PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Long-Wavelength Variations in Earth's Geoid: Physical Models and Dynamical Implications

B. H. Hager and M. A. Richards

Phil. Trans. R. Soc. Lond. A 1989 **328**, 309-327
doi: 10.1098/rsta.1989.0038

Phil. Trans. R. Soc. Lond. A **328**, 309–327 (1989)
Printed in Great Britain

Long-wavelength variations in Earth's geoid: physical models and dynamical implications

By B. H. Hager and M. A. Richards†

Seismological Laboratory, California Institute of Technology, Pasadena, California 91125, *U.S.A.*

The seismic velocity anomalies resolved by seismic tomography are associated with variations in density that lead to convective flow and to dynamically maintained topography at the Earth's surface, the core–mantle boundary (CMB), and any interior chemical boundaries that might exist. The dynamic topography resulting from a given density field is very sensitive to viscosity structure and to chemical stratification. The mass anomalies resulting from dynamic topography have a major effect on the geoid, which places strong constraints on mantle structure. Almost 90% of the observed geoid can be explained by density anomalies inferred from tomography and a model of subducted slabs, along with the resulting dynamic topography predicted for an Earth model with a low-viscosity asthenosphere (*ca.* 10^{20} Pa s) overlying a moderate viscosity (*ca.* $10^{22.5}$ Pa s) lower mantle. This viscosity stratification would lead to rapid mixing in the asthenosphere, with little mixing in the lower mantle. Chemically stratified models can also explain the geoid, but they predict hundreds of kilometres of dynamic topography at the 670 km discontinuity, a prediction currently unsupported by observation. A low-viscosity or chemically distinct D″ layer tends to decouple CMB topography from convective circulation in the overlying mantle. Dynamic topography at the surface should result in long-term changes in eustatic sea level.

1. Introduction

To understand the dynamics of mantle convection and plate motions, we must probe Earth's interior by using a variety of techniques. In the past decade, seismologists have begun to produce intriguing, albeit often fuzzy, global maps of the interior of the planet. Higher-resolution studies have revealed variations in the shape and state of stress in subducted slabs. Geochemists have analysed rocks outcropping at the surface that possess isotopic patterns indicative of isolation and incubation in isolated interior reservoirs for time periods of order 1 Ga.

Long-wavelength variations in Earth's geoid provide fundamental constraints on the interior density structure and dynamics complementary to those provided by other disciplines. Geoid undulations are primarily the result of density variations associated with mantle convection. In this paper, we present physical models of the geoid that make extensive use of recent results from seismic tomography to explain nearly 90% of the variance in the observed geoid at wavelengths greater than 4000 km. A component vital to the success of these models is the inclusion of the effects on the geoid of the dynamic topography driven by mantle convection. The amplitude of this dynamic topography, and hence the total geoid, is strongly dependent upon the mechanical and compositional stratification of the mantle. Thus the geoid produced in a dynamic Earth by a given distribution of density heterogeneities, for example as inferred

---

† Present address: Department of Geological Sciences, University of Oregon, Eugene, Oregon 97403, U.S.A.

indirectly from seismic tomography, can be used to constrain the interior structure of our planet.

A map of the observed long-wavelength geoid (Lerch *et al.* 1983) is shown in figure 1, along with continents, plate boundaries, and active hotspots. The *ca.* 20 km hydrostatic oblateness (Nakiboglu 1982) has been subtracted to reveal the *ca.* 200 m non-hydrostatic signal of tectonic origin. The non-hydrostatic geoid is dominated by polar lows and a band of equatorial highs, the latter broken by a moderate geoid low centred south of India.

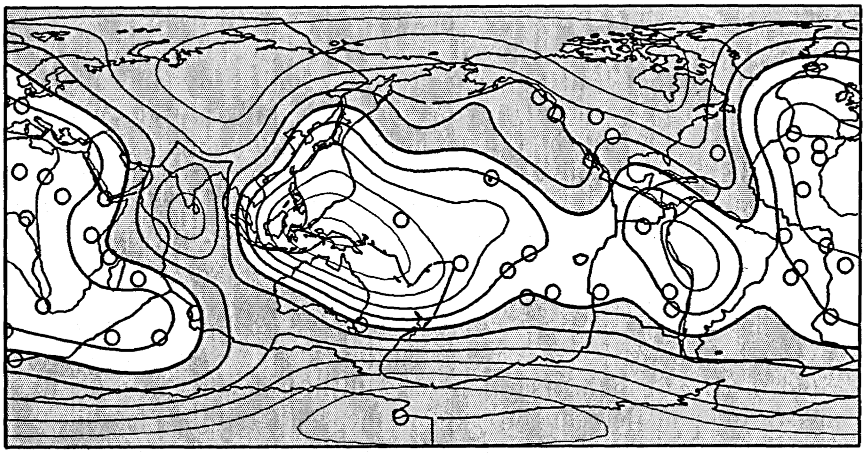

FIGURE 1. The observed long-wavelength geoid referred to the hydrostatic figure ($f = 1/299.63$). Plate boundaries and hotspots (circles) are indicated. The contour interval is 20 m and geoid lows are shaded. Cylindrical equidistant projection.

Empirical associations of the geoid with various geological and geophysical features are discussed in detail elsewhere (Richards & Hager 1988*a*). There are, in general, geoid highs associated both with convergence zones (see, for example, Kaula 1972; Chase 1979; Crough & Jurdy 1980; Hager 1984; Richards & Hager 1988*a*) and with regions having a concentration of hotspots ('hotspot provinces') (see Chase 1979; Crough & Jurdy 1980; Richards *et al.* 1988). But with the exception of convergence zones, there is little association of geoid features with plate tectonics or with continents. In particular, the spectacular association of geoid anomalies with lithospheric features so apparent at shorter wavelengths (see review by Douglas *et al.* 1987) is absent in the wavelength band shown here. Isostatic models of the geoid effects expected from density variations in the lithosphere are an order of magnitude smaller than total observed variations (Chase & McNutt 1982; Hager 1983); the primary source of the variations in figure 1 is deep.

An apparent paradox is that geoid highs are associated both with the cold, dense subducted slabs that plunge into the interior at convergence zones and with the hot, low-density mantle plumes thought to feed the hotspot volcanism at Earth's surface. This paradox has defeated successful empirical interpretation of the geoid in the context of plate tectonics and mantle convection.

Recent work in the fields of seismic tomography and fluid dynamical modelling of the mantle has made substantial progress towards resolving this apparent paradox. Studies of lower-mantle seismic velocity heterogeneity using P-wave travel-time tomography have revealed long-wavelength heterogeneities with patterns similar to the observed geoid (Dziewonski *et al.*

1977; Dziewonski 1984; Clayton & Comer 1984; Hager & Clayton 1988). In particular, regions of the lower mantle beneath the surface hotspot provinces and associated geoid highs over Africa and the central Pacific are characterized by anomalously slow mantle. If one assumes that slow mantle is hot and buoyant, and fast mantle is cold and dense, and solves Poisson's equation for the geoid predicted by the density fields inferred from lower-mantle seismic tomography, the match to the observed geoid pattern is remarkable (Dziewonski *et al.* 1977; Dziewonski 1984; Hager *et al.* 1985; Hager & Clayton 1988). Also remarkable is that the calculated geoid has the opposite sign to that observed, with geoid lows predicted over Africa and the central Pacific and geoid highs over the Poles.

This mismatch in sign is what would be expected based on the pioneering work by Pekeris (1935) over a half century ago. Pekeris pointed out that in a convecting planet, the effects of dynamically maintained topography on the geoid would be important. In fact, for a uniform-viscosity planet, the mass anomalies resulting from uplift over hot upwellings and subsidence over cold downwellings would have such a large effect on the geoid that they would reverse its sign, giving geoid highs over mantle upwellings and lows over downwellings.

Runcorn (1964, 1967) attempted to use Pekeris's theory to relate the observed geoid to plate motions and mantle convection, but had little success. The reason became clear when the theory was extended to include the effects of viscosity variation with depth, the effects of possible chemical stratification at the 670 km seismic discontinuity, and self-gravitation (Richards & Hager 1984; Ricard *et al.* 1984). Stratification of either the viscosity or composition has a major impact on the dynamic topography and resulting geoid anomalies produced for a given interior density distribution. Hence, if the density distribution can be estimated by geophysical observations or models, the geoid provides a sensitive probe of mantle structure.

Hager *et al.* (1985) showed that most of the geoid at wavelengths longer than 10000 km can be explained as the result of heterogeneities in the lower mantle inferred from seismic tomography, along with the associated dynamic topography. Using simple two-layer models to parametrize the viscosity distribution, they obtained a good agreement with the observed geoid for a model in which the lower mantle had a viscosity a factor of 10 greater than that of the upper mantle.

Hager (1984) showed that the geoid in the wavelength 4000–10000 km is dominated by the signature of subducted slabs. He also used a simple two-layer parametrization of viscosity to calculate model geoids; these models required a lower-mantle viscosity a factor greater than 30 larger than that of the upper mantle to match the observed geoid.

In this paper, we extend these models to include a more detailed and realistic parametrization of the viscosity structure of the mantle. This parametrization refines the structure of the upper-mantle to include a high-viscosity lid, a low-viscosity asthenosphere, and an increase in viscosity through the transition zone. There is increasing evidence that the D″ layer at the base of the mantle is anomalous in viscosity and/or composition, and we include these possible effects. We also include the effects of heterogeneity inferred from surface-wave studies of upper-mantle structure. We are now able to explain about 90% of the variance in the observed geoid in the wavelength band represented by spherical harmonic degrees 2–9.

In the following section, we briefly review the fluid mechanics of geoid anomalies in a convecting planet. We derive a Green function approach whereby the gravitational effects for any interior density field, including the effects of dynamic topography, can be calculated for a given model of viscosity and compositional layering of the mantle. We then present model geoids calculated for a few particular Earth models for density fields inferred from lower-mantle P-wave tomography, from upper-mantle surface-wave tomography, and from a density model for subducted slabs.

Successful models have a substantial viscosity increase (*ca.* 300) from the asthenosphere to the lower mantle. The geoid is sensitive to relative variations in viscosity, but it is not sensitive to the absolute level of the viscosity. The heat advected by the flow associated with the inferred density contrasts is, however, sensitive to the absolute viscosity, from which a lower bound to lower-mantle viscosity of *ca.* 3×10^{22} Pa s, an order of magnitude higher than that inferred by some recent studies of postglacial rebound, can be inferred (O'Connell & Hager 1984; Hager & Clayton 1988).

We discuss the geodynamical consequences of this viscosity stratification. These include a stratification in the style of mantle convection, with the upper mantle much more quickly mixed than the lower mantle. The slow circulation time for the lower mantle, compared with the faster timescale for the reorientation of plate boundaries, could explain why there is not a more obvious association of the geoid with surface tectonics, as well as the long residence of isotopic reservoirs. The viscosity stratification could also explain the shape and state of stress and seismicity distribution observed for subducted slabs (Vassiliou *et al.* 1984; Gurnis & Hager 1988). Consideration of recent constraints on the topography of the core–mantle boundary (CMB) suggests that there may be a layer of molten silicate at the top of the core.

2. Fluid mechanics of geoid anomalies

(*a*) *Physical effects*

The geoid is the equipotential surface corresponding to mean sea level. Variations in δN, the height of the geoid, relative to that expected for a hydrostatic planet, result from variations in gravitational potential δV by

$$\delta N = \delta V(a)/g(a), \tag{1}$$

where $g(a)$ is the gravitational acceleration at the Earth's mean radius a. Variations in gravitational potential are caused by non-hydrostatic variations in density $\delta\rho$. The two are related by Poisson's equation

$$\nabla^2 \delta V = -4\pi\gamma\,\delta\rho. \tag{2}$$

Here γ is the gravitational constant.

To determine the total gravitational potential, Poisson's equation must be integrated over the volume of the Earth. Convection in the mantle results in dynamic topography (Pekeris 1935). (Mid-oceanic rises, which result from cooling of the upper thermal boundary layer, and deep sea trenches are familiar examples of convectively maintained topography.) It is crucial to include the density anomalies resulting from this dynamic topography in determining the total geoid anomaly resulting from a given internal density field. At long wavelengths, the total mass displaced by dynamic topography at the top and bottom of a convecting system is comparable to the mass excess of the internal anomalies (Richards & Hager 1984). A sort of 'dynamic isostasy' approximately holds. The gravitational effects of the interior density field and the resulting dynamic topography are comparable; the total geoid is a small difference between relatively large numbers.

It is, unfortunately, difficult to observe directly the amplitude and pattern of dynamically maintained topography resulting from deep-seated density variations. Topography generated by density variations within the lithosphere, including crustal-thickness variations, dominates Earth's surface relief. There are indications that the dynamic surface topography is of order 1 km. For example, over geologic time, continental shields might be expected to be reduced by erosion to similar hypsometries. The African Shield stands hundreds of metres higher than other shields, presumably as the result of recent dynamic uplift. The ridge south of Australia is over 1000 m deeper than the average ridge crest, while the central Pacific is elevated up to 1 km (Crough 1983; Cazenave *et al.* 1988). However, although qualitative analysis indicates the existence of dynamic topography, the density structure of the lithosphere is not yet known well enough to remove its effects quantitatively. Instead we use fluid-dynamical models of mantle flow to calculate the dynamic topography.

Recently, geodetic, magnetic, and seismological observations have been used to infer the topography of the CMB. Because the CMB is hot, it is likely that any topography there is dynamically maintained. Gwinn *et al.* (1986) inferred from nutation amplitudes given by VLBI (very-long-baseline interferometry) measurements that the CMB has an excess ellipticity of *ca.* 500 m. Hide (1986) estimates that topography on the CMB is of order of a few hundred metres from models of the coupling of flow in the core to changes in length of day. Models of CMB topography from seismic tomography range from 2–12 km (Morelli & Dziewonski 1987; Gudmundsson *et al.* 1987), however. These discrepancies are discussed below.

(*b*) *Governing equations, assumptions and method of solution*

Over geologic timescales, the rocks in the Earth's mantle respond to stresses by slow, creeping flow. Inertial forces are negligible and the equilibrium equation becomes

$$\nabla \cdot \boldsymbol{\tau} + \rho \boldsymbol{g} = 0, \tag{3}$$

where $\boldsymbol{\tau}$ is the stress tensor, ρ the density, and $\boldsymbol{g}$ the gravitational acceleration.

The stress tensor $\boldsymbol{\tau}$ is related to $\boldsymbol{\varepsilon}$, the strain rate tensor, by the constitutive law:

$$\boldsymbol{\tau} = -p\boldsymbol{I} + 2\eta\boldsymbol{\varepsilon}, \tag{4}$$

where $\boldsymbol{I}$ is the identity matrix, p is the pressure, and η is an effective viscosity, in general a function of temperature, pressure, composition, strain rate, and total strain. For the models presented here, for the sake of mathematical tractability, we assume η to be spherically symmetric. This assumption is critically examined in Richards & Hager (1988*b*). They find that for half wavelengths on the order of the thickness of the mantle or greater, the effects of radial variations of viscosity are more important than the effects of lateral variations.

Finally, as discussed in Richards & Hager (1984), we make the approximation that the flow is incompressible. Zhang & Yuen (1987) have addressed this issue quantitatively. Their results show that although compressibility has some effect, this effect is small compared to effects of, for example, radial variations in viscosity. For an incompressible fluid, the continuity equation becomes

$$\nabla \cdot \boldsymbol{v} = 0 \tag{5}$$

with $\boldsymbol{v}$ the flow velocity.

Under the assumption that the viscosity is spherically symmetric, the coupled set of equations (2)–(5) can be solved analytically. A full discussion is given in Hager & Clayton (1988), where an updated (O'Connell *et al.* 1984), somewhat cleaner derivation of the equations used by

Richards & Hager (1984) is presented. The approach is to separate the radial and azimuthal variations of all variables by expanding azimuthal variations in terms of scalar or vector spherical harmonics and to utilize the orthonormality properties of the spherical harmonic basis functions to isolate the radial dependence for each spherical harmonic degree and order. The radial dependence is then solved using the propagator matrix technique (Gantmacher 1960).

The flow and resulting stresses and surface deformations are driven by interior density contrasts specified *a priori*, for example, from a density distribution inferred from seismic tomography. For mathematical simplicity, volumetric density contrasts are collapsed to form surface-mass anomalies spaced at 100 km intervals or less.

The physical boundary conditions applied are continuity of velocity and tractions across the CMB, the surface, and any interior compositional boundaries that are included in a given model (e.g. 670 km discontinuity, top of D″). Because the propagator matrix technique assumes spherical boundaries, these physical boundary conditions are analytically continued from the physical (deformed) boundaries to the mathematical (spherical) boundaries by using a first-order Taylor's series expansion. The result is that the mathematical boundary conditions become continuity of velocity and shear traction, with a jump condition in radial normal stress τ_{rr}:

$$\tau_{rr}]^{+}_{-} = \Delta\rho g\,\delta h. \tag{6}$$

Here δh is the dynamic topography of the boundary and $\Delta\rho$ is the density contrast across the boundary. At the surface and CMB, the radial velocity and shear traction vanish. At an intermediate boundary in composition, the radial velocity vanishes, while the shear tractions are non-zero but continuous.

For a given τ_{rr}, δh is inversely proportional to $\Delta\rho$. We will see later that at the surface and CMB, where $\Delta\rho$ is large, the inferred δh is small compared to the depth of the convecting mantle. For models with chemical stratification at 670, δh is large and the approximation discussed above is not very good.

(*c*) *Kernels*

In this paper, we focus on the values of the gravitational potential at the surface, along with the dynamic topography at the surface, the CMB, and any interior compositional boundaries that might exist. Values are predicted for various models of the interior density field. It is instructive to construct response functions for these quantities for a suite of models of Earth structure. These response functions, or kernels, are expressed for an interior density contrast of unit amplitude of a given spherical harmonic degree l and order m at a specified radius r, i.e. they are mixed spectral–spatial Green's functions. The calculation is linear, given the assumptions stated above, so the total response is obtained by convolving these kernels with a particular distribution of density contrasts.

Because we assume azimuthal symmetry of viscosity, the kernels depend upon spherical harmonic degree l, but not upon order m (Hager & Clayton 1988). We define normalized potential kernels $G^{l}(r)$, normalized surface displacement kernels $A^{l}(r)$, normalized CMB displacement kernels $C^{l}(r)$, and normalized chemical discontinuity topography kernels $D^{l}(r)$ such that for a spherical harmonic coefficient and density contrast $\delta\rho^{lm}(r)$

$$\delta V^{lm}_{(a)} = \frac{4\pi\gamma a}{2l+1}\int_{c}^{a} G^{l}(r)\,\delta\rho^{lm}(r)\,\mathrm{d}r, \tag{7}$$

$$\delta a^{lm} = \frac{1}{\Delta\rho_a}\int_c^a A^l(r)\,\delta\rho^{lm}(r)\,\mathrm{d}r, \tag{8}$$

$$\delta c^{lm} = \frac{1}{\Delta\rho_{\mathrm{CMB}}}\int_c^a C^l(r)\,\delta\rho^{lm}(r)\,\mathrm{d}r, \tag{9}$$

$$\delta d^{lm} = \frac{1}{\Delta\rho_a}\int_c^a D^l(r)\,\delta\rho^{lm}(r)\,\mathrm{d}r. \tag{10}$$

Here δa^{lm} is the dynamic surface topography (at $r = a$), δc^{lm} the dynamic topography of the CMB (at $r = c$), and δd^{lm} the dynamic topography of the interior chemical discontinuity (at $r = d$). With the normalization for the potential kernel, if the Earth were static, with no dynamic topography, $G^l(r) = (r/a)^{(l+2)}$.

Geoid kernels for two-layer models have been discussed in Hager (1984) and Hager *et al.* (1985). Topography kernels for two-layer models were presented in Richards & Hager (1984). An extensive discussion of the effects of adding more layers of varying viscosity is given in Hager & Clayton (1988). For reasons of space limitations, we present here kernels for only a limited set of models that are successful in matching the observed geoid.

Recognizing that there are many ways in which a model could be parametrized, we feel it important to explain our particular choice. First, the propagator matrix formulation forces us to parametrize the viscosity as a series of layers, each with constant viscosity. We have attempted to keep the number of layers small, yet at the same time we have been influenced by *a priori* expectations as to where the viscosity might be expected to change substantially.

Probably the major rheological stratification in the mantle is the contrast between the strong lithosphere and the underlying asthenosphere. While in plate interiors, the lithosphere has effectively an infinite viscosity, viewed on a global scale, the lithosphere does deform, resulting in relative motions between the plates. While the effective viscosity of the lithosphere is certainly heterogeneous, with deformation concentrated at presumably weak plate boundaries, our mathematical technique forces us to ignore this heterogeneity. We model the lithosphere as a viscous fluid of thickness 100 km; operationally, this viscosity can be thought of as the average stress divided by the average strain rate (Kaula 1980; Hager & O'Connell 1981). There is, of course, a tradeoff between the thickness of the lithosphere and its effective viscosity.

Increases in pressure with depth lead to continuous increases in viscosity, although these are partially offset by increases in temperature along the adiabat (Ashby & Verrall 1977) or reversed in thermal boundary layers. Localized phase changes and compositional changes are expected to lead to discontinuous changes in viscosity that can be large (Sammis *et al.* 1977). For this reason, we place boundaries between our viscous layers at the seismic discontinuities at 400 km and 670 km depth and at the top of D″. There will be a tradeoff between the viscosity jump at layer boundaries and continuous increases in viscosity within layers (Revenaugh & Parsons 1987). We search a model space in which the relative viscosity in a given layer can vary only in steps of $10^{0.5}$.

In our previously published models of the geoid, we have not included the effects of structure at the base of the lower mantle, primarily because it is possible to obtain excellent fits to the geoid without the added model complexity that results from including an additional layer. We include it here for several reasons. First, D″ forms the boundary region between the lower mantle and the molten core; globally averaged seismic models (Dziewonski & Anderson 1981) show a region of decreased velocity gradient in D″. D″ is almost certainly a region of increased temperature gradient, i.e. a hot thermal boundary layer, and hence is expected to have a low viscosity. More-detailed regional models (Lay & Helmberger 1983*a*, *b*; Young & Lay 1987) show a sharp boundary, so D″ may well be distinct in composition as well. Second, there is now great interest in constraining the properties of the CMB. The predicted dynamic topography at the CMB is a strong function of the structure of D″.

Kernels for several models that allow us to explain most of the variance in the long-wavelength geoid are shown in figures 2–4. For these models, the first letter refers to the style of convection: 'W' for whole mantle and 'C' for a mantle chemically stratified at 670 km depth. The second letter refers to D″: 'L' for low viscosity and 'C' for chemically distinct (also with low viscosity). Only relative viscosity variations affect the geoid. The distribution of relative viscosity η with depth is shown in the figures. For example, for our preferred model WL, the distribution is as follows: 0–100 km depth, $\eta = 1$; 100–400 km, $\eta = \frac{1}{30}$; 400–670 km, $\eta = 1$; 670–2600 km depth, $\eta = 10$; 2600 km depth – CMB, $\eta = \frac{1}{10}$.

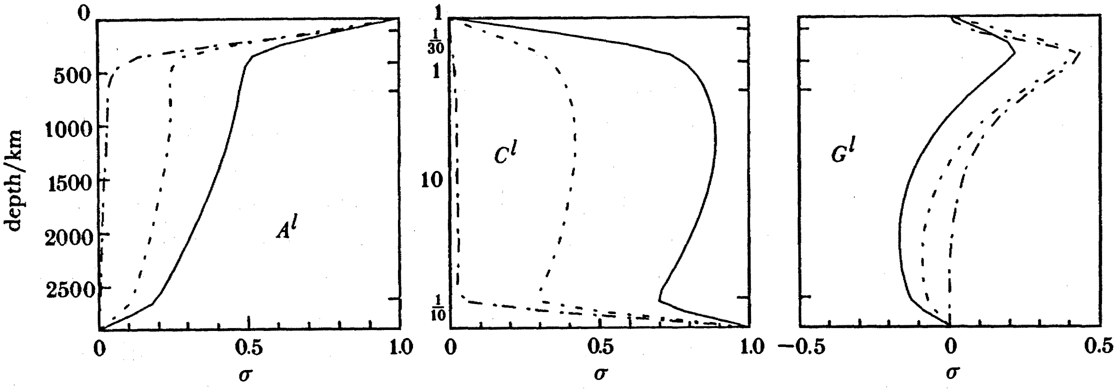

FIGURE 2. Normalized kernels for model WL, showing the effects of a surface mass σ placed at a given depth in the mantle for $l = 2$, 4, and 8. The masses displaced by surface deformation and by deformation of the CMB are normalized by dividing by $(-\sigma)$. A value of unity implies perfect dynamic compensation. The geoid kernel, which includes the effects of the dynamic topography as well as σ, is normalized by the potential due to a mass σ at the surface of the Earth. The solid line is for degree two, short dashes for degree four, and long dashes for degree eight.

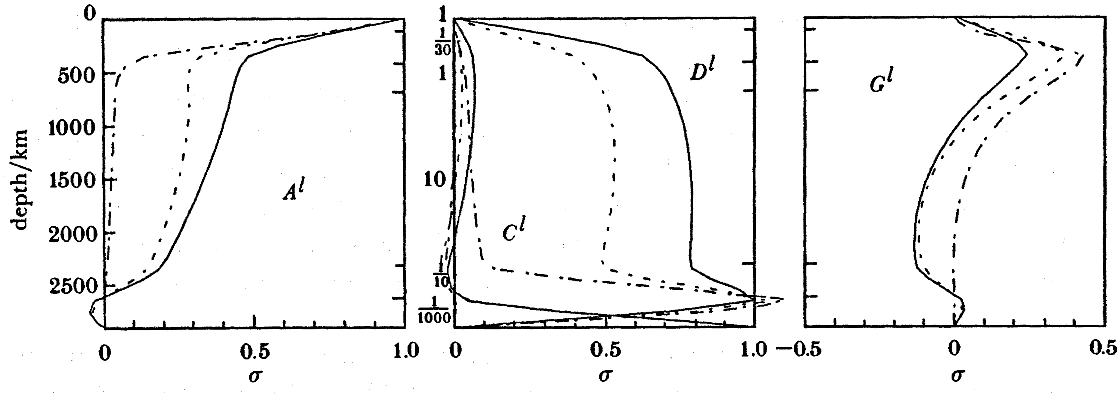

FIGURE 3. As in figure 2, but for model WC, which differs from model WL by having a chemically distinct D″ layer with a low-viscosity layer above it in the depth range 2300–2600 km to simulate the effects of an additional hot thermal boundary layer.

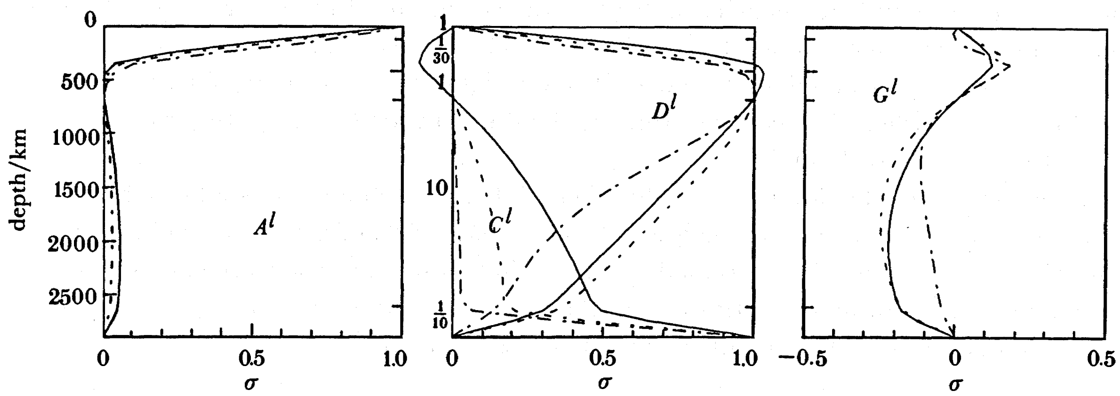

Figure 4. As in figure 2, but for model CL, which differs from model WL by having a boundary in chemical composition at the 670 km discontinuity.

3. Dynamical models of the geoid

(a) *Forward modelling with inversion for parameters*

Our previous modelling has shown that density contrasts associated with lower-mantle heterogeneity inferred from seismic tomography under the assumption that these velocity variations are thermal in origin (Hager *et al.* 1985) and upper-mantle density contrasts associated with subducted slabs (Hager 1984) can explain much of the long-wavelength undulation in the observed geoid. There are, however, substantial uncertainties associated with these models. For example, the ratio of density perturbation to velocity perturbation for the lower mantle is not known *a priori*. The actual mass anomalies associated with subducted slabs are relatively poorly constrained. There is a possible tradeoff between these parameters and the viscosity structure that best explains the observed geoid. Large-amplitude variations in upper-mantle shear-wave velocities have recently been mapped (Masters *et al.* 1982; Woodhouse & Dziewonski 1984; Nataf *et al.* 1984, 1986; Tanimoto & Anderson 1984; Tanimoto 1986; Grand 1987). These velocity anomalies could have substantial density contrasts associated with them and therefore have an effect on the geoid and dynamic surface topography. Although the various tomographic inversions tend to agree at long wavelengths, this agreement disappears at shorter wavelengths, opening the question of reliability of the inferences of mantle structure.

To investigate some of the tradeoffs and uncertainties associated with modelling the geoid, we have chosen to fit the geoid by solving the following set of linear equations (Hager & Clayton 1988):

$$Rs^{lm} + Pf_{\mathrm{p}} V_{\mathrm{p}}^{lm} + Sf_{\mathrm{s}} V_{\mathrm{s}}^{lm} = g^{lm}. \tag{11}$$

The equations are solved for R, the density contrast of mature subducted slabs (assumed distributed over a thickness of 125 km and dipping at 60°), P, the ratio of density perturbation to velocity perturbation for lower-mantle compressional-wave velocity anomalies, and S a similar ratio for upper-mantle shear-wave velocity anomalies, all assumed to be constants independent of radius. In this equation, s^{lm} is the geopotential coefficient obtained from (7) using the subducted slab model described below and V_{p}^{lm} is a similar integral involving the convolution of the potential kernel with the spherical harmonic expansion of P-wave velocity heterogeneity in the lower mantle. V_{s}^{lm} is obtained in the same way using a model of upper-mantle shear-velocity heterogeneity, and g^{lm} is the coefficient of the observed geoid. Because there is more agreement among tomographic models at long than at short wavelengths, we include the weighting functions f_p and f_s. These weights are set to unity for $l \leqslant l_{\max}$ and to zero for $l > l_{\max}$, where $l_{\max}$ can, in principle, be different for f_p and f_s. In practice, we find $l_{\max} = 4$ works best for both.

Of the viscosity models for which we show kernels here, we obtain the best fit to the observed geoid for the model shown in figure 2, which can explain 87% of the variance of the observed geoid. The predicted geoid for this model is shown in figure 5*a*; it bears a striking resemblance to the observed geoid, figure 1. In this and the following sections, we show plots only for this viscosity model. Results for the other models are very similar.

(*b*) *Individual contributions*

The model geoids shown (discussed separately in the next sections) were computed simultaneously by solving (11). The geoid contributions from the three sources are not totally orthogonal and there is some tradeoff among the parameters in (11), as discussed below in (*c*). However, this tradeoff is not so large that the individual contributions are qualitatively affected by it. The individual contributions do not change very much if, for example, they are solved for sequentially, as in our previously published results (Hager 1984; Hager *et al.* 1985).

(i) *Lower-mantle contributions to the geoid*

The contribution to the geoid calculated with the lower mantle P-wave tomographic inversion of Clayton & Comer (1984) is shown in figure 5*b*. As discussed in more detail in Hager & Clayton (1988), although the tomographic inversion is a global one, the ray coverage in the upper mantle and near the base of the mantle for this and other P-wave inversions is very poor. For this reason, we evaluate the integral in equation (7) only between the depths of 800–2800 km. We also set $l_{\max}$ at 4; above this value, the overall fit to the geoid degrades slightly. The density perturbation to velocity perturbation P has a value of 0.3 (Mg m^{-3})/(km s^{-1}).

The agreement between this predicted long-wavelength geoid and the observed geoid is striking. Both are dominated by polar lows and equatorial highs. Both encompass most of the hotspots within their highs. In fact, the correlation between the degree two component of the spherical harmonic distribution of hotspots and this predicted geoid is almost perfect ($r = 0.99$), suggesting that surface hotspots and lower-mantle thermal structure are strongly coupled (Hager *et al.* 1985; Richards & Hager 1988*a*; Richards *et al.* 1988).

The longest-wavelength features, which contribute most of the power to the geoid, appear to be explained primarily by structure in the lower mantle. The predicted geoid does not have the highs over the Andes and the western Pacific seen in the observed geoid. These can be explained by density contrasts associated with subducted slabs in the upper mantle.

(ii) *Subducted-slab contributions to the geoid*

The foci of deep earthquakes at subduction zones mark positions of the cold, central cores of subducted slabs. If we knew the density contrasts associated with these subducted slabs, we could calculate the associated geoid anomalies for a given model of Earth structure using (7). Unfortunately, whereas seismicity constrains the location of subducted slabs to some extent, the associated density contrasts are not well constrained. Although the integrated density contrast between subducted slabs and the asthenosphere in the shallow mantle is constrained by the subsidence of the lithosphere with age as it moves away from mid-oceanic ridges, it is not well constrained at depth. Complicating factors include the effects of phase changes (see, for example, Schubert *et al.* 1975; Anderson 1987), the effects of slab extension and compression (Isacks & Molnar 1971), and the likelihood that slab material is present even when there is no seismicity (see Hager (1984) and Hager & Clayton (1988) for an extended discussion). Thus, in calculating the geoid anomaly associated with subducted slabs, the distribution of mass anomalies, as well as the mantle structure, must be assumed. The validity of assumed mass distributions can then be tested by comparing the model geoid predicted to the observed.

For simplicity, we assume that the surface mass anomaly associated with subducted lithosphere is conserved with depth. We ignore slab extension and compression and assign all slabs a dip of 60°. Although seismicity cuts off at some depth for each subduction zone, probably as the result of warming of the core of the slab beyond the temperature at which brittle failure can occur for the relevant phase, the mass anomaly associated with the slab probably extends beneath the deepest earthquakes. We assume that all slabs penetrate to the same depth. For models with flow permitted into the lower mantle, we choose the lower limit in (7) to be at a depth of 800 km to avoid double counting with the lower-mantle tomographic model. For chemically stratified models, we choose 670 km as the natural lower limit for slab penetration, in accord with the small boundary deformation approximation in §2*b*. For slabs where seismicity is confined to depths shallower than 300 km, probably as the result of shallower reheating of young or slowly subducting lithosphere, we assume that the mass anomaly and depth of seismicity are related and use a mass anomaly a factor of two smaller. In the future, more detailed tomographic inversions for slab structure (Zhou & Clayton 1987) should allow us to remove many of these assumptions.

The geoid predicted for the Earth model shown in figure 2 is shown in figure 5*c*. The slab model reproduces the geoid highs associated with subduction zones. The predicted geoid is quite similar to the model obtained with a simpler two-layer viscosity model by Hager (1984), but in this case, the same viscosity model that fits the lower-mantle density model is able to fit the geoid resulting from subducted slabs. The factor R in (11) has a value of 0.10 Mg m^{-3}.

(iii) *Upper-mantle contributions to the geoid*

Inversions for upper-mantle shear-wave velocity heterogeneities reveal large variations that, if associated with large density variations, might have an observable effect on the geoid. There is a fundamental problem in comparing the results of models from surface-wave tomography and the geoid, however. Geoid kernels are zero at the surface and tend to peak at depth. On the other hand, surface waves are most sensitive to near-surface structure, and models tend to smear the effects of near-surface variations over extensive depth ranges (Tanimoto 1986). (This smearing effect is probably why subducted slabs are not visible in the uppermost mantle using surface-wave tomography; they are camouflaged by the overlying, slow, arc region. It is because of this camouflaging effect that it is necessary to include subducted slabs separately as an additional source of density contrast.) Thus the noisiest (deep) regions of the surface-wave tomographic models are highlighted by the geoid kernels.

Probably for this reason, although all the recent upper-mantle seismic models have striking similarities, we have had the most success matching the observed geoid using the longest

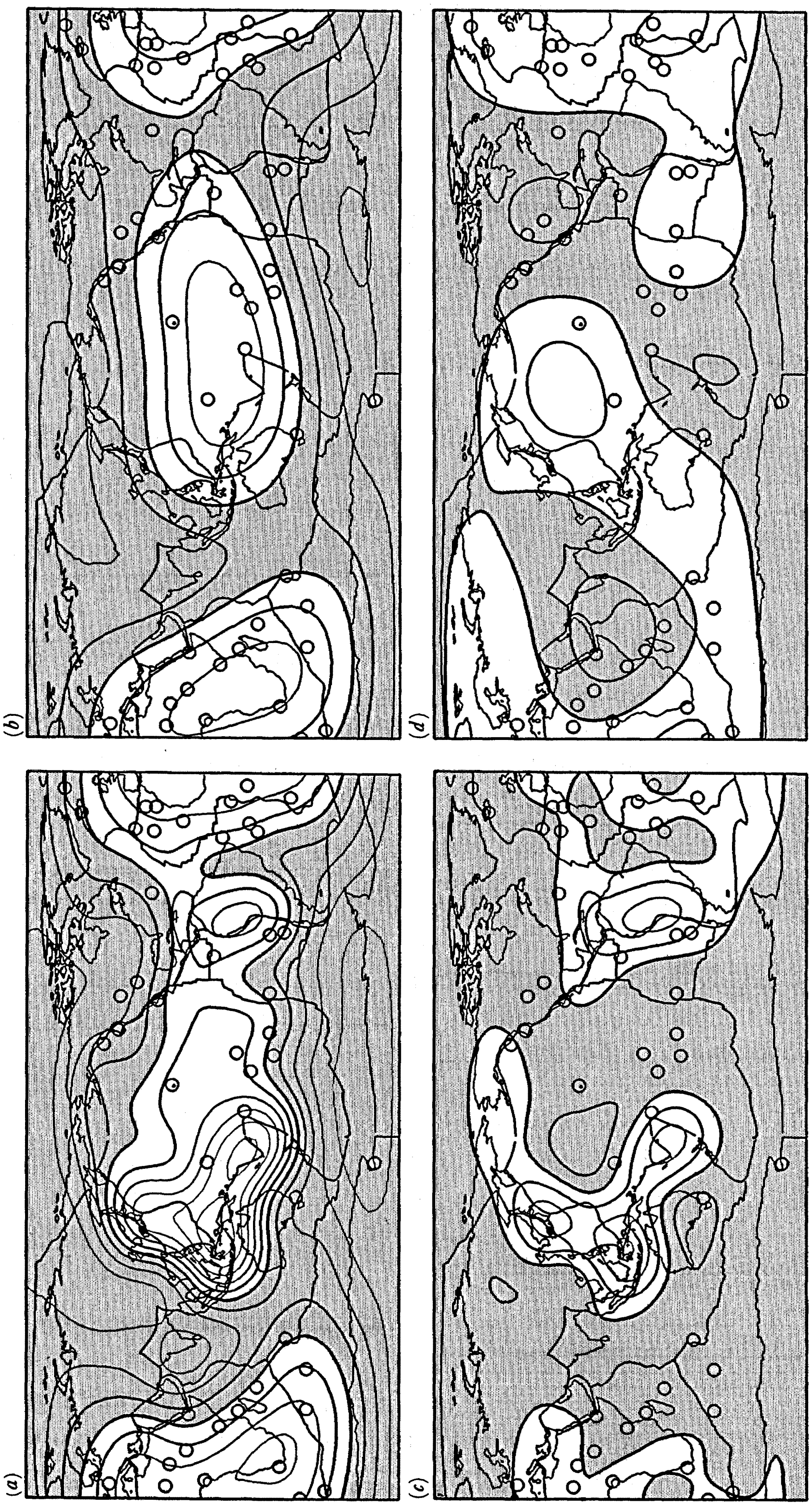

FIGURE 5. (*a*) The total dynamic geoid for model WL, obtained by adding the contributions from lower and upper mantle tomographic models for degrees 2–4 and from slabs for degrees 2–9. (*b*) The geoid predicted using the dynamic geoid kernels for model WL with the lower mantle tomographic model. (*c*) The dynamic geoid predicted for mantle wide flow model WL and the slab density contrast model discussed in the text. (*d*) A dynamic geoid predicted using density contrasts inferred from Tanimoto's (1986) upper mantle structure derived from a tomographic inversion of surface wave data and geoid kernels for model WL. For each, the contour interval is 20 m and geoid lows are shaded.

wavelengths of Tanimoto's (1986) upper-mantle model, a model which does not permit the large rapid radial variations of some other models. The predicted model, with $l_{\max}$ set to 4, is shown in figure 5*d*. The main features in this model geoid are the degree-two high over the western Pacific and the Atlantic, associated with fast material in the transition zone (Masters *et al.* 1982), and lows over North America, Antarctica, and the Carlsberg Ridge. The parameter S in (11) has a value of 0.05 $(\mathrm{Mg\ m^{-3}})/(\mathrm{km\ s^{-1}})$.

(iv) *The residual geoid and postglacial rebound contributions to the geoid*

The residual geoid, obtained by subtracting the model geoid in figure 5*a* from the observed geoid, is shown in figure 6. It is characterized by a number of geophysically interesting features. These include residual geoid highs over most of the more prominent hotspots, a phenomenon addressed in Richards *et al.* (1988). Hudson Bay and West Antarctica, loci of recent major deglaciation, remain as residual lows; indeed, the signature of delayed rebound has been recognized and discussed before (Peltier & Wu 1982). Including the gravitational effects of incomplete postglacial rebound allows over 91% of the geoid variance to be explained (Hager & Clayton 1988).

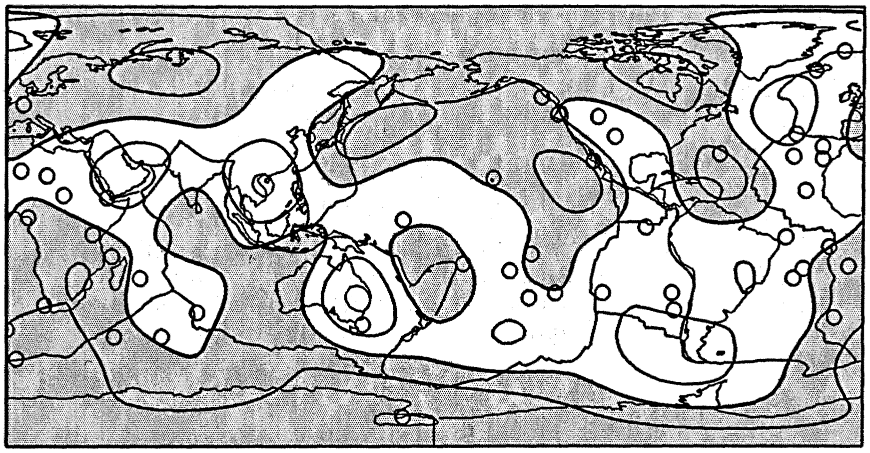

FIGURE 6. The residual geoid obtained by subtracting the model geoid in figure 5*a* from the observed geoid in figure 1. The contour interval is 20 m and geoid lows are shaded.

(*c*) *Discussion of models and tradeoffs*

We have shown results for one particular Earth model. However, as can be seen from figures 2–4, the kernels for a number of significantly different structural models are similar. For example, the positive geoid kernels required by subducted slabs for the upper mantle can occur for either mantle-wide flow models (e.g. WL and WC) or models where the 670 km discontinuity is a chemical boundary (CL). The main difference is in the amplitude of the kernels. Also, similar geoid kernels can be obtained for D″ parametrized as a low-viscosity thermal boundary layer (WL) or a layer of distinct composition (WC); quite good fits to the geoid can also be obtained by ignoring D″ altogether (Hager & Clayton 1988).

Results for a number of models are given in table 1, where values of R, P and S, along with uncertainties and covariances, are given. Also given are the percentages of the variance in the observed geoid explained by the model. Results are given for the three models with kernels shown in figures 2–4, for a hybrid model WL′ that uses slightly different near-surface viscosity models for the slabs and upper-mantle contributions, and for three models, discussed in Hager & Clayton (1988), that do not include D″ (models W, W′, and C).

Model WL′ uses a 'lithospheric' viscosity that is one tenth that in WL for calculating the geoid due to slabs and $10^{-0.5}$ that of WL for the upper-mantle surface-wave model, for reasons discussed below and in Hager & Clayton (1988). Models W, W′, and C are like WL, WL′, and CL, but lack the D″ layer.

The lithosphere at subduction zones is cut completely through by faults, with the strain associated with plate convergence highly localized. One of the major sources of buoyancy driving plate motions is, of course, localized there as well in the form of high-density subducted slabs. We view models WL′ and W′, which have weaker 'lithospheric' layers used in the geoid models for subducted slabs, as regionalized Earth models more appropriate for subduction zones (Hager & Clayton 1988). In these models, we calculate the geoid signatures of the lower mantle, the upper mantle, and subducted slabs, all with different (spherically symmetric) models for the uppermost layer, or lid. Finite-element calculations that include a lower viscosity for the lid near subduction zones perturb the geoid kernels in a similar way to a global low-viscosity lid (Richards & Hager 1988*b*).

All of the models in table 1 give an adequate fit to the geoid. Given the assumptions involved, there is probably not much significance to the small differences among them in variance reduction. Other geophysical considerations must be used to discriminate among them.

TABLE 1. MODEL RESULTS

model	WL	WL′	WC	CL
$R/(\mathrm{Mg\,m^{-3}})$	0.10 ±0.01	0.071±0.007	0.095±0.011	0.33±0.03
$P/(\mathrm{Mg\,m^{-3}})/(\mathrm{km\,s^{-1}})$	0.29 ±0.02	0.19 ±0.01	0.40 ±0.03	0.17±0.01
$S/(\mathrm{Mg\,m^{-3}})/(\mathrm{km\,s^{-1}})$	0.046±0.006	0.047±0.006	0.046±0.007	0.10±0.01
$10^4\ C_{RP}$	−0.15	−0.10	−0.42	−0.49
$10^5\ C_{RS}$	−0.17	−0.31	−1.1	−1.7
$10^4\ C_{PS}$	−0.15	−0.10	−0.30	−0.55
CMB non-hydrostatic ellipticity/m	880	440	80	230
% variance explained	87	87	81	86

model	W	W′	C
$R/(\mathrm{Mg\,m^{-3}})$	0.16±0.02	0.064±0.006	0.36±0.09
$P/(\mathrm{Mg\,m^{-3}})/(\mathrm{km\,s^{-1}})$	0.49±0.03	0.48 ±0.02	0.35±0.03
$S/(\mathrm{Mg\,m^{-3}})/(\mathrm{km\,s^{-1}})$	0.11±0.02	0.041±0.005	0.15±0.02
$10^4\ C_{RP}$	−0.48	−0.94	−1.4
$10^5\ C_{RS}$	−0.13	−0.50	−7.1
$10^4\ C_{PS}$	−0.65	−0.20	−0.60
CMB non-hydrostatic ellipticity/m	2000	2000	1800
% variance explained	84	89	82

There are substantial differences among the model predictions in other respects, however. The models that ignore D″ all have too large an excess ellipticity associated with the CMB to be in accord with the value of *ca.* 500 m inferred from nutation observations (Gwinn *et al.* 1986). Models WL, WL′, and CL all have excess ellipticities on the order of 500 m, whereas WC has a much smaller excess ellipticity. Because most of D″ is not sampled in the tomography model used, and the CMB topography kernels are large in D″, an exact match should not be expected.

There is also substantial variation in the constants R, P and S for the different models, variation that is much greater than the formal statistical estimates for any given model. This is an illustration of a common phenomenon in geophysical modelling; the most important uncertainties are the result of model parametrization, and formal error estimates are not very meaningful.

The most stable parameter is P, the ratio of density to P-wave velocity variation in the lower mantle, which varies from 0.2 to 0.5 (Mg m^{-3})/(km s^{-1}). These values are reasonable based upon limited laboratory measurements of upper-mantle minerals; few direct measurements of lower mantle minerals are available. The value of this ratio obtained from differencing values at the top and bottom of the lower mantle in Earth model PREM (Preliminary Reference Earth Model) (Dziewonski & Anderson 1981) is 0.4, within the range calculated.

For subducted slabs, R ranges by more than a factor of five, from 0.064, the value obtained from the subsidence of oceanic lithosphere with age (Hager & Clayton 1988), to 0.36 Mg m^{-3}. The high value could be indicative of changes in phase (Anderson 1987). Compressive shortening of the slab as it approaches the 670 km discontinuity would also tend to increase the inferred value of R.

There is also substantial variation in S. The lower values are consistent with variations in shear-wave velocity in the upper mantle being in part of compositional origin or caused by relaxation effects. The larger values, when combined with the observed amplitude of velocity variations, would require temperature variations of *ca.* 1000 K (see Hager & Clayton 1988 for discussion).

The relatively small covariance between slab density R and upper-mantle factor S provides justification for considering slabs and upper-mantle tomography separately. We feel that *a priori* modelling is more reliable in estimating slab structure than surface-wave tomography.

4. GEOPHYSICAL IMPLICATIONS

(a) *Viscous or chemical stratification?*

Although our preferred model allows flow through the 670 km seismic discontinuity, we cannot discard the hypothesis of a compositional barrier at that depth based on geoid models, as acceptable fits can be obtained for this type of model. The situation is similar to the state of stress in deep subducted slabs, where both models that have a compositional barrier at 670 km depth and models that have uniform chemistry but a substantial increase in viscosity at this depth can explain the observations (Vassiliou *et al.* 1984).

There are other observations that favour flow through the 670 km discontinuity. The association of low-velocity lower mantle with regions of extensive hotspot activity at the surface is suggestive of mantle-wide flow. Travel-time anomalies from deep earthquakes interpreted in terms of deeply extending high-velocity anomalies are also suggestive of transport of at least some material across the 670 km discontinuity (Jordan 1977; Creager & Jordan 1984, 1986). However, both these examples of apparent continuity of temperature fields across the upper-mantle–lower-mantle boundary could conceivably be the result of thermally coupled convection in a layered system.

The best way of discriminating between these two types of convection might be by using the dynamic topography of the 670 km discontinuity predicted to exist if the mantle were chemically stratified. Given the large density contrasts inferred for subducted slabs in this type of model to explain the observed geoid, the dynamically maintained topography would be several hundred kilometres or more in the vicinity of subducted slabs (Hager & Raefsky 1981; Christensen & Yuen 1984). The deepest earthquakes are not resolvably deeper than the average depth of the 670 km discontinuity, suggesting that a nearly isobaric phase change, rather than a chemical barrier to flow, shuts off seismicity there. Detailed observations of phases converted at the '670' beneath the Tonga slab show no indication of substantial topography (Richards & Wicks 1987). The great success that we have had in matching the observed geoid by including the effects of dynamic topography gives us confidence in the general principles involved. The lack of topography observed on the 670 km discontinuity seems difficult to reconcile with its being a compositional barrier to convection. One possible way out would be if an approximately isobaric phase change at *ca.* 670 km depth camouflages the dynamic topography there.

Although the geoid places constraints upon relative variations in viscosity, it is insensitive to the absolute value of mantle viscosity, so long as the viscosity is high enough that the Reynolds number of the flow is negligible. But the velocity of the flow associated with a given density field is, of course, inversely proportional to the absolute viscosity. O'Connell & Hager (1984) (see also Hager & Clayton 1988) used the heat flux inferred to result from flow driven by the density fields used in geoid models to bound the viscosity of the lower mantle. So that the flow does not advect more heat than is observed to pass through Earth's surface, they argued that the lower mantle viscosity must be 3×10^{22} Pa s or more. This value of mantle viscosity is consistent with the gravitational signature of postglacial rebound discussed in §3*b*iv.

Several other lines of evidence support the idea of a substantial viscosity increase in the lower mantle. Gurnis & Hager (1988) have recently investigated fluid-dynamical models of slabs sinking through a mantle with a viscosity structure similar to that inferred from geoid modelling. They found that in addition to matching the observed state of stress, slabs in their models often showed kinks as they passed into the lower mantle similar to those inferred from analysis of seismic travel times (Jordan 1977; Creager & Jordan 1984, 1986; Zhou & Clayton 1987).

A substantial increase in the viscosity of the lower mantle results in a substantial difference between the upper and lower mantle in the style of convection. So that stresses are continuous across the boundary, strain rates must vary inversely with the viscosity. Thus the upper mantle should be much more thoroughly sheared and mixed than the lower mantle, in accord with geochemical inferences.

There is also geophysical evidence supporting the concept of slow flow velocities, presumably the result of high viscosity, in the lower mantle. Grand (1987) interpreted a high-velocity anomaly beneath North America in the lower mantle in his regional tomographic inversion as a fossil remnant of the Farallon plate, suggesting a residence time in the lower mantle of over 100 Ma. This inferred slow flow is consistent with Anderson's (1982) and Chase & Sprowl's (1983) empirical interpretations of the geoid high associated with the African hotspot province being a fossil relict of the effects of the Pangean supercontinent on the lower mantle temperature field.

(*b*) *Dynamic topography*

The dynamic topography predicted at Earth's surface for our preferred model has a peak to peak amplitude of about 1 km, with uplifted regions generally associated with long-wavelength geoid highs (see Hager & Clayton 1988 for figures). Although there is no global data base of dynamically maintained topography for comparison, in regions where it is available, there is relatively good agreement between the predicted dynamic topography for our mantle-wide flow models and residual bathymetry not explained by plate cooling (Crough 1983; Cazenave *et al.* 1988). There is no such agreement for the flow models with a change in compositions at 670 km depth. The predicted dynamic topography is nearly independent of the present locations of continents and oceans. If these patterns are fixed relative to a high-viscosity lower mantle, relative motions of the continents and ocean basins should lead to changes of eustatic sea level.

Observations of the effects of core–mantle coupling on nutation amplitudes (Gwinn *et al.* 1986) and changes of length of day (Hide 1986) suggest that the topography of the CMB has an amplitude of hundreds of metres. It was to suppress the CMB topography calculated for our earlier models (Hager *et al.* 1985; Hager & Clayton 1988) that we included D″ in the models presented here. We have also argued that a low viscosity, perhaps stratified, D″ is to be expected *a priori*. The 10 km of relief at the CMB reported by Morelli & Dziewonski (1987) seem at first inconsistent with our expectations for much smaller amplitude dynamically maintained topography there.

There are at least two possible resolutions to the apparent paradox. The first is that substantial topography would be expected on the top of a chemically distinct D″. Equation (6) indicates that the dynamic topography is inversely proportional to the density contrast across a boundary. Because the density contrast between D″ and the overlying mantle is much smaller than the density contrast across the CMB, the top of a chemically distinct D″ would be expected to have very large (*ca.* 100 km) dynamic topography. The tradeoff between D″ velocity structure and thickness and CMB deflection should be considered for seismological inferences of CMB topography (Gudmundsson *et al.* 1987).

Alternatively, if there were a layer of molten silicate between the solid mantle and molten iron core, substantial dynamic topography of the solid–molten silicate boundary would result. For a given dynamic stress, boundary deflection is inversely proportional to the density contrast across a boundary (equation (6)). A layer of molten silicate at the top of the core would have a much smaller density contrast with the mantle than molten iron has, allowing much larger deflection of the boundary between solid mantle and fluid 'core'. (It is this high-impedance contrast boundary that would be seen by seismic reflection studies.) The boundary between molten silicate and molten metal would be difficult to detect using seismology, so the existence of such a layer would be difficult to detect directly. Its effects on nutation have not yet been worked out. Filling in of upwarpings of the solid boundary by 'anti-oceans' of molten silicate would shield the solid mantle from mechanical interaction with the molten iron core, allowing the change in length of day constraints to be satisfied.

5. CONCLUSIONS

Our ability to successfully model the observed geoid by using fluid-mechanical models with input of density fields inferred from seismic tomography gives us confidence both in the tomographic inversions, at least as smoothed by the long-wavelength geoid kernels, and in the importance of dynamic topography. We can fit the geoid either with models with a chemical barrier to flow at the 670 km seismic discontinuity or with models with a substantial viscosity increase with depth. The lack of observed large topography at this discontinuity leads us to prefer the latter model, i.e. deep-mantle convection.

A high-viscosity lower mantle would lead to a stratification of the style of mantle convection, without totally separating the upper and lower mantle. The slow mixing and long residence times inferred for such a high-viscosity lower mantle are consistent with a number of geochemical and geophysical observations.

Although our models are fairly successful, they are based upon a number of simplifying assumptions that should be removed in future work. Most notably, the assumption of spherical symmetry of the effective viscosity should be removed, particularly in the lithosphere. This will require substantial computing resources, but there are indications of systematic effects in the residual geoid (figure 6) that suggest that lateral variations are somewhat important. For example, residual geoid highs are associated with groups of hotspots in the central Indian Ocean, the North Atlantic, the Southeast Pacific, and the Basin and Range Province, perhaps indicative of locally negative geoid kernels resulting from lateral variations in viscosity in these regions.

Better constrained slab locations are becoming available through seismic tomography (Zhou & Clayton 1987). These can be used to improve the *ad hoc* slab model used. Considerations of solid-state theory may allow more realistic parametrizations of variations of the ratios P and S with depth. Self compression should also be included.

Finally, although the models presented here were computed by using trial-and-error forward modelling, we probably now understand the parametrization of the problem well enough to attempt a nonlinear inversion. This approach should enable us to explain more of the variance in the observed geoid, as well as to refine our models of mantle viscosity structure.

Close interaction with R. W. Clayton, R. P. Comer, and R. J. O'Connell was essential to this research. Critical reviews by D. L. Anderson, U. R. Christensen, W. Kiefer, and B. Parsons improved the manuscript substantially. This work was supported by NASA grants NAG5-315 and NAG5-842. This is contribution number 4625, Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California 91125.

REFERENCES

Anderson, D. L. 1982 *Nature, Lond.* **297**, 391–393.
Anderson, D. L. 1987 *J. geophys. Res.* **92**, 13968–13980.
Ashby, M. F. & Verrall, R. A. 1977 *Phil. Trans. R. Soc. Lond.* A **288**, 59–95.
Cazenave, A., Dominh, K., Rabinowicz, M. & Ceuleneer, G. 1988 *J. geophys. Res.* **93**, 8064–8077.
Chase, C. G. 1979 *Nature, Lond.* **282**, 464–468.
Chase, C. G. & McNutt, M. K. 1982 *Geophys. Res. Lett.* **9**, 29–32.
Chase, C. G. & Sprowl, D. R. 1983 *Earth planet. Sci. Lett.* **62**, 314–320.
Christensen, U. R. & Yuen, D. A. 1984 *J. geophys. Res.* **89**, 4389–4402.
Clayton, R. W. & Comer, R. P. 1984 *Terra Cognita* **4**, 282–283.
Creager, K. C. & Jordan, T. H. 1984 *J. geophys. Res.* **89**, 3031–3049.
Creager, K. C. & Jordan, T. H. 1986 *J. geophys. Res.* **91**, 3573–3589.
Crough, S. T. 1983 *A. Rev. Earth planet. Sci.* **11**, 165–193.
Crough, S. T. & Jurdy, D. M. 1980 *Earth planet. Sci. Lett.* **48**, 15–22.
Douglas, B. C., McAdoo, D. C. & Cheney, R. E. 1987 *Rev. Geophys.* **25**, 875–880.
Dziewonski, A. M. 1984 *J. geophys. Res.* **89**, 5929–5952.
Dziewonski, A. M. & Anderson, D. L. 1981 *Physics Earth planet. Int.* **25**, 297–356.

Dziewonski, A. M., Hager, B. H. & O'Connell, R. J. 1977 *J. geophys. Res.* **82**, 239–255.

Gantmacher, F. R. 1960 *The theory of matrices*, vols 1 and 2 (transl. from Russian by K. A. Hirsch). New York: Chelsea.

Grand, S. P. 1987 *J. geophys. Res.* **92**, 14065–14090.

Gudmundsson, O., Clayton, R. W. & Anderson, D. L. 1987 *Eos, Wash.* **68**, 1378.

Gurnis, M. & Hager, B. H. 1988 *Nature, Lond.* **335**, 317–321.

Gwinn, C. R., Herring, T. A. & Shapiro, I. I. 1986 *J. geophys. Res.* **91**, 4755–4765.

Hager, B. H. 1983 *Earth planet. Sci. Lett.* **63**, 97–109.

Hager, B. H. 1984 *J. geophys. Res.* **89**, 6003–6015.

Hager, B. H. & Clayton, R. W. 1988 In *Mantle convection* (ed. W. R. Peltier), pp. 657–763. New York: Gordon and Breach.

Hager, B. H., Clayton, R. W., Richards, M. A., Comer, R. P. & Dziewonski, A. M. 1985 *Nature, Lond.* **313**, 541–545.

Hager, B. H. & O'Connell, R. J. 1981 *J. geophys. Res.* **86**, 4843–4867.

Hager, B. H. & Raefsky, A. 1981 *Eos, Wash.* **62**, 1074.

Hide, R. 1986 *Q. Jl R. astr. Soc.* **27**, 30–20.

Isacks, B. & Molnar, P. 1971 *Rev. Geophys. space Phys.* **9**, 103–174.

Jordan, T. H. 1977 *J. Geophys.* **43**, 473–496.

Kaula, W. M. 1972 *The nature of the solid Earth* (ed. E. C. Robertson), pp. 386–405. New York: McGraw-Hill.

Kaula, W. M. 1980 *J. geophys. Res.* **85**, 7031–7044.

Lay, T. & Helmberger, D. V. 1983*a* *J. geophys. Res.* **88**, 8160–8170.

Lay, T. & Helmberger, D. V. 1983*b* *Geophys. Jl R. astr. Soc.* **75**, 799–837.

Lerch, F. J., Klosko, S. M. & Patch, G. B. 1983 NASA Tech. Memo 84986.

Masters, G., Jordan, T. H., Silver, P. G. & Gilbert, F. 1982 *Nature, Lond.* **298**, 609–613.

Morelli, A. & Dziewonski, A. M. 1987 *Nature, Lond.* **325**, 678–683.

Nakiboglu, S. M. 1982 *Phys. Earth planet. Int.* **28**, 302–311.

Nataf, H.-C., Nakanishi, I. & Anderson, D. L. 1984 *Geophys. Res. Lett.* **11**, 109–112.

Nataf, H.-C., Nakanishi, I. & Anderson, D. L. 1986 *J. geophys. Res.* **91**, 7261–7307.

O'Connell, R. J. & Hager, B. H. 1984 *Eos, Wash.* **65**, 1093.

O'Connell, R. J., Hager, B. H. & Richards, M. A. 1984 *Terra Cognita* **4**, 261.

Pekeris, C. L. 1935 *Mon. Not. R. astron. Soc. geophys. Suppl.* **3**, 343–367.

Peltier, W. R. & Wu, P. 1982 *Geophys. Res. Lett.* **9**, 731–734.

Revenaugh, J. & Parsons, B. 1987 *Geophys. Jl R. astro. Soc.* **90**, 349–368.

Ricard, Y., Fleitout, L. & Froidevaux, C. 1984 *Anns Geophysicae* **2**, 267–286.

Richards, M. A. & Wicks, C. W. 1987 *Eos, Wash.* **68**, 1379.

Richards, M. A. & Hager, B. H. 1984 *J. geophys. Res.* **89**, 5987–6002.

Richards, M. A. & Hager, B. H. 1988*a* *The physics of planets* (ed. S. K. Runcorn), pp. 247–277. Chichester: John Wiley.

Richards, M. A. & Hager, B. H. 1988*b* *J. geophys. Res.* (In the press.)

Richards, M. A., Hager, B. H. & Sleep, N. H. 1988 *J. geophys. Res.* **93**, 7690–7708.

Runcorn, S. K. 1964 *J. geophys. Res.* **69**, 4389–4394.

Runcorn, S. K. 1967 *Geophys. Jl R. astr. Soc.* **42**, 375–384.

Sammis, C. G., Smith, J. C., Schubert, G. & Yuen, D. A. 1977 *J. geophys. Res.* **82**, 3747–3761.

Schubert, G., Yuen, D. A. & Turcotte, D. L. 1975 *Geophys. Jl R. astr. Soc.* **42**, 705–735.

Tanimoto, T. 1986 *Geophys. Jl R. astr. Soc.* **84**, 49–69.

Tanimoto, T. & Anderson, D. L. 1984 *Geophys. Res. Lett.* **4**, 287–290.

Vassiliou, M. S., Hager, B. H. & Raefsky, A. 1984 *J. Geodyn.* **1**, 11–28.

Woodhouse, J. H. & Dziewonski, A. M. 1984 *J. geophys. Res.* **89**, 5953–5986.

Young, C. J. & Lay, T. 1987 *A. Rev. Earth planet. Sci.* **15**, 25–46.

Zhang, S. & Yuen, D. A. 1987 *Geophys. Res. Lett.* **14**, 899–902.

Zhou, H. W. & Clayton, R. W. 1987 *Eos, Wash.* **68**, 1379.