Journal of Marine Research, 44, 1–34, 1986

Journal of
MARINE RESEARCH

Volume 44, Number 1

A variational method for inverting hydrographic data

by Christine Provost[1] and Rick Salmon[2]

ABSTRACT

We present a new method for estimating the three-dimensional field of geostrophic velocity from hydrographic station data. Very simply, we ask for the smoothest velocity field (in the sense of an arbitrarily defined norm) which is consistent with the data and with selected approximate dynamical constraints to within prescribed misfits, which, we will argue, should never be zero. The misfits represent errors in the data and in the approximate dynamical constraints. By varying the misfits relatively to one another, we explore the full envelope of physically plausible estimates of the average geostrophic flow. We illustrate the method by application to hydrographic measurements in the Labrador Sea.

1. Introduction

In this paper we address the classical problem of estimating the large-scale, time-averaged ocean circulation from hydrographic measurements of temperature, salinity, and pressure (and hence density). As every oceanographer knows, the density field determines only the vertical shear of the horizontal geostrophic velocity, and thus leaves the absolute geostrophic velocity undetermined by a constant of vertical integration. This integration constant generally varies from one horizontal location to another. Classically, the indeterminacy is removed by assuming that the geostrophic velocity vanishes at some great and usually constant depth. Provided that the deep flow is actually small, the classical method gives reasonable estimates for the geostrophic velocity in upper waters. However, faith in the classical method has crumbled as

1. CNRS, Tour 15.5E, Universite Pierre et Marie Curie, 4 Place Jussieu, 75230 Paris Cedex 05, France.
2. Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, A-025, La Jolla, California, 92093, U.S.A.

technological advances have made possible the direct measurement of deep velocities, which are sometimes found to be quite large.

In recent years, the classical assumption of a level-of-no-motion has been replaced by other, more sophisticated assumptions which incorporate more of the physics. A typical assumption of this type is that the average velocity vector is everywhere tangent to surfaces of constant average potential density. See, for example, Stommel and Schott (1977), Wunsch (1978), Davis (1978), and Stommel and Veronis (1981). Unfortunately, although additional assumptions like the foregoing inevitably reduce (but can never eliminate) the indeterminacy discussed above, these assumptions are often far less justifiable than the more 'ordinary' assumptions of geostrophic and hydrostatic balance. For example, the assumption of mean potential density conservation by the mean flow could be grossly inconsistent with the data if mixing is locally important. In such cases, the extra assumptions may help to make the answer 'more unique,' but the agreement with reality may actually be worse.

In this paper we describe a procedure for incorporating assumptions like the mean density conservation in a way that allows their compatibility with the data and with other dynamical constraints to be assessed. Very simply, we ask for the spatially smoothest velocity field (in the sense of an arbitrarily defined norm) which is consistent with the data and dynamical constraints to within prescribed misfits, which, we will argue, should never be zero. The misfits represent 'errors' in the data and dynamics (in the sense explained below) and can be estimated from a scaling analysis. By varying the misfits relatively to one another, we can explore the full envelope of physically plausible estimates of the large-scale, time-averaged flow. Similar methods have been used by (for example) Wahba and Wendelberger (1980) for the interpolation of meteorological data; by Bennett and McIntosh (1982) to incorporate observations into tidal models with open boundaries; and by Shure *et al.* (1982) to estimate the magnetic field inside the Earth.

2. General method

Our goal is to explore the envelope of geostrophic flow fields which are consistent with the data and the imposed dynamical constraints. Since the data contain no very useful information about horizontal flow scales which are smaller than the separation between hydrographic stations, we shall be solely interested in the average velocity field, defined (for example) as the running mean average of the exact velocity over horizontal length scales comparable to the station separation, and over time scales comparable to the time required for the occupation of all stations. We regard the hydrographic measurements as measurements of this average field. The data therefore contain aliasing 'errors' caused by internal waves, mesoscale eddies, and other phenomena which escape our definition of average. Similarly, dynamical constraints like mean density conservation are 'in error' by the amount by which they fail to apply to this average flow. The 'errors' in the data and in the dynamical constraints can be

estimated from a scaling analysis. The sizes of these errors determine the allowable misfits between the data-dynamics and the average flow.

Variational calculus offers a simple and flexible method for obtaining average flow fields which are consistent with the data and the approximate dynamical constraints. Very simply, we will ask for the smoothest velocity field which agrees with the data and the approximate dynamics, to within prescribed misfits. To illustrate the method in a simple context, let $\psi(\mathbf{x})$ be an abstract flow variable which has been measured at the discrete locations $\{\mathbf{x}_i,\ i = 1, \ldots, N_d\}$ in three-dimensional space. A simple (and arbitrary) measure of the roughness of $\psi(\mathbf{x})$ is the functional

$$R[\psi] = \int\int\int d\mathbf{x}(\nabla^2\psi)^2 \tag{2.1}$$

where the integration runs over the domain of the fluid. The error in the measurement d_i of $\psi(\mathbf{x}_i)$ is

$$e_i \equiv d_i - \psi(\mathbf{x}_i). \tag{2.2}$$

If $\psi(\mathbf{x})$ approximately obeys the dynamical equations,

$$C_j[\psi] = 0, \qquad j = 1, \ldots, M, \tag{2.3}$$

(where C_j is any operator) then our method is to minimize the roughness $R[\psi]$ subject to the constraints that

$$\sum_{i=1}^{N_d} e_i^2 = q^2 N_d \tag{2.4}$$

and

$$\int\int\int d\mathbf{x}\, C_j[\psi]^2 = Q_j^2 V \tag{2.5}$$

where q is the expected error in a single measurement of $\psi(\mathbf{x})$, Q_j is the expected error in the j-th dynamical constraint (2.3), and V is the volume of the fluid. More concisely, we minimize the functional,

$$J[\psi] \equiv \int\int\int d\mathbf{x}\{(\nabla^2\psi)^2 + \gamma\sum_i (\psi - d_i)^2\delta(\mathbf{x} - \mathbf{x}_i) + \sum_j \gamma_j C_j[\psi]^2\} \tag{2.6}$$

where γ and $\{\gamma_j\}$ are the Lagrange multipliers corresponding to the constraints (2.4) and (2.5). The equation determining $\psi(\mathbf{x})$ is

$$\delta\psi: \quad \nabla^4\psi = -\gamma\sum_i (\psi(\mathbf{x}) - d_i)\delta(\mathbf{x} - \mathbf{x}_i) - \sum_j \gamma_j C_j[\psi]\delta C_j/\delta\psi. \tag{2.7}$$

The Lagrange multipliers γ and γ_j are uniquely determined by the misfits q and $\{Q_j\}$. In principle, we estimate the misfits by a scaling analysis, and determine the Lagrange multipliers from (2.4), (2.5), and (2.7). In practice, it is easier to guess values for the Lagrange multipliers, solve (2.7) for $\psi(\mathbf{x})$, and then determine the resulting misfits

from (2.4) and (2.5). The correct misfits can be reached by iterations, using the fact that larger values of the Lagrange multiplier correspond to smaller values of the corresponding misfit.

The variational formulation has several advantages. First and foremost, the relative emphasis on smoothing, agreement with the data, and agreement with the dynamical constraints can easily be changed by adjusting the Lagrange multipliers. This makes it possible to assess the quality of the various constraints. If, for example, we discover that some one of the constraints requires a very large (order one) misfit in order to achieve an adequate agreement with the data and the other dynamical constraints, then this particular constraint is obviously inappropriate and can be dropped. We thus learn something about the physics.

Secondly, the estimate $\psi(\mathbf{x})$ is produced from (2.7) by a global optimization. In contrast to other inverse methods, which pose dynamical conservation laws in the form of flux balances across a few arbitrarily-chosen closed curves, our method requires that these conservation laws be satisfied (approximately) in their differential form, at every point in the flow domain.

Finally, only the total misfit is specified for each of the data-dynamics constraints. Thus, for example (2.7) is free to concentrate all of the misfit with observations at only a few of the data points, if those data are incompatible with all the others, or with the dynamical constraints. The variational method therefore tends to isolate and disregard spurious data and dynamical constraints. Of course, the weight attached to each data-dynamics constraint can be made to vary with location in the flow. For example, instead of (2.4) we may write

$$\sum_i e_i^2/\sigma_i^2 = q^2 N_d/\sigma_0^2 \tag{2.8}$$

where σ_i^2 is the statistical average of e_i^2 and σ_0^2 is the average of the $\{\sigma_i^2\}$. For a statistically homogeneous flow (or in the absence of other information) we would assume that $\sigma_i = \sigma_0$.

The most arbitrary element of our formulation is the choice of a measure of roughness. For example, instead of (2.1) we could as easily take

$$R[\psi] = \int\int\int d\mathbf{x}\, \nabla\psi \cdot \nabla\psi \tag{2.9}$$

which attaches less penalty to very small-scale variability. The solution $\psi(x)$ to (2.7) with roughness measure (2.1) has continuous zeroth and first derivatives, whereas the corresponding solution with roughness measure (2.9) has logarithmic singularities at the data locations $\mathbf{x}_i$. In general, roughness measures containing higher order derivatives yield estimates with less singular behavior near the data locations. For a rigorous discussion, refer to Wahba and Wendelberger (1980). In this paper, we use discrete

numerical analogues of both (2.1) and (2.9) on the sphere. The numerical representation guarantees a finite solution in either case. If the solution is sensitively dependent on the choice of $R[\psi]$ or on the method of discretization, then our method has little value. However, we find that the results are not very dependent on the precise choice of the measure of roughness provided that there are a sufficient number of data and dynamical constraints. Shure *et al.* (1982) report a similar conclusion for the geomagnetic inverse problem.

3. Choise of dynamical constraints

We want to estimate the large-scale, time-average flow. To a very good approximation, this flow obeys the geostrophic equations,

$$f v = \frac{1}{\rho_0 a \cos\theta} \frac{\partial p}{\partial \lambda} \tag{3.1a}$$

$$f u = -\frac{1}{\rho_0 a} \frac{\partial p}{\partial \theta} \tag{3.1b}$$

the hydrostatic equation,

$$\frac{\partial p}{\partial z} = -\rho g \tag{3.2}$$

and an approximation to the exact continuity equation,

$$\frac{1}{a \cos\theta} \frac{\partial u}{\partial \lambda} + \frac{1}{a \cos\theta} \frac{\partial}{\partial \theta} (\cos\theta \cdot v) + \frac{\partial w}{\partial z} = 0. \tag{3.3}$$

Here (u, v, w) is the velocity in the (east, north, up) direction, p is the pressure, (λ, θ) is the (longitude, latitude), ρ the *in situ* density, ρ_0 a constant reference density, g is gravity, a is the radius of the Earth, and $f = 2\Omega \sin\theta$ where $\Omega = 2\pi$ day^{-1}. Let (U, L, H) be scales for (horizontal velocity, horizontal distance, vertical distance). If $U = 10$ cm sec^{-1}, $L = 100$ km, $H = 5$ km, then the Rossby number

$$\epsilon \equiv \frac{U}{\Omega L}$$

is of order 10^{-2}. Outside narrow boundary layers, the fractional error in (3.1) is ϵ, the error in (3.2) is $\epsilon^2 H/L \sim 10^{-5}$, and the error in (3.3) is $\delta\rho/\rho_0 \sim 10^{-3}$ where $\delta\rho$ is a typical change in the *in situ* density. These errors are all very small compared to the 'errors' in the data and in the additional dynamical constraints introduced below. Let

$$\psi(\lambda, \theta, z) \equiv p/g. \tag{3.4}$$

Then by (3.1)

$$u = \frac{-g}{a\,\rho_0 f}\frac{\partial \psi}{\partial \theta} \tag{3.5a}$$

$$v = \frac{g}{a\,\rho_0 f \cos\theta}\frac{\partial \psi}{\partial \lambda} \tag{3.5b}$$

and by (3.2)

$$\rho = -\frac{\partial \psi}{\partial z}. \tag{3.6}$$

By (3.3) and (3.5) we also have Sverdrup's equation,

$$\frac{\partial w}{\partial z} = \frac{g}{\rho_0 a^2}\frac{2\Omega}{f^2}\frac{\partial \psi}{\partial \lambda}. \tag{3.7}$$

Thus, u, v, w, and ρ are all expressible in terms of ψ. The hydrographic data are essentially aliased measurements of ρ. Thus

$$d_i = -\frac{\partial \psi}{\partial z}(\lambda_i, \theta_i, z_i) + \text{aliasing error} \tag{3.8}$$

are also expressible in terms of ψ.

The fractional error in (3.7) is dominated by eddy contributions to the vorticity equation, and is of the order of perhaps 10%. Therefore, (3.7) is a usable constraint on ψ, but only if we can eliminate the vertical velocity w. There are two ways to do this. The first is to integrate (3.7) between levels at which w is known or can be related to ψ by a boundary condition. If this method is adopted, then the corresponding misfit is relatively small because (3.7) is relatively accurate. The second method is to combine (3.7) with another equation containing w and ψ. The most obvious other equation is some form of the 'density equation,' i.e., the first law of thermodynamics. Unfortunately, all usable approximations to the density equation have relatively large errors, and hence require a relatively large misfit.

We apply the constraint (3.7) in both the vertically integrated form and also in the nonintegrated form in combination with the density equation. Since (3.7) is accurate outside boundary layers, we integrate from the top of the bottom Ekman layer to the bottom of the top Ekman layer and obtain

$$\frac{g}{\rho_0 a^2}\frac{2\Omega}{f^2}\frac{\partial \bar{\psi}}{\partial \lambda} = w_T - w_B \tag{3.9}$$

where

$$\bar{\psi}(\lambda, \theta) \equiv \int dz\, \psi(\lambda, \theta, z) \tag{3.10}$$

and

$$w_T = \frac{1}{\rho_0 f} \operatorname{curl} \tau \tag{3.11}$$

is the usual Ekman pumping approximation. Here, τ is the average wind stress at the ocean surface. If the Ekman suction into the bottom Ekman layer can be safely neglected, then the bottom vertical velocity is

$$\begin{aligned} w_B &= -\mathbf{u} \cdot \nabla H \\ &= \frac{g}{\rho_0 a^2} \frac{1}{f \cos\theta} \frac{\partial(H, \psi_B)}{\partial(\lambda, \theta)} \end{aligned} \tag{3.12}$$

where $\psi_B \equiv \psi(\lambda, \theta, -H)$, and $H(\lambda, \theta)$ is the depth of the ocean. Eq. (3.12) expresses the kinematic condition of no normal flow through the bottom. Eqs. (3.9–3.12) may be combined into a single constraint on ψ, namely

$$\frac{2\Omega}{f} \frac{\partial \bar{\psi}}{\partial \lambda} = \frac{a^2}{g} \operatorname{curl} \tau - \frac{1}{\cos\theta} \frac{\partial(H, \psi_B)}{\partial(\lambda, \theta)}. \tag{3.13}$$

This constraint applies to an accuracy of roughly 10%, depending strongly on the size of L, the representative length scale for the average flow. Note that the constraint (3.13) does not necessarily imply that the wind stress drives the local mean flow. In fact, in the application described in this paper, the wind stress curl is small of the same order as the eddy terms already neglected from (3.13), and we use (3.13) with $\tau = 0$. The equation then expresses the integrated conservation of average potential vorticity.

To apply the constraint (3.7) in nonintegrated form we invoke a form of the density equation. A relatively accurate form of the density equation is

$$\frac{D\rho^*}{Dt^*} = \frac{1}{c^2} \frac{Dp^*}{Dt^*} \tag{3.14}$$

where the asterisks denote exact, unaveraged flow variables. In (3.14), c is the sound speed, and only the effects of molecular heat and salt diffusion have been ignored. If ρ^* is replaced by the exact potential density, then the right side of (3.14) is zero. The most accurate usable approximation to (3.14) is

$$\frac{u}{a \cos\theta} \frac{\partial \rho}{\partial \lambda} + \frac{v}{a} \frac{\partial \rho}{\partial \theta} + w \frac{\partial \rho}{\partial z} = -w \frac{\rho_0 g}{c^2} \tag{3.15}$$

where, as in all other equations except (3.14), the variables represent the sought-for average flow. The largest errors in (3.15) result from eddy fluxes of density by the unaveraged flow. The sizes of these eddy fluxes are rather difficult to estimate in general, but they could easily be as large as the terms retained in (3.15). We therefore

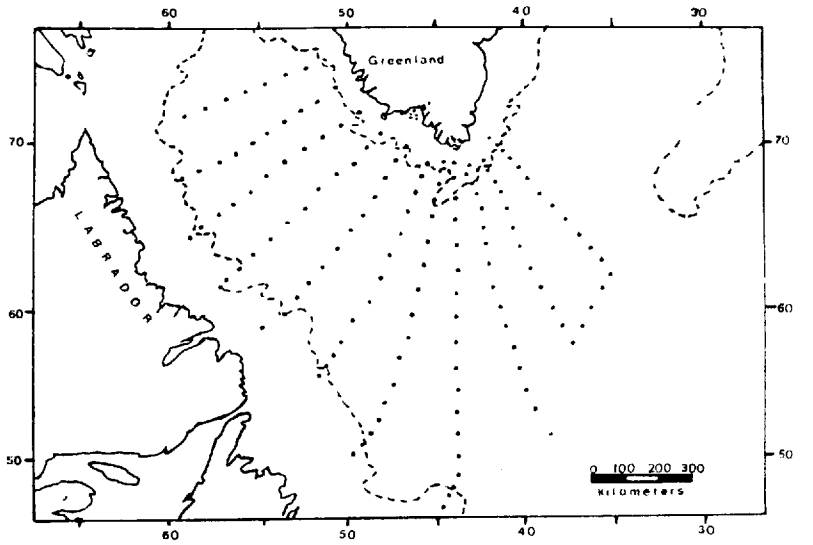

Figure 1. The 143 stations on 11 sections occupied by *Hudson* from 12 March to 12 May 1966.

regard the constraint obtained by eliminating w between (3.7) and (3.15) as rather weak, and we will apply it with a correspondingly large misfit.

We can increase the number of independent constraints by considering tracers other than density. If, for example, α is the average value of a conserved tracer which has been measured at the hydrographic stations, then we may use

$$\frac{u}{a\cos\theta}\frac{\partial\alpha}{\partial\lambda}+\frac{v}{a}\frac{\partial\alpha}{\partial\theta}+w\frac{\partial\alpha}{\partial z}=0 \tag{3.16}$$

as an additional constraint. Again, the dominant errors in (3.16) arise from eddy fluxes of tracer. If α is nonconservative (like oxygen), then explicit source terms must be added to (3.16), or the sizes of the source terms must be taken into account in deciding the misfit on (3.16).

4. Application to the Labrador Sea

C.C.S. *Hudson* occupied 143 hydrographic stations in the Labrador Sea between 12 March and 12 May 1966 at the locations shown in Figure 1. The Hudson measurements (of temperature, salinity and oxygen) have been described by Grant (1968), Lazier (1973), and Ivers (1975). In this section we apply the general method described in Sections 2 and 3 to the Hudson data, to estimate the large-scale field of horizontal and vertical velocity averaged over the two-month period of the cruise.

Figure 2 summarizes the generally accepted picture of Labrador Sea circulation. Both the shallow and deep currents are cyclonic and generally follow the bathymetric

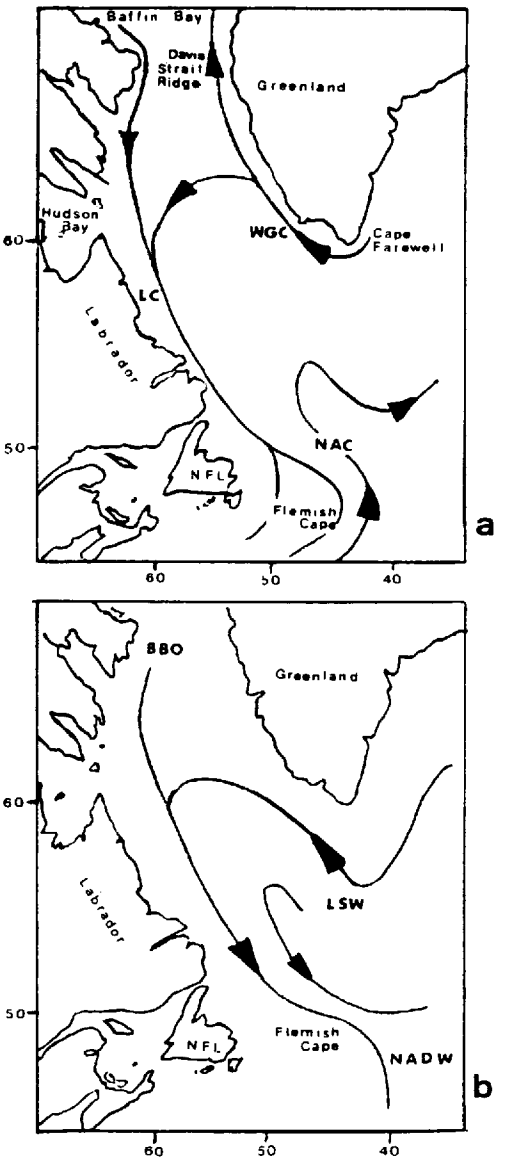

Figure 2. Schematic diagram of the circulation in the Labrador Sea near the surface (a) and in deep water (b). For explanation, refer to text.

contours. The shallow circulation (Fig. 2a) consists of the West Greenland Current (WGC), the Labrador Current (LC), and the North Atlantic Current (NAC). The West Greenland Current is relatively narrow near Cape Farewell, but it widens as it flows northwestward, and divides into two branches. One branch crosses the Davis Strait into Baffin Bay. The other branch turns westward and joins the Arctic water

flowing out of Baffin Bay to produce the Labrador Current, which flows south to Grand Banks. The North Atlantic Current meanders northeastward as far as 55N. This surface circulation was recognized in the early work of Smith *et al.* (1937).

The deep circulation in the Labrador Sea was underestimated until directly measured by Swallow and Worthington (1969) using neutrally buoyant floats. They estimated a deep transport of about 10 Sverdrups. Swallow and Worthington proved that the deep waters of the Labrador Sea originate from Norwegian Sea overflow and not from deep convection, as had previously been thought. However, intermittent deep convection does occur in the Labrador Sea in winter. Lazier (1973) documents an apparent renewal to 1500 m in 1967. The Labrador Sea water thus formed spreads at mid-depth and is found throughout the North Atlantic north of 40N and along the western boundary to 18N (Talley and McCartney, 1982). However, the Labrador Sea is a region of great inter-annual variability, and the 1966 data used in this paper show no evidence for deep convection during the previous winter.

Following the procedure described in Sections 2 and 3, we seek the smoothest field of $\psi(\lambda, \theta, z)$, defined by (3.4), which agrees with the Hudson data and with the approximate dynamical constraints. However, rather than using all the available constraints at once, we consider instead a hierarchy of variational problems, in which the number of constraints is gradually increased, in order of decreasing importance of the constraint. In the simplest member of the hierarchy, only the roughness measure and the data misfit are minimized.

a. Data-only models. We assume *a priori* that the vertical dependence of $\psi(\lambda, \theta, z)$, can be fit as closely as desired by a low-order polynomial in z. Specifically,

$$\psi(\lambda, \theta, z) = \sum_{m=1}^{M} \psi_m(\lambda, \theta)\phi_m(z) \tag{4.1}$$

where

$$\phi_m(z) = z^{m-1} \tag{4.2}$$

and $\psi_m(\lambda, \theta)$ remain to be determined. The assumption (4.1) (with M a small integer) is suppported by numerous observations of large-scale low-frequency flow and by the results described below. In the simplest model of our hierarchy, we minimize

$$J_1[\psi] \equiv R_1[\psi] + \gamma \sum_{i=1}^{N_d} \left(\frac{\partial \psi}{\partial z} (\lambda_i, \theta_i, z_i) + d_i \right)^2 \tag{4.3}$$

where

$$R_1[\psi] \equiv \int\int\int dz\, d\lambda\, d\theta \cos\theta\, \nabla_H \psi \cdot \nabla_H \psi \tag{4.4}$$

is a roughness measure and

$$\nabla_H = \left(\frac{1}{\cos\theta}\frac{\partial}{\partial\lambda}, \frac{\partial}{\partial\theta}\right) \tag{4.5}$$

is the horizontal gradient operator in spherical coordinates. In (4.3), d_i is the density measurement at $(\lambda_i, \theta_i, z_i)$, and λ is the Lagrange multiplier corresponding to the data misfit q, defined by

$$\sum_{i=1}^{N_d}\left(\frac{\partial\psi}{\partial z}(\lambda_i, \theta_i, z_i) + d_i\right)^2 = q^2 N_d. \tag{4.6}$$

Note that (4.4) does not penalize vertical variability, which is already limited by (4.1) (with M a small integer). The variational equations corresponding to

$$\delta J_1[\psi] = 0 \tag{4.7}$$

are

$$\cos\theta \cdot \nabla_H^2\left[\sum_m \psi_m \int \phi_m\phi_n\, dz\right] = \gamma\sum_i\left[\sum_m \psi_m(\lambda_i, \theta_i)\,\phi_m'(z_i) + d_i\right] \phi_n'(z_i)\,\delta(\lambda - \lambda_i)\,\delta(\theta - \theta_i), \quad \text{all } n \tag{4.8}$$

with boundary conditions

$$\nabla_H \psi_m \cdot \mathbf{n} = 0 \tag{4.9}$$

where

$$\nabla_H^2 = \frac{1}{\cos^2\theta}\frac{\partial^2}{\partial\lambda^2} + \frac{1}{\cos\theta}\frac{\partial}{\partial\theta}\left(\cos\theta\frac{\partial}{\partial\theta}\right) \tag{4.10}$$

is the Laplacian operator in spherical coordinates and $\mathbf{n}$ is the unit normal to the horizontal boundary.

Eqs. (4.8) are a coupled set of elliptic equations in the variables $\{\psi_m\}$. Eqs. (4.9), corresponding to no tangential flow at the computational boundary, are the 'natural' boundary conditions arising from the contributions to δJ_1 from $\delta\psi$ at the boundary. If alternative boundary conditions are desired, they may be inserted as constraints. For example, if the data extend up to the coastal boundaries, and if there are no boundary layers associated with unaveraged motions, then it may be preferable to require that the average flow have no normal component at the boundary. In our case these assumptions appear unjustified, and we regard the final estimate as meaningless outside the region of data coverage. Of course, both the variational equations and their boundary conditions change as we add dynamical constraints containing spatial derivatives of ψ.

The above equations must be solved numerically, and, because of the complicated

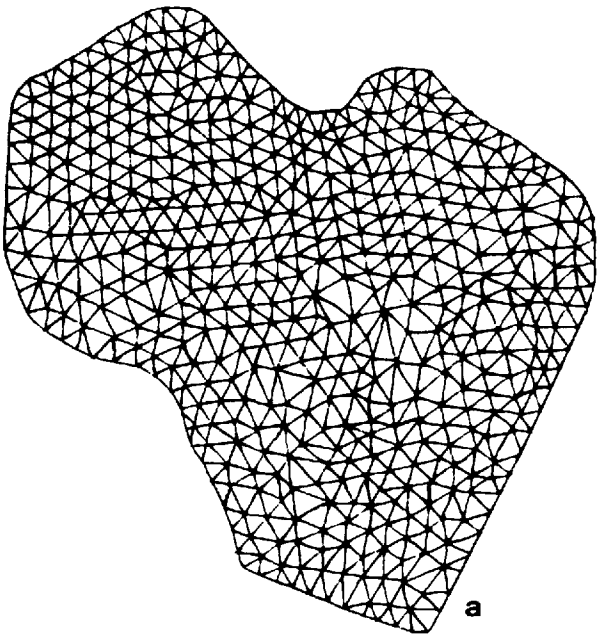

Figure 3. The finite element mesh (a) and the locations of the Hudson stations (b) and the bathymetry (in meters) (c) within the domain covered by the mesh.

geometry, we have used a method based upon finite elements and the Rayleigh-Ritz procedure. The Labrador Sea was covered by the triangular mesh shown in Figure 3. Each triangle is an element and the triangle vertices are called nodes. The node locations bear no relation to the locations of hydrographic stations. Within each element we assume that each $\psi_m(\lambda, \theta)$ is a linear interpolate of its three nodal values. Concisely,

$$\psi_m(\lambda, \theta) = \sum_{n=1}^{N} \psi_m^n S_n(\lambda, \theta) \tag{4.11}$$

where N is the number of nodes, and the 'shape function' $S_n(\lambda, \theta)$ is a tent-shaped function of λ and θ which: (a) is piecewise linear in both λ and θ on each element; (b) has unit value at the n-th node; (c) is zero at every other node. We substitute (4.11) into (4.3) and perform the integrations. Then J_1 is an ordinary function of the NM variables $\{\psi_m^n\}$. The equations

$$\frac{\partial J_1}{\partial(\psi_m^n)} = 0, \quad \text{all } m, \quad \text{all } n, \tag{4.12}$$

are finite-element analogues of (4.8) and (4.9). As the element size shrinks to zero, the solutions (4.11) of (4.12) converge to the analytical solutions of (4.8) and (4.9). Since

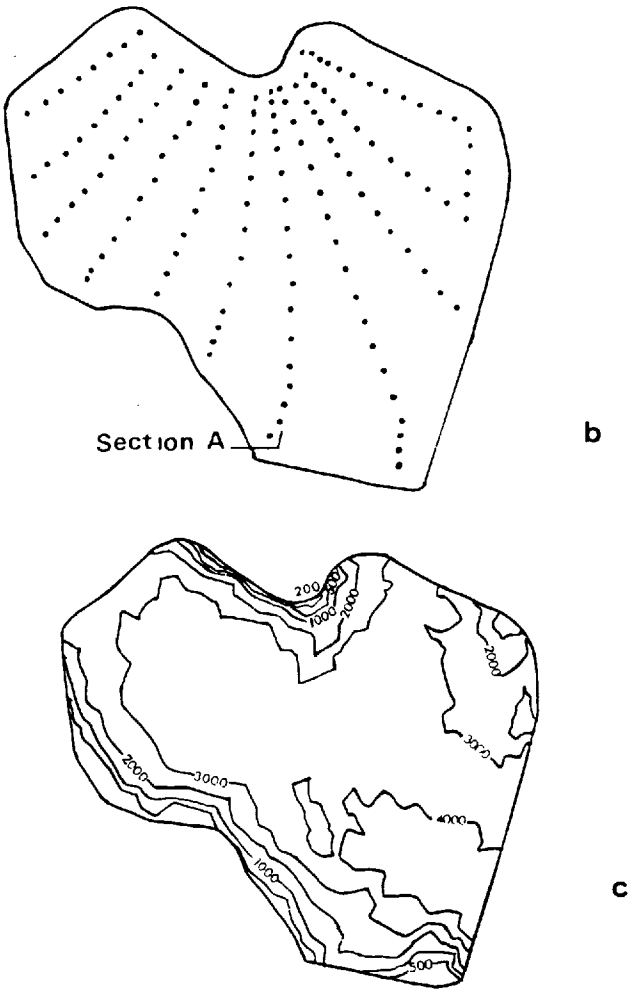

Figure 3. (Continued)

(4.12) are linear in the $\{\psi_m^n\}$, they may be written in the standard form

$$\mathbf{A}\psi = \mathbf{d} \tag{4.13}$$

where ψ is the column vector of unknowns ψ_m^n, $\mathbf{d}$ depends on the measurements d_i, and $\mathbf{A}$ is a very sparse square matrix with dimension NM. Our representation contains 600 nodes and (typically) 5 vertical modes. Thus $\mathbf{A}$ is 3000 by 3000. We solve (4.13) by block-symmetric over-relaxation, accelerated by the conjugate gradient method (Concus *et al.*, 1976). For details, refer to Provost (1983). The function $\psi_0(\lambda, \theta)$ remains undetermined by a physically irrelevant constant. As explained in Section 2, it is

Table 1.

M	3	4	5	6
q_M	.097	.042	.017	.014

easiest to guess a value for the Lagrange multiplier γ, and then adjust the guess to obtain the desired misfit q. Since

$$\gamma = -\frac{1}{N} dR / d(q^2) \tag{4.14}$$

we expect $\gamma > 0$, if decreases in the prescribed misfit cause increases in the minimum roughness. It can then be shown (Provost, 1983) that γ determines q uniquely and vice versa. The values $\gamma = \infty$ and $\gamma = 0$ correspond respectively to $q = 0$ and to

$$q^2 = \frac{1}{N_d} \sum_i d_i^2. \tag{4.15}$$

The data-only model contains no real dynamical information, and the levels-of-no-motion are decided solely by the extremely arbitrary criterion of minimum horizontal kinetic energy (4.4). The solutions $\psi(\lambda, \theta, z)$ of (4.6–4.9) are not therefore superior to dynamic topography maps drawn by subjective methods. However, the data-only solutions serve as interesting benchmarks for the more sophisticated members of the hierarchy, and they illustrate ideas of general importance.

We have solved the data-only model with various amounts of vertical resolution ($M = 3, 4, 5, 6$). The choice $M = 3$ represents the vertical profile of density by a straight line, while $M = 6$ represents the density by a fourth order polynomial in z. For all M we find that, as the Lagrange multiplier γ is increased from zero, the roughness (4.4) increases gradually until a critical misfit, q_M, is reached. Further increases in γ cause the sudden appearance of energetic small eddies in ψ (of size comparable to the station separation), with no significant further decrease in q_M. We call q_M the 'minimum achievable misfit' with M vertical modes. Table 1 gives q_M in sigma units of density. For comparison, instrumental errors alone would correspond to a misfit of .005 sigma units. But again, instrumental errors should be small compared to the aliasing errors caused by eddies and other unaveraged motions. We see from Table 1 that q_M decreases as the vertical resolution increases, but that the difference between $M = 5$ and $M = 6$ is very slight. Figure 4 compares the data-only streamfunction difference between the surface and 1000 m (with $M = 5$) with a subjective map of dynamic topography relative to 1000 m. The resemblance is very close. In all of the following, we therefore use $M = 5$.

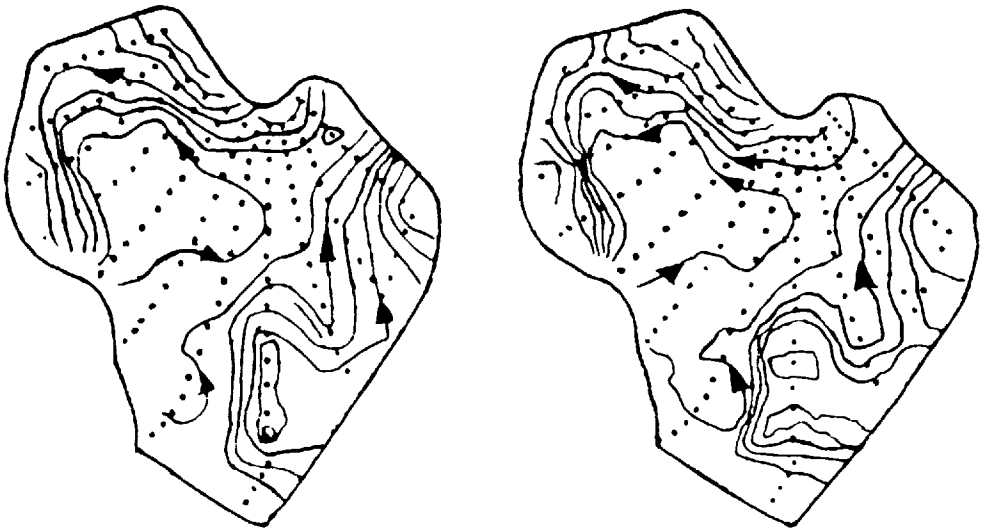

Figure 4. The surface streamfunction $\psi(\lambda, \theta, 0)$ minus the streamfunction at 1000 m, obtained from (a) the data-only model described in the text with five vertical modes, and (b) a subjective method based upon a level-of-no-motion at 1000 m. Each contour is one dynamic centimeter.

Figure 5 compares sections of horizontal velocity obtained from the data-only model using two distinct measures of roughness: the horizontal kinetic energy R_1, and the 'enstrophy',[3]

$$R_2[\psi] = \int\int\int dz\, d\lambda\, d\theta \cos\theta\, (\nabla_H^2 \psi)^2. \tag{4.16}$$

Although the data misfits are comparable ($q = .020$ and $.021$), these two sections are rather different. (We will show that these differences disappear as the number of constraints increases.) The minimum energy section (Fig. 5a) has smaller velocity extrema but greater small-scale variability than the minimum-enstrophy section (Fig. 5b). Again, however, neither solution contains any physics which is not implicit in the measurements. In particular, neither model has a physical basis for deciding the constants of vertical integration in the velocity. To remedy this situation, we begin to add the dynamical constraints discussed in Section 3.

b. Data-Sverdrup models. The second member of our variational hierarchy includes both the data constraint (4.6) and an approximation to the Sverdrup constraint (3.9).

3. Since (4.16) contains second derivatives, we cannot directly substitute the piecewise linear representation (4.11) into (4.16). Instead we replace (4.16) by a finite-difference estimate, and we find that the results are insensitive to the precise choice of difference estimate.

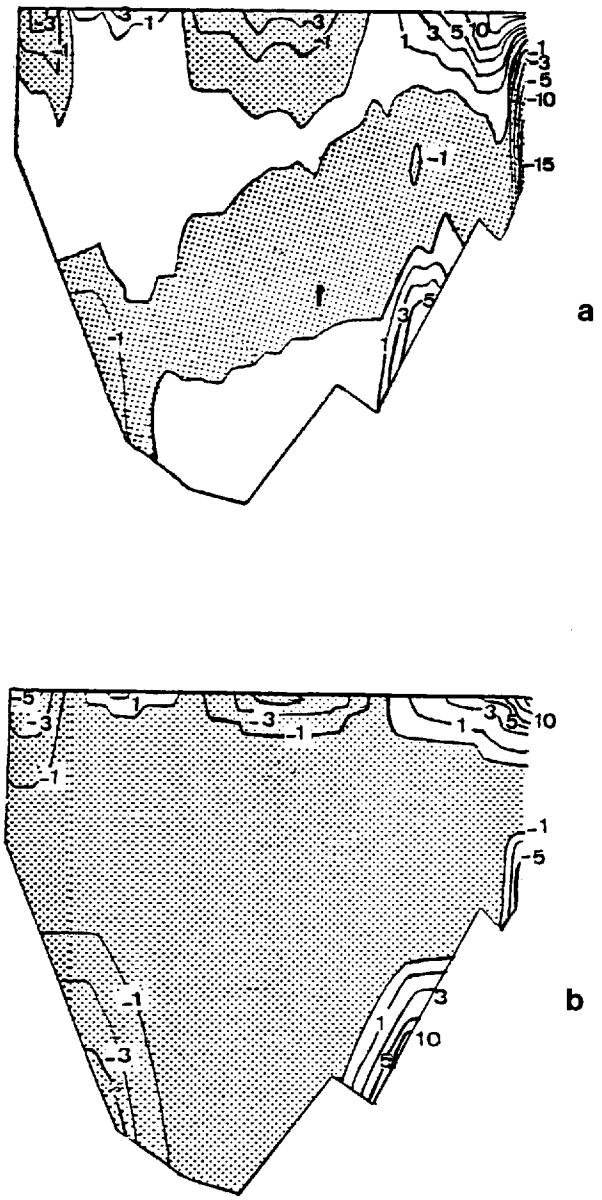

Figure 5. Normal velocity (positive northward) through section A (Figure 3b) in cm sec^{-1} obtained from the data-only model with (a) data misfit $q = .020$ sigma units and roughness measure R_1; and (b) $q = .021$ and measure R_2.

The Sverdrup constraint is very sensitive to the constant of vertical integration in the horizontal velocity, because the absolute velocity enters both the advection of planetary vorticity (the left side of (3.9)) and the bottom torque of (3.12). The primary inaccuracy in (3.9) results from the neglect of eddy vorticity fluxes. These fluxes are no greater than about 10% of the planetary advection on the large spatial scales of interest. The left side of (3.9) scales as $HU/a \sim 10^{-3}$ cm sec^{-1}. Therefore (3.9) applies with a misfit Q_s of about 10^{-4} cm sec^{-1}. The upper Ekman velocity can be estimated from the average wind stress curl and (3.11). According to Leetmaa and Bunker (1978), the average wind curl is relatively weak, and w_T is less than 10^{-4} cm sec^{-1} over much of the Labrador Sea. Thus w_T has the same size as the misfit Q_s and can be neglected with no additional loss in accuracy.

The bottom vertical velocity w_B depends on the magnitude of the horizontal velocity at the bottom, and on the angle between the horizontal velocity vector and the isobaths. However, w_B scales no larger than sU, where s is the scale for the bottom slope. Near the boundaries of the Labrador Sea $s \sim 10^{-3}$ to 10^{-2} so that $sU \sim 10^{-2}$ to 10^{-1} cm sec^{-1}, i.e., ten to one hundred times larger than the left side of (3.9). Therefore, on these steep slopes, the Sverdrup constraint will either force the bottom horizontal velocity to be near zero or to parallel the isobaths. In the central region of the Labrador Sea $s \sim 10^{-5}$ to 10^{-4} and $sU \sim 10^{-4}$ to 10^{-3} cm sec^{-1}. Therefore in the flattest regions, the Sverdrup constraint will not inhibit cross-contour bottom flow, but the constraint will tend to choose a level-of-no motion for which the vertically integrated transport across latitude lines is relatively small.

We now consider solutions to

$$\delta J_2[\psi] = 0 \tag{4.17}$$

where

$$J_2[\psi] = J_1[\psi] + \gamma_s \cdot \text{(4.19)} \tag{4.18}$$

and γ_s is the Lagrange multiplier corresponding to the Sverdrup constraint

$$\int\int d\lambda\, d\theta \cos\theta \left\{ \left[\frac{g}{\rho_0 a^2} \frac{2\Omega}{f^2} \frac{\partial \bar{\psi}}{\partial \lambda} + w_B \right]^2 - Q_s^2 \right\} = 0 \tag{4.19}$$

with misfit Q_s. Each solution $\psi(\lambda, \theta, z)$ to (4.17) is uniquely determined by a value for the data misfit q and the Sverdrup misfit Q_s, and thus corresponds to a point on Figure 6. We regard the region D_S of Figure 6 as the domain of physically plausible solutions to (4.17). The domain D_S is bounded by:

(a) the vertical line $q = .035$ sigma units. This value of data misfit is (we believe) the largest which could be explained by aliasing errors in the data. All points to the right of $q = .035$ represent solutions which have unrealistically large disagreement with the data.

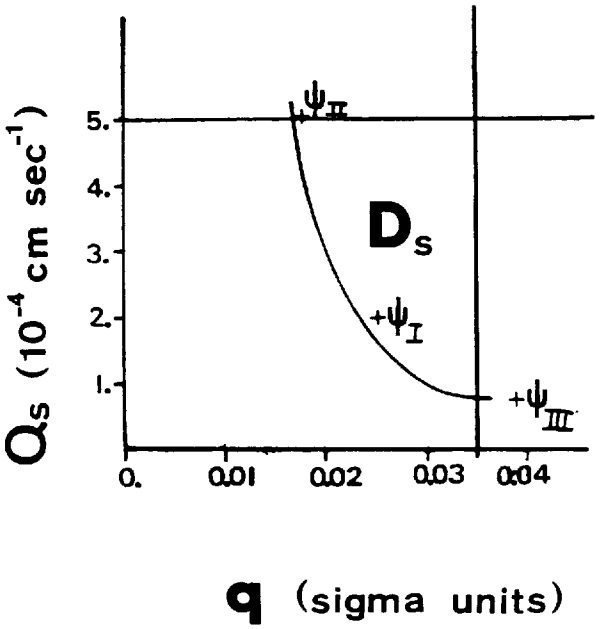

Figure 6. The domain D_S of physically plausible solutions for the data-Sverdrup model.

(b) the horizontal line $Q_s = 5 \times 10^{-4}$ cm sec^{-1}. This value of the Sverdrup misfit is the largest which can be accounted for by the neglected eddy fluxes of vorticity, the Ekman pumping at the surface, and errors in the bottom slope. All points above $Q_s = 5 \times 10^{-4}$ represent solutions which violate the Sverdrup constraint by an unrealistically large amount.

(c) a curved line of constant $R_1[\psi]$, corresponding to the roughness value at which small-scale eddies appear in $\psi(\lambda, \theta, z)$. All points below and left of this curved boundary represent solutions with unrealistically short horizontal scales (of the order of the station separation).

The three solutions ψ_I, ψ_{II}, $\psi_{III}(\lambda, \theta, z)$ shown on Figure 6 represent the extremes encountered within the domain D_S of plausible solutions. All three lie near the boundary of minimum acceptable horizontal smoothness. The solution ψ_{II} has the smallest misfit with the data, but it has an almost unacceptably large misfit with the Sverdrup constraint. ψ_{III} agrees well with the Sverdrup constraint, but its data misfit is relatively large. The solution ψ_I is a compromise between agreement with the data and agreement with the Sverdrup constraint. Figures 7, 8, and 9 show the streamfunction $\psi(\lambda, \theta, z)$ corresponding to the three models ψ_I, ψ_{II}, and ψ_{III} at various depths from 100 to 3500 m. These figures clearly show that a decrease in the Sverdrup misfit causes the flow over steep slopes to intensify and to flow more nearly parallel to the isobaths.

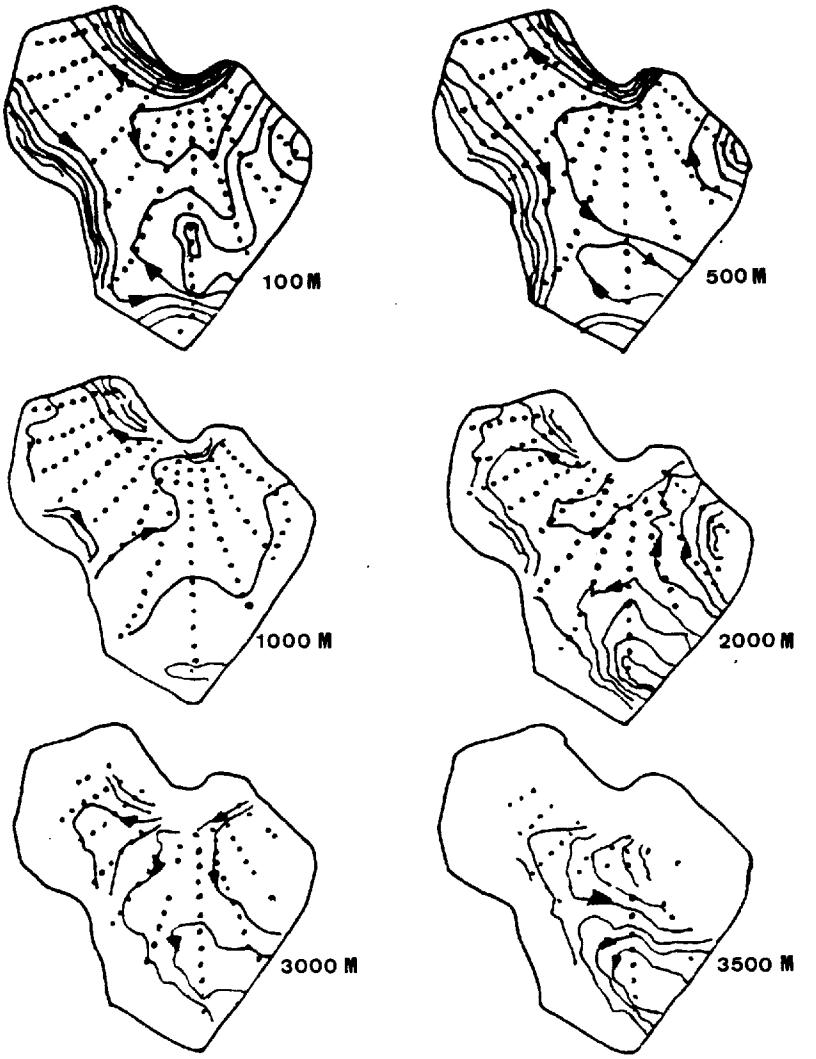

Figure 7. The data-Sverdrup steamfunction ψ_I at various depths. ψ_I is computed with a data misfit $q = .025$ sigma units and a Sverdrup misfit $Q_s = 2 \times 10^{-4}$ cm sec^{-1}.

We have examined other solutions from the interior of the domain D_S. These generally resemble ψ_I (but with greater horizontal smoothness) except near Cape Flemish in the southwestern Labrador Sea. In the latter region of steep slopes and poor data coverage, the solutions are very sensitive to the size of the Sverdrup misfit.

The streamfunction maps corresponding to ψ_I reproduce the major surface currents known from observations: the West Greenland current, the Labrador current, and the North Atlantic current. The Labrador and West Greenland currents generally follow

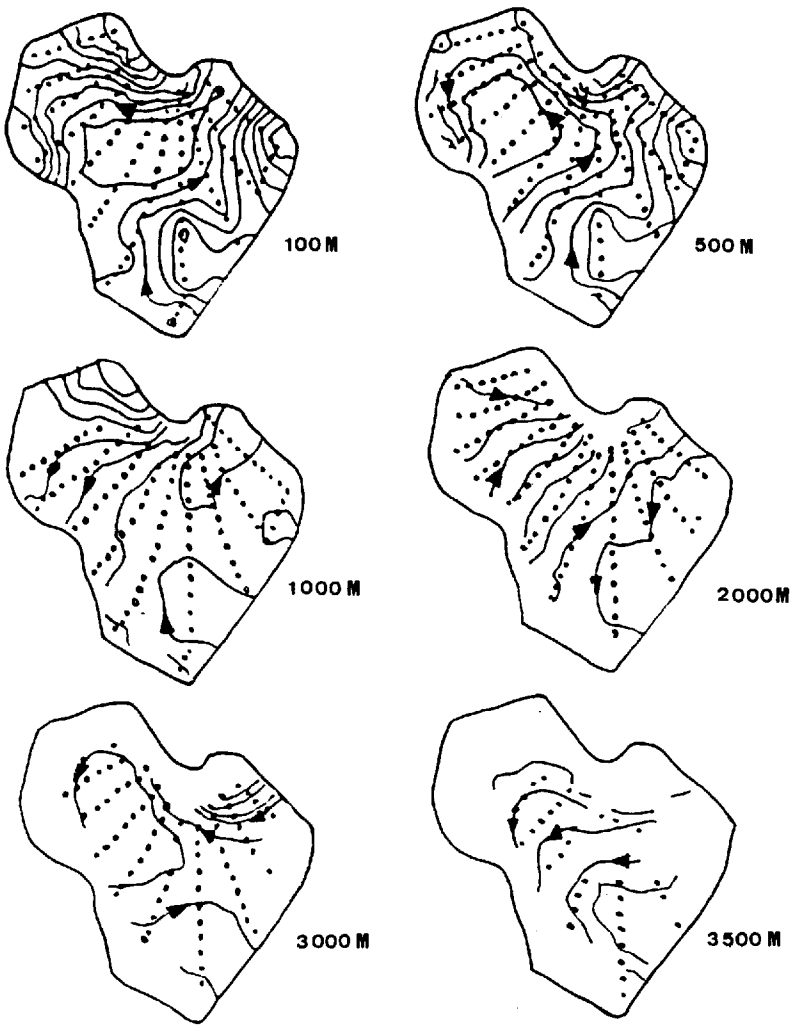

Figure 8. The same as Figure 7 but for ψ_{II}, with $q = .018$ and $Q_s = 5 \times 10^{-4}$.

the isobaths and appear to be steered by the bathymetry. The North Atlantic current is not closely correlated with bathymetry except over the steep slopes of the Reykjanes ridge. When we demand a close agreement with the Sverdrup constraint (Fig. 9) the meanders in the North Atlantic current disappear.

The flow at mid-depths in the Labrador Sea has been controversial. Ivers (1975) used a subjective method based upon the distributions of salinity and oxygen on neutral surfaces to infer a relatively strong cyclonic flow at mid-depths, whereas Swallow and Worthington (1969) argued that currents are near zero throughout the Labrador Sea in the depth range 1000–1200 m. In Figure 10 we compare the horizontal velocity field

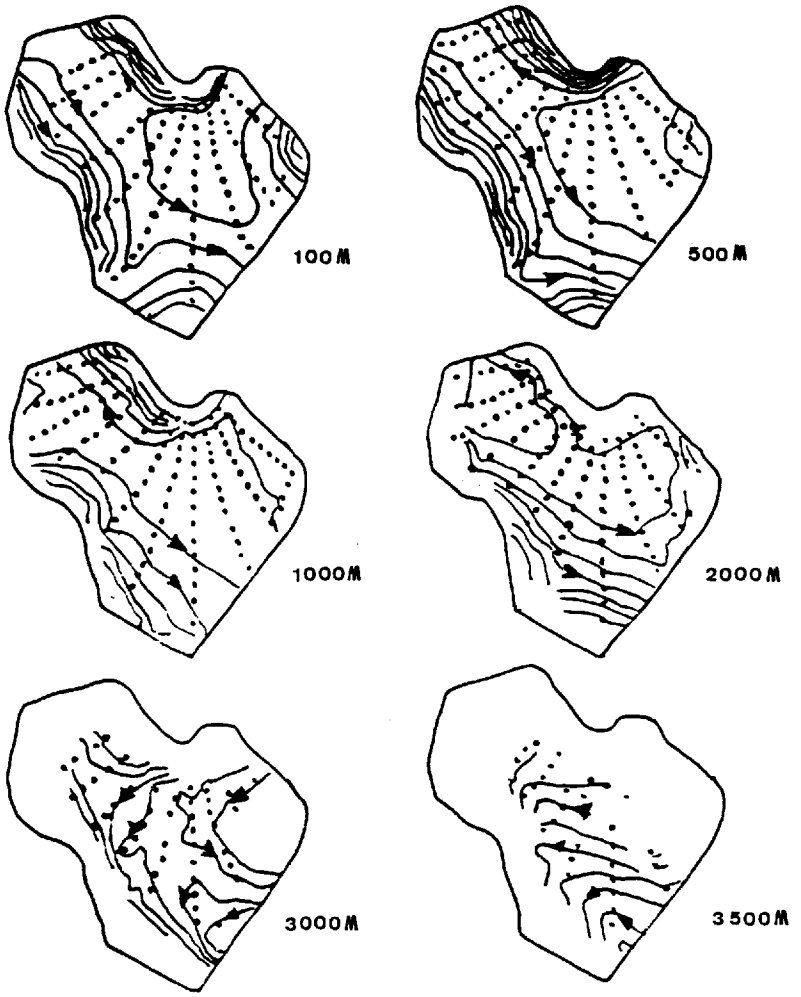

Figure 9. The same as Figure 7 but for ψ_{III}, with $q = .039$ and $Q_s = 8 \times 10^{-5}$.

through section A (Fig. 3b) from our solution ψ_I (Fig. 10a) with the corresponding result obtained by Ivers (Fig. 10c) and a third section (Fig. 10b) based upon a level-of-no-motion at 1000 m. Our section is similar to Ivers' in the western part of the section but more similar to Figure 10b in the eastern part.

The deep flow in ψ_I is also cyclonic, and generally follows the bathymetry. The large horizontal velocities (up to 30 cm sec^{-1}) agree with Ivers' conclusions.

We have computed the vertical velocity from (3.7) and (3.12) with $\psi = \psi_I$. Figure 11a shows the vertical velocity at 100 m depth. The velocity is small and positive (upward) over much of the area, and resembles the map of Ekman vertical velocity obtained by Leetmaa and Bunker (1978) from wind stress data (Fig. 11b). Although

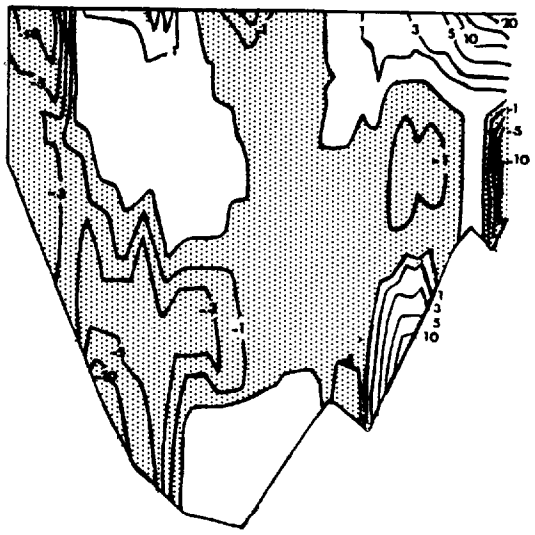

Figure 10. Northward velocity through section A predicted by (a) ψ_1, (b) subjective analysis with a level-of-no-motion at 1200 m and (c) Ivers (1975). Units of cm sec^{-1}.

these surface velocities are of the same order as the misfit $Q_s = 2 \times 10^{-4}$ cm sec^{-1}, the resemblance between Figures 11a and 11b supports the idea that the density measurements contain information about long-term wind driving. Figure 12 shows vertical velocity in the plane of section A. The bottom velocities are significantly larger than those at the surface. The deep water is rising on the Labrador side and sinking on the Greenland side. These signs of vertical velocity correspond to a conservation of potential velocity by water columns as the Greenland Current flows northward and into deeper water, and the Labrador Current flows southward into shallower water.

To test the sensitivity of data-Sverdrup models to the choice of roughness measure, we have solved the variational problem with $R_1[\psi]$ replaced by

$$R_2[\psi] = \int\int\int dz\, d\lambda\, d\theta \cos\theta\, (\nabla_H^2 \psi)^2. \tag{4.20}$$

The misfits q and Q_s are comparable to those used for the model ψ_1 above. The corresponding horizontal velocity through section A is shown in Figure 13. The differences between Figures 10a and 13 are significant, but much less pronounced than between Figures 5a and 5b. This reinforces our belief that the choice of smoothing operator is not crucial if the data and dynamics control the flow sufficiently. On the whole, the flow in Figure 13 obtained with the enstrophy measure (4.20) more closely resembles Ivers' section (Fig. 10c) than did ψ_1.

c. Data-Sverdrup-density models. To further reduce the domain of plausible solutions, and to reduce the differences caused by the choice of roughness measure, we

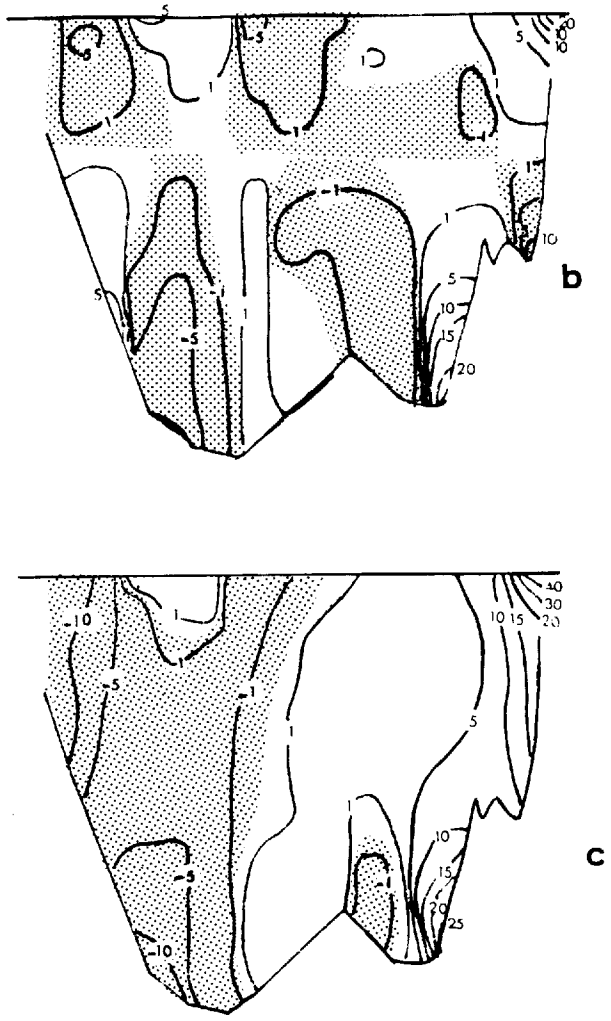

Figure 10. (Continued)

introduce our final constraint, the conservation of potential density in the form

$$\mathbf{u} \cdot \nabla \rho_\theta = 0 \tag{4.21}$$

where $\mathbf{u}$ and ρ_θ are the average velocity and potential density. Again, the errors in (4.21) result from neglected eddy fluxes of potential density, which may sometimes be as large as the terms retained. We therefore apply (4.21) with a relatively large misfit.

Eq. (4.21), unlike previous constraints, is nonlinear in the average flow variables, and would therefore require that the resulting variational problem be solved by

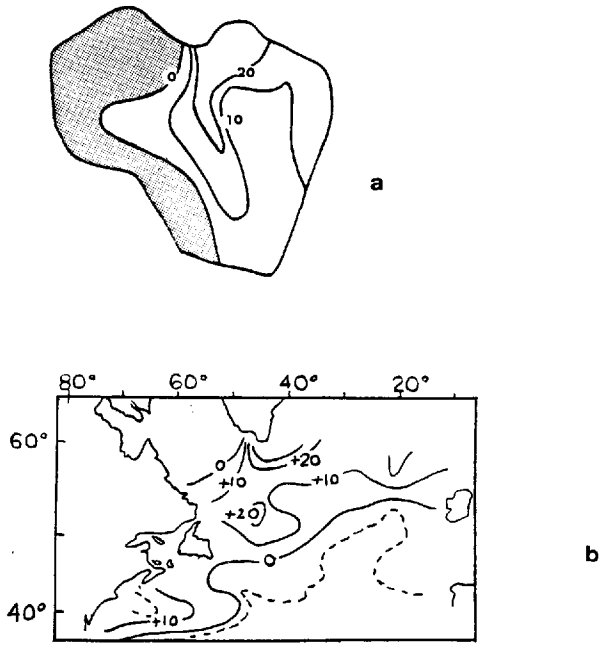

Figure 11. Vertical velocities in 10^{-5} cm sec^{-1} at 100 m depth predicted by (a) the data-Sverdrup solution ψ_I, and (b) Leetmaa and Bunker (1978).

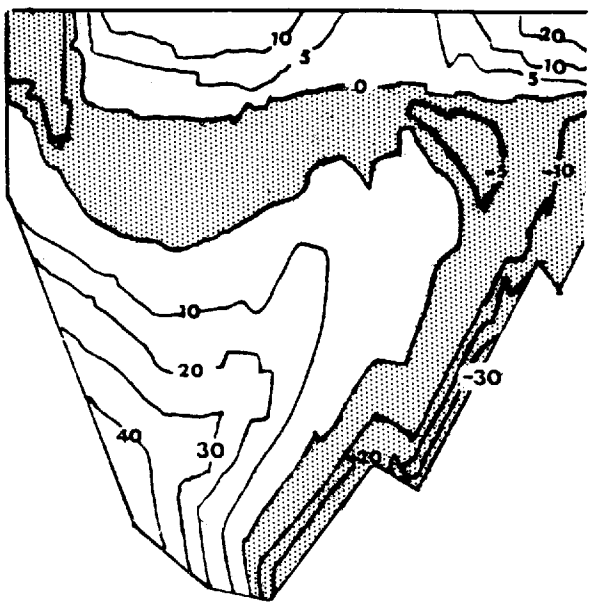

Figure 12. Vertical velocity in 10^{-5} cm sec^{-1} in the plane of section A (Fig. 3b) predicted by the data-Sverdrup solution ψ_I.

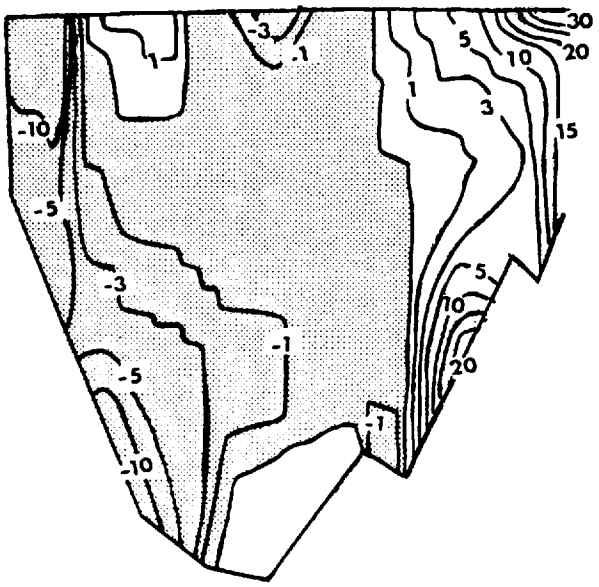

Figure 13. Northward velocity through section A in cm sec^{-1} predicted by the data-Sverdrup model with $q = .026$ sigma units, $Q_s = 1.8 \times 10^{-4}$ cm sec^{-1} and the enstrophy roughness measure R_2.

iterations. In the interest of economy, we *prescribe* ρ_θ *a priori,* from the data alone. Specifically, we determine ρ_θ as the smoothest potential density field (by roughness measure R_1) which is consistent with the data and a misfit of .020 sigma units. We then define

$$J_3[\psi] \equiv J_2[\psi] + \int\int\int d\lambda\, d\theta\, dz \cos\theta\, [(\mathbf{u} \cdot \nabla \rho_\theta)^2 - Q_\rho^2] \tag{4.22}$$

where $\mathbf{u}$ is given in terms of ψ by (3.5, 3.7, 3.12) and Q_ρ is the misfit on (4.21). The variational equation

$$\delta J_3[\psi] = 0 \tag{4.23}$$

is then linear in $\psi(\lambda, \theta, z)$ and can be solved by the same numerical methods as in previous cases. The numerical representation of ρ_θ uses the same number of vertical modes and horizontal elements as used to represent $\rho = -\partial\psi/\partial z$. Because we do not iterate, our final density estimate $-\partial\psi/\partial z$ is not perfectly consistent with ρ_θ. However, our procedure is justifiable both because the difference between ρ_θ and the potential density consistent with our variational solution is relatively small, and because the misfit Q_ρ on (4.21) is relatively large.

The misfit Q_ρ is more conveniently characterized by the "cancellation parameter"

$$p_c \equiv \frac{|\mathbf{u} \cdot \nabla \rho_\theta|}{\text{Max}\left|\frac{u}{a\cos\theta}\frac{\partial\rho_\theta}{\partial\lambda}, \frac{v}{a}\frac{\partial\rho_\theta}{\partial\theta}, w\frac{\partial\rho_\theta}{\partial z}\right|} \tag{4.24}$$

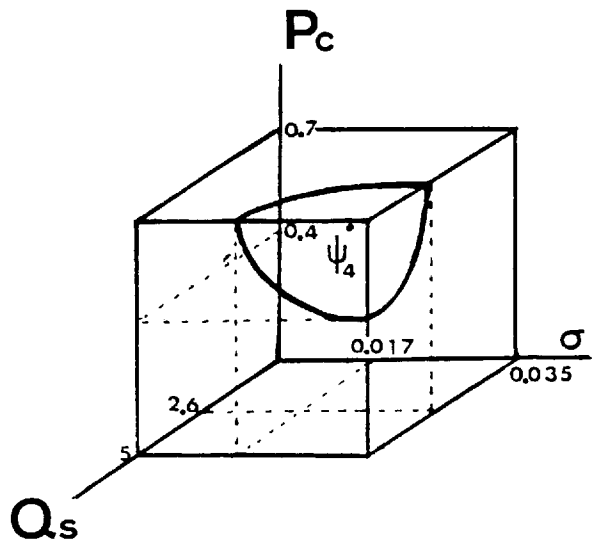

Figure 14. Domain $D_{S\rho}$ of plausible solutions in the space of misfits corresponding to the data-Sverdrup-density models.

whose smallness is a measure of the cancellation between terms in (4.21). Note that $0 < p_c < 3$. Every solution of (4.23) is characterized by a value of data misfit q, Sverdrup misfit Q_s, and P_c, the volume average of p_c, and hence corresponds to a point within the space depicted by Figure 14. The domain $D_{S\rho}$ of acceptable solutions is now a convex volume bounded by the three planes $q = .035$ sigma units, $Q_s = 5 \times 10^{-4}$ cm

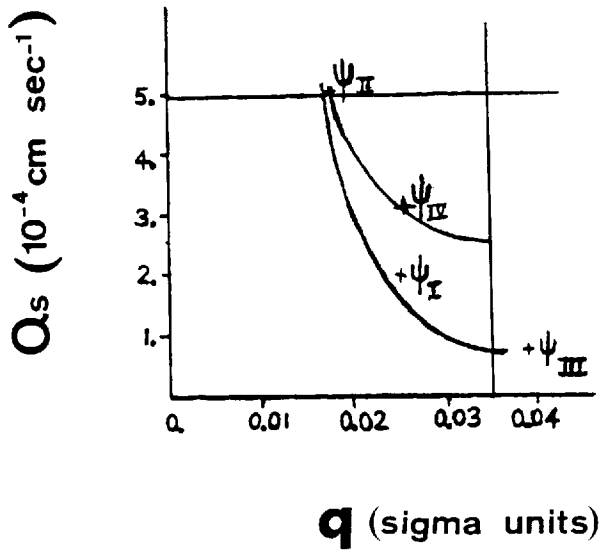

Figure 15. The projection of $D_{s\rho}$ onto the $q - Q_s$ plane. For explanation, refer to the text.

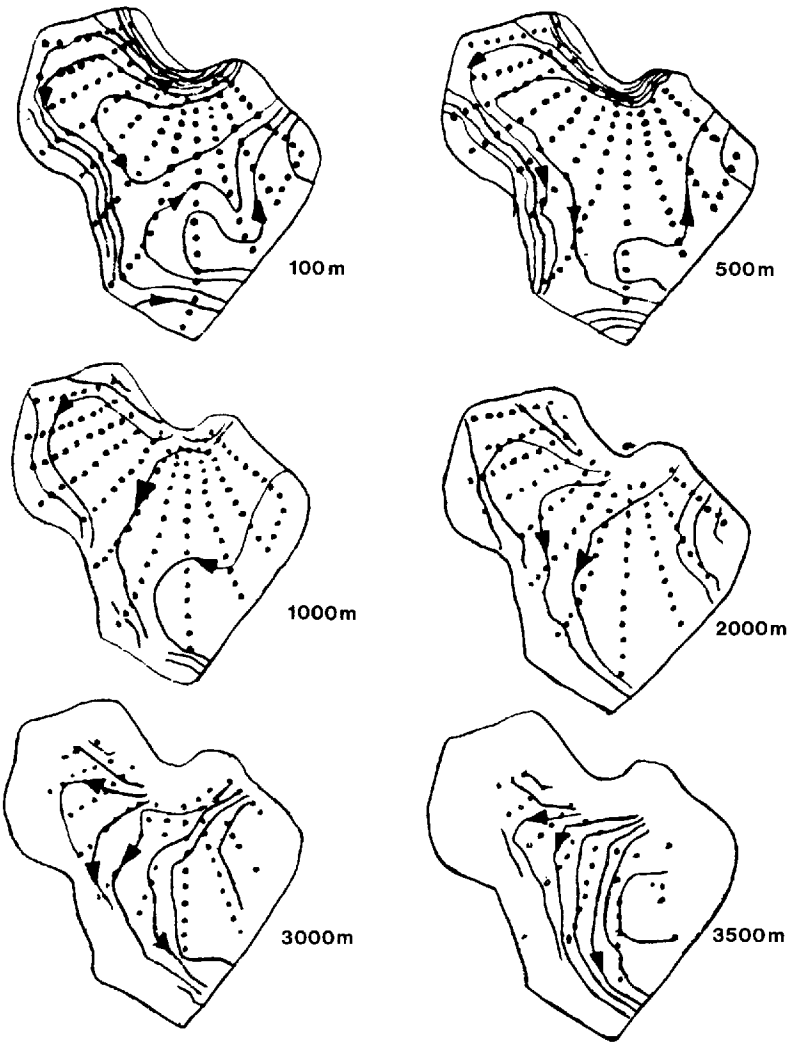

Figure 16. The same as Figure 7 but for the data-Sverdrup-density solution ψ_{IV}.

$\sec^{-1}$, and $P_c = .7$, which we regard as the maximum acceptable average noncancellation between the terms in the potential density equation. The curved boundary of $D_{S\rho}$ is a surface of minimum smoothness, corresponding to the same value of roughness R_1 as before. We find that the cancellation parameter P_c cannot be made smaller than $P_c = .4$ without exceeding the maximum misfit on the data or the Sverdrup constraint.

Figure 15 superimposes the "top face" of Figure 14 (i.e., the intersection between $D_{S\rho}$ and $P_c = .7$) on the domain D_S of Figure 6. Figure 15 shows that there must be a tradeoff between the misfit on the Sverdrup constraint and the misfit on the potential density equation. That is, we must tolerate a somewhat larger Q_s in order to reduce P_c

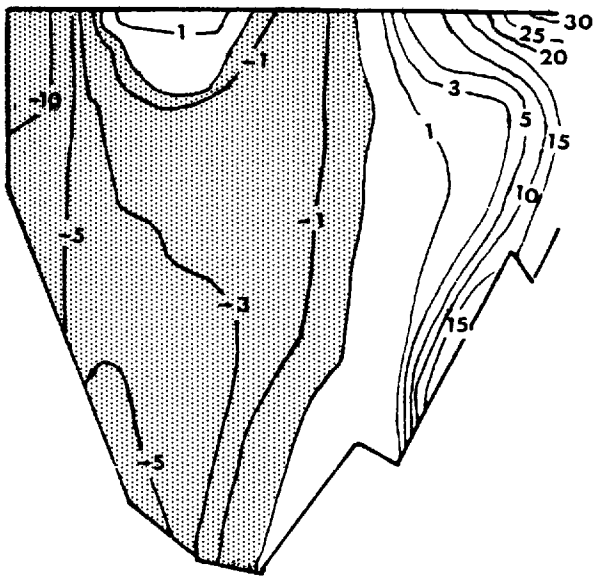

Figure 17. Northward velocity in cm sec^{-1} through section A corresponding to the data-Sverdrup-density solution ψ_{IV}.

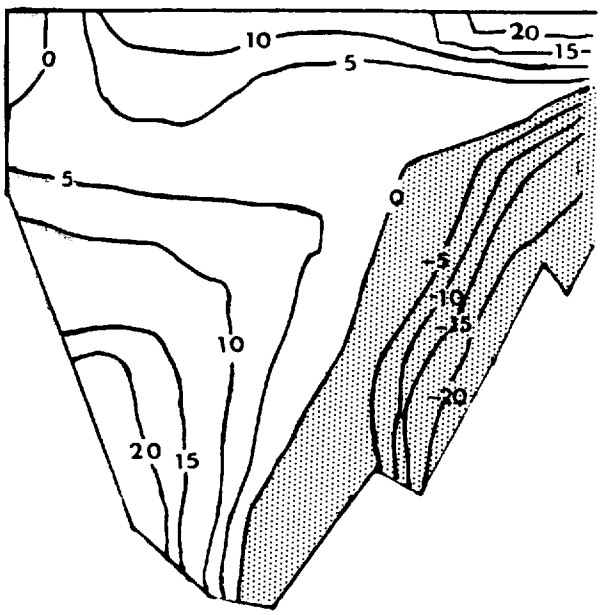

Figure 18. Vertical velocity predicted by ψ_{IV} in the plane of section A. Units of 10^{-5} cm sec^{-1}.