Deep-mantle high-viscosity flow and thermochemical structure inferred from seismic and geodynamic data

Alessandro M. Forte* **& Jerry X. Mitrovica†**

** Department of Earth Sciences, University of Western Ontario, Biology & Geology Building, London, Ontario, N6A 5B7 Canada*
† Department of Physics, University of Toronto, 60 St George Street, Toronto, Ontario, M5S 1A7 Canada

Surface geophysical data that are related to the process of thermal convection in the Earth's mantle provide constraints on the rheological properties and density structure of the mantle. We show that these convection-related data imply the existence of a region of very high effective viscosity near 2,000 km depth. This inference is obtained using a viscous-flow model based on recent high-resolution seismic models of three-dimensional structure in the mantle. The high-viscosity layer near 2,000 km depth results in a re-organization of flow from short to long horizontal length scales, which agrees with seismic tomographic observations of very long wavelength structures in the deep mantle. The high-viscosity region also strongly suppresses flow-induced deformation and convective mixing in the deep mantle. Here we predict compositional and thermal heterogeneity in this region, using viscous-flow calculations based on the new viscosity profile, together with independent mineral physics data. These maps are consistent with the anti-correlation of anomalies in seismic shear and bulk sound velocity in the deep mantle. The maps also show that mega-plumes in the lower mantle below the central Pacific and Africa are, despite the presence of compositional heterogeneity, buoyant and actively upwelling structures.

A quantitative model of thermal convective flow in Earth's mantle should satisfy a number of fundamental geophysical observations. In particular, such a model must accurately predict the principal surface data sets associated with the mantle convection process, namely: the global-scale free-air gravity anomalies which are accurately constrained by satellite data[1]; the observed tectonic-plate motions[2]; the dynamic topography on the core–mantle boundary (CMB), in particular the excess or dynamic ellipticity of the CMB inferred from space geodetic observations of the free-core nutation period[3,4]; and the dynamic topography at the surface which may be extracted from the observed topography by removing the contribution due to isostatic compensation of crustal thickness variations[5–8]. A successful mantle flow model must also be consistent with the three-dimensional structure in global seismic tomographic models, especially the observed dominance of very long wavelength heterogeneity in the bottom 1,000 km of the lower mantle[9–13]. This last observation has served as a basis for recent proposals of compositional stratification in the bottom half of the mantle[14] and thermochemical convection in the mantle[15–17]. For a recent survey of the arguments for deep-mantle compositional heterogeneity, from a mainly geochemical perspective, see ref. 18.

Thermal and compositional effects on seismic anomalies

The possible presence of significant chemical heterogeneity in the deep mantle may well bias our inferences of the temperature and density structure of convection from seismic tomographic models. This possibility is highlighted by recent seismic studies of the deep lower mantle, which reveal an unusually high ratio of shear velocity to compressional velocity anomalies[19] and a clear anti-correlation between perturbations of seismic shear velocity and acoustic (bulk-sound) velocity[20–23]. To understand the implications of this anti-correlation we first estimated the temperature and compositional derivatives of density and seismic wave velocities at 2,740 km depth, which corresponds to the top of the D″ layer in the PREM[24] seismic reference model (see Methods). We restricted our analysis of major-element compositional heterogeneity to the simplified chemical system MgO–FeO–SiO_2 and hence did not consider possible effects[25,26] due to variations in the less abundant oxides CaO and Al_2O_3, which each constitute 3–4% (by weight) of the bulk mantle composition[27]. From the estimated temperature derivatives (Table 1) it is clear that the seismic shear velocity (V_s) displays the greatest sensitivity to temperature perturbations (especially when anelastic effects are included), whereas the bulk-sound velocity (V_ϕ) is an order of magnitude less sensitive to the temperature perturbations.

Inferences of mantle viscosity and density perturbations

Given the high sensitivity of seismic shear velocity to temperature variations, we selected two recent three-dimensional mantle models[13,28] that provide lateral variations of shear velocity and used them to construct convective-flow models which satisfy the convection-related surface data described above. These tomographic models are derived from distinct, independent data sets, using different parameterizations of the heterogeneity, and using different inversion procedures. Thus we may view these two three-dimensional seismic models as representative of a wide spectrum of possible models, with one[13] (here referred to as the 'Grand' model) based on the inversion of travel-time data and the other[28] (here referred to as the 'Ek&Dz' model) based mainly on the inversion of full seismic waveforms. We assumed initially that the seismic anomalies in the two models are thermal in origin. To begin, we therefore used a velocity-to-density scaling coefficient (long-dashed line, Fig. 1a) derived by Karato[29] and subsequently modified on the

Table 1 Seismic temperature and compositional derivatives at 2,740 km depth

Elastic property	Temperature derivatives, $\partial/\partial T$ ($\times 10^5$ K^{-1})		Compositional derivatives	
	Anharmonic	Anelastic	$\partial/\partial X_{Fe}$	$\partial/\partial X_{Pv}$
$\ln\rho$	−1.0	–	+0.32	$+4.3\times10^{-3}$
$\ln V_s$	−4.7	−2.4	−0.22	+0.045
$\ln V_p$	−2.2	−0.9	−0.18	+0.047
$\ln V_\phi$	−0.7	0	−0.16	+0.048

For details on the construction of this table see the Methods section.

© 2001 Macmillan Magazines Ltd

basis of geodynamic constraints[5]. The density perturbations we obtained using this scaling are summarized in Fig. 1b, which displays the average amplitude of $\delta\rho/\rho$ as a function of depth. With the exception of the top 300 km, we note that the two seismic models yield rather different amplitudes for the density perturbations. We use these inferences of density heterogeneity to determine directly the buoyancy forces that drive the mantle convective flow.

The theory we use to calculate the mantle flow that is expected on the basis of the tomography-derived density anomalies requires a model of the rheological properties of the mantle, which we represent here in terms of a depth-dependent effective viscosity. In this flow theory, the buoyancy-induced flow in the mantle is coupled to the 13 principal surface tectonic plates[2], treated as rigid bodies, whose motions are determined by the underlying flow field[30]. The viscous-flow theory assumes an effectively linear, newtonian viscosity which is strongly suggested for the lower mantle by high-temperature creep experiments on perovskite-analogue minerals[31].

The required model of effective viscosity has been obtained using a nonlinear, iterative, Occam-style inversion[32–35] of the main convection-related geophysical data. Figure 2a shows results from inversions based on the two three-dimensional seismic models and Table 2 (rows indicated by asterisks) summarizes the fit of the corresponding flow models to the data. The viscosity profiles (Fig. 2a) are very similar, despite the non-negligible differences in the amplitudes of the density contrasts delivered by the two seismic models (Fig. 1b). The inferred viscosity variations display a pronounced low-viscosity channel in the upper mantle (between 100 and 300 km depth) and two viscosity maxima in the lower mantle, one near the top of the lower mantle and the other at 2,000 km depth. These lower-mantle viscosity maxima are not artefacts of the inversion procedure because the Occam approach explicitly penalizes radial roughness in the viscosity profiles. We carried out tests in which we removed either of the two viscosity peaks, or simply joined them to create a single high-viscosity region, and in each case the fit to the convection data (in particular the gravity anomalies) was significantly reduced.

An alternative demonstration of how the convection data constrains mantle viscosity at different depths is provided by a resolving power analysis. The resolution kernels shown in Fig. 2b clearly indicate that the viscosity maxima near 1,000 km and 2,000 km depth are well resolved, with little trade-off with viscosity structure elsewhere. The viscosity minima near 200 km and 1,400 km depth are similarly well constrained. The viscosity resolution in the top half of the lower mantle is mainly provided by the plate-motion data whereas in the bottom half of the mantle (for example, at 2,000 km depth) the resolution is strongly controlled by the gravity-anomaly data (Fig. 2b).

Independent verification of the viscosity profiles in Fig. 2a is provided by a comparison with radial profiles of seismic attenuation or Q factor[36]. We calculated the average viscosity in each of the five mantle layers used to parameterize a recent seismic Q model[36]. Assuming the frequency (ω) dependence[37], $Q \approx \omega^{\alpha}$, we translated the five-layer viscosity model into an equivalent Q model and obtained the best least-squares fit to the actual seismic Q model using $\alpha = 0.275$, which falls in the currently accepted range[37] $\alpha = 0.1$–0.3.

Next, we consider the possibility of further improving the fit to the convection data by inverting the data to derive an optimal $d\ln\rho/d\ln V_s$ scaling coefficient, rather than adopting the *a priori* scaling (long-dashed line, Fig. 1a) derived from mineral physics data. We carried out Occam inversions for $d\ln\rho/d\ln V_s$, and the two profiles we obtained are displayed in Fig. 1a. The fits to the convection data provided by the new inferences are summarized in Table 2 (rows indicated with daggers). The match we obtained with the excess CMB ellipticity (Table 2) is essentially perfect and, in the case of Ek&Dz, this is mainly achieved by adjusting the density heterogeneity inferred in the lowermost mantle using the $d\ln\rho/d\ln V_s$ scaling (solid line, Fig. 1a). We note that with the new $d\ln\rho/d\ln V_s$ profiles, differences in the average amplitudes of the density perturbations (Fig. 1c) are diminished relative to Fig. 1b.

Dynamics of lower-mantle flow and implications for mixing

To explore the dynamical impact of the high viscosity peak at 2,000 km depth (Fig. 2a) we constructed a reference two-layer viscosity model by separately averaging the viscosity variations in the upper and lower mantle. This averaging procedure was applied to the viscosity profile inferred using the heterogeneity in the Ek&Dz model and the resulting profile (long-dashed line, Fig. 2a) is characterized by a factor of 50 jump in viscosity at 670 km depth.

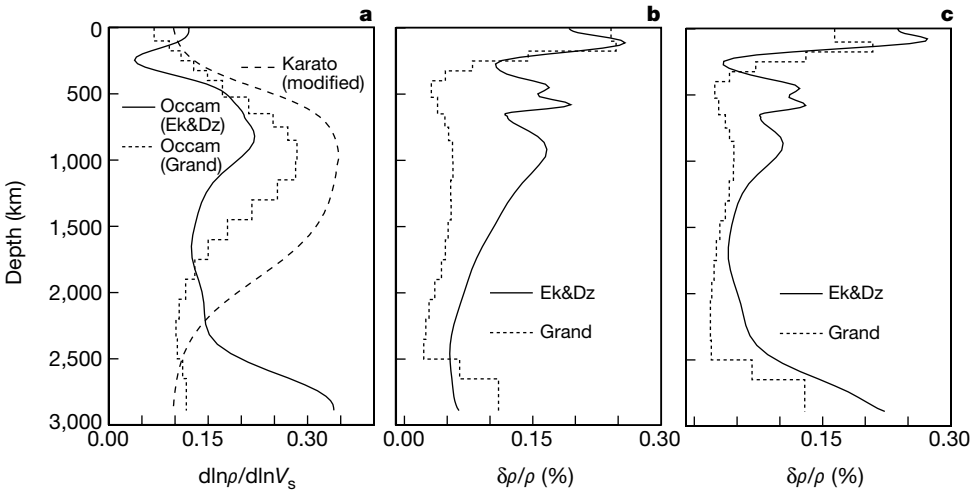

Figure 1 Geodynamic inferences of density structure in Earth's mantle. **a**, The long-dashed line represents the $d\ln\rho/d\ln V_s$ (velocity-to-density) scaling coefficient originally estimated by Karato[29] on the basis of mineral physics data that we subsequently modified on the basis of geodynamic data. The short-dashed and solid lines represent the $d\ln\rho/d\ln V_s$ scaling inferred for the three-dimensional seismic models of Grand[13] and Ekström and Dziewonski[28], respectively, after performing Occam inversions of the geodynamic data. (In these inversions the corresponding viscosity profiles in Fig. 2a were employed.) **b**, The root-mean-square (r.m.s.) amplitude of the density perturbations, as a function of depth, obtained from the three-dimensional seismic models of Grand[13] (dashed line) and Ekström and Dziewonski[28] (solid line). In both cases the $d\ln\rho/d\ln V_s$ (modified) scaling inferred by Karato[29] (long-dashed line in **a**) is employed. **c**, The r.m.s. amplitude of the density perturbations derived from the three-dimensional seismic models of Grand[13] (dashed line) and Ekström and Dziewonski[28] (solid line) on the basis of the respective Occam-inferred $d\ln\rho/d\ln V_s$ scalings in **a**.

© 2001 Macmillan Magazines Ltd

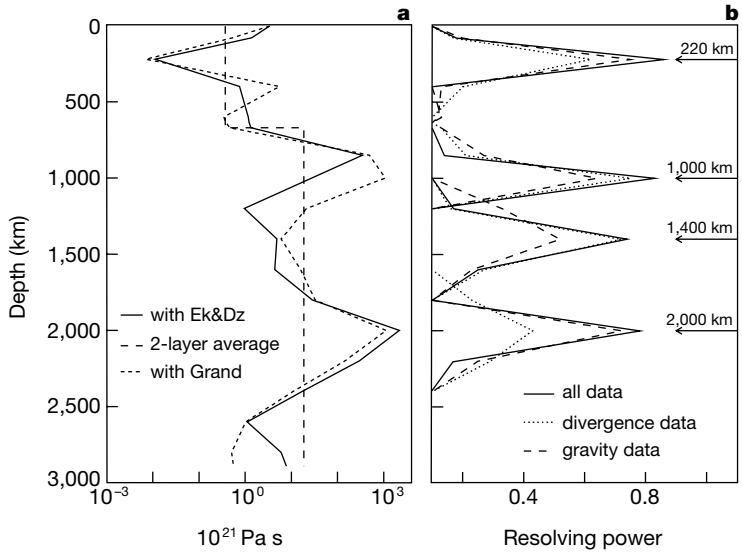

Figure 2 Depth-dependent effective viscosity of Earth's mantle. **a**, The short-dashed and solid lines represent the viscosity profiles inferred in Occam inversions of the geodynamic data (listed in Table 2) based on the seismic models of Grand[13] and Ekström and Dziewonski[28], respectively. In these Occam inversions the $d\ln\rho/d\ln V_s$ (modified) scaling of Karato[29] was employed (long-dashed curve in Fig. 1a). The long-dashed curve is a two-layer viscosity profile obtained by separately averaging the upper-mantle and lower-mantle viscosity variations in the profile (solid line) inferred on the basis of the Ekström and Dziewonski seismic model. **b**, The viscosity resolving power of the geodynamic data at four depths indicated by the horizontal arrows (for example, at 220 km). These curves, called 'resolution kernels', show the vertical averages of the mantle viscosity that are inherent in the viscosity inferences at these four depths. Perfect resolution would correspond to kernels with peak values of 1 and zero width. The solid lines show the resolution obtained with the full set of geodynamic data (gravity anomalies, plate divergence, and CMB ellipticity). The dotted lines show the resolution obtained when we ignore the gravity data and the dashed lines show the resolution in the absence of plate-divergence data.

The lower-mantle flow field predicted on the basis of this two-layer viscosity is depicted in Fig. 3a, c at depths of 1,300 and 1,800 km, respectively. Clearly, in this flow calculation the pattern of both the horizontal and vertical flow fields is unchanged across this depth range.

In contrast to this simple behaviour, consider Fig. 3b, d, where we plot flow predicted on the basis of the viscosity profile (solid line, Fig. 2a) derived by inversion of the convection data. We now observe a remarkable change in the flow regime between 1,300 and 1,800 km depth. At 1,300 km depth (Fig. 3b), the flow is organized on relatively short horizontal length scales, with clearly defined upwellings below the Pacific, especially below Hawaii and French Polynesia in the central Pacific. These upwelling centres are being fed by converging horizontal flow which, in the case of the Hawaiian plume, seems to originate from as far north as the Aleutian subduction zone. This observation seems to provide a natural explanation for the observed signature of recycled oceanic crust in ocean-island volcanic rocks[38]. In contrast, at 1,800 km depth (Fig. 3d) the individual upwellings are now absent and both the vertical and horizontal flow are dominated by very long horizontal wavelengths. Although the viscosity increases rapidly towards its maximum near 2,000 km depth, its control on the horizontal organization of flow extends to shallower depths and is already discernible at 1,800 km depth. The flow character at this depth is closely reminiscent of the oscillatory 'doming regime' observed in recent laboratory convection experiments[17].

A useful measure of the degree of mixing induced in laminar fluid flows is provided by the time-dependent separation or 'stretching' of originally adjacent particles in the fluid[39,40]. To quantify this stretching we determined the direction and magnitude of the principal (that is, maximum) strain rate at a large number of points distributed throughout the mantle, employing the strain-rate tensor $d\epsilon_{ij}/dt$ associated with the predicted flow fields. The largest of the principal stretching rates predicted at any given depth is shown in Fig. 4, where we immediately note a strong minimum at 2,000 km depth. Another prominent minimum is predicted at the top of the lower mantle, near 1,000 km depth. These zones of very low deformation rate, and hence reduced mixing, reflect the two high-viscosity regions in the inverted models; they are absent in the flow predicted using the simple two-layer viscosity profile (dashed line, Fig. 4).

Tomography-based studies of convection-related surface data have long argued for an increase of 1–2 orders of magnitude in the average viscosity between the upper and lower mantle[33,34,41–43]. These variations affect both the mixing properties[44–46] and the planform of mantle convective flow[47]. Recent investigations[46] suggest that viscosity stratification alone may not be sufficient to explain the trace-element geochemical constraints on mantle

Table 2 Fits to global convection data

Three-dimensional mantle model	Free-air gravity, $l = 2$–20	Plate divergence, $l = 1$–32	Excess CMB ellipticity	r.m.s. topography $l = 1$–20
Grand*	31% (54%)	60%	0.50 km	0.91 km
Grand†	32% (56%)	59%	0.52 km	0.71 km
Ek&Dz*	37% (77%)	57%	–0.16 km	1.32 km
Ek&Dz†	43% (71%)	69%	0.52 km	0.95 km
Ek&Dz‡	9% (–65%)	68%	0.76 km	1.04 km

All fits, with the exception of the surface and CMB topography, are expressed in terms of variance reduction. The observed free-air gravity anomalies are calculated using the (non-hydrostatic) geopotential derived from satellite data[1]. The plate-divergence data are represented here in terms of the horizontal divergence of the plate velocities given by the NUVEL-1 model[2]. The most recent inference[4] of excess or dynamic CMB ellipticity suggests a value of about 0.4 km, rather than 0.5 km as determined in earlier studies[3]. (This 20% difference is well within the uncertainties associated with the three-dimensional seismic models used in these flow predictions.) The dynamic surface topography is not employed in the Occam inversions for the mantle viscosity profile. The root-mean-square (r.m.s.) amplitude of the dynamic topography is used as an independent *a posteriori* check of the plausibility of the convective-flow models. The most recent inference[5] of the r.m.s. amplitude of the dynamic surface topography is 0.9 km. The numbers in parentheses are the variance reductions to the equivalent non-hydrostatic geoid anomalies derived from satellite data[1]. The wavelength-dependent amplitude spectra of the geoid and gravity anomalies are quite different; the latter has a relatively 'flat' spectrum, which explains why the fits to the anomalous gravity and geoid fields differ.

*The density perturbations in the convective-flow models are derived from the shear-velocity anomalies assuming a thermal origin (using the $d\ln\rho/d\ln V_s$ 'Karato' profile in Fig. 1a).

†These density perturbations are instead derived on the basis of the associated $d\ln\rho/d\ln V_s$ profiles (Fig. 1a) inferred in Occam inversions of the convection data.

‡The two-layer viscosity model (Fig. 2a) along with the Occam-inferred $d\ln\rho/d\ln V_s$ profile (solid line, Fig. 1a) are used in this viscous flow calculation.

© 2001 Macmillan Magazines Ltd

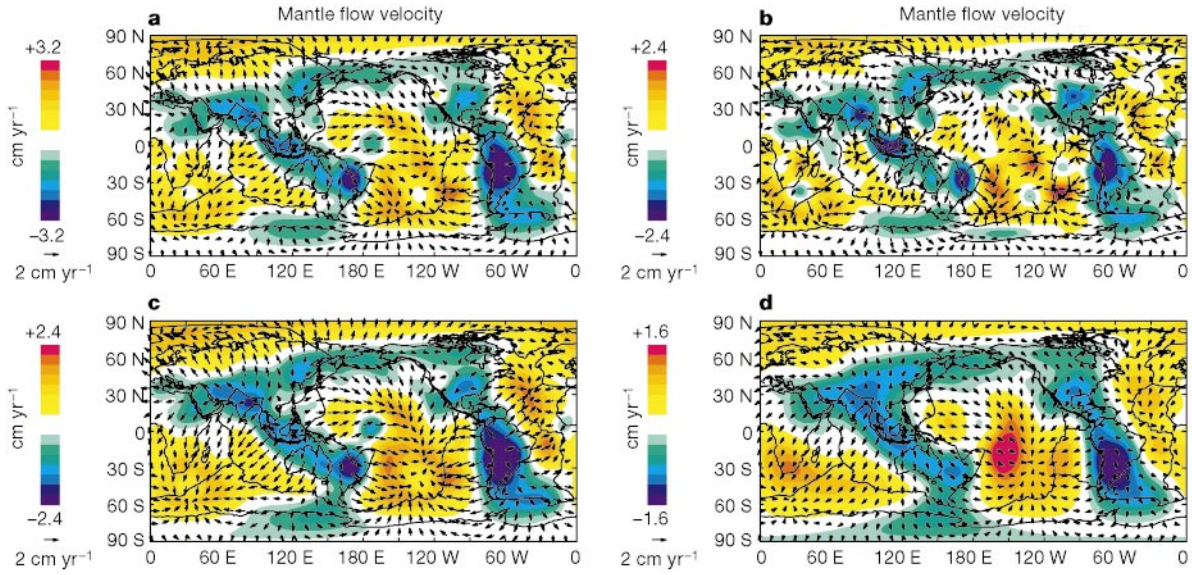

Figure 3 Predicted convective flow in Earth's lower mantle. In all maps the colour contours represent the speed of vertical flow (see the colour scale bars in cm yr^{-1}) with the blue colours representing regions of downward flow and the orange colours representing regions of upward flow. The black arrows represent the speed and direction of the horizontal component of the flow (see the horizontal velocity scale arrow). All maps depict the convective flow predicted on the basis of the mantle density anomalies derived from the three-dimensional seismic model of Ekström and Dziewonski[28] (using the $d\ln\rho/d\ln V_s$ scaling represented by the solid line in Fig. 1a). The maps in **a** and **c** represent the flow field predicted on the basis of the two-layer viscosity profile (long-dashed curve in Fig. 2a). The maps in **b** and **d** represent the flow field predicted on the basis of the Occam-inferred viscosity profile (solid curve in Fig. 2a). **a**, **b**, Depth of 1,300 km. **c**, **d**, Depth of 1,800 km.

mixing. However, the viscosity profiles employed in these numerical convection studies were not constrained to fit the full suite of surface geophysical constraints (in particular the global gravity anomalies and the pattern of present-day tectonic-plate velocities).

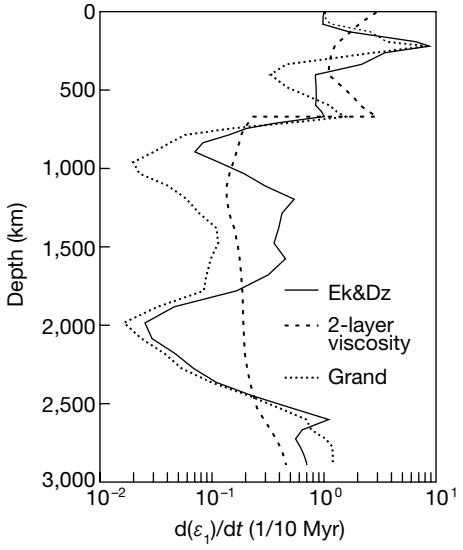

Figure 4 Flow-induced deformation or 'stretching' in the mantle. The maximum value, at any depth, of the principal rate of strain (defined as the largest positive eigenvalue of the strain-rate tensor $d\epsilon_{ij}/dt$) is plotted as a function of depth. The dashed line represents the extension rate associated with the flow field (Fig. 3a, c) predicted with the two-layer viscosity model (Fig. 2a) and the density contrasts (Fig. 1c) derived from the three-dimensional seismic model of Ekström and Dziewonski[28]. The solid line represents the extension rate associated with the flow field (Fig. 3b, d) predicted with the Occam-inferred viscosity (Fig. 2a) and the density contrasts (Fig. 1c) derived from the Ekström and Dziewonski seismic model. The dotted line is the flow-induced extension rate obtained on the basis of the density contrasts (Fig. 1c) derived from the three-dimensional seismic model of Grand[13] and the corresponding Occam-inferred viscosity (Fig. 2a).

Indeed, an idealized two-layer representation of the mantle viscosity, commonly employed in the numerical simulations, provides a poor fit to the convection data (see Table 2) and is unable to produce complex depth-dependent mixing behaviour (Fig. 4).

The flow calculations illustrated in Fig. 3 show that the transition from a domed pattern of flow in the deep lower mantle to a pattern of individual upwellings in the mid-mantle, long suggested by tomographic models[9–13], may largely be explained in terms of viscosity stratification. We should not conclude, however, that viscosity alone is responsible for maintaining this doming and that compositional heterogeneity is thus unnecessary or absent in the deep lower mantle, especially in view of the observed anti-correlation of seismic shear and bulk-sound velocity anomalies in this region[20–23]. We next consider the implications of this anti-correlation.

Thermochemical heterogeneity in the deep lower mantle

Recent tomographic studies which yield a joint description of seismic shear and bulk-sound velocity anomalies[20,22] provide a new means for determining the relative contributions of thermal and compositional heterogeneity in the deep mantle. Using Table 1, we have derived the following expressions (see Methods) relating specific linear combinations of shear and bulk-sound velocity anomalies to effective measures of compositional and thermal heterogeneity at 2,740 km depth:

$$-2.3\delta\ln V_s + 23.0\delta\ln V_\phi = \delta X_{Pv} - 3.2\delta X_{Fe} \equiv \delta X_{eff} \quad (1)$$

$$(-15.5\delta\ln V_s + 14.5\delta\ln V_\phi)\times 10^3 = \delta T + 1086\delta X_{Fe} \equiv \delta T_{eff} \quad (2)$$

The effective compositional heterogeneity δX_{eff} reflects lateral variations in the content of silica (via changes in the molar fraction of perovskite, δX_{Pv}) and iron (δX_{Fe}). The effective thermal heterogeneity δT_{eff} is a combined representation of temperature perturbations and lateral variations in iron content; however, we demonstrate below (see also Methods) that δT_{eff} provides an accurate approximation for δT.

Estimates of the effective compositional and thermal anomalies at 2,740 km depth, derived from two independent models of

© 2001 Macmillan Magazines Ltd

combined shear and bulk sound velocity anomalies[20,22], are in excellent agreement (Fig. 5). These estimates clearly show that δX_{eff} is strongly correlated to bulk-sound velocity anomalies while the shear velocity anomalies are strongly anti-correlated to δT_{eff}. We have tested the robustness of these results by considering a third, recently derived seismic model of lower mantle structure[48], and we obtained maps of δX_{eff} and δT_{eff} (see Supplementary Information) which also agree well with those shown here.

Our estimates of δX_{eff} and δT_{eff} were based on results from seismology and mineral physics. To extend this analysis to estimate lateral variations of iron or perovskite content requires an independent constraint on density perturbations. Using Table 1, we can write (see Methods):

$$\delta X_{\mathrm{Fe}} = 2.9(\delta \ln \rho + 10^{-5}\delta T_{\mathrm{eff}}) - 0.013\delta X_{\mathrm{eff}} \qquad (3)$$

We will use the $\delta\ln\rho$ field determined from geodynamic modelling (Fig. 1c). This approach assumes that the estimated amplitude of $\delta\ln\rho$ variations is independent of the initial $\delta\ln V_s$ model adopted in the geodynamic modelling and that a radially dependent scaling from shear velocity to density is reasonably valid. The former is demonstrated in Fig. 1c, while the latter is supported by the fits to the geodynamic observables[5].

The maps of iron heterogeneity (Fig. 6c–f) all show peak anomalies of about $\delta X_{\mathrm{Fe}} = \pm 0.01$, equivalent to about 10% variation relative to the bulk average value for X_{Fe}. Following equation (2), this level of iron heterogeneity suggests an error of the order of 10 K in using δT_{eff} as a proxy for δT. Accordingly, we have estimated the effective temperature derivative of shear velocity, $(\mathrm{d}\ln V_s/\mathrm{d}T)_{\mathrm{eff}}$, which provides an optimal fit (in the least-squares sense) between the δT_{eff} and $\delta\ln V_s$ fields in Fig. 5. Our estimate of $-5.2\times10^{-5}\,\mathrm{K}^{-1}$ is smaller than the purely thermal value of $-7.1\times10^{-5}\,\mathrm{K}^{-1}$ (Table 1) because of the 'masking' effect of compositional heterogeneity.

Lateral variations in chemical composition represented by δX_{Fe}

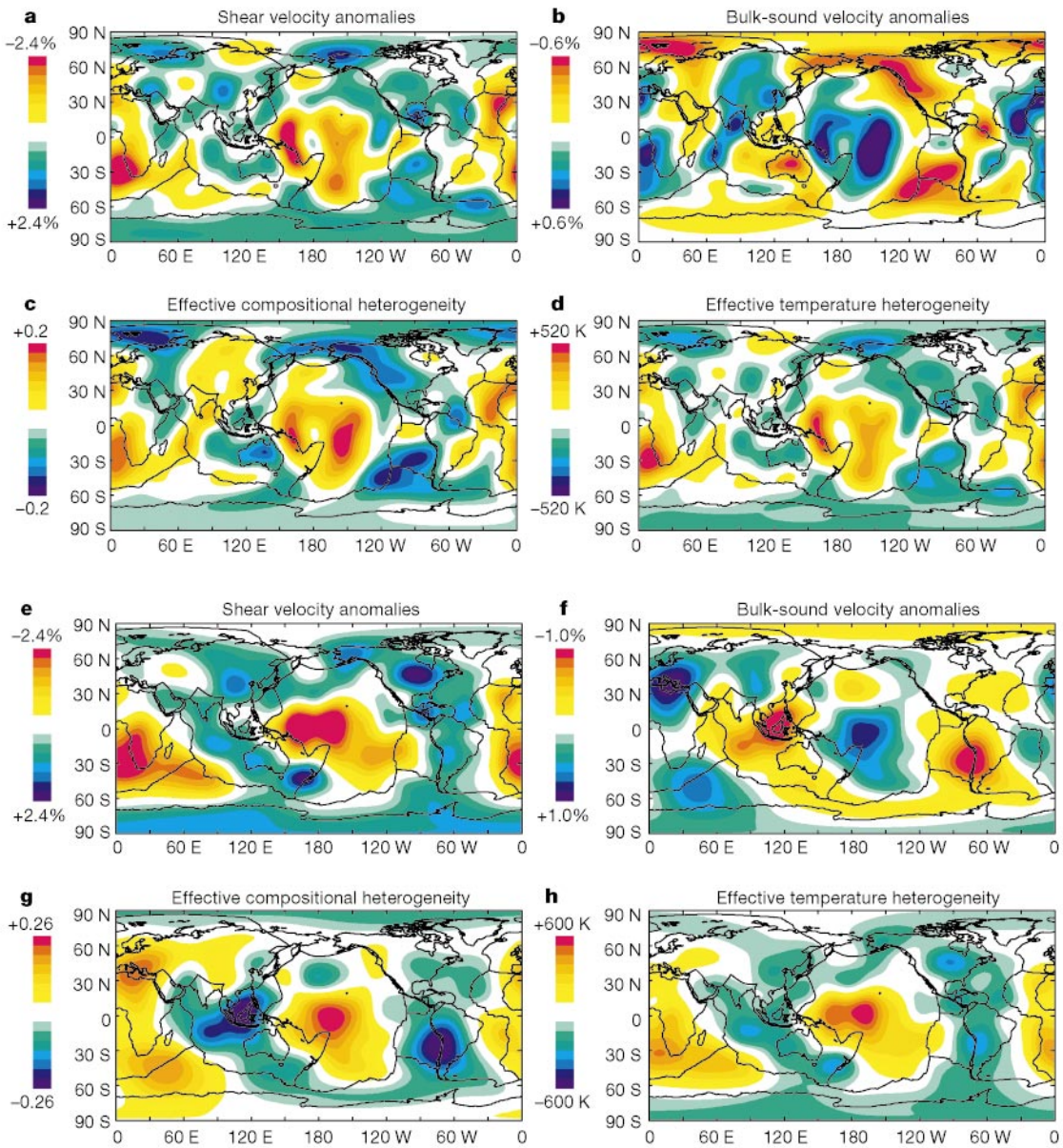

Figure 5 Compositional and thermal heterogeneity at 2,740 km depth. **a**, Relative shear velocity anomalies $\delta V_s/V_s$ obtained from the Su and Dziewonski (Su&Dz) tomography model[20]. **b**, Relative bulk-sound velocity anomalies $\delta V_\phi/V_\phi$ from Su&Dz. **c**, The effective compositional heterogeneity δX_{eff} derived from the Su&Dz model using equation (1). **d**, The effective temperature perturbations δT_{eff} derived from the Su&Dz model using equation (2). The maps in **e**–**h** are analogous to the maps in **a**–**d**, respectively, except that in **e**–**h** we used the joint description of shear and bulk-sound velocity anomalies given by the tomography model of Masters *et al.*[22]

© 2001 Macmillan Magazines Ltd

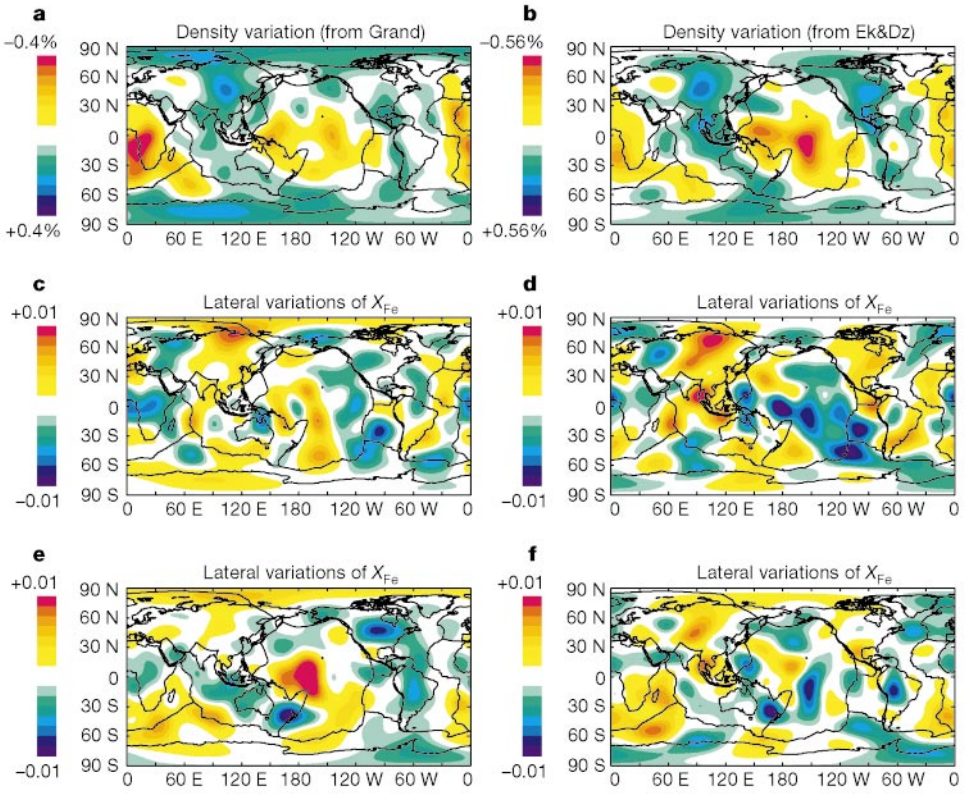

Figure 6 Lateral variations in iron content at 2,740 km depth. The inferences of iron heterogeneity δX_{Fe} shown here are obtained on the basis of equation (3), using δX_{eff} and δT_{eff} in Fig. 5, and using the geodynamic inferences of density heterogeneity (Fig. 1c). The density heterogeneity maps **a** and **b**, were obtained from the tomography models of Grand[13] and Ekström and Dziewonski[28], respectively. In maps **c** and **e** are shown δX_{Fe} derived using the tomography models of Su and Dziewonski[20] (Su&Dz) and Masters *et al.*[22] (Scripps), respectively. In both cases the inferences of δX_{Fe} are based on the density heterogeneity to map **a**. In maps **d** and **f** are shown δX_{Fe} again derived using the Su&Dz and Scripps models, respectively, except that the density heterogeneity in map **b** is used.

(Fig. 6) produce associated density anomalies $(\delta \ln\rho)_{chem}$ that may oppose the pure thermal buoyancy $(\delta \ln\rho)_{th} = -\alpha\delta T$, where α is the thermal expansivity. The magnitude of this effect, quantified by the buoyancy ratio $R_\rho = -(\delta \ln\rho)_{chem}/(\delta \ln\rho)_{th}$, has been shown to exert fundamental control on the dynamics of convection in the lower mantle[15–17]. Using Figs 5 and 6 we calculated $(\delta \ln\rho)_{chem}$ and $(\delta \ln\rho)_{th}$ and then determined the buoyancy ratio R_ρ which provides an optimal fit between these two fields. For δX_{Fe} extracted from density anomalies derived from the Grand (Fig. 6a) and Ek&Dz (Fig. 6b) models we find approximate values for R_ρ of 0.3 and 0.1, respectively. These values are significantly smaller than the $R_\rho \approx 1$ which has been assumed in some numerical convection simulations[15,16], but they fall well within the parameter space associated with a distinct doming regime[17] for mantle convection.

In view of the large-amplitude lateral temperature contrasts we have inferred (Fig. 5d, h), a legitimate concern is the extent to which the associated lateral variations in viscosity may affect our interpretations based on a purely depth-dependent effective viscosity. We constructed a tomography-based mantle flow model which explicitly incorporates lateral viscosity variations at all depths, using a previously developed theoretical formulation[30]. The lateral viscosity heterogeneity we employ is derived using a homologous temperature scaling[31], yielding large-scale lateral variations which range over three orders of magnitude (see Supplementary Information). These lateral variations are comparable to the magnitude of the geodynamically inferred radial variations in viscosity (Fig. 2a). We find that the relative mean-square amplitude of the differences between predictions (for example, geoid anomalies, CMB topography) obtained with and without these lateral viscosity variations (see Supplementary Information) are not more than 10%, which is smaller than the current misfit to the convection-related data (Table 1). Thus the error which may be incurred by neglecting large-scale lateral viscosity heterogeneity is no larger than other uncertainties (for example, in the seismic tomographic models) inherent in the current tomography-based viscous-flow models.

We conclude from Fig. 6a, b that the combined thermal and chemical anomalies in the deep lower mantle are such that the hot mega-plumes (Fig. 5d, h) below the central Pacific, and below Africa, are positively buoyant and hence actively ascending. Our analysis indicates that their buoyancy is partially reduced by the presence of compositional heterogeneity (Fig. 6c–f), but the temperature effect is dominant. This result is in agreement with other recent investigations of the dynamics of these deep mantle plumes. For example, the uplift rates for southern Africa, estimated from geological data, suggest[49] that the deep-mantle African mega-plume must be positively buoyant in order to match the estimated time-dependent uplift during the Cenozoic era. The global geophysical constraints on the dynamics of these deep mantle mega-plumes thus seem to suggest consistently that these structures are buoyant and that their dynamics (Fig. 3) are strongly controlled by the high-viscosity region that we infer near 2,000 km depth. □

Methods

Construction of Table 1

The (anharmonic or purely elastic) temperature (T) derivatives were estimated on the basis of the equation-of-state modelling by Jackson[50], in particular the results (see Fig. 9c, d in ref. 50) concerning the temperature-dependence of density ρ and seismic parameter $\phi = V_\phi^2$, where V_ϕ is the bulk-sound velocity. For example, at pressure $P = 127$ GPa (2,740 km depth), a variation in potential temperature from $T_0 = 1{,}600$ K to $T_0 = 1{,}850$ K results in a 0.44% change in density. Because $\partial/\partial T \approx (T_0/T)\partial/\partial T_0$, as shown by Stacey[51], we estimated a thermal expansion coefficient $\alpha = 1 \times 10^{-5}\,\mathrm{K}^{-1}$. From the variation in ϕ determined by Jackson[50], we similarly obtained the logarithmic T-derivative of the bulk-sound velocity $\partial \ln V_\phi/\partial T$. To determine the T-derivative of seismic shear (V_s)

© 2001 Macmillan Magazines Ltd

and compressional velocity (V_p) we also required an estimate of the T-dependence of the elastic shear modulus μ. Using Stacey's[51] derivation for $\Gamma = -(\alpha\mu)^{-1}(\partial\mu/\partial T)_p$ (see equations (70)–(71) in ref. 51) we determined the T-derivatives of V_s and V_p.

Following Karato[29], we estimated the anelastic contribution to the T-derivatives by assuming a frequency-independent seismic attenuation, or Q factor: $(\partial \ln V_s/\partial T)_{anelastic} = -(\pi Q_\mu T)^{-1}(H/RT)$, where H is the (P-dependent) activation enthalpy. Using Weertman's rule, the activation enthalpy is related to melting temperature T_m using $H = gRT_m$, where R is the gas constant and g is estimated to be approximately 30, a value appropriate for magnesium silicates such as olivine[52]. For T_m we employed a recent laboratory estimate of the deep-mantle pyrolite solidus[53]. We also used $Q_\mu = 312$, according to PREM[24], and assumed that all attenuation occurs in shear deformation (that is, the bulk attenuation is assumed to be zero).

The compositional derivatives represent the changes in density and seismic velocity due to changes in the molar fraction of iron, $X_{Fe} = [FeO]/([FeO] + [MgO])$, and the molar fraction of perovskite, $X_{Pv} = [Pv]/([Pv] + [Mw])$, where Mw denotes magnesio-wüstite. The molar fraction of perovskite is equivalent to the silica ratio: $X_{Pv} = [SiO_2]/([FeO]+[MgO])$. Variations in density and bulk-sound velocity caused by variations in iron content are (to a very good approximation) independent of pressure, and hence depth (see Fig. 9e, f in ref. 50). We thus calculated the derivatives $\partial/\partial X_{Fe}$ from standard temperature and pressure laboratory data (for example, Table 1 in ref. 50). The derivatives with respect to perovskite or silica content, $\partial/\partial X_{Pv}$, do vary with depth and we estimated the derivatives for ρ and V_ϕ directly from results in Fig. 9e, f in ref. 50. Derivatives of the shear modulus with respect to X_{Pv} are derived on the basis of the expression $\mu = (\mu/\kappa)_0[\kappa - \kappa'_\infty P]$ obtained by Stacey[54]. This relation allows us to express the compositional derivatives of shear modulus in terms of the derivatives of the bulk modulus, thereby yielding values for $\partial \ln V_s/\partial X_{Pv}$ and $\partial \ln V_p/\partial X_{Pv}$.

Estimating thermal and compositional variations

We begin with the following general expressions, based on the parameters in Table 1:

$$\delta \ln V_s = \frac{\partial \ln V_s}{\partial T}\delta T + \frac{\partial \ln V_s}{\partial X_{Pv}}\delta X_{Pv} + \frac{\partial \ln V_s}{\partial X_{Fe}}\delta X_{Fe}$$

$$\delta \ln V_\phi = \frac{\partial \ln V_\phi}{\partial T}\delta T + \frac{\partial \ln V_\phi}{\partial X_{Pv}}\delta X_{Pv} + \frac{\partial \ln V_\phi}{\partial X_{Fe}}\delta X_{Fe}$$

$$\delta \ln \rho = \frac{\partial \ln \rho}{\partial T}\delta T + \frac{\partial \ln \rho}{\partial X_{Pv}}\delta X_{Pv} + \frac{\partial \ln \rho}{\partial X_{Fe}}\delta X_{Fe}$$

where $\delta \ln V_s$, $\delta \ln V_\phi$ and $\delta \ln \rho$ are relative perturbations to shear velocity, bulk sound velocity, and density, respectively. We rewrite this system of equations in the following symbolic form:

$$\begin{pmatrix} \delta \ln V_s \\ \delta \ln V_\phi \\ \delta \ln \rho \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix} \begin{pmatrix} \delta T \\ \delta X_{Pv} \\ \delta X_{Fe} \end{pmatrix} \quad (4)$$

As joint seismic models for V_s and V_ϕ presently exist, we initially focus on the first two rows of equation (4). Eliminating δT in these rows yields:

$$\delta X_{eff} = \frac{c_{21}}{c_{12}c_{21} - c_{11}c_{22}}\delta \ln V_s - \frac{c_{11}}{c_{12}c_{21} - c_{11}c_{22}}\delta \ln V_\phi \quad (5)$$

where the 'effective' lateral variations in composition, δX_{eff}, are defined as:

$$\delta X_{eff} \equiv \delta X_{Pv} + \frac{c_{21}c_{13} - c_{11}c_{23}}{c_{12}c_{21} - c_{11}c_{22}}\delta X_{Fe} \quad (6)$$

We can furthermore show that these two rows can be expressed as:

$$\begin{pmatrix} \delta \ln V_s \\ \delta \ln V_\phi \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \begin{pmatrix} \delta T_{eff} \\ \delta X_{eff} \end{pmatrix}, \quad (7)$$

where we define the 'effective' temperature variation as:

$$\delta T_{eff} \equiv \delta T + \frac{c_{12}c_{23} - c_{13}c_{22}}{c_{12}c_{21} - c_{11}c_{22}}\delta X_{Fe}. \quad (8)$$

Inverting the matrix in equation (7) yields:

$$\begin{pmatrix} \delta T_{eff} \\ \delta X_{eff} \end{pmatrix} = \frac{1}{c_{12}c_{21} - c_{11}c_{22}} \begin{pmatrix} -c_{22} & c_{12} \\ c_{21} & -c_{11} \end{pmatrix} \begin{pmatrix} \delta \ln V_s \\ \delta \ln V_\phi \end{pmatrix}. \quad (9)$$

For a depth of 2,740 km we find, using Table 1:

$$\delta X_{eff} \equiv \delta X_{Pv} - 3.2\delta X_{Fe} = -2.3\delta \ln V_s + 23.0\delta \ln V_\phi \quad (10)$$

and

$$\delta T_{eff} \equiv \delta T + 1086\delta X_{Fe} = (-15.5\delta \ln V_s + 14.5\delta \ln V_\phi) \times 10^3 \quad (11)$$

The term $1086\delta X_{Fe}$ in equation (11) is the error introduced if we were to adopt δT_{eff} as an estimate for δT. The current best estimate[50] for the average mole fraction of iron in the lower mantle is $X_{Fe} = 0.11$, so we see from equation (11) that even if we assumed extreme lateral variations in iron content which approach 100% (that is, $\delta X_{Fe} = \pm 0.1$) the resulting contribution to δT_{eff} would be about 100 K, which is significantly smaller than the overall amplitude of δT_{eff} (Fig. 5d, h). The analysis described in the text (see Fig. 6) suggests peak variations of $\delta X_{Fe} \simeq 0.01$, and therefore the error in assuming $\delta T_{eff} \approx \delta T$ is only about 10 K. Variations in temperature and 'effective' composition in Fig. 5 are obtained by adopting the combination of seismically determined models for $\delta \ln V_s$ and $\delta \ln V_\phi$ specified in equations (10) and (11).

To separate the contributions to the effective variation δX_{eff} from δX_{Pv} and δX_{Fe} we incorporate density variations $\delta \ln \rho$ into the analysis. Using the last row of equation (4), together with the transformations (6) and (8), one can show that:

$$\delta \ln \rho = c_{31}\delta T_{eff} + c_{32}\delta X_{eff} - \frac{\Lambda}{c_{12}c_{21} - c_{11}c_{22}}\delta X_{Fe} \quad (12)$$

where Λ is the determinant of the matrix in equation (4). Given δT_{eff} and δX_{eff} from equation (9), we solve equation (12) for δX_{Fe}. Using the values in Table 1, the form of equation (12) appropriate for a layer at depth 2,740 km is:

$$\delta \ln \rho = -10^{-5}\delta T_{eff} + 0.0043\delta X_{eff} + 0.34\delta X_{Fe} \quad (13)$$

We have applied equation (13) to generate maps of δX_{Fe}, shown in Fig. 6.

The estimation of δT, δX_{Fe} and δX_{Pv} at a depth of 2,740 km is possible, in principle, because the matrix appearing in equation (4) is well conditioned. In practice, this exercise requires three independent inputs. The first two inputs are provided by the joint seismic models for $\delta \ln V_s$ and $\delta \ln V_\phi$. The final input, $\delta \ln \rho$, is determined on the basis of our geodynamic modelling (Fig. 1). Although we estimate the field $\delta \ln \rho$ by scaling $\delta \ln V_s$ to yield a fit to the geodynamic observables, we have shown that its amplitude is relatively insensitive to the starting model for $\delta \ln V_s$ (Fig. 1c). Thus, the geodynamic data provide the dominant constraint on the amplitude of $\delta \ln \rho$.

Received 10 April 2000; accepted 16 March 2001.

1. Marsh, J. G. *et al.* The GEM-T2 gravitational model. *J. Geophys. Res.* **95,** 22043–22071 (1990).
2. De Mets, C. R., Gordon, R. G., Argus, D. F. & Stein, S. Current plate motions. *Geophys. J. Int.* **101,** 425–478 (1990).
3. Gwinn, C. R., Herring, T. A. & Shapiro, I. I. Geodesy by radio interferometry: studies of the forced nutations of the Earth. 2. Interpretation. *J. Geophys. Res.* **91,** 4755–4765 (1986).
4. Mathews, P., Buffet, B. A. & Herring, T. A. What do nutations tell us about the Earth's interior? *Eos Trans. AGU* **80,** (Fall Meeting suppl.) 19 (1999).
5. Forte, A. M. & Woodward, R. L. Global 3D mantle structure and vertical mass and heat transfer across the mantle from joint inversions of seismic and geodynamic data. *J. Geophys. Res.* **102,** 17981–17994 (1997).
6. Mooney, W. D., Laske, G. & Masters, T. G. CRUST 5.1: A global crustal model at 5°×5°. *J. Geophys. Res.* **103,** 727–747 (1998).
7. Panasyuk, S. V. The effect of compressibility, phase transformations, and assumed density structure on mantle viscosity inferred from Earth's gravity field. PhD dissertation, Massachusetts Institute of Technology (1998).
8. Forte, A. M. & Perry, H. K. C. Geodynamic evidence for a chemically depleted continental tectosphere. *Science* **290,** 1940–1944 (2000).
9. Woodward, R. L. *et al.* in *Evolution of the Earth and Planets. Geophys. Monogr. Ser.* (eds Takahashi, E., Jeanloz, R. & Rubie, D.) Vol. 74, 89–109 (AGU, Washington, DC, 1993).
10. Masters, T. G. *et al.* A shear-velocity model of the mantle. *Phil. Trans. R. Soc. Lond. A* **354,** 1385–1411 (1996).
11. Li, X.-D. & Romanowicz, B. Global mantle shear-velocity model developed using nonlinear asymptotic coupling theory. *J. Geophys. Res.* **101,** 22245–22272 (1996).
12. van der Hilst, R. D., Widiyantoro, S. & Engdahl, E. R. Evidence for deep mantle circulation from global tomography. *Nature* **386,** 578–584 (1997).
13. Grand, S. P., van der Hilst, R. D. & Widiyantoro, S. Global seismic tomography: A snapshot of convection in the Earth. *Geol. Soc. Am. Today* **7,** 1–7 (1997).
14. van der Hilst, R. D. & Kárason, H. Compositional heterogeneity in the bottom 1000 km of Earth's mantle: Towards a hybrid convection model. *Science* **283,** 1885–1888 (1999).
15. Tackley, P. J. in *The Core-Mantle Boundary Region.* Geodyn. Ser. (ed. Gurnis, M. *et al.*) Vol. 28, 231–253 (AGU, Washington, DC, 1998).
16. Kellogg, L. H., Hager, B. H. & van der Hilst, R. D. Compositional stratification in the deep mantle. *Science* **283,** 1881–1884 (1999).
17. Davaille, A. Simultaneous generation of hotspots and superswells by convection in a heterogeneous planetary mantle. *Nature* **402,** 756–760 (1999).
18. Albarède, F. & van der Hilst, R. D. New mantle convection model may reconcile conflicting evidence. *Eos Trans. AGU* **80,** 535–539 (1999).
19. Roberston, G. S. & Woodhouse, J. H. Constraints on lower mantle physical properties from seismology and mineral physics. *Earth Planet. Sci. Lett.* **143,** 197–205 (1996).
20. Su, W. & Dziewonski, A. M. Simultaneous inversion for 3-D variations in shear and bulk velocity in the mantle. *Phys. Earth Planet. Inter.* **100,** 135–156 (1997).
21. Kennett, B. L. N., Widiyantoro, S. & van der Hilst, R. D. Joint seismic tomography for bulk-sound and shear wavespeed in the Earth's mantle. *J. Geophys. Res.* **103,** 12469–12493 (1998).
22. Masters, G., Laske, G., Bolton, H. & Dziewonski, A. M. in *Earth's Deep Interior: Mineral Physics and Tomography From the Atomic to the Global Scale.* Geophys. Monogr. Ser. (eds Karato, S.-I. *et al.*) Vol. 117, 63–87 (AGU, Washington, DC, 2000).
23. Ishii, M. & Tromp, J. Normal-mode and free-air gravity constraints on lateral variations in velocity and density of Earth's mantle. *Science* **285,** 1231–1236 (1999).
24. Dziewonski, A. M. & Anderson, D. L. Preliminary reference Earth model. *Phys. Earth Planet. Inter.* **25,** 297–356 (1981).
25. Sinelnikov, Y. D., Chen, G. & Liebermann, R. C. Elasticity of $CaTiO_3$-$CaSiO_3$ perovskites. *Phys. Chem. Mineral.* **25,** 515–521 (1998).
26. Zhang, J. & Weidner, D. J. Thermal equation of state of aluminum-enriched silicate perovskite. *Science* **284,** 782–784 (1999).

© 2001 Macmillan Magazines Ltd

27. McDonough, W. F. & Sun, S.-S. The composition of the Earth. *Chem. Geol.* **120,** 223–253 (1995).
28. Ekström, G. & Dziewonski, A. M. The unique anisotropy of the Pacific upper mantle. *Nature* **394,** 168–172 (1998).
29. Karato, S. Importance of anelasticity in the interpretation of seismic tomography. *Geophys. Res. Lett.* **20,** 1623–1626 (1993).
30. Forte, A. M. & Peltier, W. R. The kinematics and dynamics of poloidal-toroidal coupling in mantle flow: The importance of surface plates and lateral viscosity variations. *Adv. Geophys.* **36,** 1–119 (1994).
31. Karato, S.-I. & Li, P. Diffusion creep in perovskite: Implications for the rheology of the lower mantle. *Science* **255,** 1238–1240 (1992).
32. Constable, S. C., Parker, R. L. & Constable, C. G. Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data. *Geophysics* **52,** 289–300 (1987).
33. King, S. & Masters, T. G. An inversion for the radial viscosity structure using seismic tomography. *Geophys. Res. Lett.* **19,** 1551–1554 (1992).
34. Mitrovica, J. X. & Forte, A. M. The radial profile of mantle viscosity: Results from the joint inversion of convection and post-glacial rebound observables. *J. Geophys. Res.* **102,** 2751–2769 (1997).
35. Forte, A. M. in *Earth's Deep Interior: Mineral Physics and Tomography From the Atomic to the Global Scale.* Geophys. Monogr. Ser. (eds Karato, S.-I. *et al.*) Vol. 117, 3–36 (AGU, Washington, DC, 2000).
36. Durek, J. & Ekström, G. A radial model of anelasticity consistent with long-period surface-wave attenuation. *Bull. Seismol. Soc. Am.* **86,** 144–158 (1996).
37. Karato, S.-I. A dislocation model of seismic wave attenuation and micro-creep in the Earth. *Pure Appl. Geophys.* **153,** 239–256 (1998).
38. Hofmann, A. W. The gabbro-harzburgite connection in OIB sources. *Eos Trans. AGU* **80** (Fall Meeting suppl.) 1182–1183 (1999).
39. Ottino, J. M. *The Kinematics of Mixing: Stretching, Chaos, and Transport* Ch. 2 & 4 (Cambridge Univ. Press, Cambridge, 1989).
40. Kellogg, J. B. & O'Connell, R. J. The effects of toroidal motion and layered viscosity on mixing in three dimensions. *Eos Trans. AGU* **80** (Fall Meeting suppl.) 949 (1999).
41. Hager, B. H. Subducted slabs and the geoid: Constraints on mantle rheology and flow. *J. Geophys. Res.* **89,** 6003–6015 (1984).
42. Forte, A. M. & Peltier, W. R. Plate tectonics and aspherical Earth structure: The importance of poloidal-toroidal coupling. *J. Geophys. Res.* **92,** 3645–3679 (1987).
43. Ricard, Y. & Vigny, C. Mantle dynamics with induced plate tectonics. *J. Geophys. Res.* **94,** 17543–17559 (1989).
44. Gurnis, M. & Davies, G. F. The effect of depth-dependent viscosity on convective mixing of the mantle and the possible survival of primitive mantle. *Geophys. Res. Lett.* **13,** 541–544 (1986).
45. Manga, M. Mixing of heterogeneities in the mantle—Effects of viscosity differences. *Geophys. Res. Lett.* **23,** 403–406 (1996).
46. van Keken, P. & Ballentine, C. J. Whole-mantle versus layered mantle convection and the role of a high-viscosity lower mantle in terrestrial volatile evolution. *Earth Planet. Sci. Lett.* **156,** 19–32 (1998).
47. Bunge, H.-P. & Richards, M. A. The origin of large-scale structure in mantle convection: Effects of plate motions and viscosity stratification. *Geophys. Res. Lett.* **23,** 2987–2990 (1996).
48. Ritsema, J., van Heijst, H. J. & Woodhouse, J. H. Complex shear velocity structure imaged beneath Africa and Iceland. *Science* **286,** 1925–1928 (1999).
49. Gurnis, M., Mitrovica, J. X., Ritsema, J. & van Heijst, H. J. Constraining mantle density structure using geological evidence of surface uplift rates: The case of the African superplume. *Geochem. Geophys. Geosystems.* [online] **1** (2000).
50. Jackson, I. Elasticity, composition and temperature of the Earth's lower mantle: A reappraisal. *Geophys. J. Int.* **134,** 291–311 (1998).
51. Stacey, F. D. Thermoelasticity of a mineral composite and a reconsideration of lower mantle properties. *Phys. Earth Planet. Inter.* **106,** 219–236 (1998).
52. Jackson, I., Paterson, M. S. & Fitzgerald, J. D. Seismic wave attenuation in Åheim dunite: An experimental study. *Geophys. J. Int.* **108,** 517–534 (1992).
53. Zerr, A., Diegler, A. & Boehler, R. Solidus of Earth's deep mantle. *Science* **281,** 243–246 (1998).
54. Stacey, F. D. Theory of thermal and elastic properties of the lower mantle and core. *Phys. Earth Planet. Inter.* **89,** 219–245 (1995).

Acknowledgements

We thank I. Jackson for a review, for suggestions used in the Methods section, and for the detailed check of the derivatives in Table 1. A.M.F. also thanks A. Davaille for many helpful discussions. We acknowledge support from NSERC, the Canada Foundation for Innovation, the Ontario Innovation Trust, and the Canadian Institute for Advanced Research—Earth Systems Evolution Program.

Correspondence and requests for materials should be addressed to A.M.F. (e-mail: aforte@uwo.ca).

© 2001 Macmillan Magazines Ltd