Assimilation of Altimeter Data into a Quasigeostrophic Model of the Gulf Stream System. 
Part I: Dynamical Considerations

ANTONIETTA CAPOTONDI

PAOLA MALANOTTE-RIZZOLI

WILLIAM R. HOLLAND
Assimilation of Altimeter Data into a Quasigeostrophic Model of the Gulf Stream System.  
Part I: Dynamical Considerations

ANTONIETTA CAPOTONDI*  
MIT-WHOI Joint Program, Cambridge, Massachusetts

PAOLA MALANOTTE-RIZZOLI  
Massachusetts Institute of Technology, Cambridge, Massachusetts

WILLIAM R. HOLLAND  
National Center for Atmospheric Research, Boulder, Colorado  
(Manuscript received 21 June 1993, in final form 14 September 1994)

ABSTRACT

The dynamical consequences of constraining a numerical model with sea surface height data have been investigated. The model used for this study is a quasigeostrophic model of the Gulf Stream region. The data that have been assimilated are maps of sea surface height obtained as the superposition of sea surface height variability deduced from the Geosat altimeter measurements and a mean field constructed from historical hydrographic data. The method used for assimilating the data is the nudging technique. Nudging has been implemented in such a way as to achieve a high degree of convergence of the surface model fields toward the observations. The assimilation of the surface data is thus equivalent to the prescription of a surface pressure boundary condition. The authors analyzed the mechanisms of the model adjustment and the characteristics of the resultant equilibrium state when the surface data are assimilated. Since the surface data are the superposition of a mean component and an eddy component, in order to understand the relative role of these two components in determining the characteristics of the final equilibrium state, two different experiments have been considered: in the first experiment only the climatological mean field is assimilated, while in the second experiment the total surface streamfunction field (mean plus eddies) has been used. It is shown that the model behavior in the presence of the surface data constraint can be conveniently described in terms of baroclinic Fofonoff modes. The prescribed mean component of the surface data acts as a "surface topography" in this problem. Its presence determines a distortion of the geostrophic contours in the subsurface layers, thus constraining the mean circulation in those layers. The intensity of the mean flow is determined by the inflow/outflow conditions at the open boundaries, as well as by eddy forcing and dissipation.

1. Introduction

Numerical models of the ocean circulation have undergone a considerable development in the last 30 years. In fact, several steps toward a higher degree of realism have been undertaken since the first idealized numerical studies of westward intensification (Bryan 1963; Veronis 1966; Holland 1967). These studies were performed with barotropic models of coarse horizontal resolution in a rectangular domain. A first fundamental improvement in model "realism" has been the increased horizontal and vertical resolution. The higher horizontal resolution has allowed the inclusion of the mesoscale eddy field and its interactions with the mean flow, processes previously parameterized often in a crude fashion.

A further step in improving the model realism has then been the inclusion of a realistic geometry for studying the circulation in specific areas of the World Ocean. Models of this type can incorporate the effect of the actual coastline and bottom topography in the area of interest and allow a more straightforward geographical correspondence between the model circulation and the observed ocean circulation.

The present state of the art in ocean modeling is thus represented by eddy-resolving models in a realistic domain. Eddy-resolving models describe the evolution and statistical equilibrium of a turbulent ocean. Therefore, the general circulation of these models includes
both characteristics of the mean circulation and eddy statistics.

In the context of this paper, the term eddies is used to define deviations of the model fields from the temporal mean, thus including ringlike structures as well as meanders and changes in the stream position.

Comparisons of some aspects of the model climatology with observations have shown the sensitivity of the model behavior to the geometry, frictional parameterizations, and boundary conditions. These factors seem to affect, in an interdependent and complicated fashion, the delicate internal processes of eddy–mean flow interactions. For this reason models often fail in reproducing some basic characteristics of the ocean circulation. A typical example of this difficulty is the Gulf Stream area, where features like the separation of the stream from the coast, its mean path downstream to the Grand Banks, the mean eddy kinetic energy level and distribution, as well as the southern and northern recirculations, are notoriously very difficult to simulate accurately. Recent efforts in the modeling of the Gulf Stream region (Thompson and Schmitz 1989; Ezer and Mellor 1992) have partially succeeded in realistically reproducing the stream separation and the eddy kinetic energy level. However, they have also shown the sensitivity of the model circulation to several interdependent factors, such as the lateral boundary conditions and the inclusion of thermodynamic forcing (Ezer and Mellor 1992), as well as the presence of a deep western boundary current (Thompson and Schmitz 1989) and horizontal resolution (Schmitz and Thompson 1993).

Although models can be “tuned” to be realistic in some of their characteristics, a clear understanding of all the physical mechanisms in play, as well as the achievement of perfectly accurate simulations, are still to be considered as work in progress. For the above reasons, one can consider the possibility of using data assimilation techniques to improve the model climatology, including characteristics of both the mean field and the eddy statistical properties. Data assimilation is a relatively new topic in oceanography, relative to the long-term experience developed in meteorology. In fact, it is only with the advent of the ocean satellite missions that datasets with a large coverage over synoptic timescales have become available to oceanographers. The limitations that ocean models have in their climatological characteristics as described above suggests the use of these datasets not only for prediction purposes, as has been the case in meteorology, but also as constraints to the model behavior from a climatological point of view. The datasets obtained from satellite missions contain information only about the ocean surface. Therefore, from the data point of view, numerical models can be seen as “dynamical extrapolators” of this surface information to areas (e.g., the deep ocean) where data are not available. The goal is to achieve, through this blending of data and models, a better description of the complete, four-dimensional ocean circulation.

A considerable amount of work has already been done to test different assimilation techniques and to determine how effective they are in constraining numerical models. Most of these investigations have been carried out in the context of the so-called twin experiments, in which the data assimilated are produced by the model itself, thus allowing easy verification of the degree of success of the data assimilation process. The use of real data represents a further step and poses several new issues that have not yet been fully examined. In particular, due to the present state of ocean models as described above we need to understand the dynamical implications involved in the process of combining data and models that have somewhat different statistical characteristics in their climatologies.

In this study we address these issues using a quasigeostrophic (QG) model of the Gulf Stream area. The data we assimilate are maps of sea surface height (SSH), which have been obtained as the superposition of SSH variability deduced from the Geosat altimeter measurements and a mean field constructed from historical hydrographic data. The surface data have been assimilated by using the “nudging” technique, a technique implemented in such a way as to achieve a high degree of convergence of the model surface field toward the observations.

A preliminary experiment of this type has been presented by Capotondi and Holland (1993, henceforth referred to as CH). They have shown the effectiveness of this surface data constraint in altering the global model behavior. A preliminary comparison with some available observations seemed to indicate an improvement in the degree of realism of the model climatology. In the present study a deeper analysis of the surface data assimilation process is carried out in order to rationalize the assimilation results from a dynamical viewpoint. Here in Part I we address the following question: How does the model respond to the surface data constraint; that is, what are the physical processes responsible for the model adjustment and for the final equilibrium state when surface data are assimilated? In particular, since the surface data have been obtained as the superposition of a time-mean component and a time-dependent component, we want to understand the relative contribution of these two components in determining the final results. The aim is to obtain a physical understanding of the assimilation process so that our findings can be generalized to other assimilation experiments and ways of improving the assimilation procedure can be devised. In Part II we will try to assess the degree of success of the assimilation procedure in improving the model climatology by considering comparisons with different types of observations.

The emphasis of the present study is in understanding the model response to the prescription of a surface streamfunction boundary condition or, equivalently,
of a surface pressure boundary condition. Therefore, in this initial investigation we adopt the approach of assuming a complete and uniformly accurate knowledge of the surface streamfunction fields. In the language of estimation theory, this study might be defined as a study "in control". To what extent can a surface pressure boundary condition control the dynamical system under consideration and why?

The presentation is organized as follows: in section 2 we describe the model that has been used for the experiments, in section 3 we discuss the data, and in section 4 we formulate the mathematical framework for understanding the model adjustment and final equilibrium state when the surface streamfunction is constrained to follow a prescribed field. The way a solution to this problem can be achieved and the characteristics of such a solution are illustrated with the aid of an analytical example in section 5. The assimilation experiments are described in section 6. Special emphasis is devoted to verify the applicability of this theoretical framework to the actual problem we are considering. To identify the relative effect of the mean and eddy components of the surface data in determining the characteristics of the model behavior, we present in this section two different assimilation experiments: in the first one only the surface mean field is assimilated, while in the second experiment the total surface streamfunction field (mean plus eddies) is used. We conclude with a discussion in section 7.

2. The numerical model

The model is based upon the closed-basin quasigeostrophic model first discussed by Holland (1978) and now used in many basic studies of eddy-resolved ocean circulation. A review of some of those studies is given in Holland (1986). A description of the physics of that model is given in appendix A. The model has five layers in the vertical with layer thicknesses of 300, 450, 750, 1300, and 2200 m, respectively, from top to bottom. The values chosen for the reduced gravity at each of the four internal interfaces are \(1.222 \times 10^{-2}\), \(1.211 \times 10^{-2}\), \(9.448 \times 10^{-3}\), and \(4.368 \times 10^{-3}\) m\(^2\) s\(^{-1}\), respectively, from top to bottom. The QG streamfunctions for the velocity fields are considered to be at the center of each layer, and the transport in each layer is given by the streamfunction times its layer thickness. The horizontal resolution is \(\frac{1}{8}\) degree of latitude and longitude, a resolution necessary for resolving the turbulent processes in the model (Schmitz and Thompson 1993). The bottom friction coefficient is \(10^{-8}\) s\(^{-1}\), corresponding to a spindown time of approximately 3.2 years. The coefficient of biharmonic friction is \(2 \times 10^{10}\) m\(^4\) s\(^{-1}\). This value corresponds to a decay time of about 22 days at the length scale of the model grid and to approximately 160 years at length scales of the order of 100 km. The model has a realistic coastline, as can be seen in Fig. 1. As is well known, the QG approximation is valid only with relatively small amplitude variation in bottom topography. The accuracy of the model simulation in the case of finite topography, such as that existing in some parts of the Gulf Stream area, is not known. We have chosen to exclude from this initial investigation the uncertainties associated with the topographic issue so that the effects of data insertion can be isolated from the effects associated with bottom topography. A flat bottom, therefore, is chosen. The regional nature of the model and its utilization for studying an actual piece of the real ocean require that the lateral open boundary conditions, as well as surface forcing conditions, be carefully considered. As the QG model is not very useful for examining local thermo-
haline forcing, those aspects of the physical problem will not be considered here. For that purpose a PE model will be needed. The surface forcing, therefore, is given only by the wind stress. Hellerman's annual mean climatological winds have been used (Hellerman and Rosenstein 1983).

The lateral boundary conditions have been constructed so that the inflows and outflows are confined to the three upper layers. The mean streamfunction fields used for the assimilation experiments have been derived from the climatological data of temperature and salinity collected by Bauer and Robinson as explained in the next section. They are shown in Fig. 1. For consistency the transport at the open boundaries in the three upper layers has been derived from the same climatological data in all the experiments considered in this study. The Bauer–Robinson dataset does not resolve the jet entering at the western boundary. The latter has been determined by mass conservation from the interior geostrophic flow. The total transport of the Gulf Stream at the western boundary corresponds to the observed value of about 52 Sv (Sv = 10^6 m^3 s^-1), which has been measured off Cape Fear by Richardson et al. (1969). This is approximately the location where 80°W intersects the North American coast. Thus, the Gulf Stream enters the domain as a western boundary current somewhat southwest of Cape Hatteras. Another problem with the Bauer–Robinson dataset is the missing data in the area of the northern recirculation gyre, including the model's northern boundary. We have chosen to neglect the inflow through the northern boundary associated with the northern recirculation gyre, since this could not be determined from the data. This choice leads to the development of unrealistic features of the mean model circulation in all experiments, with and without assimilation, as discussed later.

The inflows and outflows determined from the climatological data at the northern, eastern, and southern boundaries are thus the following: a branch of the stream exits the northern boundary east of 43°W with a transport of 7 Sv; the southern boundary intersects the subtropical gyre of the climatological data (see Fig. 1); a transport of 9 Sv enters the domain west of 63°W; while 15 Sv leave the domain east of the same longitude. Most of the outflow (39 Sv) occurs at the eastern boundary. At this boundary the model streamfunction is relaxed toward the climatological values within a sponge layer of 12 grid points. This allows some degree of meandering of the model outflow around the outport defined by the climatological values. To prevent reflection of wave energy, a radiation boundary condition is also implemented at the eastern boundary. Additional details and discussion about the treatment of the open boundary conditions are given in Capotondi (1993). The mean circulation produced by the model when the above boundary conditions are used is shown in Fig. 2. The three upper layers of the model are initialized with the climatological fields in Fig. 1, while layers 4 and 5 are initially motionless. The temporal mean has been defined over a 4-year period. The mean circulation developed by the model in the three upper layers deviates noticeably from the initial conditions. In the model solution the Gulf Stream leaves the coast north of Cape Hatteras and then flows eastward almost zonally. The broad nature of the eastern boundary outflow, as well as the lack of a northern recirculation gyre inflow, seems to be responsible for the flow tendency to “fill” the northern half of the domain. This is consistent with the “diffuse separation” found by Ezer and Mellor (1992) when no slope water inflow is prescribed at their northeastern boundary. It is also consistent with the sensitivity experiments performed by Holland and Schmitz (1989, unpublished manuscript) with a quasigeostrophic model similar to the one used for this study. At the northern boundary a “Fofonoff-type” gyre is formed. These features are clearly unrealistic and suggest a reexamination of the boundary conditions used. However, although improving ocean models is a necessary, never-ending process, this is beyond our purpose here. The aim of this study is to analyze how effective assimilation of surface data can be in modifying the behavior of a given model and why.

3. The data

The absolute sea surface topography, which will be used to constrain the surface model fields in the assimilation experiments discussed below, has been constructed by obtaining the sea surface height variability from the Geosat altimeter and the mean SSH from climatological temperature and salinity data. In fact, as is well known, the geoid error in the altimeter measurements is larger than the oceanic signal that we want to measure; so to eliminate this steady error, the temporal mean is subtracted from the altimeter measurements.

The Geosat dataset has been supplied by the oceanographic group at the Jet Propulsion Laboratory. It consisted of 26 months of data, spanning the period from 7 November 1986 to the end of December 1988. The raw data consist of observations made by the altimeter along the satellite tracks at the times of the satellite passage, approximately every 17.05 days. The temporal mean has been subtracted from the data and the standard corrections (ionospheric and atmospheric path delay, inverted barometer, ocean tides, removal of blunder points, orbit error corrections, and time-mean removal) have been applied. See Zlotnicki et al. (1989) and Holland et al. (1991) for a discussion of these errors. The resulting signals are referred to as “sea surface height variability” or “eddy fields” throughout this paper.
The approach adopted in this study is to consider the model response when complete surface information is supplied. Therefore, a statistical interpolation of the data onto the model grid at fixed time intervals has been performed. The algorithm used for the statistical interpolation, which has been carried out in both space and time, is the successive corrections method formerly used in meteorology (Tripoli and Krishnamurti 1975). The description of the algorithm and the considerations that led to the parameter choices are given in appendix B.

The final interpolated dataset consists of a series of SSH maps at 2-day intervals, spanning the period November 1986 to May 1988. During the last seven months of the Geosat mission (June 1988 to December 1988), the data coverage degrades considerably. In fact, not only are the descending tracks missing in most of the domain but the ascending tracks are also very often absent, leading to large areas without altimetric information. Therefore, the data after May 1988 have not been used.

The time interval of two days has been chosen because we want to assimilate data continuously in time. Timescales shorter than a few days do not seem to be relevant in the model dynamics (Holland and Malanotte-Rizzoli 1989) so that we can interpolate linearly between our bidaily maps to obtain data to assimilate at every time step (one hour) of our model run. The characteristics of the eddy field that emerge from this interpolated dataset are discussed in Part II.

The mean streamfunction field in the upper layer, the initial conditions for the first three layers, and the inflows and outflows at the open boundaries have been computed from the climatological data of temperature and salinity analyzed by Bauer and Robinson and presented in version VIII of their atlas (Bauer and Rob-
The implementation of nudging in the QG model used is straightforward. The upper-layer model equation is altered by adding a relaxation term in the form:

$$\frac{\partial \nabla^2 \psi_i}{\partial t} = \text{\textquoteleft physics\textquoteright} - R(\nabla^2 \psi_i - \nabla^2 \psi_{\text{obs}}).$$

(4.1)

The equations for the lower layers are left unchanged. Here \textquoteleft physics\textquoteright includes the rate of change of vortex stretching, the advection of potential vorticity by the surface flow, the steady wind forcing, and the biharmonic friction term. The relaxation coefficient \(R\) is, in general, a function of space and time to account for irregular data and data error distributions in space and time (Ghil and Malanotte-Rizzoli 1991; Malanotte-Rizzoli and Young 1992). Our emphasis in this study is to analyze the model response to the prescription of a complete and uniformly accurate surface information. Surface data, in fact, are available at each grid point and at time intervals frequent enough to allow a continuous assimilation in time. Therefore, the nudging coefficient can be chosen to be a constant. The value we have chosen is \(R = (0.5 \text{ day})^{-1}\), corresponding to a relaxation time scale of \(1/2\) day. This value is close to the upper limit dictated by numerical stability considerations and thus constitutes a \textquoteleft strong\textquoteright nudging. In fact, it can be shown by scale analysis (Capotondi 1993) that the time scale of \(1/2\) day is at least one order of magnitude shorter than the timescales associated with the terms in physics. As a consequence of this implementation of nudging, the model upper-layer streamfunction \(\psi_i\) will become very close to the \textquoteleft observed\textquoteright streamfunction \(\psi_{\text{obs}}\). How do the model subsurface fields respond to this surface constraint? Since the surface data include a time-mean component and an eddy component and we are interested in separating the effect of the time-averaged flow from the effect of the eddies in controlling the model dynamics, we formally split all the model variables, the streamfunction \(\psi_k\) and the potential vorticity \(q_k\), in the same fashion:

$$\psi_k = \bar{\psi}_k + \psi'_k$$

(4.2a)

$$q_k = \bar{q}_k + q'_k;$$

(4.2b)

subscripts indicate the layer. The temporal mean is supposed to be computed over a time interval much longer than the eddy timescales. The expression for the QG potential vorticity is given by Eq. (A4) in appendix A. In the presence of a strong nudging the time-averaged version of the model equations (A5) can be rewritten in the form:

$$\frac{\partial \bar{\psi}_k}{\partial t} + J(\bar{\psi}_k, \bar{q}_k) = \bar{F}_k - J(\bar{\psi}_k, \bar{q}_k), \quad k = 2, 5.$$  

(4.3b)

The upper-layer equation expresses the data constraint we are imposing. The surface wind forcing has thus
been replaced by a surface pressure boundary condition. The equations for the subsurface layers describe the slow variation of the mean potential vorticity in the presence of mean advection, dissipation by biharmonic friction, and eddy advection of eddy potential vorticity. Biharmonic friction is negligible at the scales typical of the mean circulation (Holland 1978). The analysis of several numerical QG simulations (Rhines and Holland 1979; Holland and Rhines 1980; Marshall 1984; Lozier and Riser 1990) have shown the existence of different flow regimes; in some regions the eddy flux divergence is large and important in the mean balance, while in other regions eddy fluxes are small in the mean. In particular, Marshall's results with a barotropic model show that in some areas close to the western boundary eddy fluxes are the driving agent of the mean flow across $\bar{q}$ contours. However, in spite of the local importance of the eddy flux divergence term, his results also show that $\bar{q}$ remains a strong constraint on $\bar{\psi}$ in the inertial recirculation area. The mean flow is so strong in this area that even a large eddy forcing causes only a relatively small deflection of $\bar{\psi}$ across $\bar{q}$. The importance of mean flow advection in shaping the mean potential vorticity distribution in the subsurface layers is also evident in the numerical model we are using for the present study when no data are assimilated (Capotondi 1993). In fact, the model domain mainly covers a region of intense mean flow. Since we are interested here in the large-scale structure of the circulation, it appears sensible to adopt an inertial approach in our theoretical considerations and neglect at first order the eddy flux divergence term in Eq. (4.3b). We will verify the validity of this inertial assumption in section 6, where the results of the assimilation experiments are qualitatively compared with the predictions derived from this dynamical framework. At steady state the system (4.3) thus becomes

$$\bar{\psi}_k \equiv \bar{\psi}_{\text{obs}} \quad (4.4a)$$

$$J(\bar{\psi}_k, \bar{q}_k) \approx 0, \quad k = 2, 5. \quad (4.4b)$$

The set of equations (4.4b) defines a generalization of the Fofonoff problem (Fofonoff 1954) to a baroclinic, four-layer "ocean," with a prescribed "surface topography" given by $\bar{\psi}_{\text{obs}}$ and inflow/outflow conditions at the boundaries. Marshall and Nurser (1986) showed how to construct analytical solutions to the baroclinic generalization of a Fofonoff problem in an idealized rectangular ocean bounded by solid walls. In the context of our experiments, the particular solution of Eqs. (4.4b) that satisfies the given lateral and surface boundary conditions cannot be determined analytically due to the irregular model geometry and to the specification of a "surface topography" and boundary conditions that are not analytically defined. Therefore, it is instructive to consider a simple analytical example to illustrate how an inertial solution to the problem (4.4) can be achieved. In particular, we are interested in identifying the influence of the "surface topography" $\bar{\psi}_{\text{obs}}$ in determining the characteristics of the global solution. This is considered in the next section.

5. An analytical example

The mathematical formalism adopted for this analysis follows the approach introduced by Rhines and Young (1982; henceforth referred to as RY) in the context of the wind-driven circulation. In the case of the wind-driven circulation, the wind stress curl and, therefore, the barotropic streamfunction are known. In the present case the given quantity is the upper-layer streamfunction. As discussed in RY the contribution of the relative vorticity to the QG potential vorticity is negligible when considering the large-scale steady circulation away from lateral boundaries. Therefore, we simplify the expression of the potential vorticity [Eq. (A4) in appendix A] by neglecting the

\[ \ldots \]

FIG. 3. (a) Surface streamfunction $\bar{\psi}_{\text{obs}}$ for the analytical example; $\bar{\psi}_{\text{obs}}$ describes an anticyclonic flow inside the disk of radius $R$. (b) Geometry of the geostrophic contours in the second layer. The dashed line indicates the disk inside which the surface flow is confined. (c) Geometry of the geostrophic contours in layer 3. All the lateral boundaries are assumed to be closed. The dashed line encloses the disk of radius $R$; the dotted line defines the area of closed geostrophic contours in layer 2.
relative vorticity $\nabla^2 \psi_2$. We consider here the simple case in which $\psi_{ob}$ is given by the anticyclonic flow shown in Fig. 3a and described by the expression

$$
\psi_{ob} = \begin{cases} 
\frac{\psi_0}{R^2} (R^2 - x^2 - y^2), & r < R \\
0, & r \geq R.
\end{cases} \quad (5.1)
$$

The surface velocity field is confined inside the disk $r \leq R (r = \sqrt{x^2 + y^2})$ with intensity increasing from zero at the center to the maximum value $2\psi_0/R$ at the periphery of the disk. The presence of this surface flow distorts the interface between layer 1 and layer 2, producing a circular depression in the interface. We consider first a two-layer model and then extend our analysis to a three-layer model. A continuously stratified case for this particular problem has been discussed in Capotondi (1993) following the considerations of RY in the context of the wind-driven circulation.

### a. Two-layer model

We consider a rectangular domain as shown in Fig. 3. The model has a flat bottom, specified boundary conditions, and prescribed surface flow. Neglecting the relative vorticity and considering the fact that the Jacobian of $\psi_2$ with itself vanishes, the equation for the second layer can be written, from (4.4), in the form:

$$
J(\psi_2, \beta y + F_2 \psi_{ob}) \sim 0. \quad (5.2)
$$

The quantity we define

$$
\hat{q}_2 = \beta y + F_2 \psi_{ob} \quad (5.3)
$$

is a known function. Therefore, the problem (5.2) is a linear problem in which the $\hat{q}_2$ contours define the mean streamlines. A general solution of (5.2) will be of the form

$$
\psi_2 = A_2(\hat{q}_2). \quad (5.4)
$$

For $\psi_{ob}$ given in (5.1), the function $\hat{q}_2$ is

$$
\hat{q}_2 = \begin{cases} 
\frac{F_2 \psi_0}{R^2} (R^2 + \gamma_0^2 - x^2 - (y - \gamma_0)^2), & r < R \\
\beta y, & r \geq R.
\end{cases} \quad (5.5)
$$

The quantity $\gamma_0$ is defined as

$$
\gamma_0 = \frac{1}{2} \frac{\beta R^2}{F_2 \psi_0}. \quad (5.6)
$$

The $\hat{q}_2$ contours are straight lines outside the disk of radius $R$ and arcs of circle inside the disk. As in the problem discussed in RY, closed contours can be found if $\gamma_0 < R$. This condition is satisfied if

$$
|U_{max}| > \frac{\beta}{F_2}, \quad (5.7)
$$

where $|U_{max}| = 2\psi_0/R$ is the maximum surface velocity. Therefore, in order to have closed contours, the surface velocity must exceed the phase speed of the long baroclinic Rossby waves supported by the present model, $c_{ph}$, where $|c_{ph}| = \beta/F_2$. In fact, in the system we are considering, where only one layer can evolve freely, these waves are the only agent that can propagate the information about the boundary values into the interior. However, if the flow is sufficiently strong to oppose the Rossby wave propagation, areas isolated from the boundaries can be created. The condition (5.7) also coincides with the requirement that the basic-state potential vorticity gradient changes sign in the lower layer (in the case $U$ is a zonal flow) (Pedlosky 1979) and can thus be interpreted as a necessary condition for instability.

Contours of $\hat{q}_2$ are shown in Fig. 3b for the case $\gamma_0 = \frac{1}{4} R$, which occurs when $\psi_0 = \beta R/F_2$. The value of $\psi_{ob}$ on the open contours is determined by the boundary values. Therefore, if no inflow or outflow is prescribed at the boundaries, motion is possible only inside the closed contours, as discussed in RY. Inside closed contours the flow is undefined at this order. This is a case of steady-state resonance in which any flow is in principle possible. A weak forcing can produce an arbitrarily large response, which can be limited only by considerations of frictional effects. The selection of a particular solution in these areas will be obtained, therefore, as a balance between forcing and dissipation. In the absence of any external forcing, as in our case, the only source of energy for the motion within the closed contours is given by the eddy flux divergence term. As shown by Holland and Rhines (1980), the component of this flux that is more effective in driving the deep gyres is the flux of interface height, corresponding to vertical propagation of horizontal momentum. This justifies the choice, adopted in RY, of parameterizing this flux as a downgradient flux of interface height displacement

$$
J(\psi_2, q_2^2) \sim -\nabla \cdot [\kappa \nabla F_2_1(\psi_1 - \psi_2)], \quad (5.8)
$$

where $\kappa$ is the diffusion coefficient, generally a function of position. If for convenience $\kappa$ is chosen to be a constant and dissipation is given by bottom friction in the form

$$
D_1 = -\epsilon \nabla^2 \psi_2, \quad (5.9)
$$

the relationship between $\psi_2$ and $\psi_2$ is found (see RY for the detailed derivation) to be linear inside closed contours:

$$
\frac{dA_2}{dq_2} = \frac{\kappa}{\kappa F_2 + \epsilon}. \quad (5.10)
$$

We should notice here that the intensity of the flow in layer 2 depends on the forcing and dissipation parameters $\kappa$ and $\epsilon$, and it is an increasing function of $\kappa$ only if $\epsilon$ is different from zero. If no explicit dissipation were
present, \( dA_2/dq^2 \) would attain its maximum value, which only depends upon the model density structure:

\[
dA_2 \quad = \frac{1}{F_{21}}. \tag{5.11}
\]

The same result would hold also with a spatially varying diffusion coefficient \( \kappa \).

The general solution for \( \varphi_2 \) over the whole domain can be written in the form:

\[
\varphi_2 = A_2q^2 + C_2. \tag{5.12}
\]

Suppose that the boundary conditions prescribed in the second layer are a uniform eastward flow both at the western and the eastern boundaries:

\[
\varphi_2 = -U_2y \quad \text{at} \quad x = -L, L.
\]

In this case we have

\[
A_2 = \begin{cases} 
-\frac{U_2}{\beta}, & \text{open contours (} \hat{\tau} > \hat{R} \text{)} \\
\frac{\kappa}{kF_{21} + \epsilon}, & \text{closed contours (} \hat{\tau} < \hat{R} \text{)},
\end{cases}
\]

where \( \hat{\tau} = \sqrt{x^2 + (y - y_{02})^2} \) and \( \hat{R} = R - y_{02} \).

In (5.12) \( C_2 \) is a constant chosen so that \( \varphi_2 \) is continuous at the edge of the closed contours:

\[
C_2 = \begin{cases} 
0, & \hat{\tau}_2 < \beta R \\
-\beta \left( \frac{U_2}{\beta} + \frac{\kappa}{kF_{21} + \epsilon} \right), & \hat{\tau}_2 > \beta R.
\end{cases} \tag{5.14}
\]

In the second layer the streamline distribution is now determined not only by the prescribed surface topography but also by the topography of the lower interface, which is a function of the flow in layer 3. However, the intensity of the flow can be expected to decrease with depth so that at first order the displacement of the lower interface in layer 2 can be considered much smaller than the displacement of the upper interface. In these conditions an approximate solution in the second layer can be obtained as before, and it is given by

\[
A_2 = \begin{cases} 
0, & \text{open contours (} \hat{\tau} > \hat{R} \text{)} \\
\frac{1}{F_{21} + F_{23}}, & \text{closed contours (} \hat{\tau} < \hat{R} \text{)},
\end{cases}
\]

and

\[
C_2 = -\beta RA_2. \tag{5.19}
\]

The corresponding potential vorticity distribution is constant inside the closed contours and coincides with the value of the planetary component at the northern rim of the gyre, \( \hat{\varphi}_2 = \beta R \). We should recall, however, that in the model used for the assimilation experiments we have biharmonic friction as a dissipative term.

b. Three-layer model

From (4.4b) and (A4) the governing equations for layers 2 and 3 at the leading order are

\[
J(\varphi_2, \beta y + F_2\varphi_2 + F_2\varphi_3) = 0 \quad \text{(5.17a)}
\]

\[
J(\varphi_3, \beta y + F_3\varphi_3) = 0. \quad \text{(5.17b)}
\]

In the second layer the streamline distribution is now determined not only by the prescribed surface topography but also by the topography of the lower interface, which is a function of the flow in layer 3. However, the intensity of the flow can be expected to decrease with depth so that at first order the displacement of the lower interface in layer 2 can be considered much smaller than the displacement of the upper interface. In these conditions an approximate solution in the second layer can be obtained as before, and it is given by

\[
A_2 = \begin{cases} 
0, & \text{open contours (} \hat{\tau} > \hat{R} \text{)} \\
\frac{1}{F_{21} + F_{23}}, & \text{closed contours (} \hat{\tau} < \hat{R} \text{)}
\end{cases}
\]

and

\[
C_2 = -\beta RA_2. \tag{5.19}
\]

The corresponding potential vorticity distribution is constant inside the closed contours and coincides with the value of the planetary component at the northern rim of the gyre, \( \hat{\varphi}_2 = \beta R \). We should recall, however, that in the model used for the assimilation experiments we have biharmonic friction as a dissipative term.

Consider now the flow that can be expected in layer 3. From (5.17b) we have

\[
\varphi_3 = A_3(\hat{\varphi}_3). \tag{5.20}
\]
where
\[ \hat{q}_3 = \beta y + F_{32} A_2 \hat{q}_2 + F_{32} C_2 \]
\[ = \beta y + \tilde{F} \hat{\psi}_{\text{obs}} - F_{32} \beta RA_2, \quad (5.21) \]
where
\[ \hat{\beta} = (1 + F_{32} A_2) \beta \]
and
\[ \tilde{F} = A_2 F_{32} F_{21}; \quad (5.23) \]

\( A_2 \) is given in (5.18). The \( \hat{q}_3 \) contours, as well as the \( \hat{\psi}_3 \) contours, can thus be related to the surface topography \( \hat{\psi}_{\text{obs}} \). The influence of the surface flow in shaping the mean circulation in layer 3 is measured by \( A_2 F_{32} F_{21} \).

Outside the closed \( q_2 \) contours, \( A_2 \) is zero, so that no surface information can be felt in layer 3. Inside the closed \( q_2 \) contours, \( \hat{F} \) is given by
\[ \hat{F} = \frac{F_{32} F_{21}}{F_{21} + F_{23}}. \quad (5.24) \]

If \( H_2 = H_3 = \hat{H} \), we have that
\[ F_{23} = F_{32} = \frac{f_0}{g'_{23}} \hat{H}. \]
Therefore, (5.24) can be rewritten in the form
\[ \frac{1}{\hat{F}} = \frac{1}{F_{21}} + \frac{1}{F_{23}} \quad (5.25) \]
so that the magnitude of \( \hat{F} \) is smaller than either \( F_{21} \) or \( F_{23} \), and it is given by
\[ \hat{F} = \frac{f_0}{g'_{23}} \hat{H}. \quad (5.26) \]

where \( \hat{g} = g'_{12} + g'_{23} \) is the reduced gravity associated with the total density difference between layer 3 and layer 1. If, more generally, \( H_3 \) is different from \( H_2 \), \( \hat{F} \) can be written
\[ \hat{F} = \frac{f_0}{g'_{23} H_3} \left( \frac{1}{1 + \frac{\rho_2 - \rho_1}{\rho_3 - \rho_2}} \right). \quad (5.27) \]

We can distinguish two limiting cases.

1) \( (\rho_3 - \rho_2) \gg (\rho_2 - \rho_1) \). In this case \( \hat{F} \sim F_{32} \) so that the penetration of the surface information is only determined by the “rigidity” of layer 3. If layer 3 is very deep and/or its density is much larger than the density of the layer above, the influence of \( \hat{\psi}_{\text{obs}} \) in determining the distribution of the \( \hat{q}_3 \) contours can be expected to be negligible with respect to the planetary term.

2) \( (\rho_1 - \rho_2) \ll (\rho_2 - \rho_1) \). In this case \( \hat{F} \sim (f_0/g'_{23} H_3) \). The influence of the surface information still depends upon the thickness of layer 3, but now it depends upon the largest density difference, which is the one between layer 1 and layer 2.

From this simple analysis we see that the penetration of the mean surface information is tightly linked to the stratification characteristics of the area under consideration.

Similar analyses can be extended to a continuous stratified case. For brevity we do not consider this case here. The interested reader can find the derivation in Capotondi (1993). As in the case of the wind-driven circulation considered by RY, the area of closed geostrophic contours becomes smaller and displaced poleward with depth. In three dimensions the region of closed contours has a “bowl-like” shape whose maximum depth, corresponding to the depth of influence of the surface streamfunction field, is a function of the vertical stratification.

c. Discussion

From this analytical example we can make the following points.

(i) Due to nonlinear effects, the prescription of a surface flow \( \hat{\psi}_{\text{obs}} \) can constrain the flow structure in the subsurface layers. The “depth of influence” of this surface constraint is very strongly affected by the vertical density profile.

(ii) The intensity of the flow is dictated by the boundary values on the open contours and by the characteristics of forcing and dissipation processes inside the closed contours.

(iii) If closed contours are present in a given layer as a consequence of the prescription of the surface field, different flow regimes can be found: the functional relationship between streamfunction and potential vorticity is dictated by the boundary values in the area of the open contours and by the characteristics of forcing and dissipation inside closed contours.

(iv) If the eddy flux of interface height is the only nonconservative mechanism present, the intensity of the flow inside closed contours will not depend on the intensity of the eddy field but only upon the model density structure. The corresponding potential vorticity fields are constant. However, if some explicit dissipation is present in each layer, as in the assimilation experiments that we are going to discuss, the amplitude of the solution will depend on the relative strength of forcing and dissipation.

(v) If no-flow conditions are specified at the boundaries, motion is possible only inside closed contours as in RY. The driving agent for this flow is the eddy flux divergence term.

We now return to the specific problem under consideration and try to verify to what extent the above theoretical ideas can explain the results of our assimilation experiments.

6. The assimilation experiments

We consider here two different experiments: in the first one only the time-mean component of the surface
data is assimilated, while in the second experiment the total upper-layer streamfunction is used in the relaxation term. The comparison between the results of these two experiments will help identify the relative effects of the mean and eddy components in determining the characteristics of the flow fields that develop in the model subsurface layers. In the first experiment the nudging term has been added to the equation for the first layer in the form

$$-R(\nabla^2 \psi_1 - \nabla^2 \psi_{\text{obs}}), \quad (6.1)$$

where $\psi_{\text{obs}}$ is the climatological field in Fig. 1a. The relaxation coefficient is $R = (0.5 \text{ day})^{-1}$. At each time step $\psi_1$ is relaxed toward the steady field $\psi_{\text{obs}}$ with a very short relaxation timescale. Therefore, any time-dependent motion that the model might try to develop will be strongly damped. We will see later, in fact, that the time-dependent motion found in this experiment is extremely weak. From the considerations developed in the previous section, eddies can be expected to be the forcing agent for the flow inside the closed geostrophic contours in the subsurface layers. In particular, they are the only source of vorticity for the motion in layers 4 and 5, where no inflows or outflows are specified at the boundaries. Therefore, if the eddy field has vanishing intensity, only a very weak time-averaged flow can be expected in the two deepest layers.

The numerical simulation has been carried out for 20 years to allow all the transient processes associated with the initial conditions to decay. The initial condition in the three upper layers is given by the climatological fields in Fig. 1, while no flow is assumed initially in the two deepest layers. We have monitored the time evolution of the total kinetic energy in each of the five layers to ensure that statistical steady state (in this case almost coincident with an absolute steady state) has been reached. The "climatology" of this numerical experiment has been computed by averaging the model fields over the last four years of the simulation.

In the second experiment the upper-layer streamfunction $\psi_1$ is relaxed toward the total "observed" streamfunction $\psi_{\text{obs}}$, where

$$\psi_{\text{obs}}(t) = \tilde{\psi}_{\text{obs}} + \psi_{\text{obs}}(t). \quad (6.2)$$

As before $\tilde{\psi}_{\text{obs}}$ is the climatological field in Fig. 1a, and $\psi_{\text{obs}}(t)$ is the sequence of eddy maps constructed from the Geosat data. As described in section 2, the total duration of the $\psi_{\text{obs}}$ dataset is 570 days. The initial conditions are the same as for the previous experiment. Since we are now imposing a time-dependent constraint at the surface, we need to define sensible criteria for assessing when the model has adjusted to the observations. The evolution of the total kinetic energy during the first 570 days of the experiment shows that in each of the five layers the energy increases from the initial value to a "steady" level during the first 10-20 days of the simulation. After this short transient the level of total kinetic energy remains practically constant in each layer. The potential vorticity distributions on the other hand show a continuous evolution. Starting from the initial distributions associated with our choice of initial conditions, potential vorticity is redistributed by advection processes. These processes now include not only mean flow advection as in the previous experiment but also eddy advection. While the mean flow tries to establish constant values of potential vorticity along streamlines, the turbulent eddy field acts as an efficient mixing agent, which tends to smooth out the $q$ gradients. The evolution of the potential vorticity fields is the result of the competition between these two processes, and a statistical steady state will be reached when mean flow advection balances the eddy mixing effect. In Figs. 4a and 4b we show instantaneous potential vorticity maps in layers 2 and 3, respectively. Notice the convoluted distribution of the $q$ contours, which is due to eddy advection and is a manifestation of the enstrophy cascade.

Eddy mixing can be expected to depend only on the statistical characteristics of the eddy field and not on the details of its instantaneous realizations. To allow these processes to evolve until statistical equilibrium
is reached, we have extended this experiment beyond the 570-day duration of our dataset by assimilating the same data in a sequence of runs, each of which is started from the final fields of the previous one. The total experiment consists of 20 of these assimilation segments, totaling 11,400 days or about 31.6 years of spinup time. The convergence of the potential vorticity fields toward an equilibrium distribution is illustrated in Fig. 5, where we show meridional profiles of mean potential vorticity in the four subsurface layers. The different curves in each diagram refer to the averages over each of the 570-day assimilation segments: “A” corresponding to the first segment, “B” to the second, and so on. These profiles have been computed as averages over 10° of longitude centered at 55°W. In layers 4 and 5 the meridional profiles are dominated by the planetary vorticity gradients, and only relatively small adjustments take place from segment to segment. In layers 2 and 3 we can see the convergence of these different profiles to a meridional distribution showing a large plateau in

![Fig. 5. Meridional profiles of mean potential vorticity at 55°W in the four model subsurface layers during the experiment in which the total surface streamfunction is assimilated. The sequence of surface eddy streamfunction fields $\psi_{ed}(t)$, covering a period of 570 days, is repeated periodically 20 times in order to allow the adjustment of the subsurface potential vorticity fields. The curves in each panel refer to the averages over each of the 570-day assimilation segments: “A” corresponding to the first segment and “B” to the second and so on.](image-url)
its central part, where the potential vorticity has been homogenized due to the very effective eddy mixing. We can also notice how the value of $q$ tends to decrease, from segment to segment, at the northern end of these meridional profiles due to advection of low potential vorticity anomalies by the mean flow.

After the 31.6 years of spinup time, variations can still be observed from segment to segment in the mean streamfunction fields. The rms differences between the streamfunctions corresponding to the last two segments are only a few percent in the upper three layers, but they can be as large as 40% in layers 4 and 5, where they are mainly associated with slight changes in the position of the gyres present in the deep mean fields. The "climatology" of this model has been computed by considering a time average over the last segment.

We are now interested in comparing the model climatologies obtained in these two experiments and interpreting them in the context of the dynamical framework developed in sections 4 and 5. To that end we first try to infer, on the basis of the dynamical considerations derived from the analytical example, the structure of the flow that can be expected in layers 2 and 3 when the climatological field $\psi_{\text{obs}}$ is imposed at the surface. We then analyze in detail the mean circulation and the mean potential vorticity fields in these two scenarios: the first one in which the model time variability is very weak, and the second one in which the model eddy field is much more energetic and constrained to follow the surface eddy observations.

a. Streamfunction distributions predicted in layers 2 and 3

In the first experiment only a vanishingly weak flow is found in the two bottom layers, as we will see below. This experiment can thus be conveniently defined as one with three moving layers over a motionless deep ocean, as in our analytical three-layer case (section 5b). Therefore, we concentrate our attention on layers 2 and 3. In the following, the climatological fields for layers 2 and 3, corresponding to Fig. 1b and Fig. 1c, will be referred to as $\psi_{\text{obs}}$ and $\psi_{\text{obs}}$, respectively. At the initial time the ratio between the standard deviations of $\psi_{\text{obs}}$ and $\psi_{\text{obs}}$ is

$$\frac{\sqrt{\sum (\psi_{\text{obs}})^2}}{\sqrt{\sum (\psi_{\text{obs}})^2}} \sim 13\%.$$  

where the indices $i$ and $j$ vary over the entire model domain. Therefore, at this time, the function

$$\tilde{q}_2 = \beta y + F_{22} \psi_{\text{obs}}$$  

defines, to a good approximation, the $\psi_2$ distribution. Contours of $\tilde{q}_2$ are shown in Fig. 6a. As in the analytical example discussed in the previous section, some of the $\tilde{q}_2$ contours in Fig. 6a are closed and do not reach the boundaries, while some others are connected with the boundaries. In the area of the northern recirculation gyre, where $\psi_{\text{obs}} = 0$, the $\tilde{q}_2$ contours coincide with the $\beta y$ contours and are given by zonal lines. In this area these contours suggest a westward flow emanating from the eastward flowing jet. In fact, as the surface topography is flat in this area, fluid particles must move along latitude circles to conserve their potential vorticity. A no-flow condition must be satisfied at the coastline, so that some higher-order physics is required there to close the circulation. Therefore, we might anticipate the formation of a boundary jet along the coastline where relative vorticity, or friction, will no longer be negligible. Away from this region, the circulation suggested in layer 2 by the $\tilde{q}_2$ contours does not show evident discrepancies with $\psi_{\text{obs}}$ (Fig. 1b). Therefore, we use $\psi_{\text{obs}}$ in (5.17b) to obtain an approximation for the structure of the flow field in layer 3. At initial time layer 4 is at rest. In the experiment in which only $\psi_{\text{obs}}$ is assimilated, the two bottom layers can be con-
considered motionless. Therefore, the lower interface does not enter into play in determining the circulation pattern in layer 3. The latter can thus be described, within the approximations made, by the function \( \bar{q}_1 \) given by
\[
\bar{q}_1 = \beta y + F_3 \bar{q}_{2,\text{obs}}. \tag{6.5}
\]

The contours of \( \bar{q}_1 \) are shown in Fig. 6b. Also in this case the absence of a northern recirculation gyre flow in \( \bar{q}_{2,\text{obs}} \) leads to zonal contours of \( \bar{q}_1 \), which imply a westward flow in this area of layer 3. Again, some higher-order physics must enter into play close to the solid boundary to satisfy the no-flow condition.

If we compare the flow pattern in layer 3, as given by the \( \bar{q}_3 \) contours, with the flow pattern in layer 2, we can notice how the recirculation gyre in layer 3 appears tighter and more elongated in the NE-SW direction with respect to the recirculation in layer 2, whose outer streamlines extend beyond the model southern boundary. In Fig. 6b the contours south of about 30°N join both the eastern and western boundaries. At these latitudes the streamfunction values specified at the western boundary are constant, and they have only minor variations along the eastern boundary. Therefore, only a weak flow can be expected south of 30°N. This variation with depth of the shape of the recirculation area represents the model analog of the results of the analytical example, where the area of closed geostrophic contours becomes smaller and displaced northward with depth.

We now compare and contrast the mean streamfunction distributions that are obtained in layers 2 and 3 during the assimilation experiments with the "theoretical" flow patterns in Fig. 6.

b. The mean streamfunction fields

The time-averaged circulation obtained in the first experiment is shown in Fig. 7 in all the five model layers. In the first layer the flow field is essentially the same as \( \bar{q}_{\text{obs}} \). As anticipated, the nudging term represents the dominant contribution in the equation for the first layer, so that the upper-layer streamfunction becomes almost identical to \( \bar{q}_{\text{obs}} \). A closer comparison between the two surface fields shows that the major differences occur in the area of the jet separation from the coast, northeast of Cape Hatteras and in the area of the Grand Banks, where one of the branches of the stream turns northward. At both locations the streamlines in the \( \bar{q}_1 \) field tend to "open" toward the coast. This deviation of \( \bar{q}_1 \) with respect to \( \bar{q}_{\text{obs}} \) is produced by the circulation that develops in the subsurface layers north of the stream, as we will see in a moment. The maximum differences between \( \bar{q}_1 \) and \( \bar{q}_{\text{obs}} \) in both places are of the order of 15%. Consider now the circulation in the second layer. If the dynamical framework developed in sections 3 and 4 captures the essential physics of this model simulation, the flow pattern obtained in layer 2 should be well described by the distribution of the \( \bar{q}_2 \) contours in Fig. 6a. Comparison between Figs. 7b and 6a shows, indeed, striking similarities. The structure of the subtropical recirculation gyre obtained in this numerical simulation is rendered in great detail by the \( \bar{q}_2 \) contours. The three anticyclonic cells, which can be observed in Fig. 6a around 70°W, 58°W, and 42°W, do appear as features of the time-averaged circulation in layer 2. The same is true for the cyclonic cell predicted by the \( \bar{q}_2 \) contours inside the curve of the stream around the Grand Banks. Also as anticipated, a westward flow can be observed in the area north of the stream. As the stream emerges from Cape Hatteras and flows eastward as a free jet, fluid particles detach from the stream and move westward. As expected, a thin jet forms along the coastline to close the circulation. This coastal jet, whose intensity tends to increase with latitude, develops instabilities. As it tries to follow the irregular and sinuous coastline, meanders develop and ringlike structures are shed, which remain trapped between the jet and the boundary. The presence of this coastal jet, not predicted by the simplified derivation of \( \bar{q}_2 \), has the effect of somehow distorting the whole flow field so that a precise agreement between \( \bar{q}_2 \) and \( \bar{q}_{\text{obs}} \) cannot be found. The presence of the flow in layer 3, which has not been considered in the derivation of \( \bar{q}_2 \), is an additional reason for discrepancies. However, the basic characteristics of the flow in the second layer are predicted by the structure of \( \bar{q}_1 \), supporting the hypothesis of the inertial nature of the circulation.

Similar considerations can be applied to layer 3 (Fig. 7c). In this case the streamline distribution should be compared with the distribution of the \( \bar{q}_3 \) contours in Fig. 6b. Also in this case the shape of the subtropical recirculation gyre, the presence of smaller-scale anticyclonic cells, the development of a westward flow in the subpolar area, and the consequent formation of the coastal jet are features predicted by the \( \bar{q}_3 \) contours. The most energetic component of the circulation is found north of 30°N, a latitude which defines the southern border of the recirculation at this depth.

In Figs. 7d and 7e we show for completeness the time-averaged streamfunction fields in layers 4 and 5, respectively. In both layers, as expected, the flow is quite small almost everywhere. The only noticeable component of the circulation is found in the proximity of the northern boundary, where eddies produced by instabilities of the boundary jet have relatively larger amplitudes. Maps of eddy kinetic energy show, in fact, values lower than a few centimeters squared per second squared in most of the domain. Only close to the northern boundary isolated values as high as 100 cm² s⁻² can be found, clearly associated with instabilities of the jet. This appears to be the only area where eddies can drive any flow in the two deepest layers.
The "climatological" streamfunction fields obtained in the experiment in which both mean and eddy components are assimilated is shown in Fig. 8. We can immediately notice the similarity of the flow patterns in the three upper layers with the results of the previous experiment (Fig. 7). The surface layer is strongly constrained by the nudging procedure so that \( \psi_1 \) cannot deviate much from \( \psi_{obs} \). However, also the circulation in layers 2 and 3, although not directly constrained, has basically the same structure as in the experiment where only \( \psi_{obs} \) was assimilated. In both layers the shape of the recirculation gyres with all their small-scale features has remained very similar. Some differences can be observed in the area north of the stream. In Figs. 8b and 8c we can still notice the tendency for fluid particles to leave the jet and move westward, but now no well-defined boundary jet develops along the coastline. The presence of the eddy field is now able to supply a potential vorticity input, which allows the mean flow to move northward also in the interior, without the need of invoking the higher-order physics of a boundary layer. However, a northern recirculation gyre is still absent. Eddies alone seem to be unable to drive this component of the circulation, thus supporting some recent findings of the thermodynamic maintenance of this gyre (Ezer and Mellor 1992). Due to the presence of an energetic eddy field, motion is now possible also in layers 4 and 5. The most energetic component of the flow is found in the western half of the domain and represents the deep expression of the in-
FIG. 8. Time-averaged streamfunction fields from the experiment in which the total surface streamfunction \( \psi_{ob} \) is assimilated at the surface. (a) Layer 1: contour interval is 5000 m\(^2\) s\(^{-1}\). (b) Layer 2: contour interval is 2000 m\(^2\) s\(^{-1}\). (c) Layer 3: contour interval is 2000 m\(^2\) s\(^{-1}\). (d) Layer 4: contour interval is 2000 m\(^2\) s\(^{-1}\). (e) Layer 5: contour interval is 2000 m\(^2\) s\(^{-1}\).

Inertial recirculation for this numerical experiment. The streamfunction fields in these two layers are rich in small-scale features. Notice, in particular, the tendency for the formation of elongated zonal gyres. The presence of these gyres could be an artifact of the neglect of bottom topography. In fact, in the absence of any topographic steering, "free" flow tends to develop along \( \beta y \) contours. We will see in Part II that evidence of zonal jets has indeed been found in observations of the deep flow in this area.

Although the structure of the circulation in layers 2 and 3 has not been noticeably affected by the assimilation of the eddy field, the intensity of the flow has been altered. To illustrate the differences in the flow strength, differences that are induced when surface eddies are assimilated, we show in Fig. 9 meridional profiles of mean zonal velocity along 55°W in all the five model layers. The profiles have been averaged over 10° of longitude. In each figure the dashed line represents the zonal velocity obtained in the experiment in which only \( \psi_{ob} \) is assimilated, while the solid line corresponds to the case in which the total \( \psi_{ob} \) is imposed at the surface. In layer 1 the two profiles are almost coincident as a consequence of the strong nudging toward the same climatological flow field. In layers 4 and 5 the flow is practically zero in the absence of eddies. In layers 2 and 3 a strong correlation can be noticed between the solid and dashed velocity profiles. The position of the eastward jet and of the westward return flow has not
been significantly altered by the assimilation of the eddy field, but the amplitude of the peak zonal velocities, both eastward and westward, has been enhanced.

c. Mean potential vorticity fields

The mean potential vorticity fields obtained in the first experiment are shown in Fig. 10. In layer 1 the potential vorticity distribution has remained very similar to the initial distribution (not shown). The major differences occur in the "subpolar area," where the stretching effect produced by the flow in the second layer determines a distortion of the $\beta y$ contours present in the initial field. In layers 2 and 3 the potential vorticity distributions show the advective control of $\tilde{q}$. This can clearly be seen in Figs. 11 and 12, where the streamfunction distributions in layers 2 and 3, respectively, are directly compared with the potential vorticity distributions in the corresponding layers. In both layers the $\tilde{q}$ contours reproduce, in fact, the shape of the recirculation gyre with large areas of reduced gradients.
the differences introduced in their potential vorticity distributions by the presence of the surface eddy field are of particular interest. The maps in Figs. 13b and 13c no longer show any closed contours reproducing the exact shape of the recirculation gyres as in Figs. 10b and 10c. Only tongues of low potential vorticity anomalies can be noticed on the southern flanks of the recirculation gyres. As before, the mean flow tries to create uniform distributions of potential vorticity along streamlines, but now the energetic eddy mixing prevents the completion of this process. In Figs. 14 and 15 we compare, as before, the \( \vec{\nabla} \) contours with the \( \bar{q} \) contours in layers 2 and 3, respectively. Although the potential vorticity distributions show clearly the effect of mean flow advection, the almost exact correspondence observed in Figs. 11 and 12 is now lost.

7. Discussion and conclusions

In this paper we have tried to understand the modifications induced in the model fields by the assimilation of surface data whose climatological characteristics are different from the climatology of the unconstrained model. In particular, we have analyzed the relative effect on the model behavior of the two components of the surface observations: the mean component and the eddy component. If the relaxation time in the nudging term is much shorter than the typical model timescales, the surface fields become practically coincident with the observations and can be considered as given. In these conditions we have shown that the structure of the subsurface circulation is mainly determined by the characteristics of the surface mean field. This structure is the result of the model nonlinearities and can be interpreted in the framework of "baroclinic Fofonoff modes" in a domain with a prescribed surface topography (the surface mean field, which is assimilated at the surface) and inflow/outflow conditions specified at the open boundaries. We have shown with the aid of an analytical example how such a solution can be achieved. The geometry of the geostrophic contours in each layer can be related to the pattern of the climatological field imposed at the surface. Some of these contours join the boundaries, while others are closed and isolated from the boundaries. On the open contours the intensity of the flow is determined by the boundary values. Inside the closed contours, on the other hand, the amplitude of the flow can be expected to be the result of a balance between forcing and dissipation. Therefore, different flow regimes may exist.
inside the closed contours. In this experiment, where the intensity of the eddy field is extremely weak, we establish the conditions hypothesized in the theory of Rhines and Young (1982). As predicted by that theory, advection of potential vorticity by the mean flow is able to establish, first, uniform values of $q$ along streamlines. At this point the eddy flux term comes into play and tends to smooth the gradients between adjacent streamlines. Plateaus of "homogenized" $q$ are thus created inside the closed contours, while the $q$ gradients are expelled toward the rim of the gyres. Scatter diagrams computed on zonal lines connecting two different points along a closed streamline (Capotondi 1993) confirm in a more quantitative manner these considerations. They show, in fact, that in both layers 2 and 3, the $q$ contours are very close to the $\tilde{\psi}$ contours in a large part of the domain, as is the case for an inertial solution.

Figure 13 shows the mean potential vorticity distributions for the second experiment. The upper-layer distribution is, again, practically unchanged with respect to the initial distribution as it was in the previous experiment. In layers 4 and 5, which are now in motion, larger deviations from the zonal contours of the planetary vorticity gradients can be observed. The second and third layers are not directly constrained by the nudging procedure and, in both experiments, they carry relatively significant components of the flow. Therefore,
The results of the assimilation experiments confirm this dynamical framework. The structure of the flow remains basically the same whether or not the eddy field is assimilated at the surface, thus confirming that the surface mean field $\psi_{obs}$ indeed defines the pattern of the mean streamlines in the subsurface layers. Consequently, this defines also the paths along which mean flow advection of potential vorticity will take place. However, the evolution and final statistical steady state of the potential vorticity fields depend also upon the intensity of the eddy field. The characteristics of the mean potential vorticity distributions at statistical steady state are determined, in fact, by the relative strength of mean flow advection and eddy advection. If the eddies are very weak as in our first experiment, mean flow advection dominates during the adjustment phase. First, uniform values of potential vorticity are established along mean streamlines. The weak eddy mixing comes then into play by "slowly" eroding the gradients between adjacent streamlines and expelling them toward the rim of the gyre. If, on the other hand, eddy advection is comparable in strength with mean flow advection, potential vorticity will be stirred and mixed by the eddies before the establishment of uniform values along mean streamlines is completely achieved. Therefore, only tongues of anomalous potential vorticity values can be observed in the final time-averaged distributions as a result of mean flow advection.

The most remarkable difference introduced by the presence of an energetic eddy field is the more efficient
downward transfer of momentum. The intensity of the mean flow in the subsurface layers, in fact, is enhanced when eddies are assimilated. This is particularly evident in layers 4 and 5, where the mean flow can only be eddy driven. In fact, hardly any noticeable flow is found in these layers in the absence of an energetic eddy field.

The results of this study also show the crucial role played by the surface mean field in shaping the global model circulation. Our choice of a climatological field as the mean component of the surface flow, which we have introduced in section 2 as a "working hypothesis", is certainly debatable due to the long-term duration of the climatological dataset with respect to the duration of the Geosat dataset. The comparison of the assimilation results with in situ observations presented in Part II confirms the limitations of our surface mean field. A successful assimilation procedure will thus require an accurate definition of the surface mean circulation. The mean fields calculated for the same period as the Geosat data in some recent studies (Glenn et al. 1991; Ezer et al. 1993) could be considered as possible alternatives for the surface mean field in future studies.

Acknowledgments. We would like to thank Drs. Victor Zlotnicki and Lee Fu for making the Geosat data available to us and Dr. Nelson Hogg for kindly supplying the current meter data at the Synop East Array. Steve Worley, NCAR data support section, helped with the dynamic height computation. Julianna Chow supplied very precious assistance in the computational aspects of this work. One of us (AC) would like to thank Drs. Joe Pedlosky, Carl Wunsch, and Mindy Hall for the stimulating ideas and constructive criticism they have supplied during the course of this work; AC would also like to thank the NCAR Advanced Study Program for giving her the opportunity to carry out large parts of this work at NCAR. This study was supported by NASA through Contract 958208 with the Jet Propulsion Laboratory, NASA Grant NAGW2711 with MIT (PMR), and subcontract W-16 351B with the Jet Propulsion Laboratory to NCAR (WRH) as a part of the TOPEX-Poseidon investigation.

APPENDIX A

The quasigeostrophic model formulation with \(N\) arbitrary layers is a straightforward extension of the two-layer case described by Holland (1978). Here we shall present the semidiscrete 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 governing equations are the vorticity and interface height perturbation equations and the thermal wind relation:
\[
\frac{\partial}{\partial t} \nabla^2 \psi_k = J(f + \nabla^2 \psi_k, \psi_k) + \frac{f_0}{H_k} (w_k - w_{k+1}) + \mathcal{F}_k + T_k: \quad k = 1 \text{ to } N \quad (A1)
\]
\[
\frac{\partial}{\partial t} h_{k+1/2} = J(h_{k+1/2}, \psi_{k+1/2}) + w_{k+1/2}:
\]
\[
 k = 1 \text{ to } N-1 \quad (A2)
\]
\[
h_{k+1/2} = \frac{f_0}{g_{k+1/2}} (\psi_{k+1} - \psi_k). \quad (A3)
\]

Here integer subscripts \((k)\) denote the vertical layers \((k \text{ increasing downward})\) in which the quasigeostrophic streamfunction is defined (nominally at the center of each of the layers), while fractional subscripts \((k + 1/2)\) denote the interfaces between layers where vertical velocity and interface height perturbation are defined. The variables are the quasigeostrophic streamfunction \((\psi_k)\) 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 \(\beta\) are chosen to represent typical midlatitude gyre values. The basic background vertical stratification is written in terms of the reduced gravity
\[
g_{k+1/2} = g \Delta \rho_{k+1/2}/\rho_0, \text{ where } \Delta \rho_{k+1/2} \text{ is the (positive) density difference between layers } k + 1 \text{ and } k. \]
Frictional effects, written symbolically in (A1) as \(\mathcal{F}_k\), are parameterized as lateral friction of the biharmonic kind \((\text{Holland 1978})\) in which \(\mathcal{F}_k = -A_k \nabla^2 \psi_k\). In addition, \(T_k\) includes a bottom friction, \(-\nu \nabla^2 \psi_k\), when \(k = N\) (the bottom layer). Note that the effect of the wind forcing \(T_k\), equal to curl \(\tau/H_1\), produces an Ekman pumping stretching tendency in the upper layer equivalent to a body force acting on the upper layer. The \(T_k\) for \(k > 1\) are zero. At the sea surface \(w_{1/2} = 0\) and at a flat sea bottom \(w_{N+1/2} = 0\). The advective velocities at the interfaces needed in (A2) are calculated from a weighted average of the velocities in the layers; that is,
\[
\psi_{k+1/2} = (\alpha_{k+1/2}) \psi_k + (1 - \alpha_{k+1/2}) \psi_{k+1}, \text{ where } \alpha_{k+1/2} = H_k/(H_k + H_{k+1}).
\]

The set of equations (A1), (A2), and (A3) express the conservation of the potential vorticity \(q_k\), defined as
\[
q_k = \nabla^2 \psi_k + f + F_{k,k-1}(\psi_{k-1} - \psi_k) + F_{k,k+1}(\psi_{k+1} - \psi_k), \quad (A4)
\]
and can be rewritten in the form
\[
\frac{\partial q_k}{\partial t} + J(\psi_k, q_k) = \mathcal{F}_k + T_k: \quad k = 1, N. \quad (A5)
\]
Here \(q_k\) is the sum of the relative vorticity \(\nabla^2 \psi_k\), the planetary vorticity \(f = f_0 + \beta y\), and the stretching term \(S = F_{k,k-1}(\psi_{k-1} - \psi_k) + F_{k,k+1}(\psi_{k+1} - \psi_k)\). The quantities \(F_k\) are the inverse of the squared interfacial Rossby radii:
\[
F_k = \frac{\Delta \rho_{k+1/2}^2}{H_k g_{k+1/2}}. \quad (A6)
\]

Therefore, the \(F_k\) carry information about the vertical background stratification.

APPENDIX B

The successive correction algorithm, first introduced by Cressman (1959), is an iterative algorithm. Covariance functions with decreasing radii are used in the different iterations in order to capture smaller and smaller scales. The procedure is similar to the one outlined by Roemmich (1983) in his estimation of hydrographic quantities in the Strait of Florida. The spatial covariance functions are Cressman functions \((R^2 - r^2)/(R^2 + r^2)\), where \(r\) is the distance between the data point and the analyzed point and \(R\) is the correlation distance. At each iteration corrections to the previous estimate are computed at each grid point according to the following formula:
\[
F_i^n(v + 1) = F_i^n(v) + \sum_{i=1}^{\pi} w_{ii}(v)(F_i^0 - F_i^n(v)), \quad (B1)
\]
where \(F_i^n(v)\) is the interpolated (analyzed) value at position \(x_i\), iteration \(v\); \(F_i^0(v)\) is an observed value at position \(i\); and \(F_i^n(v)\) is the estimate of the field at position \(i\) for the \(v\)th iteration. The weights \(w_{ii}(v)\) involve separate space and time factors:
\[
w_{ii}(v) = \left\{ \begin{array}{ll}
\exp\left[-(t_x - t_i)^2/1.44T^2\right] \\
\times (R_i^2 - r_{i,1}^2)/(R_i^2 + r_{i,1}^2), & R > r, \\
0, & r \geq R.
\end{array} \right. \quad (B2)
\]
where \(t_i\) is the time at observation point \(i\), \(r_{i,1}\) is the spatial distance between interpolation point \(x\) and observation point \(i\), and \(T\) and \(R_0\) are correlation time and space scales, respectively. Thirty-four days of altimeter data centered on the time of the analysis are introduced with an \(e\)-folding scale, \(T\), of 5 days for each of the analyses, performed every 2 days; \(T\) was chosen to be 5 days, a time much shorter than the 34-day search window. This value for the Gaussian \(e\)-folding scale was intended to include altimeter data points from tracks west (-3 days) and east (+3 days) of any particular track, while preserving ocean signals with
periods in the mesoscale band and longer (V. Zlotnicki 1990, personal communication). The spatial radii \( R \) are 2°, 1.5°, 1.25°, and 1° for each of the four iterations, respectively. These values were determined by trial and error to capture scales as small as possible while still producing smooth maps.

The successive correction algorithm is not an optimal algorithm since the weights \( w_j \) in (B1) are not chosen so to minimize the expectation value of the interpolation error (Bretherton et al. 1976). The reason chosen so to minimize the expectation value of the algorithm since the weights would introduce a computational load comparable with the one of the optimal interpolation algorithm itself. Since we are not using an assimilation scheme that can rigorously account for the data error distribution, we delay for the moment the computation of the optimal weights (isotropy and homogeneity of the statistics, knowledge of the error statistics, etc.) are often met in the actual implementation of the algorithm, thus reducing its optimal character.

The successive correction scheme does not automatically supply error maps for the interpolated fields. Interpolation error maps should be considered together with the interpolated fields to be able to assess the relative reliability of different estimated values. One can compute error fields for the successive correction algorithm as outlined by Wunsch (1989). However, this would introduce a computational load comparable with the one of the optimal interpolation algorithm itself. Since we are not using an assimilation scheme that can rigorously account for the data error distribution, we delay for the moment the computation of the error maps.

REFERENCES


