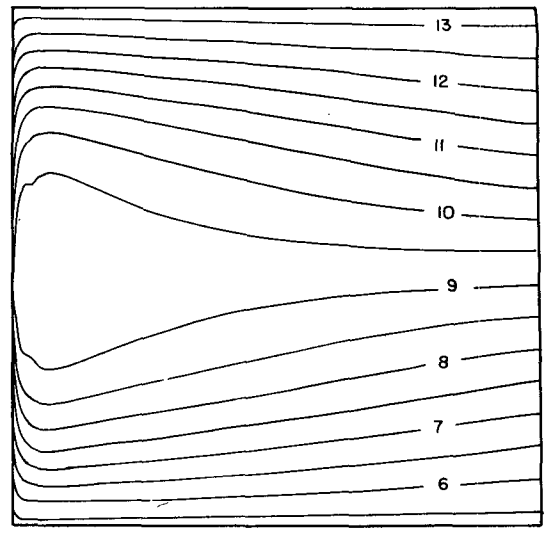

FIG. 8. Shape of the characteristics (18c) for the motion in the surface layer; units are 10^{-5} sec^{-1}.

This steering effect exerted by the effective height $\psi_2^{(3)}$ upon $\psi_1^{(I)}$ dominates also the effect of the frictional term $(-\epsilon/H_1)\nabla^2\psi_1^{(I)}$. In the frictional problem, the east-west frictional decay scale is given by $L\propto\epsilon/\beta$, apart from a proportionality factor. With vortex stretching, however, the effective β is

$$\beta_e=\beta+\frac{f_0^2}{g'_{3/2}H_1}\frac{\partial\psi_2^{(3)}}{\partial y}.$$

Figure 9 shows the intermediate layer streamfunction $\psi_2^{(3)}$ for experiment E_3. In the southern half of the subtropical gyre, where $\psi_2^{(3)}>0$, we have

$$\partial\psi_2^{(3)}/\partial y>0 \quad \text{and} \quad \beta_{e\text{-south}}>\beta.$$

For the northern half:

$$\partial\psi_2^{(3)}/\partial y<0 \quad \text{and} \quad \beta_{e\text{-north}}<\beta.$$

with $L_{\text{north}}\propto\epsilon/\beta_{e\text{-north}}$ and $L_{\text{south}}\propto\epsilon/\beta_{e\text{-south}}$, the above inequalities imply: $L_{\text{north}}>L_{\text{south}}$, that is the frictional decay scale from the data insertion section is greater in the northern than in the southern half of the subtropical gyre. This frictionally induced decay scale, however, would imply a north–south asymmetry which is just the opposite of what is observed in Fig. 4b. On the other hand the steering produced by the effective height $\psi_2^{(3)}$, due to its zonal variation $\psi_2^{(3)}(x)$, produces the pattern of the characteristics of Fig. 8 which gives the correct north–south asymmetry of the influence field of Fig. 4b.

It is thus evident how crucial is the torque exerted by the intermediate-layer motion upon the circulation in the surface layer which dominates even the effects of the high vertical friction. This remarkable feature is likely to persist even in more realistic, fully eddy resolving simulations in which the interfacial eddy stresses would play the role of vertical friction in driving the circulation in the lower layers.

5. Data constraints and model improvement

In this section we address the third proposed question, that is whether we can learn something about what is wrong with the model internal physics and/or external forcing from the effect of data insertion. To do this we make a sensitivity study scanning the two-dimensional parameter space defined by ϵ, the internal friction parameter, and τ, the wind stress amplitude. The problem then is whether it is possible to infer the "correct" values of (ϵ, τ) from data insertion alone.

Consider the one-layer, linearized problem already studied in Section 4. The reference ocean E_1 obeys the equation

$$\beta\psi_x^{(1)} + \epsilon_1\nabla^2\psi^{(1)} = +\tau_1 \sin(2\pi y/L_y). \quad (20a)$$

An "imperfect" model E_2 will in general have a wrong friction, $\epsilon_2 \neq \epsilon_1$, and a wrong wind stress amplitude, $\tau_2 \neq \tau_1$, because of our imperfect knowledge of the actual physics and surface forcing:

$$\beta\psi_x^{(2)} + \epsilon_2\nabla^2\psi^{(2)} = +\tau_2 \sin(2\pi y/L_y). \quad (20b)$$

We now carry out a data insertion experiment analogous to E_3, with a data section crossing the gyre interior in the north–south direction. In addition we consider a single gyre for simplicity. The E_3 streamfunction obeys the equation

$$\beta\psi_x^{(3)} + \epsilon_2\nabla^2\psi^{(3)} = +\tau_2 \sin(2\pi y/L_y)$$
$$\psi^{(3)} = \psi^{(1)} \text{ at } x = x_0; \quad y_1 \leq y \leq y_2. \quad (20c)$$

The influence streamfunction defined by (12a) is given by the solution of the problem

$$\beta\psi_x^{(I)} + \epsilon\nabla^2\psi^{(I)} = 0 \quad (21)$$

with

$$\psi^{(I)} = \psi^{(1)} - \psi^{(2)} \text{ at } x = x_0; y_1 \leq y \leq y_2$$
$$\psi^{(I)} = 0 \quad \text{on lateral boundaries.}$$

The influence problem (21) can be solved analytically. Let

$$\psi^{(1)} = \sin(2\pi y/L_y)\left[-\left(\frac{L}{\pi}\right)^2\frac{\tau_1}{\epsilon_1} + \alpha_1 e^{X_1 x} + \beta_1 e^{X_2 x}\right] \quad (22)$$

be the solution of problem (20a), where $L = L_y/2$;

$$X_1 = -\frac{\beta}{2\epsilon_1} + [(\beta/2\epsilon_1)^2 + (\pi/L)^2]^{1/2}$$
$$X_2 = -\frac{\beta}{2\epsilon_1} - [(\beta/2\epsilon_1)^2 + (\pi/L)^2]^{1/2} \quad (22a)$$

and

$$\alpha_1 = \left(\frac{L}{\pi}\right)^2\frac{\tau_1}{\epsilon_1}\frac{1 - e^{2X_2 L}}{e^{2X_1 L} - e^{2X_2 L}}$$
$$\beta_1 = \left(\frac{L}{\pi}\right)^2\frac{\tau_1}{\epsilon_1} - \alpha_1. \quad (22b)$$

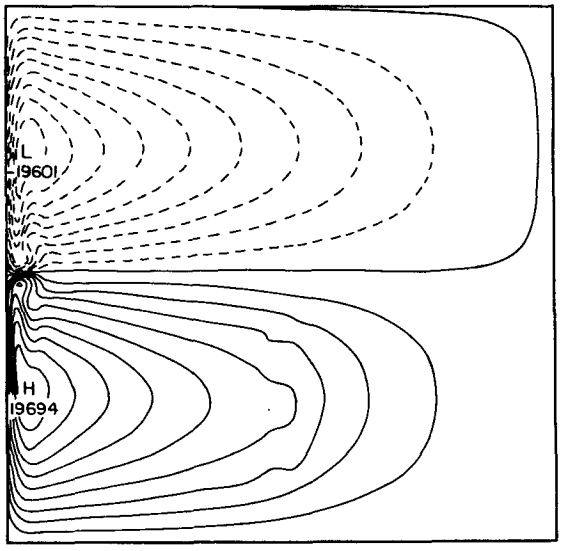

FIG. 9. Intermediate layer streamfunction $\psi_2^{(3)}$ for experiment E_3 of Table 2; contour interval: 2000 $m^2 s^{-1}$.

Similarly

$$\psi^{(2)} = \sin(2\pi y/L_y)\left[-\left(\frac{L}{\pi}\right)^2\frac{\tau_2}{\epsilon_2} + \alpha_2 e^{X_3 x} + \beta_2 e^{X_4 x}\right] \quad (23)$$

is the solution of (20b), with (X_3, X_4) analogous to (X_1, X_2) but with ϵ_2 replacing ϵ_1; and (α_2, β_2) analogous to (α_1, β_1), but with (X_3, X_4) replacing (X_1, X_2).

The solution to the influence problem (21) can be divided into two parts, the western solution $\psi_W^{(I)}$ from the data insertion section to the western boundary, and the eastern solution $\psi_E^{(I)}$ from the section to the eastern boundary. Both solutions are straightforward to evaluate. Thus, we write explicitly only the western solution, the most important one from the point of view of improved model estimates:

$$\psi_W^{(I)} = \sin(2\pi y/L_y)\alpha_{IW}(e^{X_3 x} - e^{X_4 x}) \quad (24)$$

$$\alpha_{IW} = \left[-\left(\frac{L}{\pi}\right)^2\frac{\tau_1}{\epsilon_1} + \alpha_1 e^{X_1 x_0} + \beta_1 e^{X_2 x_0} + \left(\frac{L}{\pi}\right)^2\frac{\tau_2}{\epsilon_2} - \alpha_2 e^{X_3 x_0} - \beta_2 e^{X_4 x_0}\right]\Big/(e^{X_3 x_0} - e^{X_4 x_0}). \quad (24a)$$

In (24), (24a) notice that the dependence upon the external forcing function, the wind stress τ, is linear while the dependence upon the frictional parameter ϵ is much more complex. In fact, if the frictional coefficient were correct, $\epsilon_2 = \epsilon_1$, and if only the wind stress intensity were incorrect, the influence streamfunction in the western sector would be

$$\psi_W^{(I)} = (\tau_2 - \tau_1)\left[\frac{1}{\epsilon_1}\left(\frac{L}{\pi}\right)^2 - \frac{\alpha_1}{\tau_1}e^{X_1 x_0} - \frac{\beta_1}{\tau_1}e^{X_2 x_0}\right] \times \frac{e^{X_1 x} - e^{X_2 x}}{e^{X_1 x_0} - e^{X_2 x_0}}. \quad (25)$$

When $\tau_2 \to \tau_1$, $\psi_W^{(I)} \to 0$. Also in the other, mathematically more complex case with $\tau_2 = \tau_1$, $\epsilon_2 \neq \epsilon_1$ one has $\psi_W^{(I)} \to 0$ as $\epsilon_2 \to \epsilon_1$.

The dependence of the interior response and the associated influence field upon the external driving force will in general be much simpler than the dependence upon the internal physics, like vertical friction in the present case. An external forcing in fact determines the particular solution to the full problem, with time and space scales imposed by the forcing itself. For an external forcing even varying in time but with only one dominant space scale (like the zonal wind stress here considered), the interior response will always depend linearly upon the forcing amplitude and so will the related error, or influence field, even in the context of the fully nonlinear, time-dependent model. An error in $\pm\Delta\tau$, $\pm 2\Delta\tau$, etc., will induce a corresponding error of $\pm\Delta\psi^{(I)}$, $\pm 2\Delta\psi^{(I)}$, etc. The dependence upon the internal physics is, on the other hand, more complicated, as evident from the simple linear, steady case solved here. In general, this dependence will be fully nonlinear, and the $\Delta\psi^{(I)}$ variations induced by an error $+\Delta\epsilon$ will be very different from those induced by $-\Delta\epsilon$, etc. Thus it will be much more difficult to correct the parameterizations of the model internal physics than to correct imperfect estimates of the external forcing.

We now vary systematically the point representing our system in the (ϵ, τ) phase space. When $(\epsilon_2, \tau_2) \to (\epsilon_1, \tau_1)$, the influence field $\psi^{(I)} \to 0$; the model $\psi^{(2)}$ converges to the reference ocean $\psi^{(1)}$ as well as the ocean with data insertion $\psi^{(3)} = \psi^{(2)} + \psi^{(I)}$. The question is then whether the above behavior holds for the unique point of "correct" values (ϵ_1, τ_1).

We first make the two cuts in the (ϵ, τ) phase space. In the first, the wind stress amplitude τ is right and kept constant; ϵ is wrong and varied. In the second, the friction coefficient ϵ is right and kept constant, τ is wrong and varied. Figure 10 shows the streamfunction $\psi^{(1)}$ for the reference ocean defined by the values $\tau_1 = 1.0 \times 10^{-4}$ m^2 s^{-2}; $\epsilon_1 = 2 \times 10^{-3}/H_1 = 0.6667 \times 10^{-5}$ s^{-1}. Figure 11a shows the model streamfunction $\psi^{(2)}$ given by (23) when the friction coefficient is wrong, $\epsilon_2 = 0.5000 \times 10^{-5}$ s^{-1}. Figure 11b shows the plot of the analytical $\psi^{(I)}$ field solution of (21), with a (negative) maximum of -1×10^4 m^2 s^{-1}, and Fig. 11c shows the model with data insertion $\psi^{(3)} = \psi^{(2)} + \psi^{(I)}$, in which two gyres are artificially created. The spurious steplike discontinuities are introduced by the boundary conditions on $\psi^{(I)}$ at the lateral boundaries. We now increase the friction. With $\epsilon_2 = 0.6333 \times 10^{-5}$ s^{-1}, "almost" right, Figs. 12a, b, c show, respectively, the model $\psi^{(2)}$, $\psi^{(I)}$ which converges to zero, with a negative maximum of -1×10^3 m^2 s^{-1}, and the $\psi^{(3)}$ field, model with data insertion, which is now very smooth and almost identical to the $\psi^{(1)}$ of Fig. 10. A completely similar behavior is observed when $\epsilon_2 = \epsilon_1$ and τ_2 is varied until approaching τ_1.

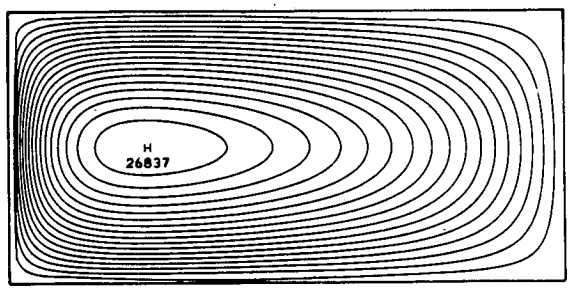

FIG. 10. Streamfunction $\psi^{(1)}$ for the "reference" ocean for the one-layer model defined by (20a). Contour interval = 1500 m^2 s^{-1}. Reference ocean values $\tau_1 = 1.0 \times 10^{-4}$ m^2 s^{-2}, $\epsilon_1 = 2 \times 10^{-3}/H_1$ s^{-1} $= 0.6667 \times 10^{-5}$ s^{-1}.

To synthesize the information contained in the above results, a norm can be defined for the error provided by the influence field $\psi^{(I)}$. Two natural measures for the error are

$$\pm|\psi^{(I)}_{\max}| \tag{26a}$$

$$\psi^{(I)}_{\text{rms}} = \frac{1}{N^2}\left[\sum_{i=1}^{N}\sum_{j=1}^{N}|\psi_{ij}^{(I)}|^2\right]^{1/2} \tag{26b}$$

(26a) is the maximum value of $\psi^{(I)}$, that is the maximum error with its sign, defining over-or-underestimation of the circulation. Equation (26b) is the root-mean-square value of the error evaluated over the whole domain of definition.

We first study the two one-dimensional cuts in (ϵ, τ) space. Figures 13a, b show the dependence of $\pm|\psi^{(I)}_{\max}|$ upon the variation of the "wrong" parameter. Specifically, Fig. 13a shows the variation in the τ-space when the parameterization of the internal physics is correct, $\epsilon = \epsilon_1$. Figure 13b shows instead the variation in the ϵ space when the amplitude of the external forcing function is correct, $\tau = \tau_1$. Figures 14a, b show the corresponding variations in $\psi^{(I)}_{\text{rms}}$ as a function of τ with $\epsilon = \epsilon_1$ (Fig. 14a), and as a function of ϵ with $\tau = \tau_1$ (Fig. 14b). In these one-dimensional scans both measures of the error go to zero at a unique point, the correct parameter value. Notice also the linear dependence of the error upon the wind intensity, τ (Figs. 13a and 14a), as shown explicitly by (25). The dependence on the internal friction ϵ (Figs. 13b and 14b) on the other hand is nonlinear even in this most simple example. A variation $\Delta\epsilon = -30 \times 10^{-8}$ s^{-1} produces a rms $\Delta\psi^{(I)}$ twice as big as that produced by $\Delta\epsilon = +30 \times 10^{-8}$ s^{-1} (Fig. 14b).

Consider now the two-dimensional scan of the (ϵ, τ) space, in which we vary simultaneously ϵ and τ. Figures 15a, b show the two-dimensional plots of $\pm|\psi^{(I)}_{\max}|$ and $\psi^{(I)}_{\text{rms}}$ respectively, in the (ϵ, τ) space. As is most clear in Fig. 15a the locus of points where $\pm|\psi^{(I)}_{\max}| \to 0$ is a straight line, obviously containing the correct point $\epsilon = \epsilon_1$; $\tau = \tau_1$. This is clear also from the analytical shape of $\psi_W^{(I)}$, and correspondingly $\psi_E^{(I)}$, given by (24). From (24), $\psi_W^{(I)} = 0$ when $\alpha_{IW} = 0$. From (24a), the condition $\alpha_{IW} = 0$ shows explicitly the linear dependence of τ_2 upon ϵ_2. Thus, in the two-dimensional pa-

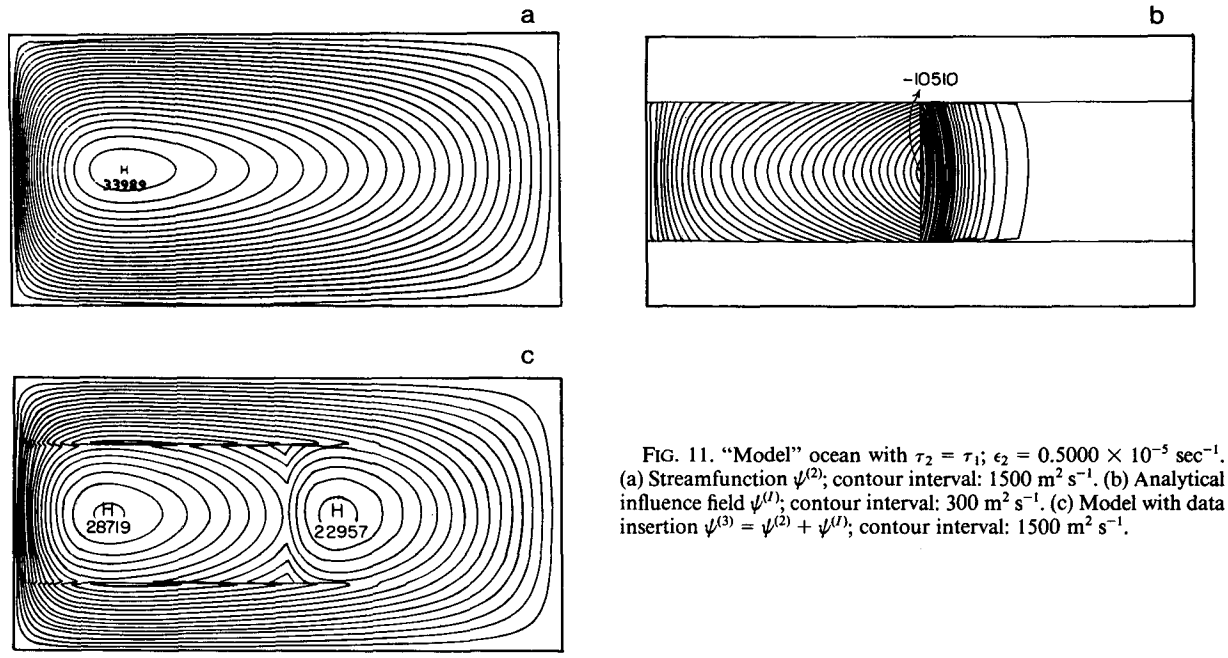

FIG. 11. "Model" ocean with $\tau_2 = \tau_1$; $\epsilon_2 = 0.5000 \times 10^{-5}$ sec^{-1}. (a) Streamfunction $\psi^{(2)}$; contour interval: 1500 m^2 s^{-1}. (b) Analytical influence field $\psi^{(I)}$; contour interval: 300 m^2 s^{-1}. (c) Model with data insertion $\psi^{(3)} = \psi^{(2)} + \psi^{(I)}$; contour interval: 1500 m^2 s^{-1}.

rameter space (ϵ, τ) and for the linear, one-layer model here considered the locus of points where the error field becomes zero is not the unique point (ϵ_1, τ_1) but rather a one-dimensional curve. It is natural to extrapolate that in a N-dimensional parameter space a multidimensional surface [$(N-1)$ dimensions for a linear model] will be in general such a locus. This nonuniqueness of the solution where the error $\psi^{(I)} \to 0$, and consequent multiplicity of (a priori) equally acceptable $\psi^{(3)}$ circulation patterns, implies that a brute-force scanning of the parameter space defined by the uncertain model parameters would not provide the "correct" solution. A quantitative criterion must instead be formulated to pick out from the ensemble of possible solutions the most probable one. A possible criterion might be, for instance, to construct the influ-

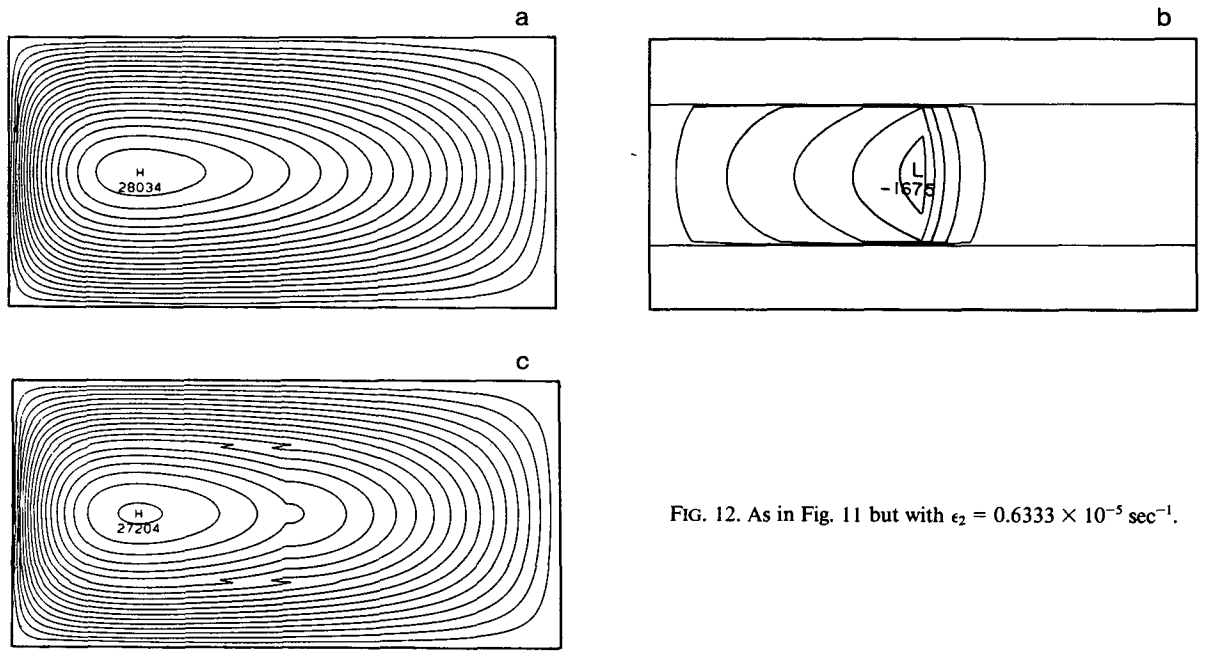

FIG. 12. As in Fig. 11 but with $\epsilon_2 = 0.6333 \times 10^{-5}$ sec^{-1}.

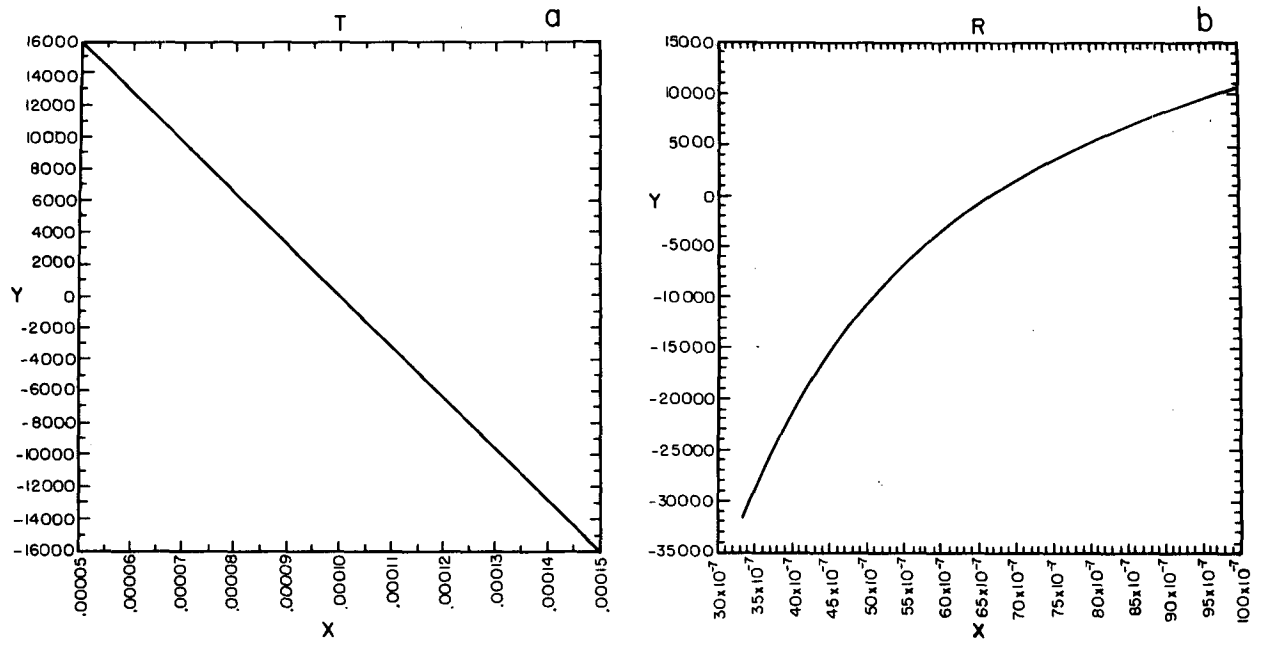

FIG. 13. One-dimensional dependence of the error norm (26a) upon the model physical parameter which is imperfectly known. (a) Dependence upon the wind intensity τ with $\epsilon = \epsilon_1$, the true value and (b) Dependence upon the vertical friction ϵ, with $\tau = \tau_1$, the true value.

ence field in which the errors in all the model's param-

Also, the above approach and discussion capitalizes on having perfect data, that is a complete knowledge of the "true" ocean, our reference ocean E_1. In reality not only will the model be imperfect but the data will contain errors, and the "true" ocean will never be known. At best, approximate values with error bounds will be available for the data to be inserted locally. The necessity of establishing an optimization procedure, or

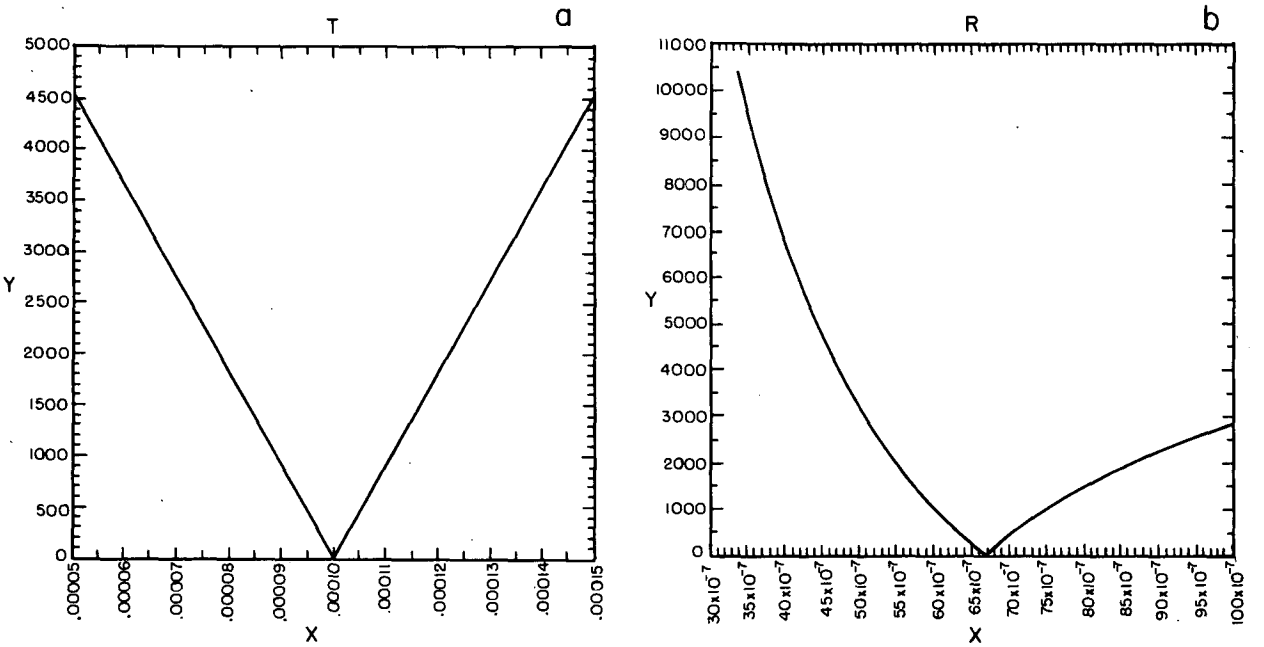

FIG. 14. As in Fig. 13 but for the error norm $\psi^{(I)}_{\text{rms}}$ defined by (26b). (a) Dependence upon τ with $\epsilon = \epsilon_1$ and (b) Dependence upon ϵ with $\tau = \tau_1$.

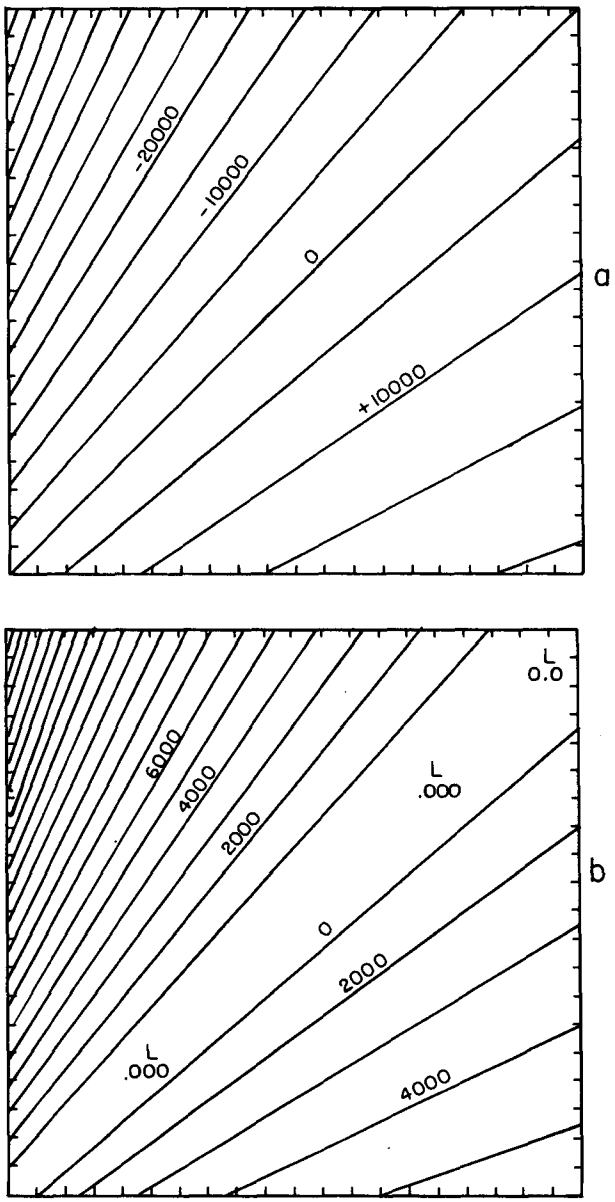

FIG. 15. Two-dimensional plots of the error norm in the (ϵ, τ) space (a) Error norm $\pm|\psi^{(I)}_{max}|$ and (b) Error norm $\psi^{(I)}_{rms}$ defined by (26b).

the inverse, has emerged already from studying the previous one-layer, linear case with imperfect model/perfect data. In the real case, with the model ocean in the nonlinear parameter range, a nonlinear inverse would be performed providing not the "true" solution (unknown) but picking out from the ensemble of possible solutions the ones consistent with the assigned data uncertainties. The data insertion process would then give not the "correct" values (ϵ_1, τ_1) (unknown) but rather a "best" estimate $\hat{\epsilon} \pm \Delta\epsilon$; $\hat{\tau} \pm \Delta\tau$ with error bounds. The simple example studied in this section is only a first step in this direction.

6. Conclusions

In this paper we have begun to explore how constraining a model calculation by inserting observations known only locally can lead to more realistic results regarding large scale ocean circulation patterns. The reasons for such a study are obvious—high quality temporal and spatial observations of the interior structure of the oceans are likely to be confined to a local region or to vertical sections for the foreseeable future. Therefore an understanding of the techniques of constraining models and the nature of the improvement of model derived estimates are important.

We have shown, for a simple part of the relevant parameter space (steady, highly frictional, nearly linear), that data constraints imposed in models lead to improvements which depend upon the location of the data in the gyre, the orientation of the observed section, and the length of that section. This has important implications for array design, especially for observational programs that have a strong modeling bias. Undoubtedly the results discussed here will need elaboration and refinement as more realistic, eddy-resolved ocean circulation experiments are examined. The nearly westward propagation of information found in these experiments will only be a part of the picture, even though important, when advection and eddy propagation effects become significant. In addition it should be kept in mind that only temperature information has been used thus far. Is this the most important internal variable? What about the barotropic mode or sea surface height or velocity information throughout the water column? The answers to these questions probably depend upon the nature of the problem being examined (e.g., steady circulation or transiently forced circulation or regions rich in mesoscale eddies). Therefore there is much work to do to unravel the important elements of inserting data into models.

Given this limitation, that the results found here are only a first step toward understanding data insertion questions, we have found some specific results that will carry over in some form to more realistic model oceans. In particular the southwestward propagation of information is likely to be an important factor also in fully eddy-resolving simulations. In this case, in fact, the interfacial eddy Reynold stresses would play the role here assigned to vertical friction and drive the circulation in the deeper layers. The mean circulation in the intermediate layer, in its turn, would act like an effective height for the upper layer and exert an important steering effect on the mean motion in that upper layer, with characteristics bending to the southwest in the subtropical gyre.

Advective effects here are minimal but a slight "downstream" spread of information shows up in the boundary currents nevertheless. We expect this effect to be much more important for cases without such large vertical friction. Note also that, as explicit friction is made less important, transients due to instabilities

will become more important and this will add another 'agent' for the horizontal spread of information.

Finally, a major goal is to learn how to improve our models, that is to make better choices of the poorly known parameters that govern the physics of the model, by using data insertion to determine the amplitude and region of influence. Given the lack of knowledge of how to parameterize small scale processes in the ocean and given the rather large uncertainty in nearly all boundary conditions (wind stress, heat and fresh water flux, inflow values across open boundaries) and in the light of the results of Section 5, a general procedure will certainly be needed. The brute force approach of running many numerical experiments is too costly to warrant the effort and, moreover, would not provide a unique answer. In a multidimensional parameter space and in more complex situations, an optimization criterion must be established and a nonlinear inverse carried out to pick out the optimal model estimates among the multiplicity of possible solutions. However it is clear that the modeling community needs a strategy for systematic improvement of models through data insertion, a strategy that can result, in the long run, in models that faithfully reproduce the observed characteristics of the oceanic circulation.

Acknowledgments. This research was carried out with the support of the National Science Foundation Grant OCE-8214821.

REFERENCES

Bennett, A. F., 1985: Array design by inverse methods. *Progress in Oceanography,* Vol. 15, Pergamon, 129–156.

Chow, J. C., and W. R. Holland, 1986: A multilayer quasigeostrophic numerical model of ocean circulation. In preparation.

Cornuelle, B., and Collaborators, 1985: Tomographic maps of the ocean mesoscale. Part I: Pure acoustics. *J. Phys. Oceanogr.,* **15,** 133–152.

Ghil, M., M. Halem and R. Atlas, 1979: Time-continuous assimilation of remote-sensing data and its effect on weather forecasting. *Mon. Wea. Rev.,* **107,** 140–171.

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., 1985: Long-range inversions for ocean acoustic tomography. *J. Geophys. Res.,* **90,** 7098–7116.

——, and W. R. Holland, 1985: Gyre scale acoustic tomography: modeling simulations. *J. Phys. Oceanogr.,* **15,** 416–438.

Ocean Acoustic Tomography Group, The, 1982: A demonstration of ocean acoustic tomography. *Nature,* **299,** 121–125.

Schroter, J., and C. Wunsch, 1986: Solution of non-linear finite-difference ocean models by optimization methods with sensitivity and observational strategy analysis. *J. Phys. Oceanogr.* (In press).