General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.

- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.

- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.

- This document is paginated as submitted by the original source.

- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

Produced by the NASA Center for Aerospace Information (CASI)

Subducted Slabs and the Geoid: Constraints on Mantle Rheology and Flow

(NASA-CR-170192) SUBDUCTED SLABS AND THE GEOID: CONSTRAINTS ON MANTLE RHEOLOGY AND FLOW (California Inst. of Tech.) 49 p
HC A03/MF A01 CSCL 08G

N83-22874

Unclas
G3/46 03360

Bradford H. Hager*

Seismological Laboratory 252-21
California Institute of Technology
Pasadena, CA 91125

April 15, 1983

*Alfred P. Sloan Foundation Fellow

Abstract

The total geoid anomaly resulting from a given density contrast in a convecting viscous earth is affected by the mass anomalies associated with the flow-induced deformation of the upper surface and internal compositional boundaries, as well as by the density contrast itself. If the internal density contrasts can be estimated, as is the case for subducted slabs, then the depth and variation of viscosity with depth of the convecting system can be constrained. The observed long-wavelength geoid is highly correlated with that predicted by a density model for seismically active subducted slabs. The (positive) sign of the correlation requires that viscosity increase with depth by a factor of 30 or more. The amplitude of the correlation can be explained if the density contrasts associated with subduction extend into the lower mantle or if subducted slabs exceeding 350 km in thickness are piled up over horizontal distances of thousands of km at the base of the upper mantle. Mantle-wide convection in a mantle that has a viscosity increasing with depth provides a simple explanation of the long-wavelength geoid anomalies over subduction zones.

Introduction

The accurate determination of the long-wavelength components of the earth's gravity field and the initial development of the concepts of plate tectonics occurred nearly concurrently. It was recognized early (e.g. Runcorn, 1967; McKenzie, 1969; Kaula, 1972) that convergence zones are generally associated with highs in the long-wavelength geoid. This association is apparent in Figure 1, which is a more recent geoid (Gaposhkin, 1979), showing the geodynamically interesting departures from the hydrostatic equilibrium figure (Nakiboglu, 1982, f = 1/299.63) superimposed on a map including plate boundaries. Although there is clearly much contributing to the geoid not associated with subduction zones, all major subduction zones are characterized either by geoid highs (Tonga and Java through Japan and Central and South America) or by local maxima in negative features (Kuriles through Aleutians).

The magnitudes of the long-wavelength ($\ell < 10$) geoid anomalies are comparable to those to be expected if only the excess density resulting from the thermal structure of subducted slabs is considered (McKenzie, 1969). At shorter wavelengths, the observed geoid anomaly is less than that predicted by thermal models, suggesting some form of regional compensation (Griggs, 1972; Chase, 1979, Crough and Jurdy, 1980; McAdoo, 1981; Davies, 1981).

At first glance, the association of geoid highs with high-density slabs provides one more interesting, but hardly vital, piece of evidence supporting the plate tectonic hypothesis, with some details of the compensation of higher harmonies yet to be worked out. But on closer scrutiny, this association provides much more important information. As

will be shown below, it places fundamental constraints on such geodynamical parameters as the variation of effective viscosity with depth. The association further suggests that there are substantial density contrasts in the lower mantle related to suducted slabs in the upper mantle. This suggests either mantle-wide flow or at least strong thermal coupling between convection cells in the upper mantle and lower mantle. An alternative explanation is that the mantle near subduction zones is nearly filled with dead slabs.

The reason that the long-wavelength geoid highs associated with subduction zones provide useful geodynamic information is that at the time scales appropriate for mantle convection, the earth responds as a viscous rather than as an elastic or rigid body. Evidence from postglacial rebound and laboratory experiments, as well as the large displacements associated with plate motions themselves, requires that mantle rocks deform by creeping flow and are describable by a viscous (though not necessarily Newtonian) rheology.

Convection in the earth causes deformation of the surface of the earth. Midoceanic ridges and deep sea trenches are readily observable manifestations of this phenomenon. The core-mantle boundary, and any intermediate chemical discontinuities that may exist, must likewise be deformed. The total geoid anomaly in a convecting system is the sum of the gravitational effects of the internal "driving" density contrasts and the effects resulting from the deformed boundaries (Morgan, 1965a; 1965b; McKenzie et al., 1974; McKenzie, 1977). The magnitude and even the sign of the total gravity anomaly as a function of wavelength depend on the spatial variations in effective viscosity (Morgan, 1965a;

McKenzie, 1977) and on the depth of the convecting system. (Richards and Hager, 1983). Thus, comparison of the total geoid anomaly with that due to the "driving" density contrasts alone can place fundamental constraints on both the variation of mantle rheology and the depth of mantle convection. This will be illustrated using some simple models in the next section.

The approach I take here is to locate subducted slabs by their seismic activity and, assuming a density contrast between slab and mantle, calculate a slab geoid anomaly δN_s^ℓ as a function of wavelength or, equivalently, of spherical harmonic degree ℓ. This geoid anomaly is then correlated with the observed geoid as a function of ℓ to see if the effect of the slab can be seen above the "noise" induced by other density contrasts. The effect of the slab is seen clearly in the range ℓ = 4-9. The ratio of the observed geoid δN_o^ℓ to the slab geoid is then compared to similar ratios calculated for a series of viscous flow models. Models of uniform composition but with viscosities increasing with depth satisfy the observations, consistent with the mantle-wide convection hypothesis. A layered flow model is acceptable only if the downwellings due to subduction in the upper mantle lie above downwellings in the lower mantle, or if our estimates of the mass anomalies associated with subducted slabs are low by at least an order of magnitude. The latter possibility might occur if return flow is so sluggish that substantial amounts of old aseismic slab material accumulate near subduction zones.

Geoid Anomalies in a Dynamic Viscous Earth

The details of the generation and computation of geoid anomalies are sketched briefly here. More details are given by Morgan (1965a), McKenzie (1977), Watts and Daly (1981), Parsons and Daly (1983), and Richards and Hager (1983).

Consider the two-dimensional flow system shown in Figure 2, consisting of a plane layer of fluid of unit depth with uniform density and viscosity overlying an inviscid halfspace of fluid with greater density. A surface mass density σ_m with the amplitude given by a cosine bell of dimensionless width $d = 1/2$ is placed at the mid-depth of the layer.

The normal and shear tractions must of course be continuous at the top and bottom of the viscous layer. If the density contrast is suddenly introduced at time $t = 0$, the boundaries will deform until a steady state develops. The sinking density contrast pulls material behind it and pushes material in front of it, causing deformation of the surfaces. In particular, this flow creates a dimple at the top of the layer. The topography of the dimple itself then sets up a second flow, which tends to fill in the dimple in a way directly analogous to postglacial rebound. Steady state ensues when the rate at which the sinking blob pulls material away from the dimple is exactly balanced by the rate at which the deformed topography fills in the dimple. An analogous analysis holds for the bulge created at the bottom of the layer.

The amplitude of the dimple is determined by the magnitude and depth of the density contrast and the depth of the layer. It is independent of the value of the viscosity of the layer (although we will

see shortly that it is affected by relative contrasts in viscosity within the layer). The time needed to reach equilibrium, governed by the rate of flow, is proportional to the viscosity.

During the time needed to reach steady state, the density contrast will sink a distance (independent of the value of viscosity) on the order of the ultimate surface deformation (Richards and Hager, 1983). This distance is observed to be a few kilometers at the earth's surface, and is greatly exaggerated in this figure. Since this distance is much less than the depth of the layer, a steady state is reached on a postglacial rebound timescale, essentially instantaneously in this problem. The steady-state flow and surface deformation in this example are obtained using the propagator techniques presented by Hager and O'Connell (1981). Deformation of the top and bottom boundaries of the system result in equivalent surface density contrasts σ_t and σ_b. These are of opposite sign from σ_m, the "driving" density contrast.

The total geoid anomaly is the sum of the contributions shown at the top of Figure 2 from σ_m, σ_t, and σ_b. Notice that because σ_b is distant from the upper surface, it contributes only at long wavelengths, as the higher wavenumber components are attenuated by distance. The geoidal contribution from the upper density contrast σ_t contains substantial short wavelength contribution, while that from σ_m is intermediate in spectral content.

The counterintuitive result of this calculation is that the total geoid anomaly is negative for a positive density contrast in a uniform viscosity fluid (Morgan, 1965a). Such a model obviously does not explain the association of geoid highs with high-density slabs. The

total geoid anomaly has a different spectral content than that due to the density contrast alone. It is smaller in magnitude than that due to the density contrast alone by about a factor of three.

Now consider a slightly more complicated model, shown in Figure 3. This model differs from that just discussed in that the lower half of the viscous layer has a viscosity higher than that in the upper half by a factor of 30. In this case, less deformation occurs at the upper boundary and more deformation occurs at the lower boundary than in the previous case with a uniform viscosity. Thus σ_t is smaller than before, while σ_b is larger.

The individual contributions to the geoid anomaly are shown at the top of Figure 3. That from σ_m has not changed; that from σ_b has increased in magnitude and is now larger than that from σ_t, which has decreased relative to the uniform viscosity case. The total anomaly is now positive, but still only a fraction of that due to the density contrast itself.

These simple models illustrate several points that are central to the interpretation presented below. First, given a density contrast in a viscous Earth, the sign of the resulting geoid anomaly is not obvious a priori. It will depend on the variation of viscosity with depth. Second, the magnitude of the total anomaly will be only a fraction of that from the "driving" density contrast alone and will also depend on the viscosity structure. Third, the amplitudes of the various wavelength components of the total geoid will differ from both those of the "driving" density contrast itself and those of the surface deformations. Both distance and viscous flow act as filters in

determining the net geoid anomaly. This suggests a spectral approach.

Dynamic Response Function for Simple Earth Models

Determination of the geoidal dynamic response function of the earth can be viewed as analogous to the determination of the instrument response of a seismometer or of the seismic response of the earth itself. The basic concepts are sketched in Figure 4.

Assume that $S(\ell,m,r)$, the density contrast between subducted slabs and the mantle as a function of spherical harmonic degree ℓ, order m, and radius r, is known (or at least can be estimated) sufficiently well to be called a signal. Other density contrasts $N(\ell,m,r)$, not associated with slabs, represent noise in this context. These density contrasts will lead to flow and surface deformation, which modify the geoid due to density contrasts S and N alone. If we define $U(\ell,m,r)$ as the geoid anomaly observed at the surface from $(S(\ell,m,r) + N(\ell,m,r))$, then $U(\ell,m,r)$ can be expressed as the convolution

$$U(\ell,m,r) = G(\ell,m,r) * (S(\ell,m,r) + N(\ell,m,r)). \qquad (1)$$

I call $G(\ell,m,r)$ the dynamic response function to emphasize that gravity depends on dynamics. Parsons and Daly (1983) refer to G as the kernel. It is the Green's function relating total geoid anomaly to driving density contrasts in a viscous Earth.

The total geoid anomaly $U(\ell,m)$ is given by the integral over radius

$$U_t(\ell,m) = \int_0^a G(\ell,m,r) * (S(\ell,m,r) * N(\ell,m,r)) \, dr \qquad (2)$$

with a the radius of the earth.

In this paper I will consider only very simple rheological models consisting of spherically symmetric layers of constant Newtonian rheology. For these models, G is a function of ℓ and r only and can be calculated analytically (Richards and Hager, 1983). For a more general, temperature-dependent, rheology, G will depend on m as well. For a stress-dependent rheology, G will also depend on the global distribution of S and N.

The response function G_r for a rigid Earth for a surface density of degree ℓ at radius r, found by solving Poisson's equation alone, is (e.g. Jeffreys, 1976):

$$G_r(\ell,r) = \frac{4\pi\gamma a}{(2\ell+1)} \left(\frac{r}{a}\right)^{\ell+2}, \qquad (3)$$

with γ the gravitational constant. To find $G(\ell,r)$ for viscous Earth models, the equations of motion as well as Poisson's equation must be solved and boundary deformation must be accounted for.

Plots of dynamic response functions $G(\ell,r)$ for various radially symmetric, Newtonian viscous Earth models for spherical harmonic degrees 2, 7, and 12 are shown in Figure 5. These response functions are normalized by dividing the dynamic response function $G(\ell,r)$ by the static response function $G_r(\ell,a)$ due to a density contrast at the surface of a rigid earth, $4\pi\gamma a/(2\ell+1)$. If $G_r(\ell,r)$ were plotted on this figure it would have the value $(r/a)^{\ell+2}$.

Models in the left column represent mantles of uniform density and composition with a ratio of upper mantle viscosity to lower mantle viscosity of 1, 0.1, and 0.01 respectively. Models in the right column are for mantles with an intrinsic density contrast at 670 km depth resulting in chemical stratification into separate flow systems above and below the 670 km discontinuity. The viscosity structure for these models is identical to that in the corresponding row for the uniform composition models.

Free-slip boundary conditions are applied at the surface and at the core-mantle boundary. Results for rigid-free boundaries are qualitatively similar and are shown (using a different nomalization) in Richards and Hager (1983). Continuity of normal and tangential stress and horizontal velocity, and zero radial velocity, are applied as boundary conditions at the (deformed) 670 km discontinuity in the chemically stratified model. Details of the solution of the governing equations are given in Richards and Hager (1983).

The dynamic response functions are identically zero at the surface and core-mantle boundary for all models, as well as at 670 km depth for the chemically stratified mantle. This is because density contrasts at

a boundary of a convecting fluid are exactly compensated by deformation of that boundary. Only density contrasts in the interior of the convecting region cause an observable geoid anomaly when surface deformation is included. The situation is similar to the geoid anomaly for isostatically compensated density contrasts in the lithosphere (e.g. Ockendon and Turcotte, 1977), which depend on the first moment of the density contrast. This moment goes to zero for density contrasts at the surface.

Model a in Figure 5 has a uniform viscosity and so is similar to the two-dimensional model shown in Figure 2. Positive density contrasts at all wavelengths and depths lead to negative geoid anomalies. Degree two geoid anomalies are excited most effectively by density contrasts in the middle of the mantle, while degree 12 density contrasts are most visible if they are located in the middle of the upper mantle. The maximum amplitude of $G(\ell,r)$ increases with increasing wave number.

The lower viscosity of the upper mantle relative to the lower mantle in model b leads to a decrease in amplitude in the deformation of the upper surface and an increase in amplitude of the deformation of the core-mantle boundary, similar to the model in Figure 3. The net result is to make $G(\ell,r)$ more positive for all ℓ and r. For degree 2, which has a wavelength much longer than the depth of the upper mantle, G is essentially zero in the upper mantle and remains negative in the lower mantle. The dynamic response function for degree 7 turns positive in the upper mantle and approaches zero in the lower mantle. $G(12,r)$ changes sign midway through the upper mantle and has an average value close to zero throughout the mantle.

In model c, the two-orders-of-magnitude viscosity contrast between the upper and lower mantle is sufficiently large to allow G to be positive throughout most of the mantle for all degrees displayed. $G(\ell,r)$ peaks at the base of the upper mantle for all degrees shown and exceeds 0.5 for $\ell = 2$.

The response functions for the chemically stratified models d, e, and f all differ significantly from the uniform viscosity models. The average values of the response function are smaller in magnitude because the boundaries are closer together. (In the limit of very thin chemical layers, geoid anomalies approach zero.)

The combination of density contrasts from the driving density perturbation and the deformed boundaries creates a mass quadrapole, with a positive density contrast σ in the interiors and negative density contrasts σ_t and σ_b at the top and bottom boundaries. For a thin layer, isostasy prevails, and $(\sigma_t + \sigma_b) \simeq -\sigma$. The total gravitational signal from this quadrapole depends upon its moment. Since σ_t and σ_b are determined through isostasy by σ, the only significant variable is the "arm length" of the quadrapole, which depends on the thickness of the convecting layer. Thus, since the upper mantle occupies about a quarter of the mantle, the peak value of G for density contrasts in the upper mantle for models d, e, and f is roughly a quarter of the peak value for G for the uniform models.

It is of interest that for model f, which has a low viscosity upper mantle, the $G(\ell,r)$ for density contrasts in the lower mantle are quite similar to those in Model a. Density contrasts in the lower mantle in f see the 670 km barrier as essentially a free-slip boundary. We can

extrapolate to estimate that a chemically stratified mantle having a lower mantle with viscosity increasing with depth would have $G(\ell,r)$ in the lower mantle qualitatively like those shown for model c, but somewhat smaller in magnitude, due to the decreased thickness of the layer and its greater distance from the surface. The dynamic response functions for density contrasts in the upper mantle for chemically stratified models are similar in form to those for Model a. Extrapolation (justified by models not presented here) leads to the conclusion that if the viscosity in the upper mantle in these stratified models increased with depth, G would change sign and be similar in topology to that shown for Model c.

The models shown in Figure 5 make it clear that the gravity field observed for a given density contrast in a viscous earth depends strongly on the variation of viscosity with depth and on whether the mantle is chemically stratified. In particular, differences in the amplitudes of the response functions between models with uniform composition and with chemical stratification at 670 km are large, about a factor of four averaged through the upper mantle. Thus, if we can estimate the density contrasts associated with slabs to within a factor of four we can place useful constraints on mantle properties.

Density Contrasts in Subduction Zones

The average density contrast between the lithosphere prior to subduction and the underlying asthenosphere is probably the best constrained parameter relevant to mantle convection. It is constrained by the observed subsidence of lithosphere as it cools moving away from

oceanic ridges. It seems safe to assume that it is possible to estimate the density contrast between the subducted slab and the asthenosphere shortly after subduction. The total excess mass per unit area of subducting slab should then be equal to the product of the density contrast between mantle and seawater and the total subsidence of the lithosphere just before it subducts.

How this excess mass varies with depth is another question. Warming the slab by conduction should have little effect on the net mass anomaly within a subduction zone region; as the slab warms, energy conservation requires the surrounding mantle to cool, leading to no net change in thermal anomaly.

Although slabs are generally assumed to subduct coherently with the speed of the overriding plate, there is the possibility that slabs sink more rapidly, leading to a net extension, with an accompanying decrease in anomalous mass. They could also sink more slowly, with the opposite effect.

Shear heating and adiabatic compression will tend to heat up the area, causing a decrease in average density. A decrease of thermal expansion coefficient with increasing pressure would also tend to decrease the average density contrast, although an increase in thermal expansion coefficient with temperature along an adiabat would tend to conteract this effect to some extent.

Elevation or depression of phase boundaries within a slab would also have an effect on the average density contrast. For example, Schubert et al. (1975) estimate that the average density contrast of a slab reaching 700 km depth would be increased by 50% by possible

elevation of the olivine-spinel transition and decreased by 15% by depression of the spinel-post spinel boundary.

Thus, the average density contrast between a slab and the surrounding mantle is somewhat uncertain. Shear and adiabatic heating, extension of the slab, a decrease of thermal expansion with depth, and depression of phase boundaries all tend to decrease this average density contrast. Compression of the slab and elevation of phase boundaries tend to increase it. Fortunately, the response functions of various mantle models are so different that we can still hope to obtain useful results. In the discussion that follows, I will assume that the average density contrast between slab and mantle does not change from the value appropriate for old lithosphere at the moment of subduction.

The next task is to locate where the slabs are. Deep seismicity is used to locate the slabs, but it should be recognized that cold slabs may be necessary, but not sufficient, for the occurrence of deep earthquakes.

The distribution of earthquakes in a Benioff zone is typically not continuous with depth, leading to uncertainty in assigning a thermal structure to a deep seismic zone, as sketched in Figure 6. The top frame shows the cold downwelling associated with a Benioff zone extending into the lower mantle. Cold slab is present even where there is no seismic activity, perhaps as the result of low stress in the upper mantle or a more ductile phase in the lower mantle (e.g. O'Connell, 1977).

The second frame shows a more conservative view - cold slabs are present only where there is deep seismicity. This frame also

illustrates the hypothesis that the mantle is chemically stratified with a barrier to convection through the 670 km discontinuity. In this model, convection in the lower mantle is not spatially correlated with that in the upper mantle.

The third frame illustrates the hypothesis that the mantle is chemically stratified, but that there is a close coupling between the upper and lower mantle. In this case the coupling is thermal, with downwelling over downwelling (Anderson, 1981). Shear coupling, with downwelling over upwelling, is also conceivable (Anderson, 1981).

In order to have a well-defined reference model, I assume for now that the second frame depicts the appropriate conceptual model - slabs are present only where there is deep seismicity. The upper mantle between 96 km and 736 km was divided into 10 shells of thickness 64 km. Each shell was divided into 1^{o} to 1^{o} blocks. The magnetic tape version of the catalogue of the International Seismological Centre (1967-1977) spanning the years 1963-1975 was scanned, and events with 30 or more P-wave arrivals were placed in the appropriate cube. The location of Benioff zones was apparent, but there was scatter of 1^{o} to 2^{o} from the planar pattern expected for subducted slabs (presumably due to hypocenter mislocation), I drew a line through the hypocenter cluster for each slab in each shell and assigned each line a density contrast of $10^{7} kg/m^{2}$ throughout the thickness of the shell. This density contrast is appropriate for mature lithosphere that has subsided 3.7 km relative to the ridge crest sinking through the mantle with a dip of 60^{o}. The density contrasts were then discretized on a 1^{o} grid and expanded in spherical harmonics.

The locations of the seismically active slabs at three representative depths are shown in Figure 7. At shallow depths (96-160 km), all slabs are active. At intermediate depths (353-416 km), there is a minimum in activity, with no earthquakes along the Eastern Pacific. At greater depths (544-608 km), activity beneath South America picks up once more. Deeper than 672 km, only a few earthquakes occur. These are in a $3^{o} \times 1^{o}$ area of the Tonga region and are most likely mislocated, although they were included. The total amount of active slabs averaged over the upper 672 km is 1/3 of the amount in the upper shell (96-160 km).

Comparison of Observed and Predicted Geoids

Once the density contrasts assumed for slabs are assigned, the resulting potential can be computed. This potential will depend on the dynamic response function $G(\ell,r)$ for the Earth model chosen. The rigid earth provides a useful (though unphysical) reference model.

The first question to address is whether the geoid anomalies from the slab are large enough to be seen. Figure 8 is a log-log plot of the root mean square value of the potential (for each degree) versus degree for the observed geopotential (Gaposhkin, 1979) and the slab potential for a rigid earth. (Fully normalized spherical harmonics are used.) The general Kaula's law, ℓ^{-2} falloff is followed by the observed gravity field. The slab potential falls off slightly less rapidly, so the spectra cross at degree 11. There are relative gaps between the amplitudes of the two spectra at $\ell = 2$, 3, and 6.

As it stands, the slab model has less power than the observed field

ORIGINAL PAGE IS
OF POOR QUALITY

for long wavelengths. By filling in the aseismic gaps in slabs, the slab spectrum would exceed the observed spectrum. Given the uncertainties discussed above, it appears that slabs have sufficient mass to make a significant contribution to the geoid. This verifies that it might be possible to recognize the slab signal S and obtain an estimate for G as a function of ℓ for the dynamic earth.

In order to retrieve an estimate $G(\ell,r)$, I first tested the correlation between the slab geoid and the observed geoid as a function of degree. Figure 9, which is a plot of correlation coefficient r versus degree, shows the result. Here r is given (e.g. O'Connell, 1971) by:

$$r = \frac{\sum_{m=0}^{\ell} \left(C_\ell^m c_\ell^m + S_\ell^m s_\ell^m \right)}{\left\{ \left[\sum_{m=0}^{\ell} \left(C_\ell^{m2} + S_\ell^{m2} \right) \right] \sum_{m=0}^{\ell} \left(c_\ell^{m2} + s_\ell^{m2} \right) \right]}^{1/2} \qquad (4)$$

with C_ℓ^m, S_ℓ^m the coefficients of the observed geoid and c_ℓ^m, s_ℓ^m coefficients of the slab geoid. Also shown are contours of the confidence level as determined from a Student's t test with $(\ell-1)$ degrees of freedom. (A confidence level of 0.99 implies that there is a

drops abruptly at $\ell = 10$ and remains poor thereafter. The relatively poor correlation at degrees 2 and 3 may not be surprising given the low power in the slab potential at those wavelengths. The poor correlation at degree 6 is anomalous. It may result from the competing effects of hotspots, which have a spectral peak at degree 6 (Crough and Jurdy, 1980), in addition to the relative minimum in the slab spectrum relative to the observed spectrum at this degree.

The sudden drop in correlation between degrees 9 and 10 is remarkable given the small (10%) difference in wavelength. (The dynamic response function is expected to vary smoothly with wavelength.) It is likely the result of the peak in global topography at degree 10, which correlates well with gravity at this degree (Phillips and Lambeck, 1979). The hotspot spectrum also has a peak at degree 10 (Crough and Jurdy, 1980).

In order to demonstrate the degree of correlation spatially, I have plotted the observed and slab geoids for degrees 4-9 in Figure 10. The slab geoid is calculated using the dynamic response functions for a rigid Earth. Both geoids show highs of comparable amplitude along the Ring of Fire and corresponding lows over Australia and the Western Pacific. The match would be improved by filling in the gap of seismic activity in South America and Java and extending the slab deeper in the Aleutians. Considering the effect of anomalously thick crust in the Andes, Tibet, and southeast Alaska, as well as the effect of the Hawaiian hotspot, would also improve the match. Still, there are significant observed geoid fluctuations in this wavelength range unrelated to present-day subduction. The slab geoid shown accounts for

50% of the variance in the geoid for degrees 4-9.

It is clear from these correlations that in the range ℓ=2-9 the average value of $G(\ell,r)$ is positive in the upper mantle. For ℓ=4-9 the correlation is sufficiently good that we can estimate its magnitude as well as its sign. The rather surprising result is that if only the seismically active part of slabs is considered, the optimum choice of G is close to that predicted for a rigid Earth!

Discussion

It is unlikely that the earth is rigid in subduction zone regions, so it is necessary to explain the surprisingly high values of G in some other way. The most straightforward approach is to abandon the model sketched in Figure 6b in favor of a model having density contrasts associated with subduction that are not limited to the seismically active parts of subducted slabs. The models shown in Figures 6a and 6c are two such models. A third possibility is that cold, dead slabs are piled up at the base of the upper mantle. Let us consider each in more detail.

Figure 6a depicts mantle-wide convection, with the thermal structure associated with slabs extending into the lower mantle. Dynamic response functions for this type of model are shown in Figures 5a-c. If we assume that slabs extend halfway through the mantle, recalling that slabs are seismically active, on average, throughout only a third of the upper mantle, we may have underestimated the total slab anomaly by a factor of about six by considering only the seismically active parts of slabs. Thus we can explain the observed magnitudes of

geoid anomalies over subduction zones with this model if $\bar{G}$, the average of $G(\ell,r)$ through the upper half of the mantle, has a value of about 1/6 or greater.

A uniform viscosity model (Figure 5a) clearly does not work - G in this case is always negative. A contrast in viscosity between the upper and lower mantle by a factor of 10 (Figure 5b) is too small to satisfy the observations, since all G's are less than 1/6 throughout the mantle. An increase in viscosity by two orders of magnitude (Figure 5c) is clearly sufficient to allow $\bar{G}$ to be positive for $2<\ell<12$ and to have a value of about 1/6 for $4<\ell<9$.

Such an increase in viscosity would lead to a decrease of slab sinking velocity in the lower mantle, causing a piling up of dense material. This would allow a somewhat smaller value of $\bar{G}$ to be acceptable. A factor of thirty increase in the viscosity between the upper and lower mantle is likely sufficient to explain the magnitude and sign of the long wavelength geoid anomalies over slabs.

Note that the sudden increase in viscosity at 670 km is used for computational simplicity. These models are not unique, in that a viscosity increasing in a continuous fashion, rather than as a sudden jump, might give similar values of $\bar{G}$. Also note that G is sensitive to viscosity contrasts and not to the absolute magnitude of viscosity, so this approach cannot give an absolute viscosity for the upper or lower mantle.

This large an increase in viscosity between the upper and lower mantle is in conflict with recent estimates of mantle viscosity from postglacial rebound (Wu and Peltier, 1982; Yuen et al., 1982). This is

not particularly surprising, since the earth's viscosity varies laterally. It is to be expected that the upper mantle in regions of subduction is less viscous than beneath shields, where postglacial rebound occurs. There is in fact evidence from the rebound of Lake Bonneville in Utah that the upper mantle viscosity is less than 10^{19} Pas (Passey, 1981). This value shows that viscosity in tectonically active regions is less than that estimated for shield areas.

McAdoo (1982) has investigated geoid anomalies over subduction zones, including the effects of deformation of the upper surface. He used an analytic corner flow model for flow in a halfspace with nonlinear rheology. He found that his model could explain the sign and approximate amplitude of geoid anomalies over subduction zones if n, the exponent in his assumed power law rheology, was greater than 2.

McAdoo's results are in many ways compatible with those presented here, and the two studies are complementary. Flow in a halfspace is the limiting case of deep flow, so both studies support the concept of flow extending deeper than 670 km. In addition, the power law corner flow model has effective viscosities increasing with depth due to the falloff of stress away from the bend region (Tovish et. al, 1978). Thus, both studies suggest that the effective mantle viscosity increases with depth. The trade-off between the variation of Newtonian viscosity with depth and the variation of effective viscosity with stress for power law rheology is not yet resolved and suggests further work.

A model like that shown in Figure 6c, with density contrasts in the lower mantle spatially correlated with those in the upper mantle, is another possibility that cannot be excluded by the gravity data at this

time. None of the dynamic response functions shown in Figures 5c-5e would satisfy the observations, but it is probably possible to construct a model with a sawtooth viscosity profile, with viscosity increasing with depth through the upper mantle, dropping in value at 670 km, and increasing through the lower mantle, which would give $\bar{G} > 1/6$. Such a viscosity profile would likely exist in a chemically layered Earth due to the competing effects of temperature and pressure (and hence, mineralogy) on viscosity.

Little work has been done investigating the dynamics of a layered convecting system. What has been done (Richter and McKenzie, 1981) shows no tendency for the thermal coupling between layers shown in Figure 6c. However, the effects of the order 100 km deformations of the 670 km boundary expected if the mantle is chemically layered (Hager and Raefsky, 1981) were not included in their models. This large boundary deformation may favor thermal coupling.

It might be argued that shear coupling between the upper and lower mantle, with hot upwellings in the lower mantle beneath subducting slabs in the upper mantle, is possible. Such a system is unlikely to satisfy the geoid observations, however. Because mantle viscosity is temperature dependent, hot thermal boundary layers are destabilized relative to cold boundary layers. Thus, the density contrast associated with a hot upwelling in the lower mantle is likely to be much smaller than the density contrast associated with a cold subducting slab. It should have a relatively small effect on the geoid, consistent with the observation that geoid anomalies over ridges are small (see Figure 1).

Is it possible to satisfy the geoid anomalies with density

contrasts confined to the upper mantle alone? Certainly not, if only density contrasts in slabs are considered. Considering possible aseismic extensions of seismic slabs, for the upper mantle alone, $\bar{G}$ must be greater than 1/3. None of the models shown for a stratified mantle show the maximum magnitude of $\bar{G}$ to be greater than 0.2, with average values less than 0.08. Thus, slabs presently subducting in the upper mantle cannot explain the geoid highs over subduction zones, even if effects of aseismic gaps in the slabs are included.

Dead slabs lying on the 670 km discontinuity are another possible source of geoid signal. The geoid signal would be relatively small, since these slabs would be isostatically compensated by deflection of the 670 km discontinuity, but since the slabs have finite thickness, there would be some effect. This is analogous to the effect on the geoid due to plate aging at the surface, which gives a small but observable geoid signal (Ockendon and Turcotte, 1977).

The effect on the geoid of a flat-lying dead slab at 670 km would be very small; a substantial piling up of slabs would be necessary in order to result in a significant geoid signal. For example, accumulation of thermally unequilibrated dead slab at the base of the upper mantle to a thickness of 350 km over a horizontal distance of 6,000 km beneath presently active slabs would be necessary to explain the observed amplitude of the geoid highs over subduction zones in a dynamic Earth. However, such an accumulation of dead slab would lead to a depression of the 670 km seismic discontinuity by at least 60 km, and likely more (Hager and Raefsky, 1981), so it should be seismically observable. No significant deflection of the 670 km discontinuity has

yet been observed, casting doubt on this model.

There has been much recent interest in features of the geoid not related to present-day subduction (e.g. Chase, 1979; Crough and Jurdy, 1980; Anderson, 1982). Chase (1979) and Crough and Jurdy (1980) have presented their estimates of the residual geoid after they subtracted their estimates of the slab contribution.

Figure 11 shows the effect of subtracting my seismic slab geoid from the observed geoid for degrees 2-10. It differs from previous residual geoids in that the highest point is now south of Hawaii, rather than near New Guinea. Negative anomalies form a continuous band, which would be made even more pronounced by removing the effects of doubled crustal thickness in Tibet and the Andes.

The residual high in the Pacific is larger than that over Africa. This is the opposite of what would be expected from Anderson's (1982) continental insulation mechanism; the primary accumulation of continents in the Mesozoic was around Africa. Correlation of residual geoid highs with hotspots (Crough and Jurdy, 1980) is obvious, with the embarrassing exception of the world's most prominent geoid low, nearly centered on the hotspot Mt. Erebus in Antarctica.

The highest values of the degree two components of the observed geoid, the slab geoid, and the residual geoid are all close to the equator, lying at (0^{o}N, 165^{o}E), (9^{o}N, 135^{o}E), and (2^{o}S, 169^{o}E), respectively. This suggests that subduction does influence the position of the pole. Jurdy (1978) concluded that subduction and polar position are not simply related primarily because she weighted all subduction zones equally, rather than according to amount of slab penetration.

Masters et al. (1982) interpreted their observations of great circle phase velocity anomalies in terms of a degree two variation in phase velocity in the transition zone. Their maximum velocity anomaly is at (5^{o}N, 145^{o}E), closer to the maximum of the degree two slab anomaly than to the maximum of the degree two observed geoid anomaly. It would not be surprising that slabs affect the velocity structure of the earth, as well as its gravity field. Masters et al. (1982), using Birch's law to relate velocity and density, also find that a value of $\bar{G}$ of about 1/5 satisfies both the gravity and the seismic data.

Conclusions

The conclusions reached in this paper are based on the assumptions 1) that the earth responds by creeping flow to stresses applied over timescales of millions of years (the timescale appropriate for subduction); 2) that boundaries marking compositional discontinuities deform to steady-state configurations on these timescales; 3) that a reasonable upper bound to the mass anomalies associated with subducted slabs is given by the subsidence of the seafloor as plates age at the surface before subducting; and 4) that radially symmetric, Newtonian viscosity models are useful in interpreting the dynamic response functions of the actual earth.

The first assumption seems unquestionable, given the observations of plate motions and postglacial rebound. The second assumption is certainly fulfilled if the earth's mantle is not chemically stratified. Mantle viscosities derived from postglacial rebound (e.g. Wu and Peltier, 1982) and rotation of the earth (e.g. O'Connell, 1971; Yuen

et al., 1982) favor relaxation of boundaries on timescales of thousands of years - much shorter than the subduction timescale of millions of years.

If the mantle is stratified chemically, however, the time needed for all boundaries to reach equilibrium can increase by many orders of magnitude. Each boundary adds a characteristic relaxation time to the system with an associated eigenmode. Richards and Hager (1983) have investigated the responses for the model in Figure 5d. For one mode, both the surface and the 670 km discontinuity deform in the same direction with comparable displacements - essentially a warping of the upper mantle. The relaxation time of this mode is approximately equal to that of the surface when no chemical interface exists at 670 km. The other mode has displacements that are opposite in direction and inversely proportional to the density contrast across the 670 discontinuity. Its relaxation time is long, since it involves diverging flow rather than simple warping. For example, the relaxation time for this mode for a 670 km thick layer of viscosity 10^{22} p and density 3.3 g/cm^3 overlying a lower mantle with the same viscosity and with a density of 3.6 g/cm^3 is 100,000 yr for the wavelength appropriate for degree 4. This is still less than the timescale for trench migration. (Relaxation times for other density-stratified Earth models are given by Wu and Peltier (1982). Their M1 mode corresponds to deflection of the 670 km seismic discontinuity.)

If an internal load is placed inside a thin layer, such as the upper mantle, initially the layer rapidly warps with both the surface and the 670 km discontinuity deflecting in the same direction. This

leads to a negative gravity anomaly for a positive density contrast, directly analogous to the example in Figure 2. As the second mode begins to relax, the anomaly slowly becomes more positive, if viscosity increases with depth, or more negative, if viscosity decreases with depth. We have already seen that the positive geoid anomalies over high-density slabs require the viscosity to increase with depth. Then if equilibrium has not yet set in, geoid anomalies over slabs must now be smaller than they will eventually become. Since they are presently too large to be explained by a simple layered model, our conclusions are not altered.

The third assumption, that our slab density estimates are appropriate, cannot be proven but seems reasonable. Heating by friction, viscous dissipation, and adiabatic compression, as well as any decrease in thermal expansion coefficient with pressure, would all tend to lower this density contrast. Elevation of phase boundaries should not lead to large enough effects to invalidate the conclusions reached here.

The assumption of radially symmetric viscosity models is potentially a serious shortcoming. Subduction zones are regions of high stress and large temperature contrasts. Large lateral variations in effective viscosity are to be expected there. Thus we must be cautious in interpreting the results in terms of specific viscosity models. For example, the slab itself has a high effective viscosity, which will tend to couple it strongly to the surface. It may require an even higher viscosity for the lower mantle than that determined here to support the slab from below sufficiently to lead to a positive geoid anomaly, given

the tendency of the slab to couple more efficiently to the upper surface, which tends to produce negative geoid anomalies. Work with more complicated models is in progress.

Many features illustrated by these models are robust, however. Stresses induced by the flow will result in dynamic deformation of the boundaries, which will in turn affect the gravity field. An increase in effective viscosity with depth will tend to make the gravity field more positive, while a decrease in viscosity with depth will make it more negative. Thin flow systems will result in smaller geoid anomalies than thicker systems.

In summary:

1) Subducted slabs are highly correlated with the long-wavelength components of the geoid and can explain 50% of the variance in the geoid in degrees 4-9.

2) The observation that geoid anomalies are positive over subduction zones requires an effective mantle viscosity that increases with depth. A factor of at least 30 is required for a Newtonian, radially symmetric viscosity profile.

3) The density contrasts in the seismically active parts of subducted slabs are insufficient to account for the observed geoid anomalies over subduction zones if dynamic deformations of the boundaries in the convecting mantle are accounted for. There are large mass anomalies in the vicinity of subducted slabs not associated with the seismically active parts of slabs.

4) The "missing" mass anomalies can be accounted for in a straightforward way if subducted slabs penetrate aseismically into the

lower mantle, consistent with the hypothesis of mantle-wide convection.

5) If the mantle is chemically stratified, either $\gtrsim$ 350 km or more of dead slab must be piled up at the base of the upper mantle (requiring a substantial and as yet unobserved deflection of the 670 km discontinuity), or the convection system in the lower mantle must be thermally coupled to that in the upper mantle. In that case, deflection of the 670 km discontinuity by the impinging slab might trigger a matching downwelling in the lower mantle.

This work suggests that a careful seismic study to place constraints on the warping of the 670 km discontinuity near subduction zones would provide useful information for discriminating between the few models compatible with the gravity data. It also suggests that a study of the possibly substantial effects of deformed boundaries on the evolution of a layered convecting system is in order.

Acknowledgements

I thank Mark Richards for the use of the dynamic response functions in Figure 5 prior to publication. D. L. Anderson has provided stimulating discussion on this topic. Supported by NSF grant EAR-8206528, and NASA grant NSG-7610. Contribution number 3878, Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California 91125.

References

Anderson, D. L., A global geochemical model for the evolution of the mantle, Evolution of the Earth, Geodynamics Series 5, AGU, 6-18, 1981.

Anderson, D. L., Hotspots, polar wander, Mesozoic convection and the geoid, Nature, 297, 391-393, 1982.

Chase, C. G., Subduction, the geoid, and lower mantle convection, Nature, 282, 464-468, 1979.

Crough, S. T., and D. M. Jurdy, Subducted lithosphere, hotspots and the geoid, Earth Planet. Sci. Lett., 48, 15-22, 1980.

Davies, G. F., Regional compensation of subducted lithosphere: effects on geoid, gravity, and topography from a preliminary model, Earth Planet. Sci. Lett., 54, 431-441, 1981.

Gaposhkin, E. M., Global gravity field to degree and order 30 from Geos 3 satellite altimetry and other data, J. Geophys. Res., 85, 7221-7234, 1979.

Griggs, D. T., The sinking lithosphere and the focal mechanism of deep earthquakes, in E. C. Robertson (ed.), The Nature of the Solid Earth, McGraw-Hill, New York, 361-384, 1972.

Hager, B. H., and R. J. O'Connell, A simple global model of plate dynamics and mantle convection, J. Geophys. Res., 86, 4843-4867, 1981.

Hager, B. H., and A. Raefsky, Deformation of seismic discontinuities and the scale of mantle convection, EOS, Trans. AGU, 62, 1074, 1981.

International Seismological Centre, Bulletin of the International Seismological Centre: Catalogue of Events and Associated Observations, 1-11, (Years 1963-1975), 1967-1977.

Jeffreys, H., The Earth, Cambridge University Press, Cambridge, 1976.

Jurdy, D. M., Ridges, trenches, and polar wander excitation, J. Geophys. Res., 83, 4989-4994, 1978.

Kaula, W. M., Global gravity and tectonics, in E. C. Robertson (ed.), The Nature of the Solid Earth, McGraw-Hill, New York, 386-405, 1972.

Masters, G., T. H. Jordan, P. G., Silver and F. Gilbert, Aspherical earth structure from fundamental spheroidal-mode date, Nature, 289, 609-613, 1982.

McAdoo, D. C., Geoid anomalies in the vicinity of subduction zones, J. Geophys. Res., 86, 6073-6090, 1981.

McAdoo, D. C., On the compensation of geoid anomalies due to subducting slabs, J. Geophys. Res., 87, 8684-8692, 1982.

McKenzie, D. P., Speculations on the consequences and causes of plate motions, Geophys. J. Roy. Astron. Soc., 18, 1-32, 1969.

McKenzie, D., Surface deformation, gravity anomalies and convection, Geophys. J. Roy Astron. Soc., 48, 211-238, 1977.

McKenzie, D. P., J. M. Roberts, and N. O. Weiss, Convection in the earth's mantle: towards a numerical simulation, J. Fluid Mech., 62, 465-538, 1974.

Morgan, W. J., Gravity anomalies and convection currents 1. A sphere and cylinder sinking beneath the surface of a viscous fluid, J. Geophys. Res., 70, 6175-6187, 1965a.

Morgan, W. J., Gravity anomalies and convection currents 2. The Puerto Rico trench and the mid-Atlantic rise, J. Geophys. Res., 70, 6189-6204, 1965b.

Nakiboglu, S. M., Hydrostatic theory of the earth and its mechanical implications, Phys. Earth Planet. Int., 28, 302-311, 1982.

Ockendon, J. R. and D. L. Turcotte, On the gravitational potential and field anomalies due to thin mass layers, Geophys. J. Roy. Astron. Soc., 48, 479-492, 1977.

O'Connell, R. J., Pleistocene glaciation and the viscosity of the lower mantle, Geophys. J. Roy. Astron. Soc., 23, 299-327, 1971.

O'Connell, R. J., On the scale of mantle convection, Tectonophysics, 38, 119-136, 1977.

Parsons, B., and S. Daly, The relationship between surface topography, gravity anomalies and the temperature structure of convection, J. Geophys. Res., 88, 1129-1144, 1983.

Passey, Q. R., Upper mantle viscosity derived from the difference in rebound of the Provo and Bonneville shorelines: Lake Bonneville Basin, Utah, J. Geophys. Res., 86, 11701-11708, 1981.

Phillips, R. J. and K. Lambeck, Gravity fields of the terrestrial planets: long wavelength anomalies and tectonics, Rev. Geophys. Space Phys., 18, 27-76, 1980.

Richards, M. A., and B. H. Hager, to be submitted, 1983.

Richter, F. M., and D. P. McKenzie, On some consequences and possible causes of layered mantle convection, J. Geophysics. Res., 86, 6133-6142, 1981.

Runcorn, S. K., Flow in the mantle inferred from the low degree

harmonics of the geopotential, Geophys. J. Roy. Astron. Soc., 14, 375-384, 1967.

Schubert, G., D. A. Yuen and D. L. Turcotte, Role of phase transitions in a dynamic mantle, Geophys. J. Roy. Astron. Soc., 42, 705-735, 1975.

Solomon, S. C., R. P. Comer and J. W. Head, The evolution of impact basins: viscous relaxation of topographic relief, J. Geophys. Res., 87, 3975-3992, 1982.

Tovish, A., G. Schubert, and B. P. Luyendyk, Mantle flow pressure and the angle of subduction: non-Newtonian corner flows, J. Geophys. Res., 83, 5892-5898, 1978.

Watts, A. B., and S. F. Daly, Long wavelength gravity and topography anomalies, Ann. Rev. Earth Planet. Sci., 9, 415-448, 1981.

Wu, P., and W. R. Peltier, Viscous gravitational relaxtion, Geophys. J. Roy. Astron. Soc., 70, 435-486, 1982.

Yuen, D. A., R. Sabadini, and E. V. Boschi, Viscosity of the lower mantle as inferred from rotational data, J. Geophys. Res., 87, 10,745-10,762, 1982.

Figure Captions

Figure 1: Observed long-wavelength geoid (Gaposhkin, 1979) referred to the hydrostatic figure of the earth ($f = 1/299.63$). The contour interval is 20 m. Cylindrical equidistant projection.

Figure 2: Illustration of the components of the geoid anomaly from a cosine-bell density contrast at the midpoint of a layer of uniform viscosity. The total anomaly (heavy solid line) is the sum of the contributions from the density contrast itself (light solid line), from dynamic deformation of the upper boundary (long dashes), and from dynamic deformation of the lower boundary (short dashes). The total geoid anomaly is negative for a positive density anomaly.

Figure 3: As in Figure 2, but now the bottom half of the layer has a viscosity a factor of thirty larger than the upper half. The sign of total geoid anomaly is now positive.

Figure 4: The total geoid anomaly $U(\ell,m,r)$ from a slab density contrast $S(\ell,m,r)$ and noise $N(\ell,m,r)$ is given by the convolution of $(S + N)$ with $G(\ell,m,r)$, the dynamic response function for a density contrast of degree ℓ, order m at radius r.

Figure 5: Dynamic response functions as functions of degree ℓ and radius r for six Earth models. Models in the left column have uniform composition, permitting mantle-wide flow. Those in the right column have a chemical discontinuity at

670 km depth, causing stratification into separate upper and lower mantle systems. Models in the top row have uniform viscosity; in the second row there is a factor of 10 contrast in viscosity η between the upper and lower mantles. This viscosity contrast is a factor of 100 in the third row. All response functions are normalized by a factor $4\pi\gamma a/(2\ell+1)$. Free-slip boundary conditions apply at the surface and at the core-mantle boundary.

Figure 6: Cartoon illustrating conceivable thermal structures (schematic isotherms) accompanying a given distribution of deep seismicity (stars).

Figure 7: The location of deep seismicity in three representative 64 km thick shells of the upper mantle.

Figure 8: Log-log comparison of the root mean square coefficient of the potential vs degree for the observed geopotential (solid line) and that for the slab model used here.

Figure 9: Correlation coefficient r between the slab potential and the geopotential vs spherical harmonic degree. Contours indicate the confidence level of the correlation, with a confidence level of 0.95 indicating a 5% chance that the correlation is random.

Figure 10: Spatial comparison of the observed geoid (above) and the rigid-earth slab geoid (below) for degrees 4-9. The contour interval is 10 m.

Figure 11: The long-wavelength residual geoid obtained by subtracting the model slab geoid from the observed geoid.

ORIGINAL PAGE IS
OF POOR QUALITY

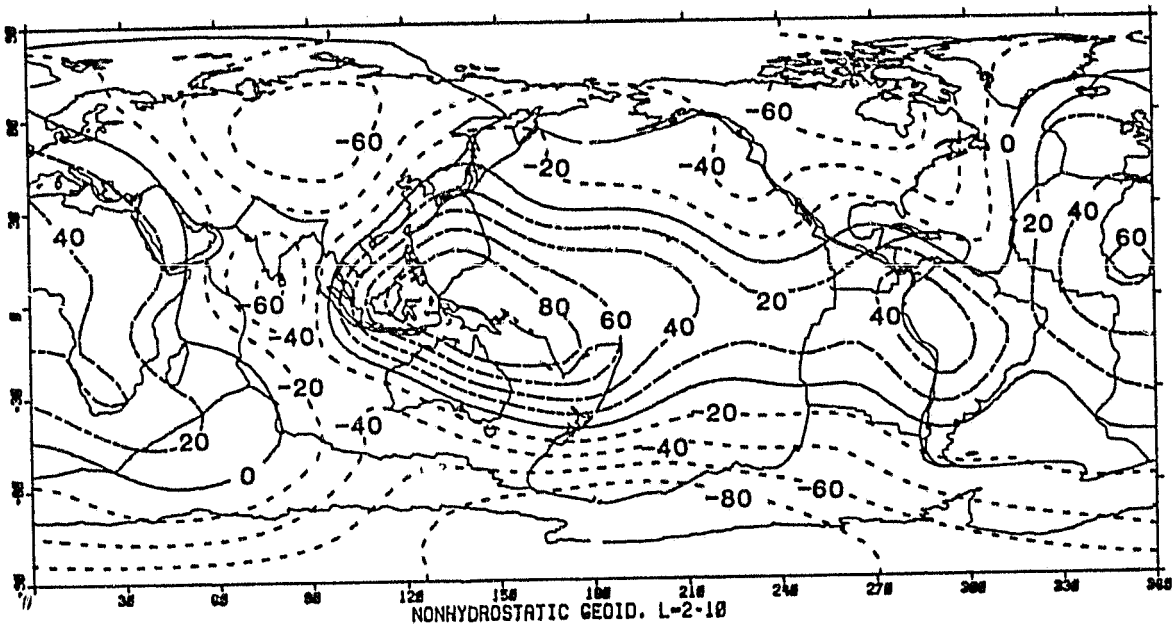


Fig. 1

ORIGINAL PAGE IS
OF POOR QUALITY

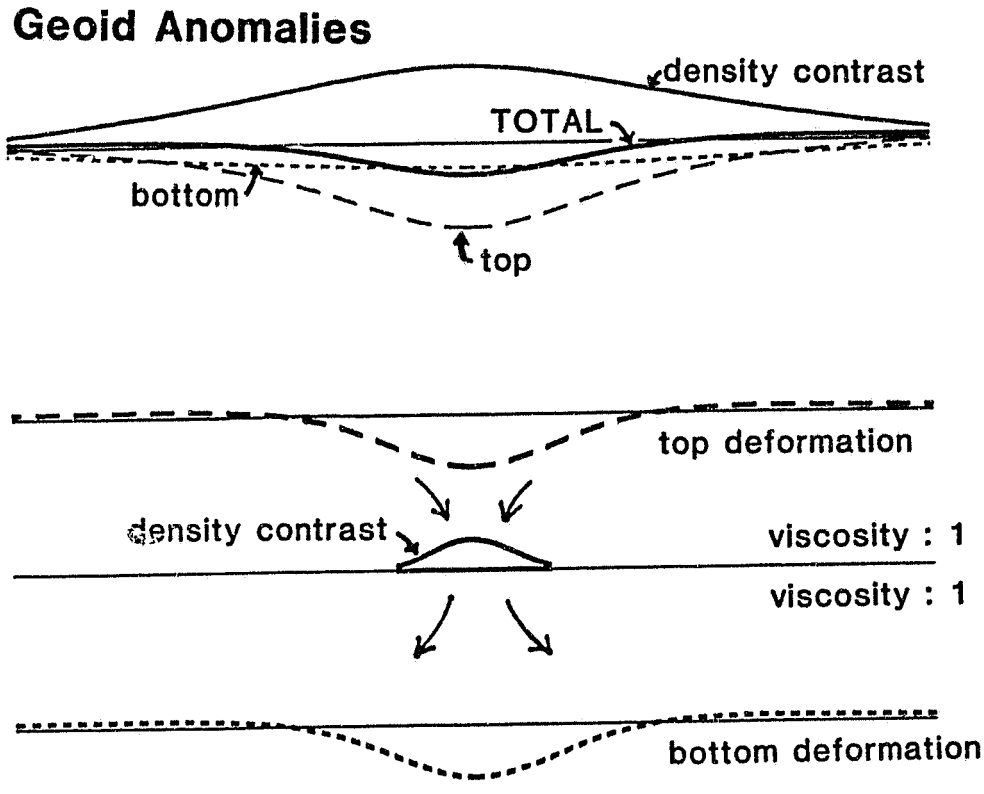


Fig. 2

ORIGINAL PAGE IS
OF POOR QUALITY

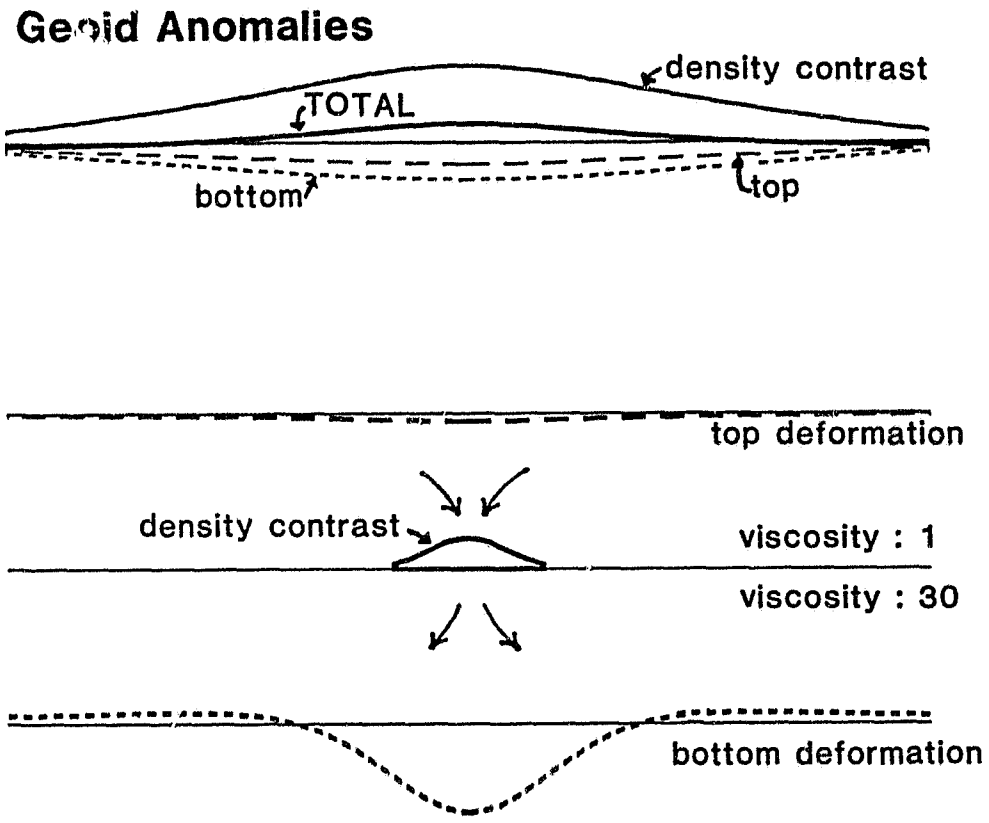


Fig. 3

ORIGINAL PAGE IS
OF POOR QUALITY

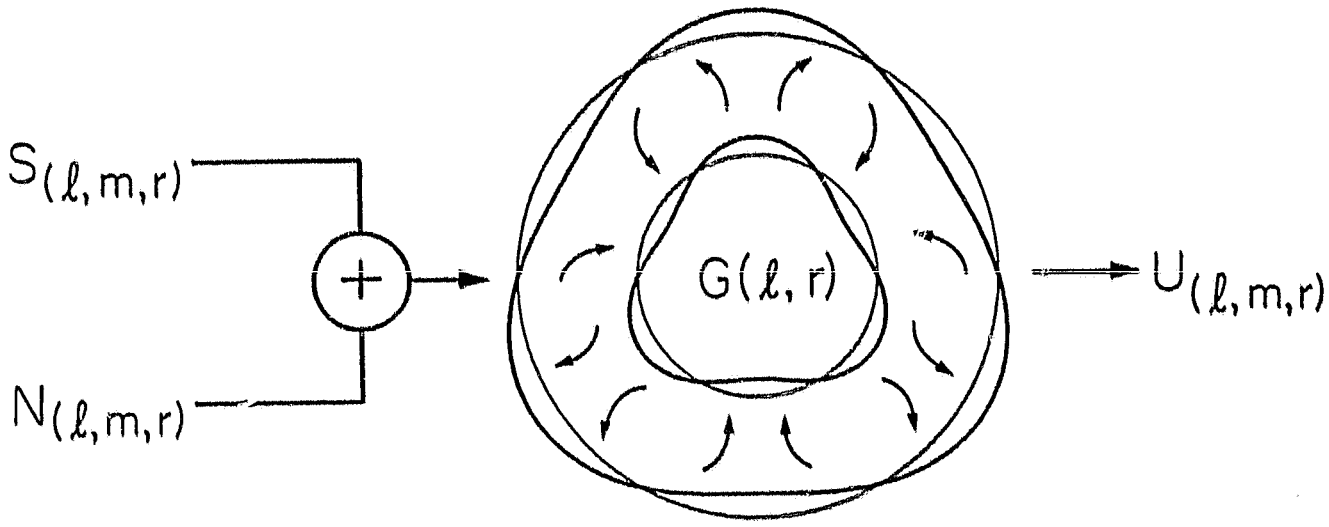


Fig. 4

ORIGINAL PAGE IS
OF POOR QUALITY

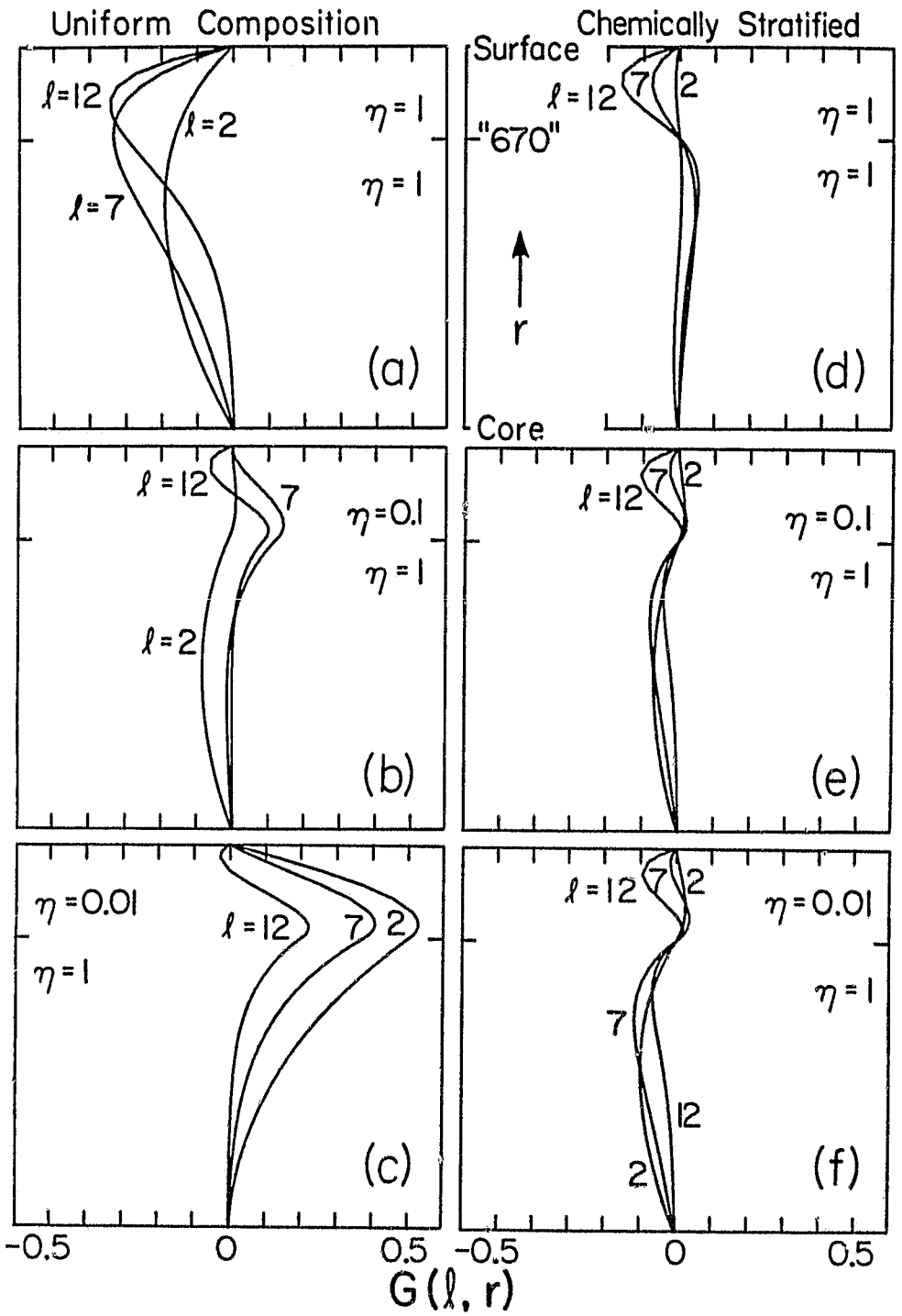


Fig. 5

ORIGINAL PAGE IS
OF POOR QUALITY

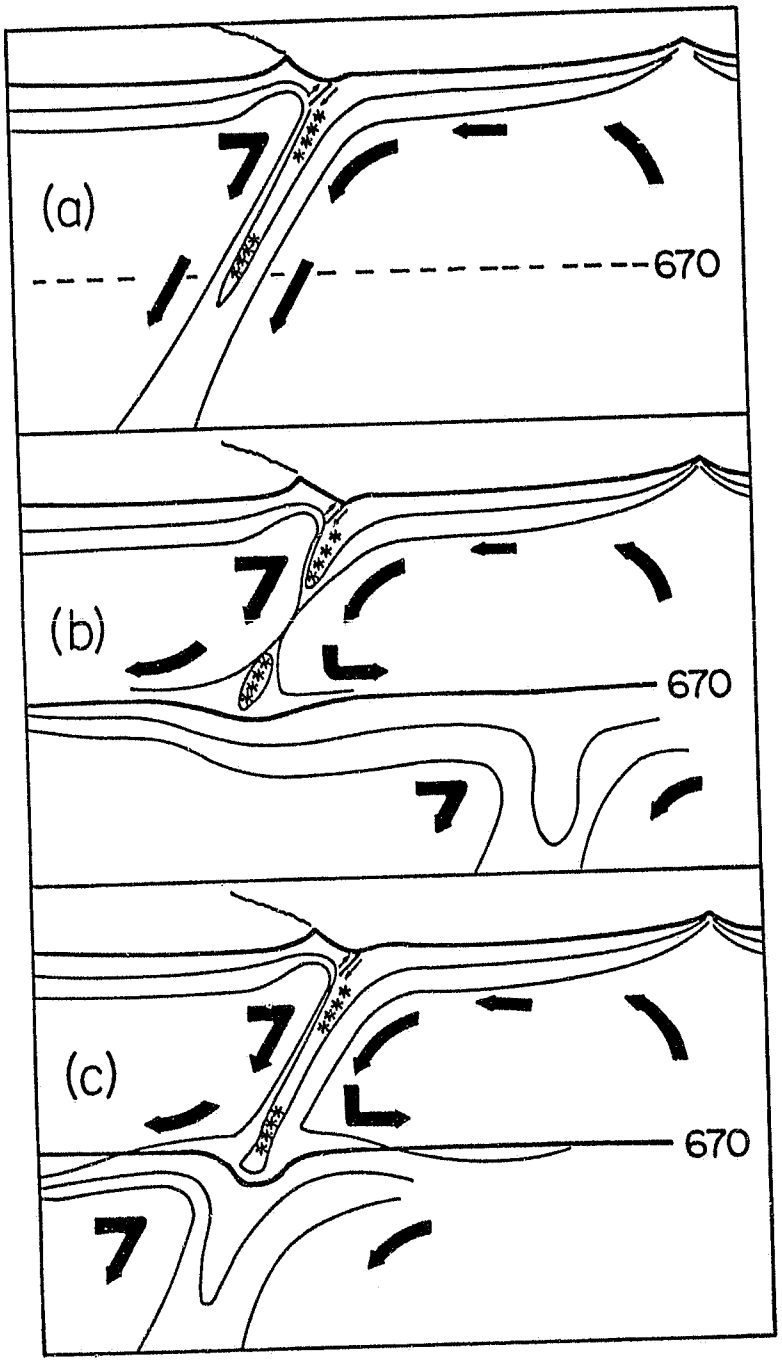
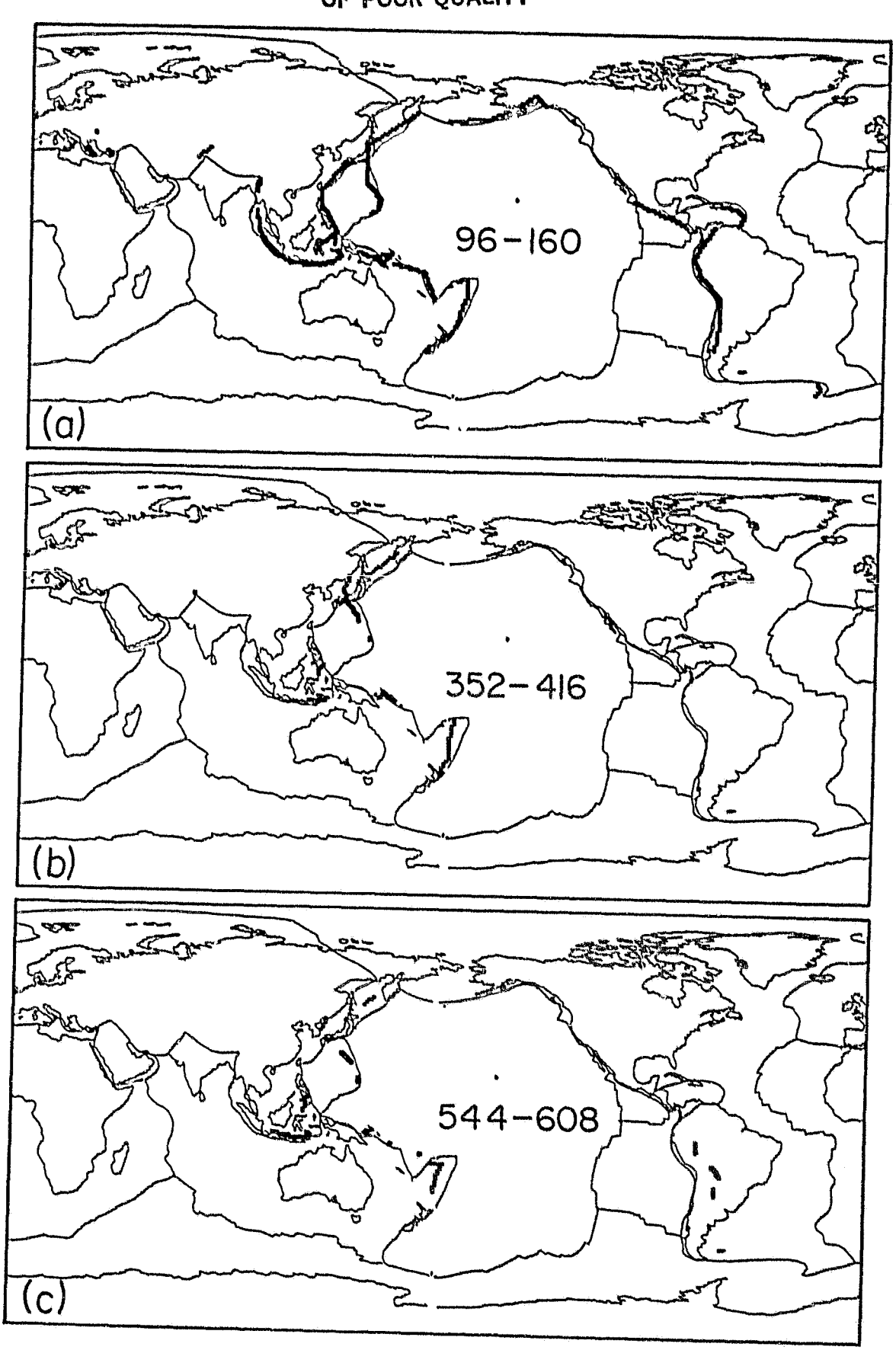


Fig. 6

ORIGINAL PAGE IS
OF POOR QUALITY



ORIGINAL PAGE IS
OF POOR QUALITY

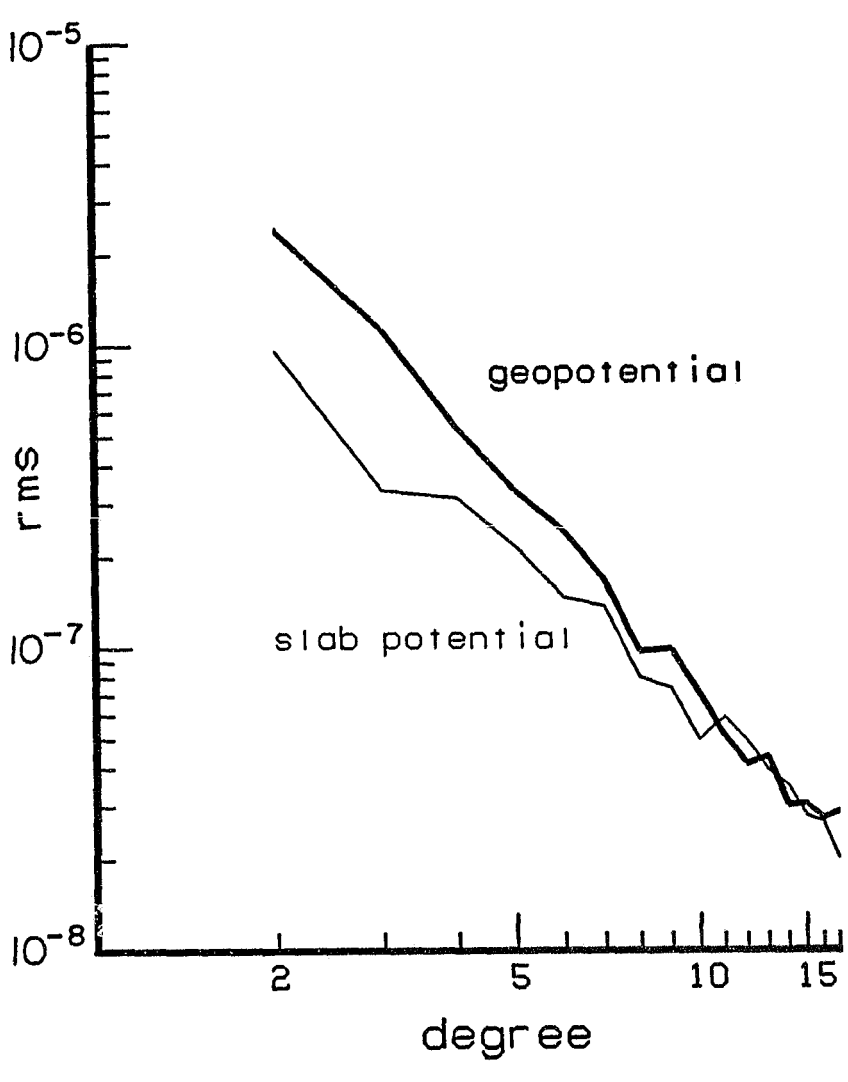


Fig. 8

ORIGINAL PAGE IS
OF POOR QUALITY

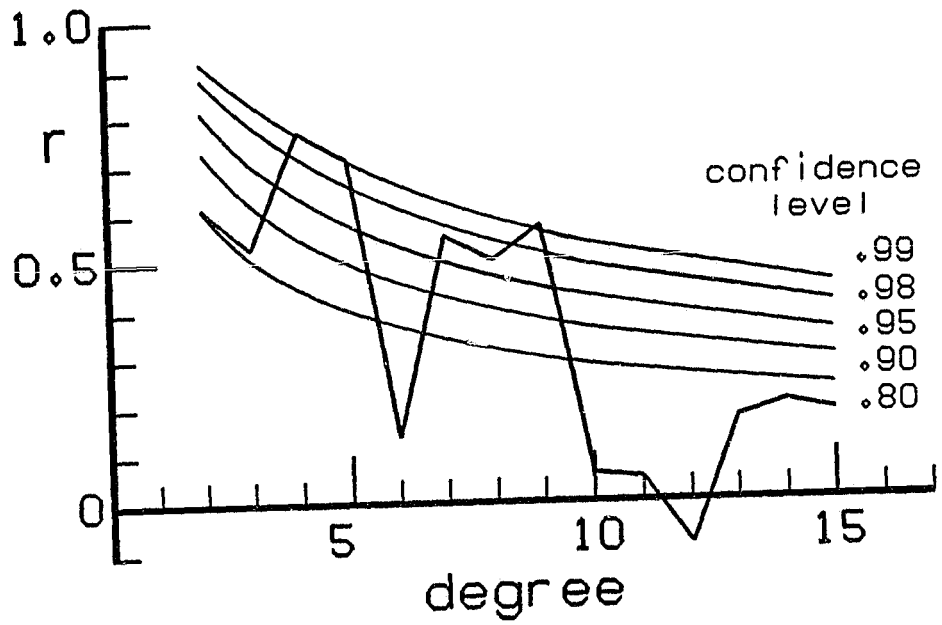
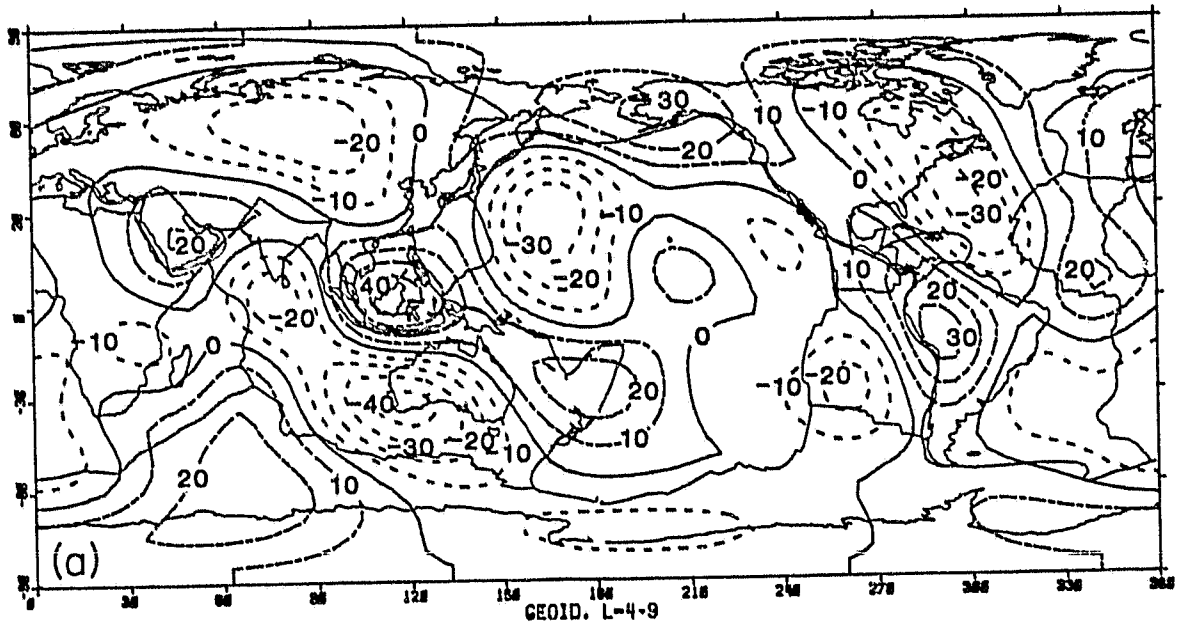


Fig. 9

ORIGINAL PAGE IS
OF POOR QUALITY



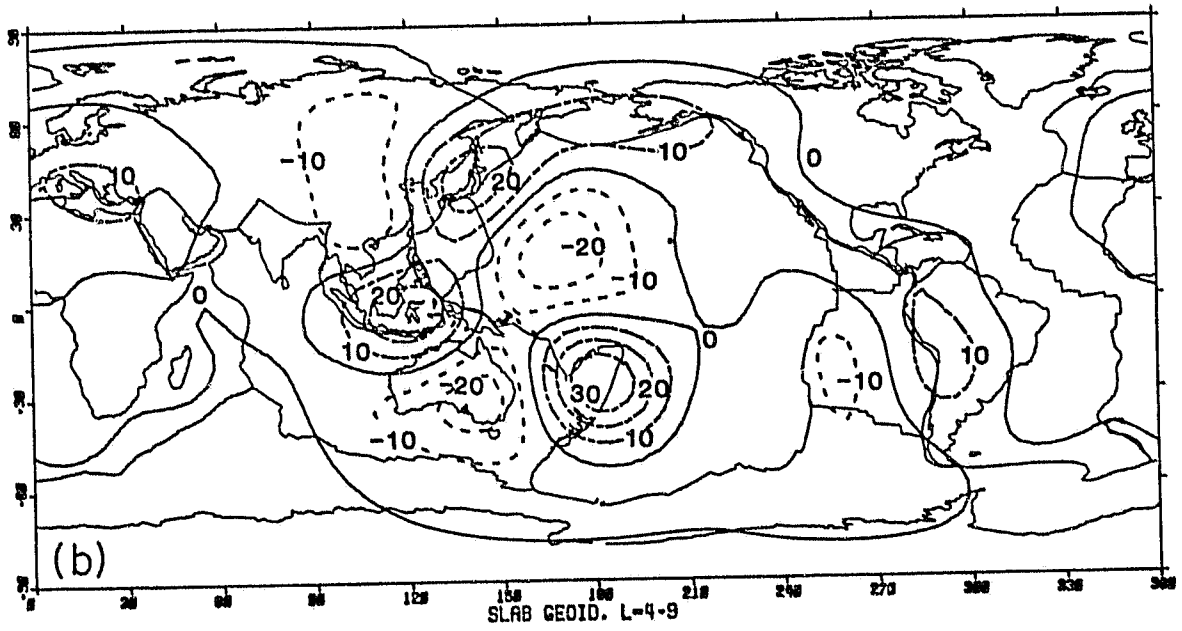


Fig. 10

ORIGINAL PAGE IS
OF POOR QUALITY

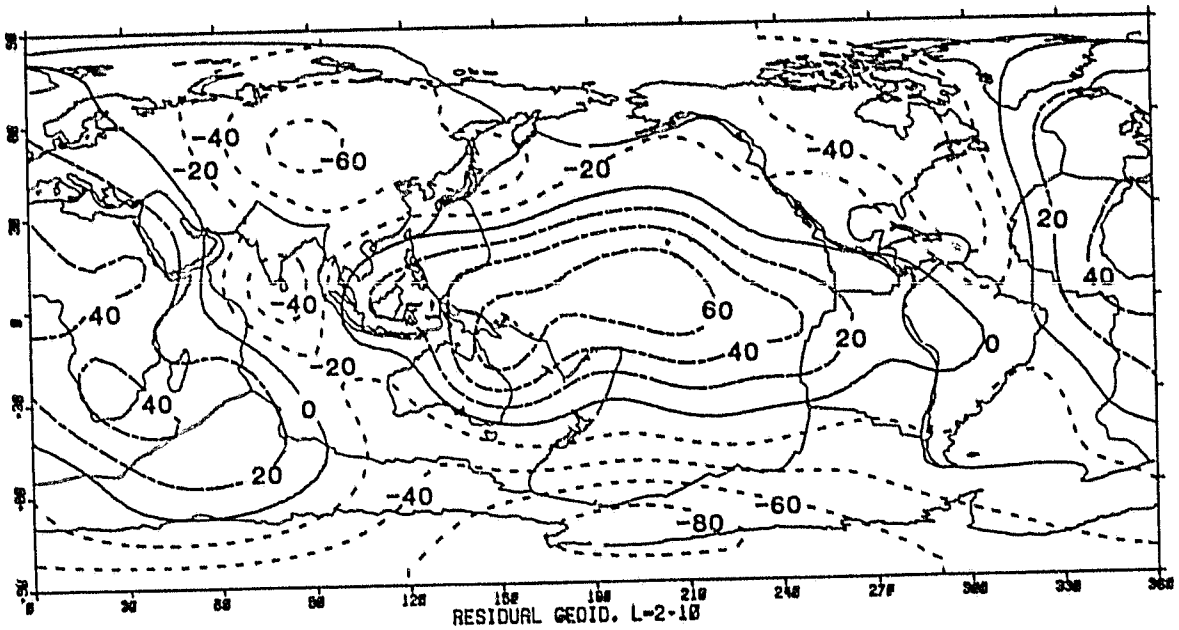


Fig. 11