Lower Mantle Heterogeneity, Dynamic Topography and the Geoid

by

Bradford H. Hager

Robert W. Clayton

Mark A. Richards

Robert P. Comer[1]

Adam M. Dziewonski[2]

Seismological Laboratory

California Institute of Technology

Pasadena, California 91125

1 now at:

Department of Earth and Space Sciences

State University of New York

Stony Brook, New York 11794

2 now at:

Department of Geological Sciences

Harvard University

Cambridge, Massachusetts 02138

June 25, 1984

ABSTRACT

Density contrasts in the lower mantle, recently imaged using seismic tomography, drive convective flow which results in kilometers of dynamically maintained topography at the core-mantle boundary and at the Earth's surface. The total gravity field due to interior density contrasts and boundary topography predicts the largest wavelength components of the geoid remarkably well. Neglecting dynamic surface deformation leads to geoid anomalies of opposite sign than are observed.

Introduction

The longest-wavelength components of the Earth's gravity field have been known from satellite geodesy for over two decades[1,2] but until recently there have been few observational constraints on the density anomalies causing them. Recent developments in seismology and theoretical geodynamics, until now relatively orthogonal branches of geophysics, now give new insight into this classical problem.

Seismological studies of lateral heterogeneity have recently determined the longest-wavelength components of seismic velocity variations in the lower mantle[3-6]. These velocity variations presumably are proportional to density variations, with both being the result of temperature differences associated with mantle convection. These inferred density variations, when taken alone, are negatively correlated with the observed geoid, seemingly in contradiction of the convection hypothesis.

Geodynamic theories of the earth's gravity field show, however, that a negative correlation is not unexpected in a convecting earth due to the counteracting effects of dynamically supported surface deformation on the gravity field[7]. For example, hot upwellings in the mantle cause uplift of the Earth's surface and of the core-mantle boundary. This dynamically maintained topography has a strong effect on the gravity field. For a given interior density structure, the total gravity field, including the effects of surface deformation, depends on the viscosity structure and on the presence or absence of chemical stratification within the mantle[8].

We show here that the longest-wavelength components of the residual geoid can be predicted from the seismically inferred density contrasts in the lower mantle using dynamic response functions for mantle convection with a moderate increase in viscosity with depth. The inferred long-wavelength temperature differences are of order 100° C. The inferred deformation of the core-mantle boundary is about 3 km. For whole-mantle convection, the calculated upper surface deformation of 2 km is in approximate agreement with observations of residual topography.

Seismic Velocity Structure of the Lower Mantle

We have used two independent methods to invert similar ISC P-wave travel time residual data sets for the velocity structure of the lower mantle. The first (due to Dziewonski,[4], hereafter referred to as method Dz) involves expressing velocity perturbations δV in terms of smooth functions – Legendre functions in radius r and spherical harmonics in colatitute θ and longitude φ :

$$\delta V(r,\theta,\varphi) = \sum_{k=0}^{K} \sum_{l=0}^{L} \sum_{m=0}^{l} f_k(r)\cdot({}_kA_l^m \cos m\varphi + {}_kB_l^m \sin m\varphi)P_l^m(\cos\theta)$$

Some 500,000 travel time residuals for teleseismic arrivals from 5,000

earthquakes were used in an iterative least-squares procedure to derive the coefficients A and B for K = 4 and L = 6, for a total of 245 coefficients.

The second method (CC, due to Clayton and Comer[5,6]) is an iterative back projection of travel time residuals along ray paths. The mantle is discretized into 29 spherical shells, each 100 km thick, with each shell divided into 1676 cells approximately 500 km on a side. At each iteration, the slowness (1/V) perturbation, δS_i , in the i^{th} cell is incremented by

$$\frac{\sum_{k=1}^{M} (l_{ki}\, \delta t_k \,/ L_k)}{\mu + \sum_{k=1}^{M} l_{ki}}$$

where, for the k^{th} ray, δt_k is the residual relative to the preceeding iteration, L_k is the total ray length, l_{ki} is the ray length in the cell (which may be zero), M is the total number of rays and μ is a damping parameter. The procedure converges to the smallest δS_i, in the sense of a Euclidean norm, where the cells intersected by more rays are given greater significance than others, that minimizes

$$\sum_{k=1}^{K} [(\delta t_k - \sum_{i=1}^{N} l_{ki} \delta S_i)^2 / L_k]$$

where N is the total number of cells and δt_k denotes a residual relative to the (spherically symmetric) starting model.

Each inversion method has advantages and disadvantages. Approach Dz gives a direct global inversion for the long wavelength features of interest for comparison with the geoid and provides a resolution matrix with formal uncertainties. It is a low order polynomial fit to somewhat sparse data, allowing model excursions in regions not well constrained by data and not allowing rapid radial variation of the model. Method CC allows determination of finer structure, and

provides a simple visualization of those regions not covered by rays, making it easy to see regions which are not constrained by observations. It does not, however, provide formal uncertainties. Since it is ultimately expanded in terms of spherical harmonics for comparison with the gravity field, it shares the problem of how best to interpolate into regions not well constrained by observations.

At this stage, the primary limitations for both approaches are the nonuniform distribution of ray coverage brought about by the uneven distribution of sources and receivers, and noise in the travel time data. It is encouraging that the two methods agree in the longest wavelength components of velocity anomalies ($l = 2–3$) although the agreement is poor at shorter wavelengths ($l \geq 4$). We confine our interpretation to those wavelengths where the methods agree, justified in part by the fact that at these wavelengths the seismic models can predict the geoid if dynamic effects are properly taken into account.

Dynamic Generation of Geoid Anomalies

Although on short timescales the mantle behaves as an elastic solid, it is commonly recognized that on geologic timescales it behaves as a viscous fluid, responding to stresses by slow creeping flow. Less commonly recognized is that the flow resulting from interior density contrasts (such as those inferred from studies of variations in seismic velocity) leads to deformation of the earth's surface, the core-mantle boundary, and any other boundary in chemical composition which may exist in the mantle. This boundary deformation occurs on the same timescale as postglacial rebound, short compared to the time it takes for the position of the interior density pattern to change appreciably.[8] Thus, boundary deformation can be taken to occur instantaneously from the standpoint of mantle convection.

The total geoid anomaly observed at the surface is the sum of the opposing effects of the interior density contrasts driving the flow and the mass anomalies caused by boundary deformations resulting from the flow. The amplitudes of the boundary deformations, and hence the sign and magnitude of the total geoid anomaly, depend on the distribution of viscosity with depth and the presence or absence of chemical stratification[8,9].

It is convenient to use spherical harmonics to describe both the gravity field and seismic velocity heterogeneities. Given the spherical harmonic coefficient of density perturbation as a function of radius, $\delta\rho_l^m(r)$, we can express the total gravitational potential U_l^m as the integral:

$$U_l^m = \frac{4\pi\gamma a}{2l+1}\int_c^a G_l^m(r)\,\delta\rho_l^m(r)dr\,. \qquad (3)$$

Here γ is the gravitational constant, a is the radius of the earth, c the radius of the core, and $G_l^m(r)$ is the dynamic response function or kernel[8,9]. This kernel includes the contributions of the boundary deformations caused by the flow induced by $\delta\rho_l^m(r)$, as well as the driving density contrast itself. Assuming a Newtonian, spherically symmetric viscosity, as we do here for mathematical tractability, the m dependence is degenerate and G is a function of l and r only.

Examples of dynamic response functions $G_l(r)$ for six simple viscous Earth models are shown in Figure 1 for spherical harmonic degrees 2, 3, and 6. U1, U10, and U30 in the left column represent mantle models of uniform composition with a ratio of lower-mantle to upper-mantle viscosity of 1, 10, and 30 respectively. (While flow velocities depend on the actual value of viscosity chosen, stresses and geoid anomalies depend only on the relative distribution of viscosity.) Models C1, C10, and C30 in the right column are for models with an intrinsic chemical density contrast at 670 km depth, with separate flow systems above and below the 670 km seismic 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.

For a given density contrast, the magnitude and sign of the resulting geoid anomaly in a dynamic Earth depends on the viscosity structure. For example, in model U1, the opposing gravitational effect of deformation of the upper boundary overwhelms that of the interior density contrast itself, leading to a negative geoid anomaly for a positive density contrast. For model U30, because of the strong increase in viscosity with depth, most of the deformation occurs at the core-mantle boundary, further away from the observer than the density anomaly driving the flow, and the effect of the density contrast dominates. Thus the response function is positive. Observation of the gravitational field of the Earth thus provides a null experiment, where the net result is a small number determined by the difference of large, nearly counterbalancing effects. The sign of the result depends on which of the effects is dominant. The anomaly also depends on the depth of the convecting system, with chemically stratified systems leading to smaller geoid anomalies for a given density anomaly. Observations of the geoid in conjunction with observations of seismic velocity heterogeneities place constraints upon the variation of mantle viscosity and the depth of mantle convection. Observation of dynamic surface deformation provides a complementary constraint.

Results

Figure 2a shows a recent long-wavelength (l = 2–6) geoid[10], referred to the hydrostatic equilibrium figure[11], superimposed on a map including plate boundaries. There is a strong association of geoid highs with subducted slabs[9,12–14] but there is much variation, particularly at the longest wavelengths, not

associated with the present plate configuration.

This can be seen more clearly in Figure 2b, which is a residual geoid determined by subtracting from the observed geoid slab effects obtained using a dynamically-consistent model[9] and the effects of crustal thickness variations and plate ages[15]. The residual geoid has a billard-ball like pattern dominated by antipodal highs over NW Africa and the Central Pacific. The highs are in regions which have been shielded from subduction for substantial parts of the recent geologic past and might be expected to represent hotter than average mantle[16,17]. Indeed, there is a strong correlation between residual geoid highs and hotspots.[13,14]

These geoids can be compared to model geoids calculated using the lower-mantle seismic models. Velocity and density perturbations are assumed to be directly proportional and the seismically inferred density anomalies are convolved with the response functions shown in Figure 1, as well as those for a static earth. If dynamic surface deformation is ignored (Figure 3a, Model CC-S, for static), the geoid anomaly is typically of opposite sign from the observed or residual geoids. For uniform composition and viscosity (Model U1, not shown) there is general agreement between the calculated and the residual geoids, but the high frequency component has too great an amplitude in the calculated geoid and the best fitting proportionality constants for degrees 2 and 3 are different. For model U10, uniform composition and a lower mantle viscosity increased by a factor of 10, there is an excellent correlation between the calculated and residual geoids at longest wavelengths for either model CC (Figure 3b) or model Dz (Figure 3c). This agreement breaks down for model U30, which has a factor of 30 increase in lower mantle viscosity.

For the chemically layered model C1, the agreement is, in general, poor. Models C10 and C30 predict geoids which are similar to each other (as can be

seen by comparing the response functions in Figure 1e and f). Either provides a reasonable match to the residual geoid, although not as good as model U10.

Correlation coefficients for degrees 2-6 are given in the Table for the observed geoid, the residual geoid, and models Dz-U10 and CC-U10. For both models, the best fitting proportionality constant is about (4 km/sec)/ (g/cm^3), within the range of values measured in the laboratory.

The dynamically maintained topography at the Earth's surface and at the core-mantle boundary for degrees 2-3, calculated for model CC-U10, is shown in Figure 4. Both the core-mantle boundary and the surface are upwarped in regions of low seismic velocity and downwarped in regions of high velocity (and presumably high density).

The total geoid is positively correlated with surface deformations and negatively correlated with seismically inferred density anomalies. This confirms the need to consider the effects of dynamic surface deformation in calculating geoid anomalies.

The map of dynamic surface topography compares favorably to long-wavelength features of the Earth's topography when features within a given tectonic province are compared. For example, depth anomalies of the order of -750 m (shallow) are observed in the S. W. Pacific, while the region S. of Australia is anomalously deep[16]. Africa stands higher than the Canadian and Siberian shields by several hundred meters[6]. Negligible dynamic topography is generated at the top surface for a chemically stratified mantle by density contrasts in the lower mantle. On the basis of predicting surface topography, the mantle wide flow model seems preferable.

Temporal variations in such large amplitude, dynamically maintained topography would be important in geologic processes such as epeirogeny and eustatic sea level changes. For example, changes in the positions of continents

and ocean basins relative to the underlying convection pattern itself could easily explain the range of eustatic sea level changes observed. Records of "eustatic" sea level from different continents would depend on their positions relative to the long-wavelength mantle flow pattern.

Dynamically maintained topography at the core-mantle boundary has a total excursion of 3 km. Such large topography has been suggested as being important in the generation of the geodynamo[19] and may be important in coupling nutation of the mantle to the core. The dynamic topography for degrees 2-3 is comparable in amplitude to the 9 km topography due to the hydrostatic ellipticity of the core-mantle boundary. Higher degree terms are expected to increase the dynamic topography substantially.

Discussion

We find the agreement between the residual geoid and the geoid calculated using the dynamic response functions for model U10 using either seismic model to be remarkable. The agreement between either seismic model and the geoid is comparable to the agreement between the two seismic models. Using the seismic models (and the dynamic slab model[9]) we can explain over 80% of the variance in the observed l = 2-3 geoid. There is less than 0.1% probability that this agreement is due to chance.

The agreement is, in fact, embarrassingly good. There are certainly large heterogeneities in the upper mantle which might be expected *a priori* to mask the signal from the lower mantle. For example, a number of seismic studies have reported a degree 2 velocity anomaly in the upper mantle which correlates well with the degree 2 geoid[20-22]. The longest-wavelength dynamic response functions for our successful model U10 are quite small in the upper mantle, however. The l= 2–3 geoid is not as sensitive to density anomalies there as in the

lower mantle; the effects of heterogeneities in the transition zone are almost totally compensated by surface deformation.

The seismic results show that at degrees 2-3, geoid lows are associated with fast, cold, lower mantle while highs are associated with slow, hot, lower mantle. This interpretation is supported by an exceedingly high correlation ($r>.99$) between the degree 2 distribution of hotspots and the average degree 2 lower mantle seismic velocity anomaly. This lends observational support to Chase and Sprowl's[17] hypothesis that geoid lows results from areas where subduction has cooled the mantle and Anderson's[16] hypothesis that shielding of the mantle by continents results in geoid highs and hotspots. Unlike those authors, we find it straightforward to reconcile the observations with mantle-wide flow, which provides a simple explanation for the fact that hotspots at the surface tend to overly hot lower mantle. The relatively higher viscosity we find for the lower mantle helps to explain the long timescales inferred to be associated with these phenomena.

Finally, although the agreement at long wavelengths is impressive, the agreement at $l \geq 4$ between the predictions of either seismic model and the geoid is not good. This may be the result of oversimplification in the flow models, which contain only two layers and which assume a spherically symmetric effective viscosity. On the other hand, the disagreement between the seismic methods at $l \geq 4$ leads to the hope that improved seismic data (such as will be provided by the proposed Global Seismic Array) and new techniques for its analysis will yield an understanding of the geoid at shorter wavelengths. The insights into mantle dynamics provided by even these relatively coarse seismic studies are extremely valuable. Further cooperative studies by seismologists and geodynamicists should lead to even better constraints on the dynamics of convection in the Earth's mantle.

Acknowledgements

This work was supported by NASA grants NAG5-315, NSG-7610, NAS5-27226, by the Alfred P. Sloan Foundation (Bradford H. Hager), and a Bantrell Postdoctoral Fellowship (Robert P. Comer). Contribution number 4065, Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California, 91125.

References

1. Kaula, W. M., J. Geophys. Res., *68*, 473-484, (1963).
2. Guier, W. H., Nature, *200*, 124-125, (1963).
3. Dziewonski, A. M., Hager, B. H., and O'Connell, R. J., J. Geophys. Res., *82*, 239-255, (1977).
4. Dziewonski, A. M., J. Geophys. Res., in press, (1984).
5. Comer, R. P., and Clayton, R. W., J. Geophys. Res., submitted, (1984).
6. Clayton, R. W., and Comer, R. P., J. Geophys. Res., submitted, (1984).
7. Pekeris, C. L., Mon. Not. Roy. Astron. Soc. Geophys. Suppl., 3, 343-367 (1935).
8. Richards, M. A., and Hager, B. H., J. Geophys. Res., in press, (1984).
9. Hager, B. H., J. Geophys. Res., in press, (1984).
10. Lerch, F. J., Klosko, S. M. and Patch, G. B., LAGE05 (GEM-L2), NASA Technical Memorandum, 84986, (1983).
11. Nakiboglu, S. M., Phys. Earth Planet. Int., *28*, 302-311, (1982).
12. Kaula, W. M., *The Nature of the Solid Earth*, 385-405, E. G. Robertson (ed.),McGraw-Hill, New York, (1972).
13. Chase, C. G., Nature, *282*, 464-468, (1979).
14. Crough, S. T., and Jurdy, D. M., Earth Planet. Sci. Lett., *48*, 15-22, (1980).
15. Hager, B. H., Earth Planet. Sci. Lett., *63*, 97-109, (1983).
16. Anderson, D. L., Nature, *297*, 391, (1982).
17. Chase, C. G., and Sprowl, D. R., Earth Planet. Sci. Lett., *62*, 314-320, (1983).
18. Crough, S. T., Ann. Rev. Earth Planet. Sci. *11*, 165 (1983).
19. Malin, S. R. C., and Hide, R., Phil. Trans. R. Soc. Lond. A 306, 281 (1982).
20. Masters, G., Jordan, T. H., Silver, P. G., and Gilbert, F., Nature, *298*, 609-613, (1982).

21. Woodhouse, J. and Dziewonski, A. M., J. Geophys. Res., in press (1984).

22. Nakanishi, I., and Anderson, D. L., Nature, 307, 117-121, (1984).

Figure Captions

Figure 1. Dynamic response functions for surface density contrasts of spherical harmonic degrees 2 (solid), 3(long dashes), and 6 (short dashes) plotted against radius for six Earth models. Models U, in the left column, have uniform composition which permits mantle-wide flow. Models C, in the right column, have a chemical discontinuity at 670 km depth, causing stratification into separate upper and lower mantle flow systems. Models in the top row have uniform viscosity; those in the middle row have a factor of 10 viscosity increase below 670 km. This viscosity increase is a factor of 30 for models in the bottom row.

Figure 2a. The observed geoid[10] for l = 2-6 referred to the hydrostatic figure of the earth. Geoid lows are shaded and the contour interval is 20m. In this map and the maps that follow we show plate boundaries and continents for reference.

Figure 2b. The residual geoid for l = 2-6 obtained by subtracting a dynamically consistent slab model[9]. Lows are shaded; the contour interval is 20 m.

Figure 3. Geoids calculated by applying dynamic response functions, such as those in Figure 1, to the seismic model. CC-S, Figure 3a, is a static model (boundary deformation is not included). CC-U10, Figure 3b, is for a uniform composition model with a factor of 10 increase in lower mantle viscosity, as is Dz-U10, Figure 3c.

Figure 4. Calculated topography for l = 2–3 at the surface (a) and core-mantle boundary (b) for the model CC-U10, which matches the geoid. Surface topography was calculated assuming a density contrast of 2.3 g/cm^3 (appropriate for suboceanic regions). Topography at the core-mantle boundary is calculated for a density contrast of 4.5 g/cm^3 across

the interface.

Table

degree	CC-U10 2	3	4	5	6	Dz-U10 2	3	4	5	6
observed	0.84	0.76	0.34	-0.17	-0.02	0.68	0.74	-0.24	0.24	0.17
residual	0.95	0.78	-0.33	-0.28	-0.29	0.83	0.82	-0.46	0.23	-0.10
Dz-U10	0.79	0.66	0.01	0.12	0.31					

Correlation coefficients for degrees 2-6 for two model geoids predicted using the seismically inferred lower mantle heterogeneity convolved with dynamic response functions for earth model U10.

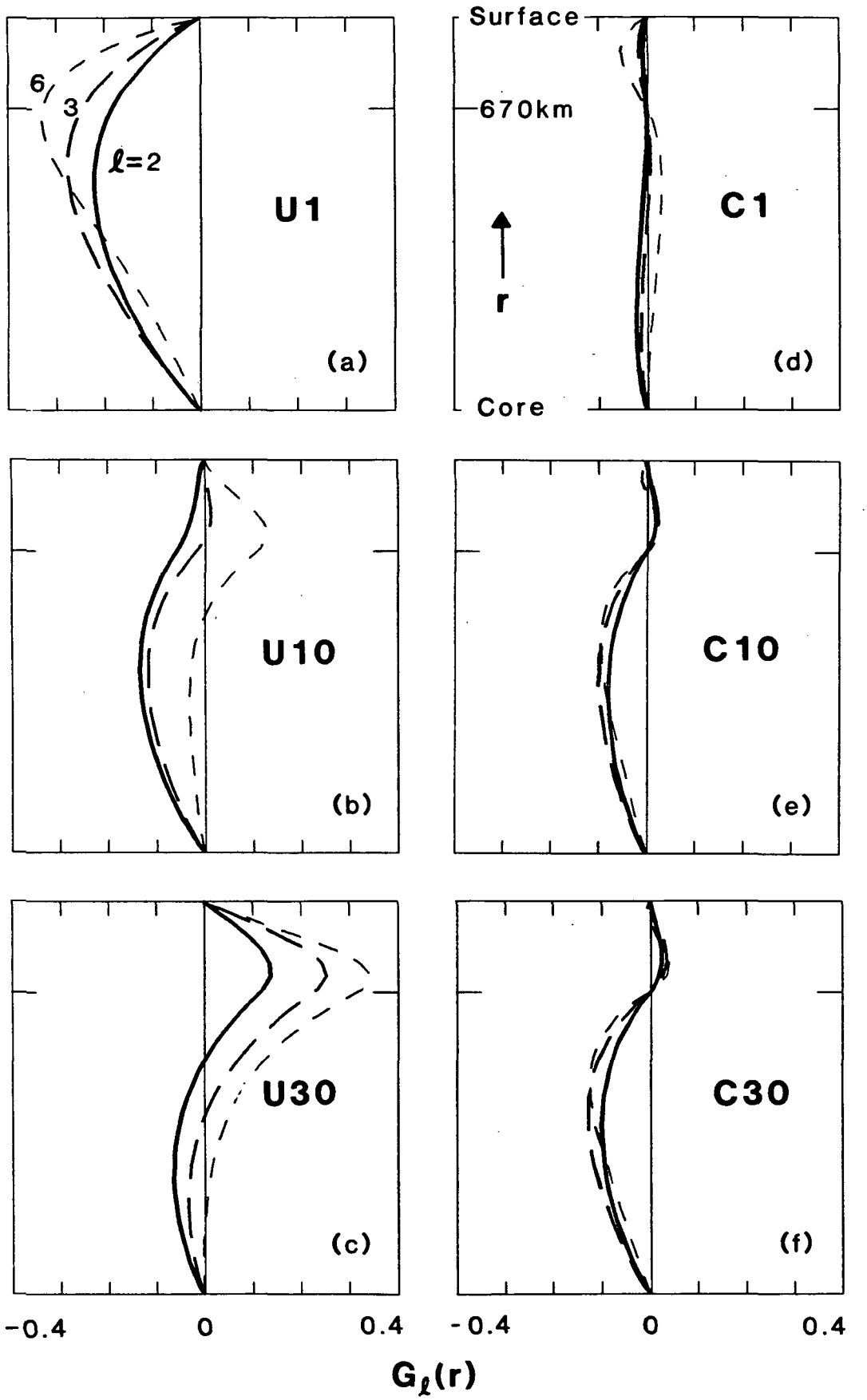
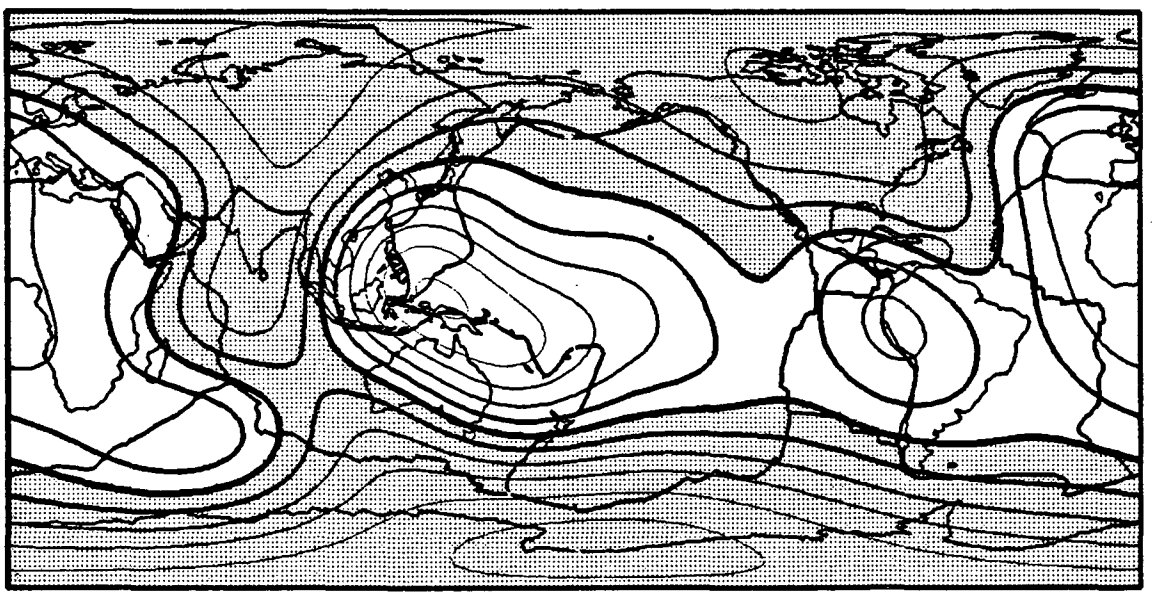

UNIFORM COMPOSITION
CHEMICALLY STRATIFIED
Surface
670km
Core
r
6
3
ℓ=2
U1
C1
U10
C10
U30
C30
(a)
(b)
(c)
(d)
(e)
(f)
-0.4
0
0.4
-0.4
0
0.4
$G_\ell(r)$

Fig. 1

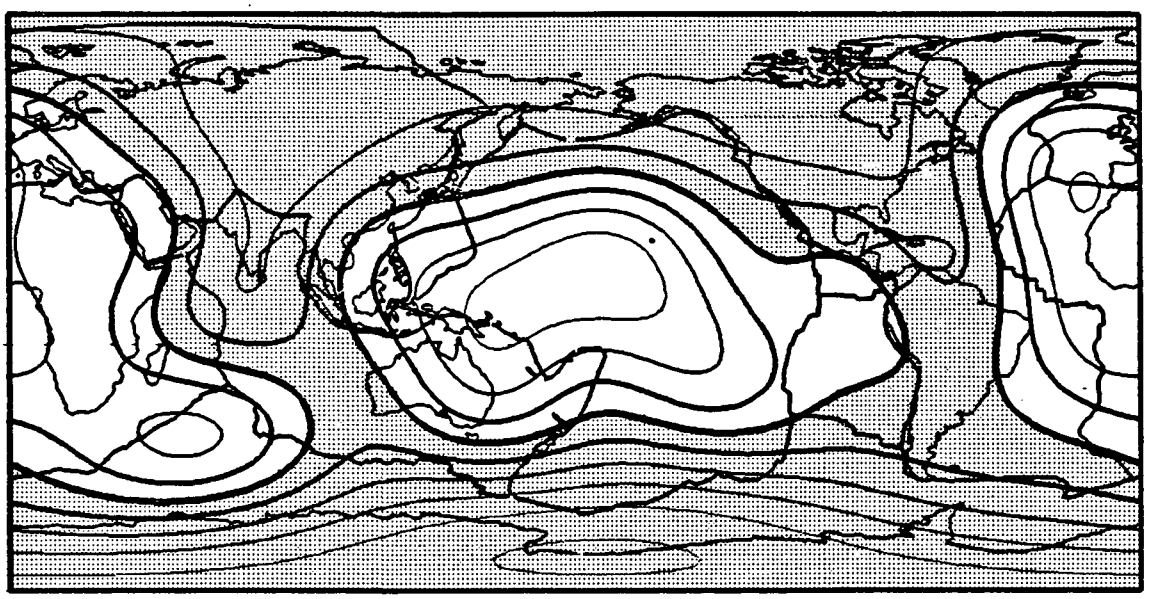
Observed Geoid: degree 2-6



contour interval: 20 m

Fig. 2a

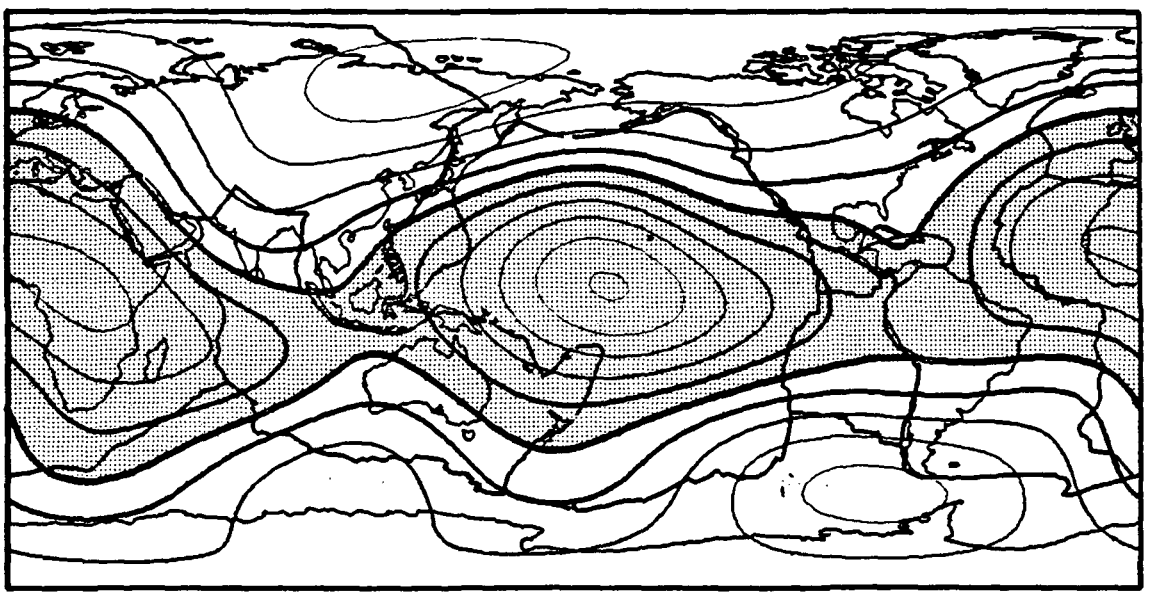
Slab Residual Geoid: degree 2-6



contour interval: 20 m

Fig 2b

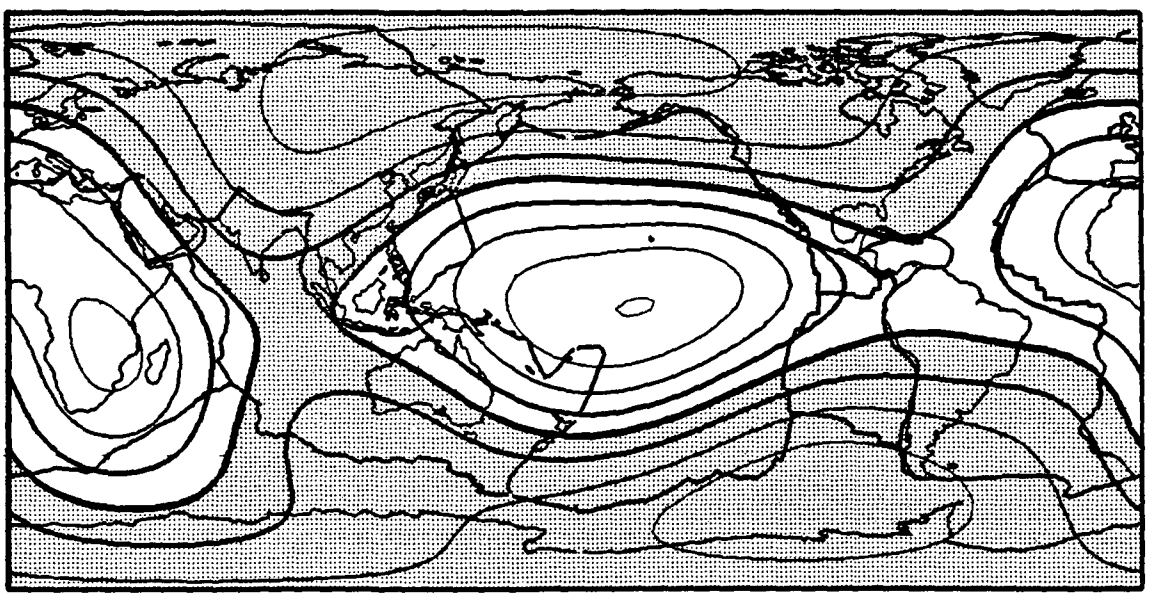
Model CC-S Geoid: degree 2-6



contour interval: 50 m

Fig. 3a

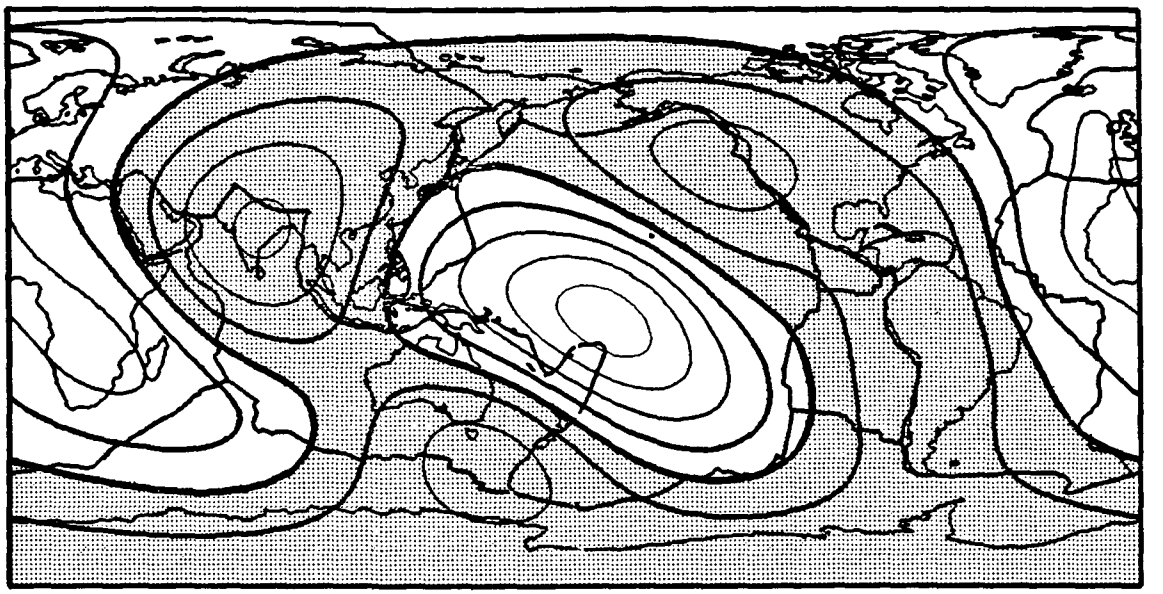
Model CC-U10 Geoid: degree 2-6



contour interval: 20 m

Fig. 3b

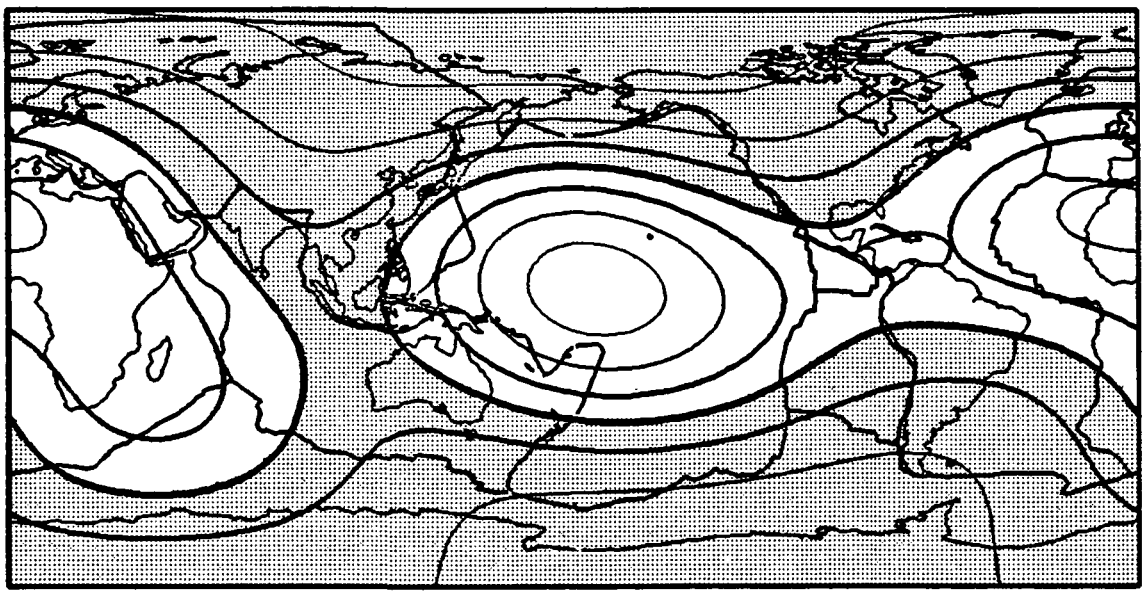
Model Dz-U10 Geoid: degree 2-6



contour interval: 20 m

Fig. 3c

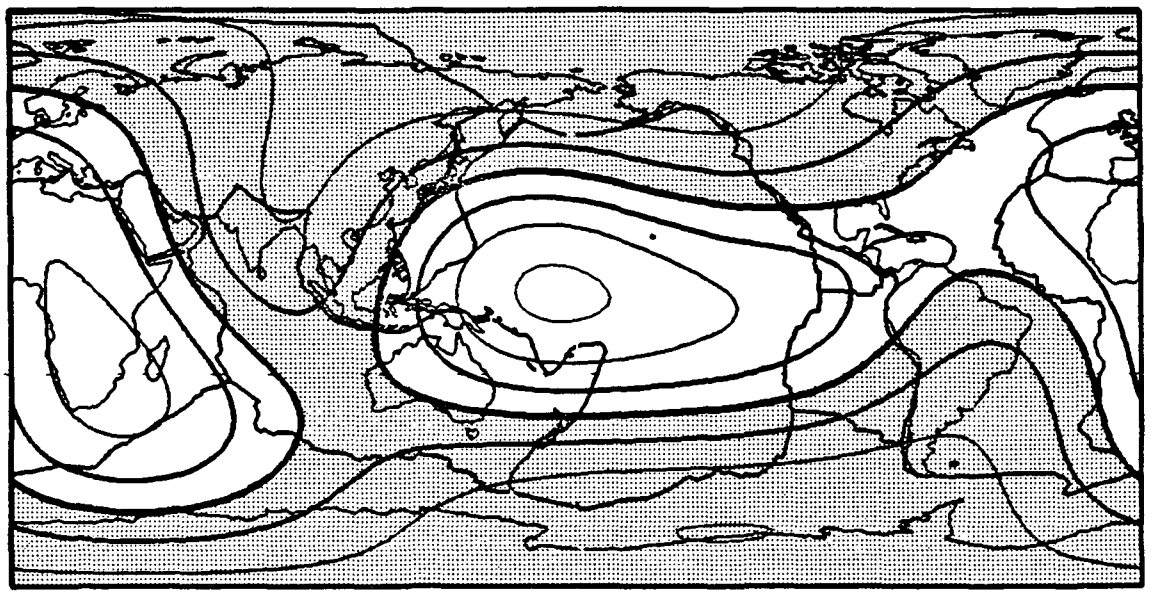
CC-U10 Surface Topography: degree 2-3



contour interval: 250 m

Fig. 4a

CC-U10 Core Topography: degree 2-3



contour interval: 500 m

Fig. 4b