A comparison of tomographic and geodynamic mantle models

Thorsten W. Becker and Lapo Boschi
Department of Earth and Planetary Sciences, Harvard University, 20 Oxford Street, Cambridge, Massachusetts 02138, USA (becker@eps.harvard.edu)

Abstract: We conduct a comprehensive and quantitative analysis of similarities and differences between recent seismic tomography models of the Earth’s mantle in an attempt to determine a benchmark for geodynamic interpretation. After a spherical harmonic expansion, we find the spectral power and radial correlation of each tomographic model as a function of depth and harmonic degree. We then calculate the correlation, at the same depths and degrees, between all possible pairs of models, to identify stable and model-dependent features (the former being usually of longer spatial wavelength than the latter). We can therefore evaluate the degree of robust structure that seismologists have mapped so far and proceed to calculate ad hoc mean reference models. Tomographic models are furthermore compared with two geodynamic subduction models that are based on plate motion reconstructions. We find systematically low intermediate-wavelength correlation between tomography and convective reconstruction models and suggest that the inadequate treatment of the details of slab advection is responsible. However, we confirm the presence of stable, slab-like fast anomalies in the mid-mantle whose geographic pattern naturally associates them with subduction. This finding, in addition to our analysis of heterogeneity spectra and the absence of strong minima in the radial correlation functions besides the one at \( \sim 700 \) km, supports the idea of whole mantle convection with slab penetration through the 660 km phase transition, possibly accompanied by a reorganization of flow.

Keywords: Tomography; seismic structure; subduction history; mantle convection; correlation analysis.

Index terms: Core and mantle; dynamics, convention currents and mantle plumes; dynamics of lithosphere and mantle general; Earth’s interior-composition and state.

Received April 10, 2001; Revised July 24, 2001; Accepted July 25, 2001; Published January 10, 2002.


1. Introduction

Over the last two decades, numerous tomographic models of the Earth’s interior have been derived from different types of seismological measurements and with different techniques [e.g., Dziewonski et al., 1977; Masters et al., 1982; Dziewonski, 1984; Inoue et al., 1990; van der Hilst et al., 1997]. These models can be seen as snapshots of the convecting mantle, thereby providing important constraints on the planet’s dynamics [e.g., Hager et al., 1985; Mitrovica and Forte, 1997]. Here we conduct a comparison between global three-dimensional (3-D) mantle models, derived from seismological data, and theoretical geodynamical models. In order
to proceed from mapping heterogeneity to testing geologically relevant hypotheses, such an undertaking is needed for a number of reasons:

1. Discrepancies between tomographic models often arise from differences in the modeling procedure [e.g., Boschi and Dziewonski, 1999]. Using a systematic computation of the correlation between different models, we attempt to distinguish stable features from those that depend on data selection and technical choices.

2. Three-dimensional models of compressional (P) and shear (S) wave velocity in the Earth’s mantle are derived from independent observations. The existence of a correlation between P and S velocity anomalies might mean that they have a common origin, generally believed to be thermal and associated with mantle convection; however, where uncorrelated P and S velocity heterogeneities are found, compositional heterogeneity can be invoked [e.g., Su and Dziewonski, 1997; Kennett et al., 1998]. It is therefore important to specifically measure the correlation between P and S models.

3. Several authors have attempted to reconstruct the convective flow of the mantle [e.g., Ricard et al., 1993; Lithgow-Bertelloni and Richards, 1995, 1998; Bunge et al., 1998; Bunge and Grand, 2000; Steinberger, 2000a, 2000b] and their results have been used to explain the current pattern of seismic heterogeneities. However, there is still some controversy on this subject; debated issues include the extent to which the phase transition at 660 km represents a barrier to mantle flow [e.g., van der Hilst et al., 1997; Puster and Jordan, 1997] and whether a layer or other large-scale structures exist in the deep mantle [e.g., van der Hilst and Kárason, 1999; Tackley, 2000]. To contribute to this discussion, we carry out a quantitative comparison of seismological versus geodynamic results with the same algorithm used to find the correlation between tomographic models.

[3] More than any previous comparative interpretation of global tomography [e.g., Grand et al., 1997; Masters et al., 1999] and geodynamics [e.g., Lithgow-Bertelloni and Richards, 1998; Bunge et al., 1998], our study includes a comprehensive, consistent, and quantitative analysis of all recently published models, allowing the reader to make an informed choice as to which features and geodynamic inferences can be considered robust. The additional online material, which can be found at http://www.geophysics.harvard.edu/geodyn/tomography, makes all correlation and power spectra plots as well as model data available in a convenient form.

[4] We follow Masters et al. [1999] and choose to measure the similarity between models in terms of the linear correlation between their spherical harmonic expansions. Only with a spectral analysis are we able to identify wavelength-dependent features and detect changes in the character of the spectrum of imaged heterogeneity (section 4.1.1). Our global measure of correlation suffers from certain drawbacks [e.g., Ray and Anderson, 1994], especially when narrow signals such as subducting slabs are studied. Wavelets might be better able to detect local features [e.g., Bergeron et al., 2000]; in section 4.3 we circumvent these problems with an additional spatial domain analysis.

[5] It is our finding that the correlation between modern global tomographic models is high at long wavelengths, even when P and S wave models are compared with each other; in particular, slab-like structures extending below the 660 km phase transition are a stable feature of all models, and no other radial discontinuity is required at larger depths. Tomographic models are less similar at shorter wavelengths, and on a global intermediate-wavelength scale, do not yet correlate well with the slab signal of geodynamic models.

2. Models

[6] Following is a brief description of all the models we study.
2.1. P Wave Tomography

[7] MIT model hwe97p is parameterized in terms of $2^\circ \times 2^\circ$ blocks of variable radial extent [van der Hilst et al., 1997] and is undefined in areas where the data coverage was considered inadequate (“gaps”).

[8] Model kh00p is derived similarly to hwe97p, but a coarser parameterization ($3^\circ \times 3^\circ$ blocks) and additional travel time data (from “core-phases”) has led to a model that is defined everywhere in the mantle [Kárason and van der Hilst, 2001].

[9] Model bdp98 denotes the Harvard equal area block model BDP98 ($5^\circ \times 5^\circ$ at the equator, constant radial extent of $\sim 200$ km) [Boschi and Dziewonski, 1999].

[10] Model bdp00 is an unpublished improvement of bdp98 based on further relocation efforts [see Antolik et al., 2001].

[11] All P wave models are based on body wave travel time measurements collected by the International Seismological Centre (ISC). ISC data can be improved by source relocation; this has been done by both the Harvard group [Su and Dziewonski, 1997] and with a different method by Engdahl et al. [1998]. Both MIT models were derived from Engdahl et al.’s [1998] data set.

2.2. S Wave Tomography

[12] Love and Rayleigh waves are mostly sensitive to anomalies in horizontally and vertically polarized shear velocity, $v_{sh}$ and $v_{sr}$, respectively, and only marginally affected by perturbations in $P$ velocity, $v_p$ [e.g., Anderson and Dziewonski, 1982]. Observations of surface waves are therefore usually taken into account in deriving S models, while $v_p$ heterogeneities are only constrained by the travel times of body waves whose ray geometry is generally nearly vertical within the upper mantle. As a result, the data coverage for $v_s$ in the upper mantle is much more uniform than for $v_p$.

[13] Model grand is the equal-area block model as of Grand’s ftp-site in fall 2000 [see Grand et al., 1997], distributed on a $2^\circ \times 2^\circ$ grid. The model was derived from a combination of body and surface wave measurements with a two step process [Grand, 1994]: first, observations are explained in terms of upper and lowermost mantle structure only. Second, the authors invert the residual travel-time anomalies to find velocity heterogeneities in the rest of the mantle.

[14] Model ngrand is an updated version of grand, as of Grand’s ftp-site in June 2001. The inversion that led to ngrand was damped more strongly in the upper mantle than that of grand (S. Grand, personal communication, 2001); as a result, the new model is different from grand mostly in amplitude, rather than pattern, of heterogeneity (see 4.1.2).

[15] Here s20rts denotes the Caltech model S20RTS, parameterized horizontally in terms of spherical harmonics up to degree $\ell_{\text{max}} = 20$ and radially with a set of cubic splines [Ritsema and van Heijst, 2000]. The model is derived from a data set that, in addition to body and Rayleigh wave measurements, includes observations of normal mode splitting functions.

[16] Here saw24b16 describes the Berkeley $v_{sh}$ model SAW24B16 [Mégnin and Romanowicz, 2000], derived by fitting body and surface wave transverse-component waveforms. Parameterized with spherical harmonics ($\ell_{\text{max}} = 24$) and cubic splines.

[17] Model sb4l18 denotes the Scripps model SB4L18, from observations of body, Love, and Rayleigh waves, and normal modes [Masters et al., 1999]. Parameterized in terms of equal-area blocks ($4^\circ \times 4^\circ$ at the equator) with 18 radial layers.
Model s20a stands for the Harvard model S20A from observations of body, Love, and Rayleigh waves [Ekström and Dziewonski, 1998]. The \( v_{sh} \) and \( v_{sv} \) anomalies were treated as independent free parameters; \( v_s \) is subsequently estimated from their Voigt average. Spherical harmonics representation (\( \ell_{\max} = 20 \)) is used horizontally; radially, upper and lower mantle are parameterized separately with two sets of Chebyshev polynomials.

Here s362d1 denotes the Harvard model S362D1, derived with a procedure analogous to s20a (including the discontinuity at 660 km) but described by a cubic spline parameterization both horizontally and vertically [Gu et al., 2001]. Lateral resolution is equivalent to \( \ell_{\max} \sim 18 \).

In addition, we will also use lower resolution joint inversions for \( v_s \) and \( v_p \) in section 4.2.3. We consider MK12WM13 [Su and Dziewonski, 1997] (spherical harmonics, \( \ell_{\max} = 12 \), Chebyshev polynomials with depth, \( v_p \) and \( v_s \) anomalies denoted by mk12wm13p and mk12wm13s, respectively), SB10L18 by Masters et al. [2000] (similar to sb4118 but \( 10^\circ \times 10^\circ \) blocks, pb10l18 and sb10l18), and Harvard model SPRD6 from normal mode splitting coefficients [Ishii and Tromp, 2001] (spherical harmonics, \( \ell_{\max} = 6 \), sprd6p and sprd6s).

2.3. Mean Tomography Models

While most tomographic models present significant discrepancies, they agree on certain, long-wavelength patterns. Efforts to define a 3-D reference Earth model from an inversion of geophysical observables are currently under way (see, e.g., the Reference Earth Model (REM) web site, http://mahi.ucsd.edu/Gabi/rem.html). A REM would be a starting point for higher resolution models and provide the much needed benchmark to evaluate geodynamical hypotheses. Here we adopt a pragmatic approach and calculate two mean models by taking a weighted average of several models, assuming that such “stacking” will enhance the “signal-to-noise” ratio (see appendix B). The result is a largest common denominator model which we intend to update as tomographic research progresses. Model pmean is our mean \( P \) wave model based on bdp00 and kh00p. Model smean denotes our mean \( S \) wave model, based on ngrand, s20rts, and sb4118.

2.4. Geodynamic Models

We compare the previous velocity models with an upper mantle slab model, two geodynamic models that account for inferred past subduction, and, in a statistical sense, with a 3-D thermal convection calculation.

Model rum is our spherical harmonic expansion of slabs in the upper mantle obtained from the RUM seismicity contours [Gudmundsson and Sambridge, 1998], which are in turn based on the Engdahl et al. [1998] catalog. We integrate along the RUM contours at each layer using them as \( \delta \)-functions such that the effective width of the slabs is determined by \( \ell_{\max} \) and the \( \cos^2 \)-taper that we apply for \( \ell > 0.75 \ell_{\max} \).

Model lrr98d denotes the density model by Lithgow-Bertelloni and Richards [1998] which is given on spherical harmonics laterally (\( \ell_{\max} = 25 \)) and layers with depth. Model lrr98d is based on “slablets,” i.e., negative buoyancy anomalies that sink at different speeds in the upper and lower mantle after starting at estimated past trench locations which are based on Mesozoic and Cenozoic plate reconstructions [Lithgow-Bertelloni et al., 1993; Ricard et al., 1993]. The sinking rate was adjusted to fit geopotential fields, tomography, and plate motions.

Model stb00d describes the density model by Steinberger [2000b], which is given on spherical harmonics (\( \ell_{\max} = 31 \)) and radial layers. Model stb00d is also based on past plate motions and subduction; Lithgow-Bertelloni et al.’s [1993] sets of plate boundaries were,
however, interpolated at 2 Ma intervals, while Lithgow-Bertelloni and Richards [1998] held boundaries fixed during individual plate-tectonic stages. Model stb00d is furthermore different from lrr98d in that it allows for lateral advective flow of slabs in addition to plate motion and slab buoyancy. Model stb00d can be considered more realistic than lrr98d with respect to the treatment of convective flow.

Model zmg00t is a temperature snapshot from a 3-D spherical convection calculation by Zhong et al. [2000]. We use the residual (RMS) temperature from their case 7 at time 9.25 × 10^{-4}. Case 7 is an incompressible, temperature and depth-dependent viscosity calculation without phase transitions that allowed for plate-like flow through the inclusion of fixed weak zones [Zhong et al., 2000, Plates 2a and 2b]. Assuming constant thermal expansivity, α, variations in nondimensional temperature, T, relate to density, ρ, as d ln ρ = −αΔT. With ΔT = 1800°K for the nonadiabatic mantle gradient and α = 1.4 × 10^{-5} K^{-1}, we scale with αΔT ≈ 0.025.

Current and past plate motions are some of the best indicators for the style of convective flow in the mantle. The derived slab sinker trajectories and density distributions of models such as stb00d should thus be among the best constrained geodynamic models. However, given the discrepancies that we observe between lrr98d and stb00d (see section 4.3), we will not attempt to explore thermal convection models [e.g., Tackley, 1998; Bunge et al., 1998; Zhong et al., 2000] in greater detail at this point but only complement power spectra of tomography with one representative pattern from zmg00t (section 4.1.1).

3. Measures of Model Correlation

Before carrying out any comparisons, we must find a consistent description. As tomo-
where $\int_{\Omega} d\Omega$ indicates integration over the unit sphere. Our experiments with different numerical integration methods and smoothly interpolated grids with spacings between 0.5° and 2° indicate that the spurious power which is introduced by the expansion of block models should be smaller than ~2.5% for degrees $\ell \leq 20$.

[31] Some tomography models are undefined in areas where the ray coverage was considered inadequate. The effect of these gaps is most extreme for hwe97p, where the areal coverage varies between ~40% at the surface and ~95% at the core mantle boundary (CMB). Since the size of gaps can be large, we set the velocity perturbations to zero in those regions before computing (3) and (4). This choice has the same effect as the imposition of a strong norm-minimization constraint in a least squares fit of \{a_{im}, b_{im}\} (i.e., underestimation of RMS heterogeneity). While this approach is not ideal, we think that a more elaborate treatment is unnecessary since the gapless kh00p has replaced hwe97p, and we will not base any of our conclusions upon hwe97p. Gaps in grand and ngrand occupy a fraction smaller than 1.1% and 0.3% at all depths, respectively; in these cases we have interpolated using the “surface” algorithm [Wessel and Smith, 1991] before expanding the fields.

3.2. Power as a Function of Wavelength

[32] We define the spectral power of the field $\delta v(\theta, \phi)$ per degree $\ell$ and unit area as [e.g., Dahlen and Tromp, 1998, B.8]:

$$\sigma_{\ell}^2 = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} (a_{im}^2 + b_{im}^2).$$

such that a $\delta$-function results in a flat spectrum (depth dependence will be assumed implicitly).

The root-mean-square (RMS) power of the expansion is then

$$\delta_{\text{RMS}} \approx \frac{1}{\sqrt{4\pi}} \sigma_{\text{RMS}} = \sqrt{\frac{1}{4\pi} \sum_{\ell=1}^{\ell_{\text{max}}} (2\ell + 1) \sigma_{\ell}^2}$$

since

$$\int_{\Omega} [\delta v(\theta, \phi)]^2 d\Omega = \sum_{\ell=0}^{\ell_{\text{max}}} \sum_{m=-\ell}^{\ell} (a_{im}^2 + b_{im}^2).$$

We note that other definitions of spectral power can be found in the literature [e.g., Su and Dziewonski, 1991]; our choice (normalization by $2\ell + 1$ in (5)) emphasizes the wavelength dependence of heterogeneity.

3.3. Cross-Model Correlation

[33] To evaluate the similarity of any two models at a given depth, we find the correlation $r_{\ell}$ between the spherical harmonic expansions $\{a_{im}, b_{im}\}$ and $\{c_{im}, d_{im}\}$ of the corresponding fields at each wavelength $\ell$,

$$r_{\ell} = \frac{\sum_{m=-\ell}^{\ell} (a_{im}c_{im} + b_{im}d_{im})}{\sqrt{\sum_{m=-\ell}^{\ell} (a_{im}^2 + b_{im}^2) \sum_{m=-\ell}^{\ell} (c_{im}^2 + d_{im}^2)}}.$$

[34] The total correlation up to $\ell_{\text{max}}$ is given by

$$r_{\text{tot}} = \frac{\sum_{\ell=1}^{\ell_{\text{max}}} \sum_{m=-\ell}^{\ell} (a_{im}c_{im} + b_{im}d_{im})}{\sigma_{\text{RMS}}^{\{a_{im}, b_{im}\}} \sigma_{\text{RMS}}^{\{c_{im}, d_{im}\}}}.$$

where $\sigma_{\text{RMS}}^{\{a_{im}, b_{im}\}}$ and $\sigma_{\text{RMS}}^{\{c_{im}, d_{im}\}}$ denote the RMS sums for each model. We have verified that the spectral formula (9) yields results that are similar to those found from applications of the common spatial description (see appendix C3).

[35] After applying (8) and (9) to find the correlation between every combination of two models at $M$ evenly spaced $z_j$ where the models have been interpolated, we take the weighted average ($0 \leq z_j \leq 2871$ km):

$$\langle r_{\text{tot}} \rangle = \frac{\sum_{j=1}^{M} w_j r_{\text{tot}}(z_j)}{\sum_{j=1}^{M} w_j} \text{ with } w_j = (a-z_j)^2,$$

where $a$ is the Earth’s radius; each layer enters the average according to its volume. While $\langle r \rangle$ is
thus biased by shallow structure, our approach seems appropriate for a global measure of similarity and we find that cross-model \( r \) does not depend strongly on this weighting.

[36] Equations (8)–(10) are estimates of the similarity of two models in terms of global heterogeneity patterns; differences in amplitude do not affect \( r' \) or \( r_{\text{tot}} \max \) (see appendix A for the statistical interpretation of \( r \)). We find and compare \( r_{\text{tot}} \max \) and \( r_{\text{tot}} \) to evaluate separately the correlation between the long and intermediate spatial wavelength components of the models.

### 3.4. Radial Parameterization

[37] We generally attempt to remain close to the original radial parameterization of all models: if models were described in terms of discrete layers, we first evaluate them at the original mid-layer depths, \( z_i \). Likewise, we evaluate spline models at, roughly, the original spline knots and models that are parameterized with Chebyshev polynomials at constantly spaced (\( \sim 150 \) km) intervals. Subsequently, if the value of the model is needed at any other depth \( z_j \), we find it by linear interpolation.

[38] To explore how results are affected by different radial parameterizations, we will additionally show results from models that were reparameterized in terms of normalized Chebyshev polynomials of order \( k \) [e.g., Su and Dziewonski, 1997]. We obtain the Chebyshev coefficients by a least squares fit of the original expansion coefficients at the \( z_i \) using a combination of norm and roughness damping. This leads to a smoothed but faithful representation of the original models (see appendix C4): with \( k_{\max} = 20 \), we have a resolution of \( \sim 150 \) km and achieve variance reductions typically better than 95%.

### 3.5. Radial Correlation

[39] Following Puster and Jordan [1997], we calculate the radial correlation matrix \( r(z_1, z_2) \) between the lateral structure of the same model at any two depths \( z_1 \) and \( z_2 \). (Correlation \( r(z_1, z_2) \) is simply the value of \( r_{\text{tot}} \max \) found from the expansions \( \{a_{\text{im}}, b_{\text{im}}\} \) and \( \{a_{\text{im}}, b_{\text{im}}\} \) at the depths \( z_1 \) and \( z_2 \), respectively.) Two derived, closely related measures of radial coherence as a function of \( z \) are (1) the correlation coefficient \( r(z - \Delta z, z + \Delta z) \) for a fixed depth bracket \( \Delta z \), and (2) the value of \( \Delta z \) associated with a contour of constant \( r(z - \Delta z, z + \Delta z) \).

[40] As outlined by Puster and Jordan [1997], radial correlation functions can be interpreted as a measure of mass flux between different depth ranges in the convecting mantle. In practice, we compute \( r(z_1, z_2) \) for all \( z_i \) and interpolate on a smooth field at \( \sim 25 \) km grid spacing. For models whose spline parameterization density varies with depth, it is difficult to obtain adequate depth spacings; our discrete \( z_i \) estimate of the correlation functions suffers therefore from some oversampling, especially for \( \Delta z > 20 \) rts. The resulting oscillations are, however, easily detected and can be avoided if we choose a Chebyshev parameterization with depth (see appendix C4).

### 4. Results

#### 4.1. Analysis of Individual Models

[41] We study the spectral signal, \( \delta v_{\text{RMS}} \), and radial correlation of each model before comparing models to one another. For consistency, all values of geodynamic models are scaled by

\[
\Lambda = \frac{d \ln v_i}{d \ln \rho} = \frac{\delta v_i}{\delta \rho} = 3.6, \quad (11)
\]

a weighted radial average (we neglect depth-dependence of \( \Lambda \) for simplicity) of Karato’s [1993] profile.

#### 4.1.1. Power spectra variation with depth

[42] After computing the spectral power per degree \( \sigma_i^2(5) \), we normalize it by its maximum
at each depth in order to emphasize the dominant wavelengths and denote the resulting quantity \( \sigma^2_\ell \). The first moment of a weighted sum of \( \sigma^2_\ell \),

\[
M(z) = \frac{\sum_{\ell=1}^{\ell_{\text{max}}} (2\ell + 1) \sigma^2_\ell(z)}{\sum_{\ell=1}^{\ell_{\text{max}}} (2\ell + 1) \sigma^2_\ell}
\]

is a measure of the \( z \) dependence of the strongest wavelengths, i.e., the “color” of the heterogeneity spectrum. (Absolute values of \( M(z) \) are only meaningful when \( \sigma^2_\ell \to 0 \) as \( \ell \to \ell_{\text{max}} \); most tomographic models approximately satisfy this condition.)

[43] Figure 1 shows the normalized spectrum (see Figure C2 for \( \sigma^2_\ell \)) and \( \delta v_{\text{RMS}} \) (equation (6)) for a selection of tomographic and geodynamic models. Most tomographic models are dominated by long wavelengths (“red,” \( \ell \leq 5 \)) at all depths [e.g., Su and Dziewonski, 1992]. These low degree patterns roughly correspond to the continent/ocean function at the surface, the circum-Pacific subduction signal in the mid-mantle, and the “mega-plumes” toward the CMB. In other words, mantle convection appears to be organized by plate-scale flow with length scales as observed at the surface [e.g., Davies, 1988].

[44] \( P \) models kh00p and bdp00 have a different spectral character in the uppermost mantle (where \( M(z) \) and \( \delta v_{\text{RMS}} \) from bdp00 are smoother functions of depth) and for 1600–2400 km depths. In both cases, there is a shift in kh00p’s spectrum from degree 2 to degree 1. Furthermore, bdp00 indicates a stronger change in spectral character at \( z \sim 800 \) km than kh00p. The \( \delta v_{\text{RMS}} \) of bdp00 is overall larger in the upper mantle than that of bdp98: most likely a result of Antolik et al.’s [2001] source relocation with subsequent improvement of the Harvard data (see the additional online material). The kh00p and bdp00 models are consistent in that the absolute \( \ell = 2 \) power has a local minimum at \( \sim 2000 \) km depth (Figure C2).

[45] The \( \sigma^2_\ell \)-spectrum of s20rts is dominated by \( \ell = 2 \) everywhere in the lower mantle but weaker at intermediate \( \ell \) than for the \( P \) models. Model s20rts’ \( \delta v_{\text{RMS}} \), minimum at \( \sim 1600 \) km depth, is focused in the uppermost mantle where \( \sigma^2_\ell \) is strongest in degrees \( \ell = 1 \) and \( \ell = 5 \), as expected in an \( S \) model with a well-constrained ocean-continent signal [Su and Dziewonski, 1991]. The spectrum of s20rts becomes continuously redder with increasing depth starting from \( \sim 1500 \) km. This is a common feature for \( S \) models while \( P \) models typically have intermediate wavelength power and a corresponding local maximum in \( M(z) \) at \( \sim 2000 \) km. We find a minimum in absolute \( \sigma^2_\ell \) for s20rts at \( \sim 1600 \) km (Figure C2), possibly related to the fading slab signal and to the uppermost boundary of the large slow anomalies that reach down to the CMB [Dziewonski, 1984]. We also observe that the spectral power of s20rts is consistently higher at even rather than odd \( \ell \), up to \( \ell \sim 12 \); one reason for this could be the \( \ell \)-dependence of the sensitivity of normal mode splitting functions (used, among other data, to derive s20rts) to the Earth’s structure. However, s20a and the subduction signal in stb00d and lnr98d indicate similar streaks in the power spectrum, while otherwise similar models, such as sb4118, have no such property (see the additional online material).

[46] The spectrum and \( \delta v_{\text{RMS}} \) of s362d1 show the effect of a 660 km deep discontinuity in the radial parameterization: the inversion shifts heterogeneity to the upper mantle and the sub 660 km spectrum gets whiter. Gu et al. [2001] find that such variations are not as pronounced when the parameterization discontinuity is placed at other depths. A change in spectral character below 660 km can also be found, to a lesser extent, in other models (e.g., bdp00); it might indicate the effect of a viscosity increase in the lower mantle [e.g., Mitrovica and Forte, 1997], leading to reor-
organization of flow and transient slab hold-up (see section 4.1.3).

The $\sigma^2_{\ell}$ of stb00d is strongest for $\ell \leq 3$ but differs from tomography in that it has relatively high power over a broad range of wavelengths as expected from the narrow slab signal. Besides a trend toward a bluer spectrum below 660 km, there is no clear tendency of $M(z)$ of the subduction signal to vary with depth but, as noted above, we find that even $\ell$ is stronger in the middle mantle than odd $\ell$ power for stb00d and lrr98d (see Figures C2 and C6 and the
additional online material). Thermal convection model zmg00t is similar to tomography regarding the low degree pattern of heterogeneity. Indeed, case 7 is Zhong et al.’s [2000] preferred model since the inclusion of “plates” lead to the characteristic red signal of seismological models in a temperature dependent viscosity calculation. Model zmg00t furthermore resembles tomography in that the signal is bluer in the middle mantle than toward the thermal boundary layers (TBLs, $z \lesssim 500$ km and at the CMB) where $\delta v_{\text{RMS}}$ variations are strongest.

Figure 2 shows depth averaged power spectra, $\sqrt{\langle \sigma_l^2 \rangle}$. Figure 2 shows depth averaged $\sqrt{\langle \sigma_l^2 \rangle}$ for a selection of models. In general, $S$ models are characterized by stronger heterogeneity than $P$ models [e.g., Anderson, 1987; Karato, 1993] (also see section 4.2.3). As noted above, spectral power for tomography is concentrated at low degrees ($\ell = 1, 2$, with a local maximum at $\ell = 5$) and rapidly decays when $\ell \gtrsim 5$ [e.g., Su and Dziewonski, 1991]. Model s362d1 is an outlier, in that it shows the most rapid decrease of $\delta v_f^2$ for $\ell \gtrsim 12$, mostly owing to a relatively weak high frequency signal in the upper mantle (Figure C2). The geodynamic models stb00d,
lrr98d, rum, and zmg00t have a stronger high frequency character than that mapped by tomography.

4.1.2. RMS heterogeneity

The aforementioned concentration of heterogeneity toward the boundary layers of the mantle and the global minima at \( \sim 1600 \text{ km} \) (\( S \) models) or \( \sim 2000 \text{ km} \) (\( P \) models) depth are common features of \( \delta v_{\text{RMS}} \) as a function of \( z \) (Figure 3). In the case of tomographic models, \( \delta v_{\text{RMS}} \) is a smooth function of depth; exceptions are grand and ngrand (whose high \( \delta v_{\text{RMS}} \) focusing is the result of the inversion procedure), as well as s20a and s362d1 (based on a discontinuous radial parameterization (section 2.2)). Model ngrand is significantly closer in \( \delta v_{\text{RMS}} \) to the other \( S \) models than its ancestor grand because of modified damping (see section 2.2). As the cross-model correlation between grand and ngrand shows (see the additional online material), patterns were only slightly affected by this modification (\( r_{20} = 0.9 \)).

The \( \delta v_{\text{RMS}} \) based on subduction models does not agree well with tomography in the

**Figure 3.** The \( \delta v_{\text{RMS}} \) versus \( z \), \( \ell_{\text{max}} = 31 \). Symbols at \( z < 0 \) denote depth averaged \( \langle \delta v_{\text{RMS}} \rangle \).
upper mantle but shows a consistent increase in heterogeneity below \( \sim 1500 \) km. We furthermore find broad agreement between the thermal convection snapshot zmg00t and \( \delta V_{RMS} \) from tomography. However, seismological models indicate stronger variations of \( \delta V_{RMS} \) with \( z \). Also, the upper boundary layer structure is generally more pronounced than the deep one for tomography. Reasons for these discrepancies are the continent/ocean differences and tectosphere (not included in any of the geodynamic models), the fact that \( \delta V_{RMS} \to 0 \) at the surface and at the CMB are boundary conditions of zmg00t, and that Zhong et al.’s [2000] Rayleigh number is smaller than Earth’s by a factor of \( \sim 10 \). We should therefore expect that the TBL thickness is overpredicted and \( T_{RMS} \) is underpredicted for zmg00t. Variations in \( \Lambda \) or \( \alpha \) (see section 4.1) and other effects such as compressibility will also affect the depth-dependence of \( \delta V_{RMS} \) and the spectrum as predicted by geodynamics [e.g., Tackley, 1996]. However, a detailed discussion of dynamic convection models is beyond the scope of this paper.

4.1.3. Radial correlation

Figure 4 shows three measures of the radial correlation of tomographic models: the corre-
lation matrix, $\Delta z$ at constant \( r(z - \Delta z, z + \Delta z) \), and \( r(z - \Delta z, z + \Delta z) \) at constant \( \Delta z \) (see section 3.5). All estimates in Figure 4 are based on \( \ell_{\text{max}} = 20 \) expansions (see section C4 for \( \ell_{\text{max}} \)-dependence). Model bdp00 (Figure 4a) is characterized by a local minimum in \( \Delta z \) at \( \sim 600 \) km; correlation is then relatively high in the midmantle and decreases toward the CMB. For kh00p (Figure 4b), local minima in radial correlation are found at \( \sim 400 \) km, \( \sim 800 \) km, \( \sim 1700 \) km, and \( \sim 2300 \) km. Our results are similar (but not identical) to those of van der Hilst and Káraison [1999] who pointed out the decrease in correlation at \( \sim 1700 \) km. Since estimates of radial correlation are parameterization dependent, we have repeated our calculation for kh00p using the original blocks or a radial Chebyshev reparameterization. The results of this exercise, summarized in section C4, are consistent with Figure 4b.

The radial correlation of $S$ models increases with increasing depth (e.g., sb4l18 in Figure 4c, with local minima in $\Delta z$ at $\sim 300$ km and $\sim 600$ km). This effect is explained by the general tendency of $S$ models to become redder with increasing depth (section 4.1.1): since long-wavelength features are better correlated, we then expect a more homogeneous radial correlation in the lower mantle. Indeed, if we damp out high frequency structure in the $P$ models for the lower mantle (typically concentrated at $z \sim 2000$ km), the resulting plots of the radial correlation function resemble those of $S$ wave models.

Figure 4. (continued)
We show in Figures 5a and 6a the radial correlation $r$ at $\Delta z = 200$ km of a selection of models as it results from our initial calculations with a discrete radial sampling. Then we repeat the calculations after reparameterizing the models over a radial Chebyshev polynomial basis and give the results in Figures 5b and 6b. Consistent features of the $P$ models in Figure 5b are a broad global minimum in correlation for $400 \text{ km} \leq z \leq 700$ km (possibly with a local maximum at $\sim 660$ km), an increase in $r$ toward $1500$ km, and a decrease to a second minimum at $1700 \text{ km} \leq z \leq 2400$ km. The latter feature is less pronounced in pb10l18 (see the additional online material). As anticipated above, $S$ models (Figure 6) are generally characterized by an increase in $r$ with increasing $z$. Other notable features are local maxima at $\sim 660$ km, artificial oscillations of the $r$ versus $z$ curve obtained from s20trs (explained in 3.5) and the anomalously large (parameterization related, section 2.2) excursions of s362d1 at 660 km.

Radial correlation estimates vary strongly with model parameterization [e.g., Ritzwoller and Lavely, 1995] and power spectra appear to be a more robust estimate of structural change than radial correlation functions. However, we
find some indication for low radial correlation in $P$ and $S$ models at $\sim 750$ km, and in $P$ models at $\sim 2000$ km, previously associated with reorganization of flow and possible deep mantle structure, respectively [van der Hilst and Kárason, 1999]. The work of Puster and Jordan [1997] and the stb00d (whole mantle flow) derived correlation profiles show that local minima in correlation are not necessarily indicative of a layered style of convection. Changes in slab morphology [e.g., van der Hilst et al., 1997; van der Hilst and Kárason, 1999] and the general flow pattern due to the phase transition with a viscosity increase at 660 km [e.g., Mitrovica and Forte, 1997] are therefore likely explanations for the first minimum in correlation. Especially transient slab flattening and possible segmentation can be expected to yield apparent structural changes below 660 km as indicated by some of the power spectra that were discussed in section 4.1.1.

The findings that $z \sim 2000$ km is a global minimum of $\delta v_{RMS}$, that the radial correlation of $S$ models shows no clear decrease at these depths, and the apparent absence of large scatterers in the lower mantle [Castle and van der Hilst, 2000], make the existence of a strong...

---

**Figure 5.** Radial correlation $r(z - \Delta z, z + \Delta z)$ at fixed $\Delta z = 200$ km for (a) $P$ wave tomography with layers and (b) on radial Chebyshev parameterization; stb00d is shown for comparison.
global structural change at ∼2000 km seem unlikely at this point. However, the local minima in absolute $\ell = 2$ power that we found at ∼1700 km (section 4.1.1) are consistent with a fade-out of the slab signal at these depths [e.g., van der Hilst and Káraason, 1999; Káraason and van der Hilst, 2000] and local structural change [e.g., Saltzer et al., 2001] cannot be ruled out.

4.2. Cross-Model Comparisons

We now quantify similarities and discrepancies between models, focusing on a representative subset. The additional online material includes correlation plots for all possible pairs of models from section 2 with the exception of zmg00t which represents the current state of Earth’s mantle only in a statistical sense.

4.2.1. P wave models

Figure 7 shows $r^\ell$ between bdp00 and kh00p, an example of the good correlation that generally characterizes P models. The pattern and $\delta v_{RMS}$ of bdp00 and kh00p are mostly consistent throughout the lower mantle and up to $\ell = 20$; significant deviations are found in the uppermost mantle (where $\delta v_{RMS}$ of kh00p is weaker) and at the CMB (where the inclusion of core phases has enhanced the $\delta v_{RMS}$ of
kh00p [Karrison and van der Hilst, 2001]). Other local minima in \( r^\text{tot}_8 \) and \( r^\text{tot}_{20} \) are at \( \sim 300 \) km, 700 km, and \( \sim 1900 \) km. All three depths show a long wavelength breakdown in correlation (\( r^\text{tot}_8 \to r^\text{tot}_{20} \)), especially at \( \ell = 4 \). The average correlation between bdp00 and kh00p is high (\( \langle r^8 \rangle = 0.71 \)), to be compared with (\( r^8 \)) = 0.85 for bdp00–bdp98, and (\( r^8 \)) = 0.69 for kh00p–hwe97p (but see section 3.1).

4.2.2. S wave models

Figure 8 shows \( r^\ell \) for s20rts and sb4l18. Especially at low harmonic degrees, the two models are very consistent, more so than the \( P \) models in Figure 7. However, at \( \ell \gtrsim 12 \), correlation degrades such that \( \langle r^20 \rangle \) is slightly lower than for bdp00–kh00p. These intermediate wavelength discrepancies that we find for most \( S \) models are likely due to the greater variety in input data, while \( P \) models are inverted from very similar data sets (section 2). The \( r^\text{tot}_8 \) and \( r^\text{tot}_{20} \) for Figure 8 have a global maximum near the surface, decrease toward \( \sim 1500 \) km and, as the low-degree portion of the spectrum becomes more important, grow monotonically thereafter while \( r^\text{tot}_{20} \to r^\text{tot}_8 \).

Figure 6. Radial correlation \( r(z - \Delta z, z + \Delta z) \) at \( \Delta z = 200 \) km for (a) \( S \) wave tomography with layers and (b) on radial Chebyshev parameterization. The continuous Chebyshev parameterization is clearly not suited to represent discontinuities in s362d1.
The $r^j$ behaves as in Figure 8 for most combinations of $S$ models (see the additional online material and section 4.2.5). Exceptions are s20a and s362d1: these models correlate well with other $S$ models in the upper mantle. Because of their inherent radial discontinuity, however, they are different from other models for $700 \text{ km} \lesssim z \lesssim 1500 \text{ km}$. Yet, between s362d1 and smean, $r^2$ and $r^3$ at $z = 800 \text{ km}$ are still $\gtrsim 0.68$ and $\langle r_8 \rangle = 0.7$. As noted by Gu et al. [2001], the data fit of their 660 km discontinuity model was not significantly better than that of their continuous parameterization inversion. This means that discontinuous descriptions of mantle structure are consistent with, but not necessarily required by, the data.

4.2.3. $\delta v_s$ versus $\delta v_p$

Figure 9 is representative of comparisons between models of $v_s$ and $v_p$. We find that those are generally not as well correlated with each other as models of the same kind. This result can partly be explained by the systematic differences in data distribution and sensitivity, especially in the upper mantle (see section 2.2). The $r_{20}^{tot}$ in Figure 9 shows local minima at $\sim 700 \text{ km}$ and at $\sim 1800 \text{ km}$, but only the former is accompanied by a broad $r_8^{tot} \rightarrow r_{20}^{tot}$...
cross-wavelength breakdown (not as pronounced for comparisons with kh00p).

Figure 10a shows $r_{20}^{\text{tot}}$ for a combination of $S$ models and bdp00 and kh00p; Figure 10b allows a comparison of our mean models, long-wavelength joint inversions, and the normal mode model SPRD6 (see sections 2.2 and 2.3). We observe that correlation between $\delta v_s$ and $\delta v_p$ is low for all models between $\sim 300$ km and $\sim 700$ km; another minimum is found at $\sim 2000$ km for some models. The $r_{20}^{\text{tot}}$ between most $P$ and $S$ models has a local maximum at $\sim 2500$ km and then decreases again toward the CMB, hinting at compositional heterogeneity at the bottom of the mantle.

Comparisons of the $\delta v_s-\delta v_p$ components of models derived from joint $v_s-v_p$ inversions (Figure 10b) can lead to quite different considerations, depending on the model [Masters et al., 2000]: in the case of MK12WM13, the $\delta v_s-\delta v_p$ correlation has a pronounced minimum at $\sim 1000$ km and is smaller at low $\ell$ than at...
The correlations of $\mu$ versus $p_{\text{mean}}$ and $\delta v_x$ versus $\delta v_p$ components of SB10L18, in contrast, are consistent, with local minima at at $\sim 400$ km and $\sim 2000$ km (as before, only the former shows $r_{\text{tot}}^8 \rightarrow r_{\text{tot}}^{20}$). Absolute correlations are, however, higher for SB10L18.

We next determine

$$R = \frac{\Delta \ln v_x}{\Delta \ln v_p} = \frac{\delta v_x}{\delta v_p},$$

first from the RMS heterogeneity ratio of the models (solid lines in Figures 10a and 10b, mean $\langle R \rangle$ indicated on the $R$ axis at the CMB). Second, we calculate $R$ from a linear regression of the expansion coefficients of the models at each depth ($\ell_{\text{max}} = 20$); an estimate whose reliability can be judged from the corresponding $r_{\text{tot}}^{20}$. In the absence of information about model uncertainties, $R$ will vary depending on the assumed standard deviations, $\Sigma$, of each model. We therefore show a range of best-fit $R$ values from an iterative linear regression [e.g., Press et al., 1993, p. 666] where $\Sigma$ of the $S$ model, $\Sigma_{\delta v_x}$, is assumed to be twice that of the $P$ model (leading to lower estimates for $R$) or vice versa, where $\Sigma_{\delta v_p} = 2\Sigma_{\delta v_x}$ (leading to

![Figure 8](Image)
higher estimates). The corresponding \( \langle R \rangle \) values are indicated with different size symbols on the upper \( R \) axis in Figure 10a and 10b. For comparison, we add the expected variation of \( R \) based on mineral physics if heterogeneity were purely thermal in origin, one estimate (black line) from Karato [1993] and the other (gray inverted triangles) from ab initio calculations for MgSiO\(_3\) perovskite by Oganov et al. [2001].

[64] Measured \( R \) in Figure 10a typically increases for \( z \geq 400 \) km, and is between 1 and 4 for RMS estimates. We note that \( R \geq 2.5 \) implies a negative correlation between bulk sound and shear wave velocity [e.g., Masters et al., 2000, equation (4)] which is usually interpreted as an indication of compositional heterogeneity. Mean \( \langle R \rangle \) values based on \( \delta v_{\text{RMS}} \) are, indeed, in general larger than the mineral physics estimates for temperature and pressure effects in Figure 10a, in agreement with findings from direct inversions for \( R \) [e.g., Robertson and Woodhouse, 1996]. However, lower-end linear regression estimates and RMS \( \langle R \rangle \) from sprd6 and SB10L18 (Figure 10b) fall close to the mineral physics values [see, also, Masters et al., 2000] and recent estimates by Karato and Karki [2001] are
larger than Karato’s [1993] R values by ~0.4. Hence the significance of the observed mid-mantle deviations from a homogeneous composition trend remains to be determined.

In synthesis, judging from \( r \) and \( R \) for \( \delta V_p - \delta V_s \), regions of global deviation from the predictions of mineralogy for a chemically homogeneous mantle are likely to be limited to the tectosphere, the transition zone (where data coverage for \( P \) is inferior), the CMB region, and, less pronounced and with all the caveats from section 4.1.3, the depth range at ~2000 km.

4.2.4 Geodynamic models

The correlation between two subduction models, sb00d and lrr98d, is shown in Figure 11. They are most consistent in the upper mantle where slab locations are well con-
strained by seismicity ($r_8$) with rum is 0.63 and 0.61 for lr98d and stb00d, respectively). Moreover, advection is only active in stb00d once slabs sink below 380 km. Steinberger’s [2000b] and Lithgow-Bertelloni and Richards’ [1998] approach should therefore lead to very similar results for shallow structure. Once lateral advection becomes more important with depth, correlation decreases with $z$ up to $\sim 1000$ km and stays low for $\ell \geq 5$ throughout the lower mantle. The finding that the geodynamic models do not agree well with each other globally ($r_8 = 0.44$) implies that differences in methodology and the effect of lateral advection on the narrow slab features affect the global measure $r$ significantly (see section 4.3). This observation can also guide us as to how to judge the correlation of slab and tomography models.

Figure 10. (continued)

[67] The $r^\ell$ between geodynamic and tomographic models is in fact high for long wavelengths ($\ell \lesssim 3$) [Ricard et al., 1993; Lithgow-Bertelloni and Richards, 1998] throughout the mantle, but $r^\ell$ is low for higher $\ell$. We find that global $\langle r_8 \rangle$ is small on average ($\langle r_8 \rangle \sim 0.3$) and there is no clear depth-dependence of $r_8^{\text{tot}}$, besides that correlation is usually largest below
the transition zone at \( \sim 800 \) km and smallest at \( z \sim 1500 \) km depth (see the additional online material). The best \( \langle r_8 \rangle \)-correlation between subduction and tomography models is \( \langle r_8 \rangle = 0.33 \) between lrr98d and smean (see Figure 23 and section 4.2.5), but lrr98d and stb00d lead to very similar \( \langle r_{20} \rangle \)-results \( (\langle r_{20} \rangle = 0.18 \) and \( \langle r_{20} \rangle = 0.21 \) with smean, respectively).

4.2.5. Summary of average cross-model correlations

We summarize our findings in Figure 12 which shows the total depth-averaged correlations \( \langle r_8 \rangle \) and the total cross-model correlations at 600 km, 1400 km, and 2750 km, for a selection of \( P \), \( S \), and geodynamic models. (For \( \langle r_{20} \rangle \), see the additional online material.) In general, agreement between tomography is poor at \( z \sim 600 \) km (where \( S \) models correlate fairly well with subduction models) and increases with larger depths. We find that our models smean and pmean correlate better \( (\langle r_8 \rangle = 0.71) \) with each other than any other \( \delta v_r \)-\( \delta v_p \) combination in Figure 12a and achieve the highest \( \langle r_8 \rangle \) with slab models lrr98d or stb00d. For the \( S \) models that were not used for the construction of smean (see appendix B), the discontinuity models s20a and s362d1 are

Figure 11. Cross-model correlation for stb00d (solid \( \delta v_{\text{RMS}} \)-line) and lrr98d (dashed \( \delta v_{\text{RMS}} \)-line). Compare with Figure C6 and see Figure 7 for description.
found to be more similar to smean than saw24b16 which is $v_{s,h}$ only. We cannot identify particular depth ranges where anisotropy in $S$ wave propagation might cause deviations. However, we note that correlation between smean and saw24b16 has a local minimum in the lower mantle (see the additional online material) where Mégnin and Romanowicz [1999] argue that their approach has led to improved resolution over other $S$ models.

From Figure 12 we can also see that subduction models correlate better with $S$ than with $P$ models. On average, lrr98d is about as similar to tomographic models (mean $\langle r_8 \rangle / r_{20}$) from Figure 12 is 0.24/0.13) as stb00d (mean
\(\frac{r_8}{r_{20}}\) is 0.23/0.16, even though stb00d is a more sophisticated model in terms of the treatment of mantle flow. We will discuss these findings further in section 4.3.

In a final step toward cross-model similarity synthesis, Figure 13 explores how similar \(P\) (bdp00 and kh00p), \(S\) (s362d1, s20a, sb418, saw24bl6, s20rts, and ngrand), and geodynamic models (stb00d and lrr98d) are on average. The highest correlations are generally associated with the longer wavelength component (\(S\) models in particular) and, for tomography, larger depths. Subduction models become progressively uncorrelated with increasing \(z\).

### 4.3. Comparison of Tomography and Geodynamic Models

We have seen that the geodynamic models stb00d and lrr98d do not correlate well globally with tomography for \(\ell \gtrsim 5\). We now argue that this does not imply that there is no slab signal in the mantle but that our understanding of flow modeling has to be improved.

#### 4.3.1. Subduction versus fast anomalies only

Slabs in the mantle will be colder than their surroundings and thus show up as fast anomalies only. Therefore we set to zero all
slow anomalies present in tomographic models, reexpand the models, thus “clipped,” and recompute their correlation with the geodynamical models. This procedure leads to some increase in the correlation between lrr98d, stb00d, rum, and tomography models, particularly in the upper and middle mantle. However, the depth-averaged correlation is still poor ($r_{8}^{tot} \leq 0.4$). We thus infer that the low global correlation between tomography and subduction models cannot be explained as an effect of the lack of independent hot upwellings in subduction models. An alternative explanation is that our knowledge of mantle viscosity and of the velocity at which slabs sink is still incomplete. We will now analyze this possibility with additional calculations.

We smooth the depth-dependence of stb00d and lrr98d by taking, for each expansion, a sliding boxcar average with depth extent 

![Figure 14](image)

**Figure 14.** The $r_{8}^{tot}$ of (a) stb00d and (b) lrr98d with $P$ and $S$ models for radial boxcar averaging. Line thickness indicates $\delta z_{box}$; the vertical gray line denotes the 99% confidence level.
$\delta z_{\text{box}}$ up to 600 km (mean of $[z - \delta z_{\text{box}}/2; z + \delta z_{\text{box}}/2]$) and then find, again, the $r^\text{tot}_8$ correlation between stb00d (Figure 14a) or lrr98d (Figure 14b) and several clipped ($\delta v > 0$) tomography models. The results, summarized in Figure 14, indicate that correlation becomes higher with increasing $\delta z_{\text{box}}$. Our radial smoothing filter therefore limits problems associated with the short radial correlation length $\Delta z$ of subduction models. We find that correlations are higher with $S$ models, and the behavior of $r^\text{tot}_8$ as a function of depth is different depending on the model (stb00d or lrr98d, see section 4.3.2). In general, $r^\text{tot}_8$ is negative in the shallow mantle ($z \lesssim 200$ km) owing to the tectosphere and continent or ocean differences imaged by tomography but not included in the subduction models. The $r^\text{tot}_8$ then increases to its maximum in the middle mantle ($z \sim 750$ km) where slabs might be more easily detected by tomography since an increase in viscosity at 660 km could lead to a broadening of the subduction signal in
the lower mantle. Correlation then decreases toward the CMB.

4.3.2. Effect of slablet sinking speed

We simulate the effect that a wrong estimate of the sinking velocity of slabs would have on subduction models: neglecting lateral motion, upper/lower mantle differences, and upwellings, this can be done by “stretching” the models by a factor \( f \), i.e., mapping the depth interval \([z_a, z_b]\) onto \([fz_a, fz_b]\) (Figure 15). For stb00d, \( \langle r_8 \rangle \) with clipped tomography can be improved by up to 43% (Figure 15a) with the best \( f_{\text{opt}} \sim 1.75 \), corresponding to a higher sinking speed. As Figure C6 shows, lrr98d and the modified stb00d are more similar \( \langle r_8 \rangle = 0.66 \) in this case, too. However, individual \( \langle r_8 \rangle \) correlations of the stretched stb00d with tomography \( \langle r_8 \rangle \leq 0.42 \) are still not much better than for lrr98d, for which we find \( f_{\text{opt}} \sim 0.75 \) with a smaller relative increase of \( \langle r_8 \rangle \) of \( \leq 20\% \) \( \langle r_8 \rangle \leq 0.38 \), most pronounced for \( P \) models (Figure 15b). This implies that Lithgow-Bertelloni and Richards’ [1998] optimization with respect to the sinking velocity was basically successful; Steinberger’s [2000b] more realistic subduction calculation with fewer free parameters and lateral advection
did not produce a better model when global correlation with tomography is used as a measure.

[75] In synthesis, we find that geodynamic models show some resemblance to tomographic models on a global scale. The correspondence is best at ~800 km depth but not as good as between tomographic models. The reason for this could be that the subduction process is not yet modeled correctly and the lateral and depth offsets that might result from slab interaction with 660 km [e.g., Zhong and Gurnis, 1995; Christensen, 1996] could explain some of the weak correlation at intermediate $\ell$.

Especially transient slab flattening or segmentation (already invoked to account for changes in spectral characteristics) will degrade the global correlation as the comparison between the two, fairly similar, subduction models shows.

[76] Inaccuracies in the advection process are, of course, not the only possible explanation for the poor global agreement between subduction models and tomography. One question that needs to be considered is the precision to which we can infer past plate motions and how subsequent modifications in the reconstructions will map themselves into the large scale density
field. However, other effects such as (possibly local) compositional heterogeneity might also be important.

4.3.3. Midmantle slabs

[77] Figures 16 and 17, including clipped tomographic and geodynamic maps above (550 km) and below 660 km (850 km), substantiate our finding that subduction models are most similar to tomographic ones at \( z \sim 800 \) km. Some of the seismically active trenches (e.g., Japan, Kurile, Solomon, and Peru-Chile) are clearly present at \( z = 550 \) km in all models (Figure 16). (We show average fields from 500 to 600 km.) However, only ngrand includes a distinct image of the Tonga and Indonesia slabs; strong fast anomalies underneath North America appear only in the \( P \) models. Also, older slab material might have accumulated in the Mediterranean [e.g., Wortel and Spakman, 2000]: we can find such a signal most clearly in kh00p.

[78] Figure 16 also shows that at 550 km depth, all the \( P \) models include other, probably not slab-related, fast anomalies beneath the cratonic regions of Canada, Africa, Eurasia, and Australia. Since the tectosphere is believed to terminate at \( z < 550 \) km and \( S \) models (well constrained in the upper mantle, see section 2.2) are not anomalously fast in the same regions, we suggest that these \( \delta v_p > 0 \) features are partly due to a fictitious effect (“smearing”) of the nonuniform coverage achieved by \( P \) body wave data. At 850 km depth (Figure 17), all models are remarkably consistent under the Americas, Indonesia, eastern Phillipines, and Tonga, with a robust slab signal below 660 km. With the exception of the mantle below Africa, s362d1 is the only model to include strong fast anomalies that are clearly not subduction-related.

[79] Figure 18 shows the correlation, at 550 and 850 km, between clipped tomography models and, alternatively, rum (Figure 18a), stb00d (Figure 18b), and lr98d (Figure 18c). The correlation values that we find are statistically significant at the 99% level for most models at 850 km and at least for \( S \) models at 550 km depth. Again, we attribute the low correlation obtained from \( P \) models in the upper mantle to the nonuniform ray coverage inherent to seismic observations of \( P \) velocity (see sections 4.2.5 and 4.3.1). This is an important consideration if \( P \) models are to be interpreted geodynamically. In most cases, the highest values of tomography versus geodynamic correlation are found at \( \sim 850 \) km (see 4.3.2), confirming that most fast seismic anomalies found in the middle-mantle are slab-related.

4.3.4. Slow velocity anomalies

[80] Identifying convective features in tomographic models is more difficult for slow than for fast wave speed anomalies. While slabs are of great importance for mantle convection [e.g., Davies and Richards, 1992], we also expect to see some trace of the upwellings be it in the form of large-scale swells or narrow plumes. It is not clear, however, if global tomography is able to image the latter features at this stage [e.g., Ritsema et al., 1999].

[81] We clip tomographic models as above, now eliminating all positive velocity anomalies, and then compare the resulting “slow-anomaly-only” models with \( \delta \)-function expansions of hot spot distributions from Steinberger [2000a] and 3SMAC [Natatak and Ricard, 1996]. The resulting values of \( r^\text{st}_\delta \) are given in appendix C7 (Figure C7). As noted by Ray and Anderson [1994], there is no clear correlation between surface hot spot locations and tomographically mapped anomalies (the correlation is statistically significant only near the surface and at \( z \sim \)
However, this does not imply that hot spots are not plume-related since plume conduits are likely to be deflected during their ascent [e.g., Steinberger, 2000a] and larger scale downwellings might affect the plume source locations in the deep mantle [e.g., Richards et al., 1988; Steinberger and O’Connell, 1998].

Figure 16. Comparison of tomographic models (positive $\delta v$ only, “+”), old cratons from 3SMAC [Nataf and Ricard, 1996], and geodynamic models with $l_{\max} = 20$ at $z = 550$ km (average from 500 km to 600 km depth, $\delta z_{\text{box}} = 100$ km, Robinson projection). Red plate boundaries are from NUVEL-1 [DeMets et al., 1990].
Certain slow anomalies that are not ridge-related appear systematically in all tomographic models at $z = 300$ km (see Figure C8): around the Afar region and Iceland (possibly related to plumes), in the southwestern Pacific (possibly related to the superswell), and in the central Pacific region. The latter anomaly is of complex structure, widespread, and
lies in a region where seismic observations are affected by strong radial anisotropy [Ekström and Dziewonski, 1998]. At $z = 1500$ km (Figure C9), all models are dominated by two large anomalies centered on southwest Africa and the central Pacific (thus characterized by a strong $\ell = 2$ component); of these, at least the African one stretches down to the CMB in more than one model [e.g., Dziewonski, 1984]. At 2500 km depth (Figure C10), we find that the African anomaly is accompanied by one underneath the Antarctic plate at 60°E, 60°S. The Pacific part of the $\ell = 2$ pattern can also be separated into three subanomalies, the westernmost lying underneath the Nazca plate (see, also, Figure C1).

5. Conclusions

[83] The spectra of seismic models of the Earth’s mantle are predominantly of long spatial wavelength [e.g., Su and Dziewonski, 1992]. We have found that the long-wavelength components of most tomographic models published within the last decade are systematically well correlated with each other, indicating a substantial agreement between different techniques. As a general
rule, correlation is highest in the lowermost mantle (see, e.g., Figures 12 and 13), where the coverage of teleseismic travel-time data is most uniform [e.g., van der Hilst et al., 1997]. Although most models are described with a fine lateral parameterization (at least up to spherical harmonic degree $\ell = 20$, or equivalent), correlation is always lower for shorter spatial wavelengths, especially for $S$ models (Figures 7–9). This suggests that, so far, attempts to image the smaller ($\gtrsim 1000$ km) scale structure of the mantle have not been equally successful. The correlation between models of the same type is significantly higher than when $P$ and $S$ models are compared with each other. This discrepancy can partly be explained in terms of different sensitivities of the $P$ and $S$ data sets to lateral structure at different depths. Correlation between $\delta v_s$ and $\delta v_p$ anomalies is lowest in the upper mantle and at the CMB (Figures 9 and 10), where a common thermal origin might not be sufficient to explain the imaged heterogeneities but compositional anomalies could be invoked.

[84] We found some evidence for a change of the spectral character of heterogeneity below 660 km and local minima in the radial correlation function at $\sim 700$ km for $S$ and $P$ models (Figures 4–6) but failed to detect strong layering or global discontinuities at other depths in the mantle. However, as the 660 km discontinuity model s362d1 and cross-model discrepancies at $\sim 2000$ km show, interesting depth ranges in the mantle coincide with those depths...
where tomographic model consistency is still limited.

Tomography does not correlate well globally with models based upon geodynamic reconstructions of mantle flow for \( \ell \approx 5 \); seismic observations and subduction history models do not yet produce identical images. We have, however, found that fast anomalies are imaged consistently in the midmantle where we would expect slabs in the absence of a long-term barrier to flow at 660 km (Figures 14 and 17). This substantiates previous studies [van der Hilst et al., 1997; Čížková et al., 1998; Bunge et al., 1998], and slab penetration is found to be a common phenomenon (Figure 18). As discussed, it can be expected that future flow models will do a better job in predicting slab locations and current discrepancies should lead to a better understanding of the nature of the subduction process.

Our results are consistent with an emerging whole mantle convection paradigm in which the phase transition, with a probable viscosity increase, at 660 km can lead to transient slab flattening and flow reorganization but, in the long term, subduction maintains a high mass flux between the upper and lower mantle.

Appendix A. Statistical Significance of Linear Correlation

Assuming a binormal distribution for the deviations from a linear trend, we can apply Student’s \( t \) test [Press et al., 1993, p. 637] to
evaluate the likelihood, $p$, that a correlation $r$ between two sets of coefficients is caused by chance. The number of degrees of freedom is $(\ell_{\text{max}} + 1)^2 - 3$; for $\ell_{\text{max}} = 20$ and $\ell_{\text{max}} = 8$, the 99% significance levels ($p = 0.01$) for $r$ are then given by 0.123 and 0.286, respectively. Most of the $r$ values we find are therefore “significant” at the 99%-level, although care should be taken when interpreting such statements [e.g., Ray and Anderson, 1994].

Appendix B. Construction of Mean Models

[88] To create an ad hoc reference against which we can compare individual models, we con-
struct “stacked” P and S wave tomography models. Models smean and pmean have been calculated by a weighted average of well-correlated S models with similar input data (ngrand, s20rts, and sb4118) and the newer P wave models (bdp00 and kh00p), respectively. We first determine depth-averaged $\delta_{\text{RMS}}$ for each input model and then scale the models.
such that they would lead to a mean model with a depth-averaged \( \delta v_{\text{RMS}} \) that corresponds to the mean \( \delta v_{\text{RMS}} \) of all input models. This procedure maintains the depth dependence of \( \delta v_{\text{RMS}} \) for each model but evens out total heterogeneity amplitude differences between models. (We have experimented with additional, \( \ell \)-dependent average correlation weighted models, results were not much different.)

The spatial expansion of the resulting mean models is shown at selected depths in Figure C1 (for the spectral and \( \delta v_{\text{RMS}} \) character, see Figures 2 and 3 and the additional online material).

**Figure C3.** Comparison of total correlations between bdp00 and kh00p based on grids (circles) with spherical harmonics, \( r_{31} \) (after equation (9), triangles), and spherical harmonics based on a \( k_{\text{max}} = 20 \) Chebyshev reparameterization (inverted triangles).
Appendix C. Additional Material

C1. Mean Model Features

Figure C1 shows maps of our mean tomographic models smean and pmean at selected midlayer depths, confirming that the structures that were discussed in sections 4.3.3 and 4.3.4 are robust features of tomography.

C2. Absolute Power Spectra

Figure C2 shows absolute power, $\sigma_i^2$, to be compared with the normalized spectral power, $\hat{\sigma}_i^2$, shown in Figure 1.
C3. Comparison of Grid and Spherical Harmonic Correlation

Figure C3 shows a comparison of grid [Press et al., 1993, p. 636] and spherical harmonics based estimates of total correlation between bdp00 and kh00p as a function of depth. The area-weighted, discrete data estimate is based on an expansion of the \( \ell_{\text{max}} = 31 \) representation of both fields on 1.4° × 1.4° blocks, and the spherical harmonics estimate is \( r_{31} \), after (9). Both methods yield similar results in general, and we find no systematic deviations with depth. We have also included a \( r_{31} \)-correlation estimate that is based on a \( k_{\text{max}} = 20 \) radial Chebyshev reparameterization; features are similar but the curve is smoother, as expected.

C4. Alternative Radial Correlation Estimates for kh00p

Figure C4 shows the radial correlation function estimate \( r \) at \( \Delta z = 200 \) km for different parameterizations of kh00p. We observe that radial features of our \( \ell_{\text{max}} = 20 \) spherical harmonics based estimate as in Figure 4b are very similar to what we would...
obtain using original grid data (absolute numbers for $r$ differ, however). The comparison of $\ell_{\max} = 8$ and $\ell_{\max} = 31$ estimates shows that small-scale structure in $r$ is found across all wavelengths. The Chebyshev radial parameterization introduces some smoothing but is otherwise able to recover the major features of the original model.

**C5. Correlation Between lrr98d and smean**

Figure 5 shows the cross-model correlation for the best $\langle r_8 \rangle$ pair of subduction and tomography models, lrr98d and smean. However, intermediate wavelength correlation with smean is slightly better for stb00d ($\langle r_{20} \rangle = 0.21$) than for lrr98d ($\langle r_{20} \rangle = 0.18$). The depth dependence of $r_8^{\text{tot}}$ and $r_{20}^{\text{tot}}$ in Figure C5 is similar to that shown in Figure 14 and shows a midmantle maximum.

**C6. Comparison Between lrr98d and Stretched stb00d**

Figure C6 shows the cross-model correlation for lrr98d and the $f_{\text{opt}} = 1.75$, stretched stb00d.

---

**Figure C6.** Cross-model correlation for lrr98d (solid $\delta v_{\text{RMS}}$-line) and $f_{\text{opt}} = 1.75$ stretched stb00d (see section 4.3.2, dashed $\delta v_{\text{RMS}}$-line). Compare Figure 11, and see Figure 7 for description.
version of stb00d which was discussed in section 4.3.2 as a simplified way to study modified slablet sinking speeds (compare Figure 11). Note that $r^\ell$ is consistently higher for even than for odd $\ell$, especially below 700 km. The even degrees, related to the circum-Pacific subduction, are prominent in the power spectra of stb00d and lrr98d (see Figures 1, C2, and the additional online material) and appear to be best constrained at depth.

C7. Comparison of Slow Anomalies With Hot Spot Locations

[95] To complement our correlation study for fast anomalies, we have expanded the hot spot
lists of Steinberger [2000a] and 3SMAC [Nataf and Ricard, 1996] as negative $\delta$-functions (damped with a $\cos^2$-taper for $\ell > 0.75\ell_{\text{max}}$) and compared them to slow $\delta\nu$ only tomography ($\nu$ should be positive if hot spots are in $\delta\nu < 0$ regions). Figure C7 shows that we find only a weak correlation between surface observations of hot spots and slow anomalies that might be connected to rising plumes, consistent with earlier results [e.g., Ray and Anderson, 1994].

The correlation is best near the surface, at $\sim 1500$ km, and at the CMB, the depths at which Figures C8 through C10 show the spatial expansions of the fields.

Acknowledgments

[96] We are indebted to all authors of the seismic and geodynamic models for sharing their results and insights. We also want to thank our reviewers Rob van der Hilst
Figure C8. Comparison of negative δν only tomography and hot spot distributions (arbitrary units). The ℓ_{max} = 20, and depth of tomography is z = 300 km. Hot spot fields are based on δ-function expansions of hot spot locations from Steinberger [2000a] and 3SMAC [Nataf and Ricard, 1996].

Figure C9. Negative δν only tomography at z = 1500 km.
and Hans-Peter Bunge as well as Bernhard Steinberger and James Kellogg for valuable suggestions on the manuscript. Richard O’Connell, Adam Dziewonski, and Göran Ekström were of great help, offering continuing support and encouragement. Funding for TWB was provided by a DAAD Doktorandenstipendium under the HSP-III program. Most figures were prepared with GMT by Wessel and Smith [1991].

References


