

Seismic and mineralogical structures of the lower mantle from probabilistic tomography

I. Mosca

Department of Earth Sciences, Utrecht University, PO Box 80021, 3508 TA
Utrecht, The Netherlands.

Present address: Department of Earth and Environmental Sciences,
Ludwig-Maximilians-University, Munich, Germany.

L. Cobden

Department of Earth Sciences, Utrecht University, Utrecht, The Netherlands
Present address: Institut für Geophysik, Westfälische-Wilhelms-University,
Münster, Germany

A. Deuss

Department of Earth Sciences, University of Cambridge, Cambridge, UK

J. Ritsema

Department of Earth and Environmental Sciences, University of Michigan,
Ann Arbor, Michigan

J. Trampert

Department of Earth Sciences, Utrecht University, Utrecht, The Netherlands

Abstract. One of the most powerful approaches for understanding the 3-D thermo-chemical structure of the lower mantle is to link tomographic models with mineral physics data. This is not straightforward because of strong trade-offs between thermal and chemical structures and their influence on seismic structures. They can be reduced by mapping simultaneously perturbations of wave-speeds and density anomalies and by the quantitative assessment of the accuracy and uniqueness of seismic and mineralogical data.

Here, we present new tomographic maps of low order even-degree seismic structures which are an improvement on earlier models. They satisfy constraints from body wave, surface wave and normal mode data simultaneously, thereby enhancing the spatial resolution. Furthermore, the seismic structure at a given location is represented by a probability density function (*pdf*) which takes into account the uncertainty and non-uniqueness of the solution due to modelling and data restrictions.

Following a robust statistical procedure, we fit heterogeneity of wave-speeds and density from hypothetical thermo-chemical models to those of our tomographic maps. We thereby constrain lateral variations of temperature as well as iron, silica and post-perovskite concentration in terms of *pdfs*. Our work shows that large scale chemical variations are likely everywhere in the lower mantle. In most of the D'' region post-perovskite is most abundant in the Circum-Pacific belt, but near the core its lateral variation is more complex. Furthermore, post-perovskite concentration trades off with the ampli-

36 tudes of temperature and silicate variations, but not with their lateral dis-
37 tribution. This might be the reason why temperature and silicate variations
38 appear not constrained by our data in the lowermost few hundred km of the
39 mantle.

1. Introduction

In the last three decades a large number of global tomographic studies have produced maps of the seismic properties (i.e., shear wave-speed, compressional wave-speed and to a lesser extent density) of the Earth's interior. Despite different datasets and different strategies for solving the inverse problem, the maps show common long wavelength variations, especially of shear wave-speed anomalies. In particular, S-wave speed perturbation models [e.g. *Ritsema et al.*, 1999; *Masters et al.*, 2000; *Megnin and Romanowicz*, 2000; *Gu et al.*, 2001; *Grand*, 2002; *Becker and Boschi*, 2002; *Antolik et al.*, 2003; *Ritsema et al.*, 2011] resolve globally the Earth's deep interior better than P-wave speed perturbation models [e.g. *Bijwaard et al.*, 1998; *Vasco and Johnson*, 1998; *van der Hilst et al.*, 1997; *Boschi and Dziewonski*, 2000; *Masters et al.*, 2000; *Fukao et al.*, 2001; *Karason and van der Hilst*, 2001; *Zhao*, 2001] which are more suitable for detecting hundred-km wide structures (e.g., subducting slabs) on the regional scale. The uppermost part of the lower mantle is dominated by higher than average wave-speed associated with down-going lithospheric slabs beneath North and South America as well as the Western Pacific [e.g. *van der Hilst et al.*, 1997; *Fukao et al.*, 2001]. Between 1500 and 2000 km depth a coherent pattern of seismic velocity heterogeneity disappears. The D'' region, i.e. the few hundred km above the CMB [*Bullen*, 1949], is dominated by lower than average shear wave-speed beneath the central Pacific Ocean and Africa and high shear wave-speed anomalies beneath the Circum-Pacific belt. The two low-velocity regions have been interpreted as upwellings of hotter material and have therefore been referred to as "super-plumes" [e.g. *Su and Dziewonski*, 1997; *Forte and Mitrovica*, 2001; *Romanowicz and Gung*, 2002] or as large

low shear velocity provinces (LLSVP) after *Garnero and McNamara* [2008]. We prefer the latter as it does not suggest any particular dynamic behaviour. Although some P-wave speed variation models display a similar pattern in the lowermost mantle [e.g., *Masters et al.*, 2000; *Karason and van der Hilst*, 2001; *Zhao*, 2001], there is no general consensus on the compressional velocity structure due to the limited sampling by P-wave phases (e.g., PcP, P_{diff}) in this region.

In the last ten years it has been widely accepted that seismic anomalies have both a thermal and a chemical origin [for review, *Romanowicz*, 2003; *Trampert and van der Hilst*, 2005; *Lay*, 2007]. Studies based on both the forward and the inverse problem have contributed towards accepting the thermo-chemical nature of Earth’s deep structure. Waveform modelling and travel-time analysis have produced results which are not biased by the ill-posed and ill-conditioned tomographic inverse problem and therefore have provided in some cases better estimations for the small-scale seismic structure. For example, they have revealed a larger amplitude and a stronger lateral gradient of shear velocity beneath the Central Pacific Ocean [*Breger and Romanowicz*, 1998; *Tanaka*, 2002; *Ford et al.*, 2006; *He et al.*, 2006] and Africa [*Ritsema et al.*, 1998; *Ni and Helmberger*, 2003a, b; *Toh et al.*, 2005] than those from tomographic studies. Such a high amplitude (between 1 and $\sim 7\%$ beneath Central Pacific and $\sim 3\%$ beneath Africa) cannot be caused only by a thermal effect.

The joint inversion for pairs of seismic observables have undoubtedly shown the existence of significant chemical variations. One of the most important contributions is the density variation model of *Ishii and Tromp* [1999] produced independently of compressional and shear wave-speed perturbations from a dataset of high quality normal-mode

measurements. The most surprising feature of their model was the anti-correlation between S-wave speed and density heterogeneity in the deep mantle. *Su and Dziewonski* [1997], *Masters et al.* [2000] and *Forte and Mitrovica* [2001] inverted for both the S-wave speed and the bulk sound velocity using different seismic datasets and found a strong anti-correlation between their anomalies. The tomographic models based on the joint inversion for shear and compressional wave-speeds allowed several authors to determine the shear to compressional velocity anomaly ratio which is between 2.5 and 4 in the D'' region [e.g. *Robertson and Woodhouse*, 1996; *Su and Dziewonski*, 1997; *Kennett et al.*, 1998; *Vasco and Johnson*, 1998; *Masters et al.*, 2000; *Romanowicz*, 2001; *Ritsema and van Heijst*, 2002; *Antolik et al.*, 2003]. The relative behaviour between seismic anomalies (i.e., $dlnv_s$ versus $dln\rho$, $dlnv_s$ versus $dlnv_p$, $dlnv_s$ versus $dlnv_c$) contradicted the thermal paradigm widely accepted for the lower mantle and implied a less dominant role of the thermal perturbations and, thus, an increasing role for chemical anomalies [*Forte and Mitrovica*, 2001; *Karato and Karki*, 2001; *Saltzer et al.*, 2001; *Deschamps and Trampert*, 2004]. If the temperature alone were to be responsible for the seismic heterogeneity, a shear to compressional wave-speed anomaly ratio would be lower than 2 [e.g. *Karato and Karki*, 2001] and the scaling factor of density to shear wave-speed anomaly would be ≈ 0.4 [e.g. *Anderson et al.*, 1984].

There is a new incentive to quantify the chemical and thermal nature of seismic heterogeneity in the mantle following the discovery of the phase change in perovskite, the most abundant mineral of the lower mantle, to post-perovskite [*Murakami et al.*, 2004; *Oganov and Ono*, 2004; *Tsuchiya et al.*, 2004b]. Mineral physics predicts that the conversion of perovskite into post-perovskite involves an abrupt increase in S-wave speed variations and

107 a smaller increase in P-wave speed perturbations [*Tsuchiya et al.*, 2004b; *Wentzcovitch*
 108 *et al.*, 2006]. This appears consistent with images of shear wave velocity as well as with
 109 the high ratio of shear to compressional wave-speed anomalies in the D'' region. Fur-
 110 thermore, seismic anisotropy, which is relatively strong in localized regions of the deep
 111 interior [e.g. *Kendall and Silver*, 1996; *Vinnik et al.*, 1998; *Panning and Romanowicz*,
 112 2004], may be consistent with the intrinsically anisotropic structure of post-perovskite.
 113 Mineral physics suggests that the post-perovskite phase transition occurs in cooler regions
 114 characterized by a strong increase in $dlnv_s$, a small increase in $dlnv_p$ and a 1-2% increase
 115 in $dln\rho$ [*Oganov and Ono*, 2004; *Tsuchiya et al.*, 2004b, c]. Hence, it is most likely that
 116 post-perovskite is present at the Circum-Pacific belt [*Wentzcovitch et al.*, 2006].

117 Geodynamic studies have found that the relatively high density beneath the central
 118 Pacific Ocean and Africa can be explained by the existence of a thick compositionally
 119 distinct piles of ancient slabs of oceanic lithosphere [*Kellogg et al.*, 1999; *Karato and*
 120 *Karki*, 2001; *Nakagawa and Tackley*, 2004]. Such piles are denser than the surrounding
 121 mantle and perhaps iron-enriched. This iron enrichment may elevate the perovskite (Pv)
 122 to post-perovskite (pPv) transition to shallower depths than in the ambient mantle [*Ohta*
 123 *et al.*, 2008]. In this context the sources of chemical heterogeneity, well known so far, can
 124 be represented by the Pv into pPv phase transition and the higher concentration of iron
 125 in the deep interior.

126 *Trampert et al.* [2004] and *Deschamps et al.* [2007] have combined seismological obser-
 127 vations with constraints from mineral physics and geodynamics. Using the sensitivities
 128 from mineral physics [*Trampert et al.*, 2001; *Deschamps and Trampert*, 2004], thermal
 129 and chemical variations in the deep Earth are related to anomalies in P-wave speed, S-

130 wave speed and density mapped simultaneously from the inversion of surface wave and
 131 normal-mode measurements. The thermal and chemical perturbation model has sub-
 132 sequently been linked to geodynamics in *Deschamps et al.* [2007] and *Deschamps and*
 133 *Tackley* [2008, 2009]. They generated thermo-chemical convection models for the lower
 134 mantle and compared the associated thermal and compositional density distributions with
 135 those from *Trampert et al.* [2004] in terms of power spectra. Using this procedure the be-
 136 haviour of the thermo-chemical convection in the lower mantle can be correlated with
 137 changing rheological, physical and compositional parameters.

138 More recently, an approach for linking seismological and mineralogical data at the base
 139 of the mantle is given in *Cobden et al.* [submitted]. This is based on a statistical compari-
 140 son of P-wave speed and S-wave speed anomalies generated (1) via tomographic inversion
 141 of seismic data and (2) via thermodynamic modelling of hypothetical thermo-chemical
 142 structures.

143 The current paper extends the works of *Trampert et al.* [2004] and *Cobden et al.* [sub-
 144 mitted]. Here we image spatial variations in temperature and composition in the lower
 145 mantle structure by combining the most recent mineral physics experimental constraints
 146 and an updated seismic dataset. It is possible, therefore, to map lateral variations in
 147 post-perovskite content in the deep interior. In Sections 2-4 we present the approach,
 148 data and results for obtaining long wavelength models of even order structures of P-wave
 149 speed, S-wave speed and density, using body wave, surface wave and normal-mode mea-
 150 surements and a Monte Carlo approach for the inverse problem. Sections 5 and 6 describe
 151 the procedure and the results of the thermo-chemical model corresponding to the observed
 152 seismic anomalies. Finally, in Section 7 we draw general conclusions.

2. Method for mapping seismic structures

Probabilistic tomography [*Trampert et al.*, 2004] is a recent technique for solving the seismic inverse problem. It is based on a full model space search method, the Neighbourhood Algorithm, to infer all the possible information from data [*Sambridge*, 1999a, b]. So far probabilistic tomography has been applied to invert normal-mode splitting functions and surface-wave phase-velocity maps [*Beghein et al.*, 2002; *Resovsky and Trampert*, 2003; *Trampert et al.*, 2004; *Visser et al.*, 2008]. For these data it is straightforward to reduce the inverse problem into a small number of unknowns (less than 20), an important condition for applying a Monte Carlo approach. Indeed, a global search technique becomes computationally too expensive with increasing the number of free parameters. The construction of phase velocity maps and splitting functions for low spherical harmonics is a well defined inverse problem which does not require regularization. The depth inversion is a more poorly constrained inverse problem and the biggest source of non-uniqueness in the final model. However, it can be represented by a limited number of parameters and, thus, is suitable for probabilistic tomography. Body wave measurements have not been incorporated into a two-step probabilistic tomography yet, but they would increase the depth resolution of the seismic velocity in the lower mantle. However, decomposing the inverse problem of body waves in radial and lateral components is not straightforward. Travel-time data are often described using path integrals (e.g., ray theory) which produce too many unknowns for a global search approach.

Mosca and Trampert [2009] proposed a uniform formalism to invert body wave, surface wave and normal-mode data using a Monte Carlo technique. Assuming a spherically layered Earth, a path-average approximation can be formulated to separate the radial

and lateral components of travel-time measurements in long wavelength structures. In this way the delay time as a function of ray parameter p is the average of the two-way vertical travel-time perturbation $\delta\tau(p, \theta, \phi)$ at a location (θ, ϕ) between source and receiver:

$$\delta T(p) = \frac{1}{\Delta} \int_0^\Delta \delta\tau(p, \theta, \phi) d\Delta \quad (1)$$

and the two-way vertical travel-time perturbation at a location (θ, ϕ) is related to the radial velocity perturbation $\frac{\delta v(r, \theta, \phi)}{v_0(r)}$ at (r, θ, ϕ) with respect to PREM [Dziewonski and Anderson, 1981]:

$$\delta\tau(p, \theta, \phi) = -2 \int_{r_{bot}}^a K(r) \frac{\delta v(r, \theta, \phi)}{v_0(r)} dr. \quad (2)$$

Here, $K(r)$ is the sensitivity kernel that depends only on depth, v_0 is the velocity in PREM, r_{bot} is the radius of the bottom point of the ray path, and a is the Earth's radius.

Equations (1) and (2) are similar to classical tomography for normal-modes [e.g., Dahlen and Tromp, 1998] and surface waves [e.g., Woodhouse and Wong, 1986]. Thus, body wave travel-time tomography can be separated into two steps. In the first step two-way vertical travel-time residual maps are constructed as a function of ray parameter, comparable to splitting functions and phase velocity maps as a function of frequency. The second step involves their local depth inversion, the less well constrained part, but represented by few unknowns and, thus, suitable for being solved by the Neighbourhood Algorithm.

Two-way vertical travel-time residual maps, splitting functions and phase velocities maps are local functions linearly related to the Earth's structure $x = (v_p, v_s, \rho)$ through the kernel $K_x(r)$:

$$\delta d(\theta, \phi) = \int_0^a K_x(r) \delta x(r, \theta, \phi) dr. \quad (3)$$

This expression corresponds to equation (2) for body wave measurements. For free oscillations and surface waves $\delta d(\theta, \phi)$ is the local splitting function and the local phase perturbation, respectively. We first expand the data δd into spherical harmonics up to degree L for lateral heterogeneities:

$$\delta d(\theta, \phi) = \sum_{l=0}^L \sum_{m=-l}^{+l} c_{l,m} \Psi_{l,m}(\theta, \phi). \quad (4)$$

where $\psi_{l,m}$ is the spherical harmonic function, l and m are the angular and azimuthal order, respectively, and $c_{l,m}$ are the harmonic coefficients. The coefficients $c_{l,m}$ are functions of depth by virtue of (3). Therefore, we expand the model parameters $\delta x(r, \theta, \phi)$ laterally into spherical harmonics up to degree L and radially into K natural cubic spline functions, $f_k(r)$:

$$c_{l,m} = \sum_{k=1}^K c_{l,m}^k \int_0^a K_x(r) f_k(r) dr \quad (5)$$

with $c_{l,m}^k$, being the model coefficients. The full problem spans the lateral and radial components simultaneously with $K(L+1)^2$ unknowns. Using equations (4) and (5), the problem can be divided into two steps: (1) the representation of the data (travel time, phase and frequency variations) as a function on the sphere and (2) the local depth inversion of these functions for depth. The latter part trivially decomposes into individual spherical harmonic components (equation (5)) solvable by the Neighbourhood Algorithm (NA).

2.1. Search and appraisal stage of the Neighbourhood Algorithm

NA is a guided Monte Carlo method because it is based on a full space search. It consists of two parts, the search and the appraisal stage. The first part of the NA samples the entire model space, null space included, and attributes a probability to each solution

compatible with the data. The second stage represents the information from the ensemble of models in terms of a Bayesian posterior probability density function (*pdf*). We refer the reader to *Sambridge* [1999a, b] for a complete explanation about this technique. Following *Tarantola* [1987] we define the *pdf* of the estimated model parameter m from a set of measurements d :

$$P(m|d) = k\rho(m)L(m|d) \quad (6)$$

with $\rho(m)$, the prior *pdf* of the model; k , the normalization factor; and $L(d|m)$, the likelihood function which measures how well the model fits the data. This is proportional to the negative exponential of the misfit function χ^2 :

$$L(m) \propto \exp\{-\chi^2\} \quad (7)$$

where χ^2 evaluates the difference between observed and predicted data. Since we use body waves (BW), surface waves (SW) and free oscillations (NM), the misfit function is:

$$\chi^2 = \frac{1}{N} \left(\sum_{BW} \frac{(\delta\tau_i^{obs} - \delta\tau_i^{th})^2}{(\sigma_i^{BW})^2} + \sum_{SW} \frac{((\delta c/c)_i^{obs} - (\delta c/c)_i^{th})^2}{(\sigma_i^{SW})^2} + \sum_{NM} \frac{(\eta_i^{obs} - \eta_i^{th})^2}{(\sigma_i^{NM})^2} \right) \quad (8)$$

with N , the total number of data. The measurements are the i -th spherical harmonic coefficient of the vertical body wave travel-time residual maps, $\delta\tau_i$, surface wave phase velocity maps, $(\frac{\delta c}{c})_i$, and normal-mode splitting functions, η_i , with the associated uncertainty, σ_i . To estimate the predicted data of equation (8) the search stage evaluates the forward problem of each generated model (equation (5)).

To sample the model space efficiently and quickly, four parameters must be defined: n_s , the number of models generated at each iteration; n_r , the number of "best" data-fitting models around which to draw new models; N_{it} , the number of iterations; and n_{init} , the number of models sampled at the first iteration. After testing different combinations of

tuning parameters, we have found that in a highly-dimensional parameter space (> 30), it is better to have many iterations and small n_s and n_r with $n_r \ll n_s$ in order to converge towards the best fitting region within a practical computational time [Mosca, 2010]. We therefore choose $n_{init} = 100$, a suitable number to start the survey without missing the region of good data-fit, $N_{it} = 10000$, $n_s = 20$, $n_r = 2$ and, thus, generate 200100 models in the first stage. For a more exploratory search, the parameters n_s and n_r should have similar values. The search becomes more restrictive and the sampling of the model space more localized around the minimum-misfit area with decreasing n_r . We have tested that with our choice of parameters we do not miss any potential solutions in the model space. The algorithm has been designed to run on parallel computers. In particular, the search stage has recently been reformulated to scale more efficiently [Rickwood and Sambridge, 2006]. In our case, we generate 200100 models in ~ 5 hours using 20 processors on a 32-node cluster.

The appraisal part of the NA is controlled by N_w and N_c which are the numbers of models used to re-sample the parameter space by Gibbs samplers. Their choice is less critical than that of n_s and n_r and we only have to ensure proper convergence of the Monte Carlo integration. We set $N_w = 5$ and $N_c = 400$ so that we re-sample 158000 models in ~ 2.5 hours with 80 processors of a 32-node cluster.

Besides the tuning parameters, the algorithm requires a choice of the boundaries of the model space. This, together with the model parametrization and the data uncertainty, represents the prior constraints on the final model. As a result of many tests [Mosca, 2010], we impose $dlnv_p$ to lie between -0.01 and +0.01, $dlnv_s$ and $dln\rho$ between -0.02 and +0.02 and the topographic variations $\frac{dr_{670}}{r_{670}}$ and $\frac{dr_{cmb}}{r_{cmb}}$ between -0.005 and +0.005.

The latter corresponds to a topographic variation between -28 and +28 km at the 670-km discontinuity and between -17 and +17 km at the CMB. Due to the strong trade-off between density and topography a smaller range would affect the estimation of $d\ln\rho$.

3. Data & Parametrization

The measurements used in this study consist of body waves, free oscillations and surface waves with associated uncertainty. Table 1 lists the type (seismic phase for body waves, the mode for surface waves and normal-modes) and the number of data.

The body wave travel-time residuals are measured from the maximum cross-correlation between low-pass filtered ($T > 16$ s) seismograms and synthetics calculated with respect to PREM within 80-s time windows around major phases [Ritsema *et al.*, 2011; Ritsema and van Heijst, 2002]. Source parameters are from the global CMT catalogue (<http://www.globalcmt.org>). Furthermore, the time delays have been corrected for crustal, ellipticity and relocation corrections [Ritsema *et al.*, 2011]. In particular, the crustal structures have been calculated using finite-frequency theory [Ritsema *et al.*, 2009]; whereas, the corrections due to the source mislocation have been obtained from fitting travel times in a smooth degree-12 velocity model [Ritsema *et al.*, 2004]. We have found that the largest correction is due to relocation errors which are between -5 and -1 s compared to 1-2 s for the other two corrections [Mosca, 2010].

To build consistent two-way travel-time residual map $\delta\tau(p, \theta, \phi)$ we select ray paths within a narrow range of ray parameters to maximize the ray coverage. Following this criteria we analyze P-, P_{diff} and PP-phases on the vertical component seismogram; and S-, S_{diff} , SS- and SSS-phases on the transverse component. The diffracted ray paths have a constant ray parameter; the ray parameters of P- and S-phases vary by 1 s/deg; and

those of PP-, SS- and SSS-phases, by $\Delta p = 5$ s/deg. We constructed 44 compressional and 46 shear travel-time residual maps. The uncertainty of the travel-time maps has two main sources. The first contribution is due to the spherical harmonic expansion of the two-way vertical travel-time perturbation (equation (4)). It measures the discrepancy between the observed travel-time residual and the theoretical delay time predicted by the harmonic expansion. The second contribution defines the deviation of the path-average approach from ray theory and finite-frequency approach [Mosca and Trampert, 2009]. This term is larger than the first one and thus contributes most to the uncertainty of travel-time measurements. The error in taking a range of p rather than a single value in the construction of the maps is negligible. We find that error-bars of shear phase data are less than 4.0 s. The standard deviation of compressional phase data is included between 0.8 and 1.5 s. We assume that the travel-time residual uncertainty is equally divided among the spherical harmonic components.

For the normal-mode data we use spheroidal splitting function measurements with corresponding uncertainty below 3mHz from *Deuss et al.* [2011] and their corresponding sensitivity kernels. The data from *Deuss et al.* [2011] include the most recent seismic events (e.g., the 2008 Wenchuan China, the 2004 Sumatra and the 2005 Pakistan earthquakes) and new global seismic stations providing a better data coverage than that of *Resovsky and Ritzwoller* [1998]. Although the amplitudes of the splitting functions between the two catalogs are comparable, the discrepancy in terms of pattern is large for spheroidal modes with a deeper sensitivity. We exclude data with a sensitivity to the inner core and neglect the interaction between multiplets. We use even-degree structure coefficients corresponding to the fundamental mode and the first five overtones (Table

1). In total, 38 spheroidal modes for degree-0 and degree-2 (5 structure coefficients), 32 for degree-4 (9 structure coefficients) and 26 for degree-6 (13 structure coefficients) are included. They are corrected for the crust using the model CRUST5.1 [Mooney *et al.*, 1998]. Crustal structures have been obtained by computing the exact local perturbation from CRUST5.1 and normal-mode equations in a radially symmetric model [Woodhouse, 1974; Bozdag and Trampert, 2008]. We then subtract the crustal correction from the corresponding splitting function. Since we are only working with structure coefficients up to degree-6, CRUST5.1 has sufficient complexity. As expected, crustal structure coefficients have larger values (up to 5.0 mHz) for modes with shallow sensitivity. We further exclude toroidal modes from our procedure to avoid an inversion of both horizontally-polarized and vertically-polarized shear wave velocity in the upper mantle. This would require to work with many more model parameters which is not feasible in a Monte Carlo approach. Furthermore, toroidal modes are less sensitive to density because they only involve horizontal motions.

Surface wave data are from Visser *et al.* [2008] who computed phase velocity maps using the approach of Yoshizawa and Kennett [2002]. We consider 53 fundamental and higher mode Rayleigh wave phase velocity maps with their uncertainty and the associated kernel, up to the fourth mode. We exclude high frequency (> 0.002 Hz) fundamental modes as well as the fifth and sixth overtones which are strongly sensitive to the crust [Bozdag and Trampert, 2008]. We apply the crustal corrections, computed by a procedure similar to that described for normal-modes, to surface wave measurements using the model CRUST5.1 which is sufficient for the correction of low degree maps. We do not consider Love wave phase velocity maps for the same reason as for the toroidal modes.

The Earth's structure is expanded laterally into spherical harmonics and vertically into natural cubic splines as explained in Section 2. The depth inversion is performed by considering one spherical harmonic coefficient at a time. Although two-way travel time residuals and phase velocities provide maps at even and odd degrees, splitting functions are only available for even degrees. We therefore discard the odd degrees and the three data sets together provide 181 coefficients for degree $l = 0$ and 2, 175 for $l = 4$ and 169 for $l = 6$ (Table 1). Since care was taken that there is no spectral leakage for the low degree expansions, neglecting the odd degrees is not a problem, except for the interpretation of the final models.

The most suitable number of splines depends on several factors. First, using the path-average approximation, the body wave coverage is only good between a depth of 670 and 1900 km and in the D'' region. Surface wave measurements have a good depth resolution in the upper mantle down to 670 km depth and free oscillations constrain only the large-scale structure of the mantle. We simultaneously invert for seismic wave-speeds and density. Surface waves and free oscillations are also sensitive to topography perturbations. Hence, we also solve topographic variations at 670 km depth as well as at the core-mantle boundary. We assume that surface topography is isostatically compensated and thus taken into account by the crustal correction. We therefore parametrize the mantle with 10 natural cubic spline functions, a compromise between expected resolution and the number of parameters (Figure 1). This means that our model space has 32 dimensions. For each of the 28 structure coefficients (1 degree-0, 5 degree-2, 9 degree-4 and 13 degree-6) we have 10 radial unknowns for $dlnv_p$, 10 for $dlnv_s$, 10 for $dln\rho$ and 2 for topographic perturbations.

The measurements described above constrain the compressional wave-speed, shear wave-speed and density throughout the mantle. This is shown in Figure 2 where we plot the partial derivatives (radial integral in equation (5)) of seismic data as a function of depth of the spline knot. The lower mantle is well resolved by body waves and normal-modes, whereas surface waves and normal modes image the upper mantle well. Furthermore, constraints on density structures are provided mainly by normal mode measurements since the density sensitivity is absent in body wave data and weak in surface waves. Specifically, the density sensitivity in the deep interior is determined by two spheroidal modes of the third overtone, five of the fourth overtone and three of the fifth overtone.

4. Seismic wave-speeds and density models

In Section 2, we described theoretically how we inverted body wave, normal-mode and surface wave data locally for depth using the NA. In this section we show the results for the lower mantle. The images for the upper mantle are displayed in Appendix 1.

Using the Neighbourhood Algorithm, we obtain per spherical harmonic coefficient a probability density function for each of the 32 components of the mantle model (10 for $dlnv_p$, 10 for $dlnv_s$, 10 for $dlnv\rho$, $\frac{dr_{670}}{r_{670}}$ and $\frac{dr_{cmb}}{r_{cmb}}$). To evaluate the model at a certain latitude, longitude and depth, we draw random values which re-sample the 1-D (Gaussian and non-Gaussian) marginal probability density function of the model parameters using the Metropolis-Hastings algorithm [Metropolis *et al.*, 1953; Tarantola, 1987; Mosegaard and Tarantola, 1995]. The algorithm generates random walks sampling an initial, unknown probability distribution. Only those walks fitting the target distribution are accepted. The advantage of this Markov Chain Monte Carlo method is that we obtain a sequence of random samples from a probability distribution for which direct sampling is difficult.

At each depth of the spline knots, we define laterally 264 grid nodes of $15^\circ \times 15^\circ$ area, and express the perturbation of v_p , v_s and ρ as a *pdf* at a specific node by summing the re-sampled *pdfs* of the spherical harmonic coefficients. By virtue of the Central Limit theorem, the *pdf* at a given node converges towards a Gaussian probability distribution. To display the models of lateral variations it is therefore sufficient to show their means (Figure 3) and standard deviations (Figure 4) at each of the 264 grid nodes and 10 spline knots. In the Bayesian framework the mean is the most likely value and the standard deviation is the width of the *pdf*. It is important to note however that the mean model does not necessarily fit the data best. To identify the best fitting model the joint probability density function has to be analysed, not the marginals. The appraisal stage of the NA also computes the 2-D marginal *pdf* between pairs of unknowns. This quantity provides information about the trade-off between model parameters. Some of the 2-D marginals are discussed in Mosca [2010]. They are not important for our conclusions and therefore not included here.

In Figure 3 at 2891 km depth our maps show negative seismic wave-speed anomalies beneath the central Pacific Ocean and Africa surrounded by an increase in S-wave speed which is particularly strong beneath South America and the eastern Asia. Low S-wave velocity provinces are correlated well with compressional wave-speed structures at the same locations. At 2600 km depth negative anomalies in v_s are more pronounced and correspond to an increase in density. The agreement between positive wave-speed anomalies and subducting slabs in the lower mantle is not clear from our images. The reason is that such small-scale heterogeneities are below the lateral resolution of our tomographic images which map structure of wavelength equal to or higher than ~ 3300 km.

The lateral distribution of uncertainties from Figure 4 is relatively uniform. The relative average standard deviation is $\sim 70\%$ of the maximum amplitude of $dlnv_p$, $\sim 55\%$ of the maximum amplitude of $dln\rho$ and between 48% and 58% of the largest $dlnv_s$. Such large standard deviations are a combination of two factors. The first is the uncertainty of the seismic data and the second is due to the broad search performed by the NA. Furthermore, $dlnv_p$ values are associated with the highest error-bar because the uncertainty of two-way travel-time residual maps of compressional phases are higher than those of shear phases. We would like to stress again that our model parameters are given as marginal *pdfs* and therefore we should not subject the mean to over interpretation.

Other more robust features of seismic heterogeneity can be detected by calculating correlation coefficients and the root-mean-square (RMS) amplitudes. Moreover, these statistical parameters can be used to relate seismic wave-speeds and density anomalies to properties of thermo-chemical variations. To evaluate them we cannot use the 1-D marginals shown in Figures 3 and 4 because after marginalizing, information related to higher dimensional correlations might be lost and the result would be biased towards lower values than they really are. The calculation must draw samples in the multi-dimensional *pdf*. This is not an easy task, but we can approximate these samples by considering the output from the first stage of the NA. This estimate will be close to drawing in the joint probability densities because we tuned the NA to converge fast (i.e., $n_r \ll n_s$). The result is displayed in Figure 5. We also show the statistics for bulk sound velocity perturbations, $dlnv_c$, calculated from shear and compressional wave-speeds:

$$v_c^2 = v_p^2 - \frac{4}{3}v_s^2, \quad dlnv_c = \frac{dlnv_p - \gamma dlnv_s}{1 - \gamma} \quad (9)$$

where $\gamma = \frac{4v_s^2}{3v_p^2}$. We compute the 1-D marginal of $dlnv_c$ for a given spherical component using the concept of convolution between two *pdfs* (i.e., the 1-D marginals of shear and compressional velocities). The first column of Figure 5 displays the probability distribution of the correlation coefficient for pairs of seismic variations as a function of depth. The correlation between P-wave and S-wave speed perturbations is mostly positive but decreases with increasing depth. Variations of v_s are anti-correlated with both density and bulk sound velocity anomalies, although the tail of their distributions is often very long. Both the anti-correlation $dlnv_s - dln\rho$ and those $dlnv_s - dlnv_c$ indicate a chemical origin for seismic heterogeneities in the deep interior [Karato and Karki, 2001; Forte and Mitrovica, 2001; Trampert et al., 2004].

The second column of Figure 5 shows the profile of the likelihood of the RMS amplitude of $dlnv_p$, $dlnv_s$, $dln\rho$ and $dlnv_c$ with varying depth and, hence, provides information about seismic anomaly amplitudes. Clearly, P-wave speed variations are smallest, while $dln\rho$ has the largest values also indicated by Figure 3. The likelihood of ratios $R = \frac{dlnv_s}{dlnv_p}$ and $\xi = \frac{dln\rho}{dlnv_s}$ at various depths are shown in the third column of Figure 5. The range of R is between 1.2 and 3.0 in the lower mantle, in agreement with previous studies which found a compressional to shear wave-speed anomaly ratio between 2 and 3.5 in the D'' region [e.g. Masters et al., 2000; Romanowicz, 2001; Deschamps and Trampert, 2003; Ritsema and van Heijst, 2002; Antolik et al., 2003]. The width of the histogram of R is particularly important. As pointed out by Deschamps and Trampert [2003], this contains valuable information about the origin of the seismic perturbations. The probability distribution of the density to shear wave-speed anomaly ratio is between 0.8 and 1.7. This range is not in agreement with earlier studies which estimated values of ξ smaller than 1.0 [Forte

et al., 1994; *Cadek and Fleitout*, 1999; *Ishii and Tromp*, 1999; *Karato and Karki*, 2001; *Romanowicz*, 2001]. Most of these studies however assume a priori temperature to be responsible for density variations or used a regularized inversion. Our density variations are much larger, because we searched all models compatible with our data. We therefore think that the density model from this work is a robust clue to the existence of significant chemical heterogeneities in the Earth’s deep interior.

4.1. Consistency with previous models

To check the consistency of our results, we compare them with two different tomographic models. The first one is the S-wave velocity model S20RTS of *Ritsema et al.* [1999] and the corresponding P-wave velocity model P12 of *Ritsema and van Heijst* [2002]. Both S20RTS and P12 are three-dimensional tomographic velocity models expressed as perturbations from PREM. The models are parameterized radially with 21 natural cubic splines between the Moho and the core-mantle boundary and laterally with spherical harmonics up to degree 20 for S20RTS and up to degree 12 for P12. We only plot their lateral heterogeneity for even degrees up to 6 to facilitate a direct comparison between the maps.

The second model for comparison is GyPSuM of *Simmons et al.* [2009, 2010]. This is a three-dimensional mantle model of compressional wave-speed, shear wave-speed and density perturbations and has been produced from a joint inversion of seismic body wave travel-time measurements, geodynamic and mineralogical data. They computed density perturbations from S-wave variations by a scaling factor calculated using a Monte Carlo approach. We have chosen to compare our results to that model because it is one of the few models to have inverted for independent density perturbations. GyPSuM consists of 22 layers radially and 275×275 km blocks laterally. To make a consistent comparison,

we expand their model into spherical harmonics and again map only degrees 0, 2, 4 and 6.

Our models are given as marginal *pdfs* at each node, while the models for comparison only take one specific value obtained by regularized inversion. The best way to compare them is to plot the comparison models on top of our marginals at specific latitude-longitude points. We have chosen the point $(0^\circ, 180^\circ)$ which is exactly in the center of the Pacific LLVZ. It is clear that S20RTS/P12 and GyPSuM intersect with our *pdfs* and therefore formally agree (Figure 6). Several features are noteworthy. S20RTS/P12 and GyPSuM are overall small, except near the boundaries of the model where they show more amplitude. This is mostly due to regularization rather than data constraints. Our *pdfs* are wide which reflects both constraints from the data and our large search bounds. Both approaches reflect different strategies: finding one model given certain data and regularization constraints (i.e., S20RTS/P12 and GyPSuM) versus finding all models within a certain region (i.e., this work). While they are both valid and complementary, they will have a large influence on thermo-chemical inferences based upon them [*de Wit et al.*, 2012].

A further test to check the robustness of our results is to compare them with the model from *Resovsky and Trampert* [2003], referred to as RT03. This is a seismic model produced with free oscillation and surface waves data using the NA. In RT03, the lower mantle is parameterized with three layers only due to computational limitations at the time. For each spherical component we therefore average our spline parameterization over those three layers to make the comparison. The *pdfs* overlap nicely (Figure 7), but there also some differences. RT03 has been constructed without body wave data, and

the surface wave and free oscillation data were from *Trampert et al.* [2001] and *Resovsky and Ritzwoller* [1998], respectively. While the surface waves of *Trampert et al.* [2001] are similar to those of *Visser et al.* [2008], the normal-mode data of *Deuss et al.* [2011] have a significantly different pattern from those of *Resovsky and Ritzwoller* [1998], as explained in Section 3. Furthermore, we also use body wave measurements which were excluded in RT03. The presence of body waves increases the depth resolution in the lower mantle and, especially, in the D'' region. For this reason, our models have a depth resolution of a few hundred km whereas the depth resolution of RT03 is ~ 1000 -km.

5. Approach for constraining the thermo-chemical structure

To quantify the thermal, chemical and mineral phase variations fitting the results of Section 4, we first compute the seismic properties of many hypothetical mineralogical mixtures, applying the thermo-dynamically self-consistent method of *Stixrude and Lithgow-Bertelloni* [2005]. This is a relatively recent approach to determine both physical properties (elastic wave-speeds and density) and the phase equilibria of an assemblage of mantle phases although for our purposes, we only use the data to calculate physical properties. The self-consistency of the method ensures a robustness of all aspects of the mantle physics because they are derived within the same formalism. We apply the third-order finite strain Birch-Murnaghan equation-of-state (EOS) from *Stixrude and Lithgow-Bertelloni* [2005]. This allows us to calculate elastic moduli as well as the density for each mineral mantle phase present. The overall bulk modulus K and shear modulus G for a given mineral assemblage are computed using the Voigt-Reuss-Hill average of the component phases, while the density is the average of the densities of minerals weighted by their volumetric proportions. Defining shear and compressional wave-speeds from these K , G and ρ for

a mineralogical assemblage is then straightforward. For a more comprehensive description of this procedure the reader can consult *Stixrude and Lithgow-Bertelloni* [2005, 2011] and *Xu et al.* [2008]. Elastic parameters used in this study are taken from *Stixrude and Lithgow-Bertelloni* [2011]. These parameters have been inferred using a large compilation of laboratory and theoretical mineral physics experiments and are derived to fit the accompanying equation-of-state [*Stixrude and Lithgow-Bertelloni*, 2011]. We select seven end-member mineral phases at depths between 950 and 2000 km: (*Mg, Fe*) perovskite, Al_2O_3 perovskite, $CaSiO_3$ perovskite, (*Mg, Fe*) ferropericlasite and SiO_2 stishovite. In the lowermost mantle we include also the post-perovskite phases: (*Mg, Fe*) post-perovskite, Al_2O_3 post-perovskite. At depths below 1700 km free silica is presented as seifertite and above this depth it is presented as stishovite [e.g. *Murakami et al.*, 2003; *Tsuchiya et al.*, 2004a].

We generate many hypothetical mineral assemblages whose temperature and composition are chosen at random within the thermo-chemical ranges using a Monte Carlo procedure. The boundaries, listed in Table 2, are larger because the average thermo-chemical structure of the mantle is relatively unknown. This also means that the contribution of the uncertainties of the mineral physics measurements on our results is negligible. Recent studies of the partition coefficient of iron between perovskite and ferropericlasite have restricted the range of this parameter relative to our chosen range [*Irifuene et al.*, 2010; *Narygina et al.*, 2011]. We find, however, that assuming narrower bounds than 0.001-2.0 for the partition coefficient does not change our thermo-chemical perturbations.

After generating a set of 100000 hypothetical reference (m_{ref}) and perturbed (m_{pert}) model mineral assemblages at each pressure, we calculate "synthetic" P-wave speed, S-

516 wave speed and density using the EOS of *Stixrude and Lithgow-Bertelloni* [2005]. For each pair of mixtures we define their "synthetic" variation as a finite difference:

$$d\ln m = \frac{m_{pert} - m_{ref}}{m_{ref}}$$

517 where m is v_p , v_s and ρ [*Cobden et al.*, submitted]. Such anomalies reflect the underlying anomalies in temperature T and composition C :

$$d\ln m = \frac{\partial m}{\partial T} dT + \frac{\partial m}{\partial C} dC \quad (10)$$

518 where C represents the minerals of the assemblage.

519 We impose the constraint that elastic bulk and shear moduli as well as density in the reference models must fit PREM within 1%. We provide thus a higher consistency between
520 mineralogical and seismic models since the latter are expressed by relative deviations from
521 PREM. The uncertainty bounds of 1%, however, allow us to accept still a broad range of
522 thermo-chemical mixtures, implying that a physical reference model for temperature and
523 composition is poorly constrained [*Deschamps and Trampert*, 2004; *Cobden et al.*, 2009].
524 Furthermore, the choice of PREM does not affect our results. *Deschamps and Trampert*
525 [2003, 2004] and *Cobden et al.* [2009] verified that with ak135 [*Kennett et al.*, 1995] rather
526 than PREM the estimates of thermal and compositional structures in the lower mantle
527 are not significantly different. Indeed, seismic velocities of different reference models only
528 show discrepancies from each other when the uncertainty bounds are less than 1%.
529

530 To avoid unrealistically large thermo-chemical variations, we restrict the chemical composition of the perturbed model with respect to the reference model by imposing two
531 conditions. First, the (Pv+pPv)-component of the perturbed model falls within 12% of
532

the (Pv+pPv)-component of the associate reference model. Second, the volume of iron of the perturbed model is within 6% of Fe of its reference model.

To find which compositions agree with the seismic data, we simultaneously fit variations of compressional wave-speed, shear wave-speed and density in the hypothetical thermo-chemical models to those of tomographic models which are therefore the observed data. To perform this, we use the Metropolis-Hastings Algorithm, described in Section 4. In this case the random samples are the hypothetical mineralogical values of $dlnv_p$, $dlnv_s$ and $dln\rho$; and the target distribution is given by the seismic *pdfs* of $dlnv_p$, $dlnv_s$ and $dln\rho$. A significant advantage of Metropolis Algorithm is the possibility to re-sample the probability distribution of accepted hypothetical models to obtain a close fit to the seismic *pdf*.

We evaluate thermo-chemical variations at six depths in the lower mantle. At a given depth our tomographically-inferred seismic variations in a grid node are described by a Gaussian *pdf* centered around a mean with a specific width, as explained in Section 4. Thus, we draw 10000 random normally-distributed deviates from the mean within one standard deviation and compare them with the 100000 hypothetical mineralogical assemblages. The Metropolis algorithm accepts those hypothetical reference- and perturbed-model assemblages whose $dlnv_p$, $dlnv_s$ and $dln\rho$ fit the seismic observations simultaneously. Figure 8 illustrates how the Metropolis Algorithm works. The left-hand column shows the normalized probability distribution of compressional wave-speed, shear wave-speed and density variations for the 100000 hypothetical models and the 10000 random deviates from the seismologically observed 1-D *pdf* at one node. Only 5% of the hypothetical mixtures fall within the observed *pdf* and therefore these ones tend to be accepted

after applying the Metropolis algorithm. Their probability distribution is then re-sampled to obtain a closer fit to the seismic distribution (right-hand column of Figure 8).

6. Thermal and compositional model

Using the Metropolis Algorithm, we determine the assemblages that best explain the seismic heterogeneity out of a broad range of thermo-chemical models. Temperature and chemical variation images derived from this approach fit the seismic structures shown in Figures 3 and 4. We calculate changes in iron, silicate and post-perovskite in the lower mantle. Changes in iron are variations in the total volume of all Fe-bearing minerals (Fe-Pv, Fe-pPv and FeO); silica perturbations are perturbations in Si-bearing post-perovskites and perovskites and free SiO_2 ; and changes in post-perovskite are expressed as percentage of pPv with respect to the sum (pPv+Pv). Mapping variations of other elements, such as aluminium and calcium, might be important in understanding the origin of seismic anomalies [Irfune and Tsuchiya, 2007]. However, we find that in our calculations these effects on seismic properties are of second order compared to the other compositional variations and in the interests of brevity are not shown here.

From a set of 100000 hypothetical thermo-chemical models, the Metropolis Algorithm extracts between 3000 and 8000 models at the top of the lower mantle and up to 13000 models in the D'' region.

Maps of lateral changes in temperature and volumetric fraction of iron, silicates and post-perovskite at six depths of the deep interior are shown in Figure 9. Images of $dpPv$ and dFe are significantly correlated with S-wave speed and density perturbations of Figure 3, respectively. The amplitude of silicate variations in D'' region are small because seismic wave-speeds are more sensitive to post-perovskite than silicate anomalies. In D'' region,

a higher content of iron and silicate corresponds to a decrease in post-perovskite beneath the central Pacific Ocean and Africa. These features are linked to an increase in density and a negative v_s anomaly. At some depths, especially at the top of the lower mantle, the temperature and the total volume of silicate from the reference model are such that the corresponding variation is shifted towards positive numbers more frequently than to negative numbers. This is due to the constraints to fit PREM imposed on the reference models. Furthermore, our range of thermo-chemical variables are not necessary centred on values which fit PREM.

Figure 10 exhibits the lateral distribution of uncertainty of the thermo-chemical structure in Figure 9. The standard deviations of iron and post-perovskite variations are around 32% and 61% of the maximum anomaly amplitude respectively. Variations in temperature and silicate have large uncertainty in the lowermost mantle, between 85% and 92% of the largest size of the perturbations. This is explained by two effects. First, the greater the number of mineralogical end-members in the assemblage is, the stronger the trade-offs among them are. Moreover, at the perovskite to post-perovskite phase transition there are many possible combinations of mantle minerals which generate the same value of the derivative of the seismic variable (seismic wave-speed and density) with respect to the composition (equation (10)). The second reason is the wide range of permitted temperatures in the D'' region (because of approaching the thermal boundary layer above the liquid core) which may trade-off with changes in composition.

The overall correlation as a function of depth is a measure of how well the elastic properties of the thermo-chemical images fit those of the tomographic data (Figure 11). This parameter is higher than 0.85 at any depth and the lowest values are associated with

compressional wave-speed structure, especially at the top of the lower mantle. This suggests that a broad range of thermo-chemical variations fit the $dlnv_p$ tomographic model because P-wave velocity perturbations are less sensitive to thermo-chemical changes. Further, Figure 11 exhibits a high agreement between seismic and mineralogical models for density anomalies.

We also investigate the correlation as a function of depth between seismic and thermo-chemical properties as well as between thermal and compositional properties. The first and second column of Figure 12 show that compressional and shear wave-speed perturbations are negatively or poorly correlated with changes in temperature and composition at any layer, with the exception of the positive correlation between $dlnv_s$ and $dpPv$. This implies that: (1) robust P and S wave-speed variations alone are not sufficient to properly map thermo-chemical structure in the lower mantle; (2) the constraints provided by a consistent model of $dln\rho$ are essential to reduce the trade-off between temperature and composition. Density variations are strongly correlated with iron heterogeneity and negatively correlated with anomalies in temperature (third column of Figure 12). This suggests that changes in iron may strongly perturb the density structure in the deep interior. The fourth column of Figure 12 shows the probability distribution between thermal and chemical heterogeneities. Looking at these histograms it seems there is a discrepancy between them and the maps in Figure 9, e.g., the positive correlation between dSi and $dpPv$ and between dT and $dpPv$. The explanation for this is that computing the direct correlation between pairs of observables is not straightforward when there are many variables (dT , dSi , dFe and $dpPv$) biasing each other. In statistics this type of correlation

is known as multiple correlation. Furthermore, our models are *pdfs* and just interpreting one realization (i.e., Figure 9) is misleading.

It is interesting to check the temperature and composition of the subset of reference models accepted by the Metropolis Algorithm. Since they fall within 1% of PREM, the average of their thermo-chemical properties should, in principle, be consistent with a reference composition. This is also confirmed by the small lateral variation of the reference mixtures. In Figure 13 we plot the temperature and composition profile of accepted reference models, together with their standard deviation, as a function of depth. The uppermost lower mantle in our model contains $25 \pm 6\%$ of (Mg,Fe)O and 60-80% of silicate where the content of SiO_2 is less than 2% and $CaSiO_3$ is between 2 and 8%. The D'' region is dominated by the post-perovskite phase whose percentage is $65 \pm 20\%$ within the sum (Pv+pPv); $CaSiO_3$ and SiO_2 (not plotted) remain almost constant and the volume of (Mg,Fe)O is $18 \pm 10\%$. Therefore, the small volume of SiO_2 and $CaSiO_3$ suggests that there is not a dominant basaltic component in the deep interior. Moreover, in the D'' region the temperature is 2485 ± 357 K at 2600 km and 2900 ± 371 K at the CMB. Experimental and theoretical studies found that the stability regime of post-perovskite is 2500 K at a pressure between 120 and 136 GPa [Murakami *et al.*, 2004; Oganov and Ono, 2004]. Due to the large uncertainty in the Clapeyron slope, our thermal calculations are compatible with such predictions. Moreover, the temperature at the CMB has been estimated to lie between 2700 and 4200 K [e.g. Williams, 1998; van der Hilst *et al.*, 2007; Cobden *et al.*, submitted], in agreement with the temperature predicted by this work.

Results shown so far are based on the hypothesis of the coexistence of (laterally varying) Pv and pPv at the D'' region. In the next subsection we will discuss different scenarios for the lowermost mantle.

6.1. The effect of post-perovskite in the D'' region

To analyse the effect of post-perovskite in explaining the observed seismic structures, we estimate thermo-chemical anomalies for three possible scenarios in the lowermost mantle. First, we assume the coexistence of Pv and pPv (Section 6). In the second case, (Mg,Fe,Al)-perovskite is completely converted into post-perovskite (pPv-model) and, thus, perovskite is only present in the form of $CaSiO_3$. In the third scenario post-perovskite is excluded from the thermodynamic calculations (pPv-free model). We notice that the time to compute 100000 hypothetical pPv-free mixtures is 3 times (or more) longer than the time for pPv-bearing models due to the difficulty in finding a pPv-free assemblage which fits PREM.

For the three cases the reference state is not the same (Figure 13) and therefore we plot the anomalies with zero-mean to make a meaningful comparison between these hypothetical D'' regions. Figure 14 shows images of dT , dFe , $dSilicate$ and $dpPv$ (if present) at 2600 and 2891 km depth. In Case 2 and 3 the variation of the content in post-perovskite is $\pm 100\%$ and 0% , respectively. Looking at these maps, we notice that the variation of (p)Pv does not change the anomalies in iron, whereas it affects the amplitude of the thermal anomalies which are higher in the second case. However, the pattern of the maps does not vary significantly in the three hypothetical lowermost mantles.

The thermal and chemical profiles of the reference thermo-chemical models for a pPv-scenario as well as a pPv-free scenario appear to be meaningful (Figure 13). In the pPv-

case the average temperature is (3313 ± 378) K in the D'' region, perhaps too high for the occurrence of pPv. In the pPv-free mantle, the temperature and silicates have a smaller standard deviation, as expected. The perovskite into post-perovskite phase transition is associated with a high $dlnv_s$ and implies a trade-off between temperature and composition. Such a trade-off explains the large uncertainty in the lateral compositional variations of the pPv-bearing scenarios.

Although accepting one of the three proposed scenarios and rejecting the other two is meaningless due to the large uncertainty of thermo-chemical variations, we favour a pPv-bearing scenario in line with the statistical analysis of *Cobden et al.* [submitted].

7. Discussion & Conclusion

Earlier probabilistic tomography works produced long wavelength models of seismic wave-speeds and density from the inversion of surface wave and free oscillation measurements [Resovsky and Trampert, 2003; Trampert et al., 2004]. In this study we implemented body waves into probabilistic tomography. We imaged seismic structures by inverting two-way vertical travel-time residual maps, normal-mode splitting functions and surface wave phase velocity maps locally for depth. We used a new catalogue of splitting functions computed from the latest large earthquakes and, hence, characterized by a better data coverage. The robustness of these measurements ensures strong constraints on seismic velocities and, especially, on density structures. The inclusion of body wave measurements improves the depth resolution in the lowermost mantle.

After we constructed independent maps of density and wave-speeds, we used them to discriminate which temperature variations and mineralogical assemblages can explain the seismic data. Using the latest mineral physics data and the method of *Stixrude and*

Lithgow-Bertelloni [2005], together with the Metropolis Algorithm, we mapped anomalies in temperature and composition. The robustness of the thermal and chemical perturbation model from this approach is indicated by the good agreement between our results and those of *Cobden et al.* [submitted]. Although they are obtained by a similar method for constraining thermo-chemical variations, the seismic data used to fit mineralogical models are different. In *Cobden et al.* [submitted] tomographically-inferred data are variations of compressional wave-speed and shear wave-speed obtained from the lateral inversion of P_{diff} and S_{diff} measurements, suitable phases to image the CMB. Our observed seismic data are given by $dlnv_p$, $dlnv_s$ and $dln\rho$ from the depth inversion of body waves, surface waves and normal-modes using a global space approach. In spite of this, our conclusions agree well with those of the stochastic analysis of *Cobden et al.* [submitted] which have found a closer fit to the seismic data for a pPv-bearing CMB than for a pPv-free CMB.

The highlight of this work is the mapping of lateral variations in post-perovskite in the D'' region. Those images suggest a geographical candidate for the Pv to pPv conversion (i.e., the circum-Pacific belt) based on not only constraints from mineral physics but also tomographic data. The main conclusion from the present paper (i.e., the existence of significant chemical heterogeneities in the deep interior) disagrees with those works which point out primarily a thermal origin of the seismic heterogeneity in the deep interior [e.g. *Steinberger and Calderwood*, 2006; *Kawai and Tsuchiya*, 2009; *Simmons et al.*, 2010]. Such a discrepancy can be due to two factors. Firstly, imposing prior constraints on the mineralogical mixtures (such as explaining as much as possible with temperature) can drive the search of the best assemblage of minerals towards different seismic and thermo-chemical perturbation models. We did not make any prior assumption on the

711 distribution of the temperature and composition which thus assume values within broad
 712 thermo-chemical ranges, including also unrealistic mineralogical assemblages rejected then
 713 by the Metropolis Algorithm. Secondly, if seismic data $dlnv_p$, $dlnv_s$ and $dln\rho$ are not
 714 independent, imposing scaling factors together with regularization can have a large effect
 715 on the solution. The use of a Monte Carlo approach ensures the least possible guidance
 716 of the solution. Both of these explain the apparent discrepancy between our most likely
 717 model and GyPSuM [*Simmons et al.*, 2010]. However, when we consider the full *pdfs* of
 718 our model, GyPSuM falls within the distribution and thus is one possible solution.

719 The lateral thermo-chemical variations shown in this paper have an uncertainty which,
 720 at worst, is 80% of the maximum amplitude of the variations. These standard deviations
 721 combine the mineral physics data available nowadays with the current lateral and radial
 722 resolution of tomographic models. In this framework the most important further devel-
 723 opment is to use a finer lateral resolution for tomographic models estimated by a Monte
 724 Carlo approach. This means including both the cross-coupling splitting functions and the
 725 harmonic degrees higher than 6. It will be interesting to see if the odd degrees of the
 726 Earth structure might reveal new features in imaging wave-speeds and density as well as
 727 thermal and compositional anomalies, as shown by *Bull et al.* [2009]. Furthermore, a lat-
 728 eral resolution lower than ~ 3300 km can display smaller scale features (e.g., subducting
 729 slabs and plumes) which have been recognized by most tomographic studies in the last
 730 twenty years. This will also allow us to image small scale features of thermo-chemical
 731 structures.

732 Last but not least, from a computational viewpoint, establishing efficient sampling
 733 techniques for highly-dimensional spaces (32-D in the present case) will allow us to exploit

a finer radial parametrization. Producing a global tomographic model from a Monte Carlo approach is time consuming because the curse of dimensionality quickly limits the size of the inverse problem [*Curtis and Lomax*, 2001].

8. Appendix 1

For completeness, we display seismic maps of P-wave speed, S-wave speed and density in the upper mantle (Figure 15), together with their uncertainty (Figure 16). Although the upper mantle is more heterogeneous than the lower mantle and cannot be adequately described by a low order expansion, the comparison between our model and S20RTS, P12, GyPSuM and RT03 confirm the same agreement we observed in Subsection 4.1 (Figures 17 and 18).

Acknowledgments.

We thank M. Sambridge for providing the Neighbourhood Algorithm code. This work and the 128-processor cluster used for several calculations were financed by the Dutch National Science Foundation under the grant number NWO:VICI865.03.007. Other computational resources for this work were provided by the Netherlands Research Center for Integrated Solid Earth Science (ISES 3.2.5 High End Scientific Computation Resources). A. Deuss was funded by the European Research Council under the European Community’s Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement number 204995. J. Ritsema was funded by NSF grant EAR/1015172.

References

Anderson, O., E. Schreiber, R. Liebermann, and N. Soga (1984), Some elastic constant data on minerals relevant to geophysics, *J. Geophys. Res.*, *89*, 5953–5986.

- Antolik, M., Y. J. Gu, and G. Ekstrom (2003), A new joint model of compressional and shear velocity in the Earth's mantle, *Geophys. J. Int.*, *153*, 443–466.
- Becker, T. W., and L. Boschi (2002), A comparison of tomographic and geodynamic mantle models, *Geochem. Geophys. Geosys.*, *3*, 2001GC000,168.
- Beghein, C., J. S. Resovsky, and J. Trampert (2002), P and S tomography using normal mode and surface wave data with a neighborhood algorithm, *Geophys. J. Int.*, *149*, 646–658.
- Bijwaard, H., W. Spakman, and E. Engdahl (1998), Closing the gap between regional and global travel-time tomography, *J. Geophys. Res.*, *103*, 30,055–30,078.
- Boschi, L., and A. M. Dziewonski (2000), Whole Earth tomography from delay times of P, PcP, and PKP phases: lateral heterogeneities in the outer core or radial anisotropy in the mantle?, *J. Geophys. Res.*, *105*, 13,675–13,696.
- Bozdog, E., and J. Trampert (2008), On crustal corrections in surface wave tomography, *Geophys. J. Int.*, *172*, 1066–1082.
- Breger, L., and B. Romanowicz (1998), Thermal and chemical 3D heterogeneity in D'', *Science*, *282*, 718–720.
- Bull, A. L., A. K. McNamara, and J. Ritsema (2009), Synthetic tomography of plume clusters and thermo-chemical piles, *Earth Planet. Sci. Lett.*, *278*, 152–162.
- Bullen, K. E. (1949), Compressibility-pressure hypothesis and the Earth's interior, *MNRAS Geophys. (Suppl.5)*, pp. 355–368.
- Cadek, O., and L. Fleitout (1999), A global geoid model with imposed plate velocities and partial layering, *J. Geophys. Res.*, *104*, 29,055–29,075.

Cobden, L., S. Goes, M. Ravenna, E. Styles, F. Cammarano, K. Gallagher, and J. A. Connolly (2009), Thermo-chemical interpretation of 1-D seismic data for the lower mantle: the significance of non-adiabatic thermal gradients and compositional heterogeneity, *J. Geophys. Res.*, *114*, doi:10.1029/2008JB006,262.

Cobden, L., I. Mosca, J. Trampert, and J. Ritsema (submitted), On the likelihood of post-perovskite at the core-mantle boundary: a statistical interpretation of seismic observations, *Phys. Earth Planet. Int.*

Curtis, A., and A. Lomax (2001), Prior information, sampling distributions and the curse of dimensionality, *Geophysics*, *66*, 372–378.

Dahlen, F. A., and J. Tromp (1998), *Theoretical global seismology*, Princeton University Press, NJ.

de Wit, R. W., J. Trampert, and R. D. van der Hilst (2012), Towards quantifying uncertainties in travel-time tomography using the null space shuttle, *J. Geophys. Res.*, pp. doi: 10.1029/2011JB008,754, in press.

Deschamps, F., and P. J. Tackley (2008), Searching for models of thermo-chemical convection that explain probabilistic tomography - I. Principles and influence of rheological parameters, *Phys. Earth Planet. Int.*, *171*, 357–373.

Deschamps, F., and P. J. Tackley (2009), Searching for models of thermo-chemical convection that explain probabilistic tomography - II. Influence of physical and compositional parameters, *Phys. Earth Planet. Int.*, *176*, 1–18.

Deschamps, F., and J. Trampert (2003), Mantle tomography and its relation to temperature and composition, *Phys. Earth Planet. Int.*, *140*, 277–291.

- 798 Deschamps, F., and J. Trampert (2004), Towards a lower mantle reference temperature
799 and composition, *Earth Planet. Sci. Lett.*, *222*, 161–175.
- 800 Deschamps, F., J. Trampert, and P. J. Tackley (2007), Thermo-chemical structure of
801 the lower mantle: seismological evidence and consequences for geodynamics, in *Super-*
802 *plumes: Beyond plate tectonics*, pp. 293–320, Springer, The Netherlands.
- 803 Deuss, A., J. Ritsema, and H. J. van Heijst (2011), Splitting function measurements for
804 Earth’s longest period normal-modes using recent earthquakes, *Geophys. Res. Lett.*, *38*,
805 L04,303, doi: 10.1029/2010JB0046,115.
- 806 Dziewonski, A. M., and D. L. Anderson (1981), Preliminary Reference Earth Model, *Phys.*
807 *Earth Planet. Int.*, *25*, 297–356.
- 808 Ford, S. R., E. G. Garnero, and A. K. McNamara (2006), A strong lateral shear velocity
809 gradient and anisotropy heterogeneity in the lowermost mantle beneath the southern
810 Pacific, *J. Geophys. Res.*, *111*, B03,306, doi: 10.1029/2004JB0003,574.
- 811 Forte, A. M., and J. X. Mitrovica (2001), Deep-mantle high-viscosity flow and thermo-
812 chemical structure inferred from seismic and geodynamic data, *Nature*, *410*, 1049–1056.
- 813 Forte, A. M., R. L. Woodward, and A. M. Dziewonski (1994), Joint inversions of seismic
814 and geodynamic data for models of three-dimensional mantle heterogeneity, *Geophys. J.*
815 *Res.*, *99*, 21,857–21,877.
- 816 Fukao, Y., S. Widiyantoro, and M. Obayashi (2001), Stagnant slabs in the upper and lower
817 mantle transition region, *Rev. Geophys.*, *39*, 291–323.
- 818 Garnero, E. G., and A. K. McNamara (2008), Structure and dynamics of Earth’s lower
819 mantle, *Science*, *320*, 626–628.

Grand, S. P. (2002), Mantle shear wave tomography and the fate of subducted slabs,
Philos. Trans. R. Soc. Lond. A, *360*, 2475–2491.

Gu, Y. J., A. M. Dziewonski, and G. Ekstrom (2001), Models of the mantle shear wave
velocity and discontinuities in the pattern of lateral heterogeneities, *J. Geophys. Res.*,
106, 11,169–11,199.

He, Y., L. Wen, and T. Zheng (2006), Geographic boundary and shear wave velocity
structure of the Pacific anomaly near the core-mantle boundary beneath western Pacific,
Earth Planet. Sci. Lett., *244*, 302–314.

Irifune, T., and T. Tsuchiya (2007), Mineralogy of the Earth: Phase transitions and
mineralogy of the lower mantle, in *Treatise on Geophysics - Mineral physics*, vol. 2, pp.
33–62, Elsevier, Amsterdam.

Irifune, T., T. Shinmei, C. A. McCammon, N. Miyajima, D. C. Rubie, and D. J. Frost
(2010), Iron partitioning and density changes of pyrolite in Earth’s lower mantle, *Sci-*
ence, *327*, 193–195.

Ishii, M., and J. Tromp (1999), Normal-mode and free-air gravity constraints on lateral
variations in velocity and density of the Earth’s mantle, *Science*, *285*, 1231–1236.

Karason, H., and R. D. van der Hilst (2001), Tomographic imaging of the lowermost
mantle with differential times of refracted and diffracted core phases (PKP, P_{diff}), *J.*
Geophys. Res., *106*, 6569–6587.

Karato, S.-I., and B. B. Karki (2001), Origin of lateral variation of seismic wave velocities
and density in the deep mantle, *Geophys. J. Res.*, *106*, 21,771–21,783.

Kawai, K., and T. Tsuchiya (2009), Temperature profile in the lowermost mantle from
seismological and mineral physics joint modeling, *Proceedings of the National Academy*

of *Sciences of the USA.*, 106, 22,119–22,123.

Kellogg, L. H., B. H. Hager, and R. D. van der Hilst (1999), Compositional stratification in the deep mantle, *Science*, 283, 1881–1884.

Kendall, J. M., and P. G. Silver (1996), Constraints on seismic anisotropy on the nature of the lowermost mantle, *Nature*, 381, 409–412.

Kennett, B. L., E. Engdahl, and R. Buland (1995), Constraints on seismic velocities in the Earth from travel-times, *Geophys. J. Int.*, 150, 37–44.

Kennett, B. L., S. Widiyantoro, and R. D. van der Hilst (1998), Joint seismic tomography for bulk sound and shear wave speed in the Earth’s mantle, *J. Geophys. Res.*, 103, 12,469–12,493.

Lay, T. (2007), Deep Earth structure - Lower mantle and D”, in *Treatise on Geophysics - Seismology and Structure of the Earth*, vol. 1, pp. 619–654, Elsevier, Amsterdam.

Masters, G., G. Laske, H. Bolton, and A. M. Dziewonski (2000), The relative behavior of shear velocity, bulk sound speed, and compressional velocity in the mantle: implications for chemical and thermal structure, in *Earth’s Deep Interior: Mineral Physics and Tomography from the Atomic to the Global scale*, vol. 117, pp. 63–87, AGU, Washington.

Megnin, C., and B. Romanowicz (2000), The 3D shear velocity structure of the mantle from the inversion of body, surface and higher mode waveforms, *Geophys. J. Int.*, 143, 709–728.

Metropolis, N., A. W. Rosenbluth, M. N. Teller, and E. Teller (1953), Equation of state calculations by fast computing machines, *J. Chem. Phys.*, 1, 1087–1092.

Mooney, W., G. Laske, and G. Masters (1998), Crust5.1: a global crustal model at 5x5 deg, *J. Geophys. Res.*, 103, 727–747.

- 866 Mosca, I. (2010), *Probabilistic tomography using body wave, normal-mode and surface*
867 *wave data*, Ph.D. thesis, Utrecht University , the Netherlands.
- 868 Mosca, I., and J. Trampert (2009), Path-average kernels for long wavelength travel-time
869 tomography, *Geophys. J. Int.*, *177*, 639–650.
- 870 Mosegaard, K., and A. Tarantola (1995), Monte Carlo sampling of solutions to inverse
871 problems, *J. Geophys. Res.*, *100(B7)*, 12,431–12,447.
- 872 Murakami, M., K. Hirose, S. Ono, and Y. Ohishi (2003), Stability of $CaCl_2$ -type and
873 PbO_2 -type SiO_2 at high pressure and temperature determined by in-situ X-ray mea-
874 surements., *Geophys. Res. Lett.*, *30*, doi:10.1029/2002GL016,722.
- 875 Murakami, M., K. Hirose, K. Kawamura, N. Sata, and Y. Ohishi (2004), Post-perovskite
876 phase transition in $MgSiO_3$, *Science*, *304*, 855–858.
- 877 Nakagawa, T., and P. J. Tackley (2004), Thermo-chemical structure in the mantle arising
878 from a three-component convective system and implications for geochemistry, *Phys.*
879 *Earth Planet. Int.*, *146*, 125–138.
- 880 Narygina, O., L. S. Dubrovinsky, H. Samuel, C. A. McCammon, I. Kantor, K. Glazyrina,
881 S. Pascarelli, G. Aquilanti, and V. B. Prakapenkac (2011), Chemically homogeneous
882 spin transition zone in Earth’s lower mantle, *Phys. Earth Planet. Int.*, *185*, 107–111.
- 883 Ni, S., and D. V. Helmberger (2003a), Seismological constraints on the South African
884 super-plume could be the oldest distinct structure of the Earth, *Earth Planet. Sci.*
885 *Lett.*, *206*, 119–131.
- 886 Ni, S., and D. V. Helmberger (2003b), Ridge-like lower mantle structure beneath South
887 Africa, *J. Geophys. Res.*, *108*, B2094, doi:10.1029/2001JB001,545.

- 888 Oganov, A. R., and S. Ono (2004), Theoretical and experimental evidence for a post-
889 perovskite phase transition of $MgSiO_3$ in Earth's D'' layer, *Nature*, *430*, 445–448.
- 890 Ohta, K., K. Hirose, T. Lay, N. Sata, and Y. Ohishi (2008), Phase transition in pyrolite
891 and MORB at lowermost mantle conditions: Implications for a MORB-rich pile above
892 the core-mantle boundary, *Earth Planet. Sci. Lett.*, *267*, 107–117.
- 893 Panning, M. P., and B. Romanowicz (2004), Inferences of flow at the base of the Earth's
894 mantle based on seismic anisotropy, *Science*, *303*, 351–353.
- 895 Resovsky, J. S., and M. H. Ritzwoller (1998), New and refined constraints on three-
896 dimensional Earth structure from normal modes below 3mHz, *J. Geophys. Res.*, *103*,
897 783–81.
- 898 Resovsky, J. S., and J. Trampert (2003), Using probabilistic seismic tomography to test
899 mantle velocity-density relationships, *Earth Planet. Sci. Lett.*, *215*, 121–134.
- 900 Rickwood, P., and M. Sambridge (2006), Efficient parallel version using the
901 Neighbourhood Algorithm, *Geochem. Geophys. Geosyst.*, *7*, Q11,001, doi:
902 10.1029/2006GC001,246.
- 903 Ritsema, J., and H. J. van Heijst (2002), Constraints on the correlation of P- and S-wave
904 velocity heterogeneity in the mantle from P, PP, PPP and PKPab traveltimes, *Geophys.*
905 *J. Int.*, *149*, 482–489.
- 906 Ritsema, J., S. Ni, D. V. Helmberger, and P. Crotwell (1998), Evidence for strong shear
907 velocity reductions and velocity gradients in the lower mantle beneath Africa, *Geophys.*
908 *Res. Lett.*, *25*, 4245–4248.
- 909 Ritsema, J., H. J. van Heijst, and J. H. Woodhouse (1999), Complex shear wave velocity
910 structure imaged beneath Africa and Iceland, *Science*, *286*, 1925–1928.

- 911 Ritsema, J., H. J. van Heijst, and J. H. Woodhouse (2004), Global transition zone tomog-
912 raphy, *J. Geophys. Res.*, *109*, B02,302, doi: 10.1029/2003JB002,610.
- 913 Ritsema, J., H. J. van Heijst, J. H. Woodhouse, and A. Deuss (2009), Long-period body
914 wave travel-times through the crust: implication for crustal corrections and seismic
915 tomography, *Geophys. J. Int.*, *179*, 1255–1261.
- 916 Ritsema, J., A. Deuss, and H. J. van Heijst (2011), S40RTS: a degree-40 shear-velocity
917 model for the mantle from new Rayleigh wave dispersion, teleseismic travel-time and
918 normal-mode splitting function measurements, *Geophys. J. Int.*, *184*, 1223–1236.
- 919 Robertson, G. S., and J. H. Woodhouse (1996), Ratio of the relative S to P velocity
920 heterogeneity in the lower mantle, *J. Geophys. Res.*, *101*, 20,041–20,052.
- 921 Romanowicz, B. (2001), Can we resolve 3D density heterogeneity in the lower mantle?,
922 *Geophys. Res. Lett.*, *28*, 1107–1110.
- 923 Romanowicz, B. (2003), Global mantle tomography: Progress status in the past 10 years,
924 *Ann. Rev. of Earth Planet. Sci.*, *31*, 303–328.
- 925 Romanowicz, B., and Y. Gung (2002), Superplumes from the core-mantle boundary to
926 the lithosphere: Implications for the heat flux, *Science*, *296*, 513–516.
- 927 Saltzer, R. L., R. D. van der Hilst, and H. Karason (2001), Comparing P and S wave
928 heterogeneity in the mantle, *Geophys. Res. Lett.*, *28*, 1335–1338.
- 929 Sambridge, M. (1999a), Geophysical inversion with a neighbourhood algorithms I- Search-
930 ing a parameter space, *Geophys. J. Int.*, *138*, 479–494.
- 931 Sambridge, M. (1999b), Geophysical inversion with a neighbourhood algorithms II- Ap-
932 praising the ensemble, *Geophys. J. Int.*, *138*, 727–746.

- 933 Simmons, N. A., A. M. Forte, and S. P. Grand (2009), Joint seismic, geodynamic and
934 mineral physical constraints on three-dimensional mantle heterogeneity: Implications
935 for the relative importance of thermal versus compositional heterogeneity, *Geophys. J.*
936 *Int.*, *177*, 1284–1304.
- 937 Simmons, N. A., A. M. Forte, L. Boschi, and S. P. Grand (2010), GyPSuM: A joint
938 tomographic model of mantle density and seismic wave speeds, *J. Geophys. Res.*, *115*,
939 B06,306, doi: 12.310/2010JB007,631.
- 940 Steinberger, B., and A. Calderwood (2006), Models of large-scale viscous flow in the
941 Earth’s mantle with constraints from mineral physics and surface observations, *Geophys.*
942 *J. Int.*, *167*, 1461–1481.
- 943 Stixrude, L., and C. Lithgow-Bertelloni (2005), Thermodynamics of mantle minerals I:
944 physical properties, *Geophys. J. Int.*, *162*, 610–632.
- 945 Stixrude, L., and C. Lithgow-Bertelloni (2011), Thermodynamics of mantle minerals II:
946 Phase equilibria, *Geophys. J. Int.*, *184*, 1180–1213.
- 947 Su, W.-J., and A. M. Dziewonski (1997), Simultaneous inversion for 3D variations in shear
948 and bulk velocity in the mantle, *Phys. Earth Planet. Int.*, *100*, 135–156.
- 949 Tanaka, S. (2002), Very low shear wave velocity at the base of the mantle under the South
950 Pacific superswell, *Earth Planet. Sci. Lett.*, *203*, 879–893.
- 951 Tarantola, A. (1987), *Inverse problem theory*, Elsevier, Amsterdam.
- 952 Toh, A., B. Romanowicz, Y. Capdeville, and N. Takeuchi (2005), 3-D effects of sharp
953 boundaries at the borders of the African and Pacific super-plumes: Observations and
954 modelling, *Earth Planet. Sci. Lett.*, *233*, 237–253.

- Trampert, J., and R. D. van der Hilst (2005), Towards a quantitative interpretation of global seismic tomography, in *Earth's Deep Mantle: Structure, Composition, and Evolution*, vol. 160, pp. 47–62, AGU, Washington.
- Trampert, J., P. Vacher, and N. Vlaar (2001), Sensitivities on seismic velocities to temperature, pressure and composition in the lower mantle, *Phys. Earth Planet. Int.*, *124*, 255–267.
- Trampert, J., F. Deschamps, J. S. Resovsky, and D. Yuen (2004), Probabilistic tomography maps significant chemical heterogeneities in the lower mantle, *Science*, *306*, 853–856.
- Tsuchiya, T., R. Caracas, and J. Tsuchiya (2004a), First principles determination of the phase boundaries of high-pressure polymorphs of silica., *Geophys. Res. Lett.*, *31*, doi:10.1029/2004GL019,649.
- Tsuchiya, T., K. Tsuchiya, and K. Umemoto (2004b), Phase transition in $MgSiO_3$ perovskite in the Earth's lower mantle, *Earth Planet. Sci. Lett.*, *224*, 241–248.
- Tsuchiya, T., K. Tsuchiya, and K. Umemoto (2004c), Elasticity of post-perovskite $MgSiO_3$, *Geophys. Res. Lett.*, *31*, doi:10.1029/2004GL020,278.
- van der Hilst, R. D., S. Widiyantoro, and E. Engdahl (1997), Evidence for deep mantle circulation from global tomography, *Nature*, *386*, 578–584.
- van der Hilst, R. D., M. V. D. Hoop, P. Wang, S.-M. Shim, P. Ma, and L. Tenorio (2007), Seismostratigraphy and thermal structure of Earth's core-mantle boundary region, *Science*, *315*, 1813–1817.
- Vasco, D. W., and L. R. Johnson (1998), Whole Earth structure estimated from seismic arrival times, *J. Geophys. Res.*, *103*, 2633–2671.

Vinnik, L., L. Breger, and B. Romanowicz (1998), Anisotropic structures at the base of the mantle, *Nature*, *393*, 564–567.

Visser, K., J. Trampert, and B. L.N. (2008), Global anisotropic phase velocity maps for higher mode Love and Rayleigh waves, *Geophys. J. Int.*, *172*, 1016–1032.

Wentzcovitch, R. M., T. Tsuchiya, and K. Tsuchiya (2006), $MgSiO_3$ post-perovskite at D'' conditions, *Proc. Natl. Acad. Sci U.S.A.*, *103*, 504–546.

Williams, Q. (1998), The temperature constrast across D'', in *The core-mante boundary region*, pp. 73–81, AGU, Washington.

Woodhouse, J. H. (1974), Surface waves in a laterally varying layered structure, *Geophys. J. R. Astr. Soc.*, *37*, 461–490.

Woodhouse, J. H., and Y. K. Wong (1986), Amplitud, phase and path anomalies of mantle waves, *Geophys. J. R. Astron. Soc.*, *87*, 753–773.

Xu, W., C. Lithgow-Bertelloni, L. Stixrude, and J. Ritsema (2008), The effect of bulk composition and temperature on mantle seismic structure, *Earth Planet. Sci. Lett.*, *275*, 70–79.

Yoshizawa, K., and B. L. Kennett (2002), Determination of the influence zone for surface wave paths, *Geophys. J. Int.*, *149*, 440–453.

Zhao, D. (2001), Seismic structure and origin of hotspots and mantle plumes, *Earth Planet. Sci. Lett.*, *192*, 251–265.

Captions

Table 1: The first and the second column list the type (phase for body waves, mode for surface waves and normal-modes) and the third and the fourth column the number of data (different ray parameters for body waves and different frequencies for Rayleigh waves and spheroidal modes) used in the inversion.

Table 2: Depth, pressure, range of temperature on the left-hand side and range of parameters for hypothetical compositional mantle assemblages on the right-hand side. Pv indicates the (Mg, Fe)-perovskite; mw, the (Fe, Mg)O and pPv, the (Mg, Fe)-postperovskite.

Figure 1: Radial parametrization of the Earth’s interior from the surface to the core-mantle boundary with 10 natural cubic spline functions.

Figure 2: Sensitivity of body wave (blue lines), surface wave (black lines) and normal-mode (red lines) measurements to compressional wave-speed (top left-hand plot), shear wave-speed (top right-hand plot) and density (bottom left-hand plot). They are projected onto spline basis and plotted as a function of the spline knots. We scaled the partial derivatives (radial integral in equation (5)) by the corresponding maximum value of each dataset in order to plot them together.

Figure 3: Compressional wave-speed (first column), shear wave-speed (second column) and density (third column) perturbation maps at six depths, together with hotspots (green circles) and plate boundaries (green lines). Red (blue) regions denote slower (faster) than average seismic wave-speed and a decrease (increase) in density. Measurements were calculated for even degrees 0, 2, 4 and 6. Perturbations are given in percent with respect to PREM. We divided each layer into 264 nodes of $15^\circ \times 15^\circ$ area and estimated the mean and standard deviation in a grid node from the 1-D marginal *pdf* provided by the Neighbourhood Algorithm.

Figure 4: As Figure 3, but for the lateral distribution of the uncertainty of $dlnv_p$, $dlnv_s$ and $dln\rho$.

Figure 5: Normalized distribution of the correlation coefficient (the first column), root-mean-square (RMS) amplitude (the second column) and seismic ratios $\frac{dlnv_s}{dlnv_p}$ and $\frac{dln\rho}{dlnv_s}$ (the third column) as a function of depth between all possible models generated by the NA.

Figure 6: Probability density function of $dlnv_p$ (first column) and $dlnv_s$ (second column) at the grid node ($0^\circ, 180^\circ$) and at various depths. We binned 10000 normally distributed random deviates from the chosen *pdf* and built the black histograms. Red and green lines represent seismic variations at ($0^\circ, 180^\circ$) from P12/S20RTS and GyPSuM, respectively.

Figure 7: As Figure 6, but for the comparison of the pdf of $dlnv_p$ (first column), $dlnv_s$ (second column) and $dln\rho$ (third column) between this study (blue histograms) and those of *Resovsky and Trampert* [2003] (red histograms).

Figure 8: Comparison of normalized probability distributions of $dlnv_p$, $dlnv_s$ and $dln\rho$ between hypothetical mineralogical models (blue histograms) and seismic models (red histograms) at a grid node of 2891 km depth. Red distributions are given by 10000 random deviates from the *pdf* described by the mean and the standard deviation of Figures 3 and 4. On the left-hand column we binned 100000 thermo-chemical mixtures generated from the method of *Stixrude and Lithgow-Bertelloni* [2005] and thus constructed the blue distributions. On the right-hand column blue histograms are built from 5143 accepted thermo-chemical mixtures fitting seismic observations. Using the Metropolis Algorithm their distribution has been modified to fit that of the seismic data.

Figure 9: Lateral variations of temperature, iron, silicates and post-perovskite at various depths of the lower mantle. Temperature variations are expressed in Kelvin, while iron, perovskite

and post-perovskite perturbations are given in percent. Furthermore, variations are measured relative to thermochemical models which fit PREM to within 1%. Red regions denote negative chemical anomalies and a higher than average temperature, while blue regions are associated with positive chemical variations and a lower than average temperature. Maps show hotspots (green circles) as well as plate boundaries (green lines).

Figure 10: As Figure 9, but for the uncertainty of thermal and chemical anomalies.

Figure 11: Correlation coefficient as a function of depth between $dlnv_p$, $dlnv_s$ and $dln\rho$ of Figure 3 and $dlnv_p$, $dlnv_s$ and $dln\rho$ of thermo-chemical mixtures accepted by the Metropolis Algorithm and, thus, fitting the seismic data. All correlations are calculated for the statistically most likely seismic and thermochemical models.

Figure 12: Probability distribution of the correlation coefficient between seismic structure and thermo-chemical anomalies (first, second and third column) as well as between anomalies in temperature and composition (fourth column).

Figure 13: Profile of the reference temperature and volume of iron, silica and post-perovskite as a function of depth within their error-bars. These values are computed as the average of the temperatures and compositions of reference models which both fit PREM to within 1% and have associated perturbed models which fit the tomographic observations. At depths equal to or higher than 2600 km we studied three different scenario: the coexistence of Pv and pPv (black lines); the total conversion of Pv into pPv (green lines); and a pPv-free chemical variation model (red lines).

Figure 14: As Figure 9, but for thermal and chemical perturbations with zero-mean in the lowermost mantle. Maps describe three possible scenario of the D'' region: the coexistence of Pv

1063 and pPv (first and second rows); the total conversion of Pv into pPv (third and fourth rows);
1064 and the exclusion of pPv (fifth and sixth rows).

1065 **Figure 15:** As Figure 3, but for seismic variations of P-wave speed, S-wave speed and density
1066 at different depths in the upper mantle.

1067 **Figure 16:** As Figure 15, but for the lateral distribution of the uncertainty of $dlnv_p$, $dlnv_s$
1068 and $dln\rho$.

1069 **Figure 17:** As Figure 6, but for various layers in the upper mantle.

1070 **Figure 18:** As Figure 7, but for the layers (24-420) and (420-670) km in the upper mantle.

Body wave measurements			
Phase	# data	Phase	# data
P	11	S	3
P_{diff}	1	S_{diff}	1
PP	32	SS	39
		SSS	3
Surface wave measurements			
Mode	# data	Mode	# data
0	9	3	10
1	14	4	7
2	13		
Normal-mode measurements			
Mode	# data l=0,2	# data l=4	# data l=6
0	14	12	11
1	8	6	5
2	6	6	6
3	2	2	0
4	5	3	1
5	3	3	3

Table 1.

			% Pv+mw component	85-100
			% Pv+pPv within Pv+mw component	60-100
Depth [km]	Pressure [GPa]	Range of temperature [K]	% Fe within Pv+pPv component	0-20
950	36.345	1000-3000	Partition coefficient Fe-Mg within Pv+mw component	0.001-2.0
1200	47.854	1000-3000	pPv within Pv+pPv component	0-100
1550	64.510	1400-3150	Al_2O_3 within $Al_2O_3 + CaSiO_3 + SiO_2$ component	0-100
2000	86.928	1600-3300	SiO_2 within $Al_2O_3 + CaSiO_3 + SiO_2$ component	0-100
2600	118.991	1800-3500		
2891	135.750	2300-4800		

Table 2.

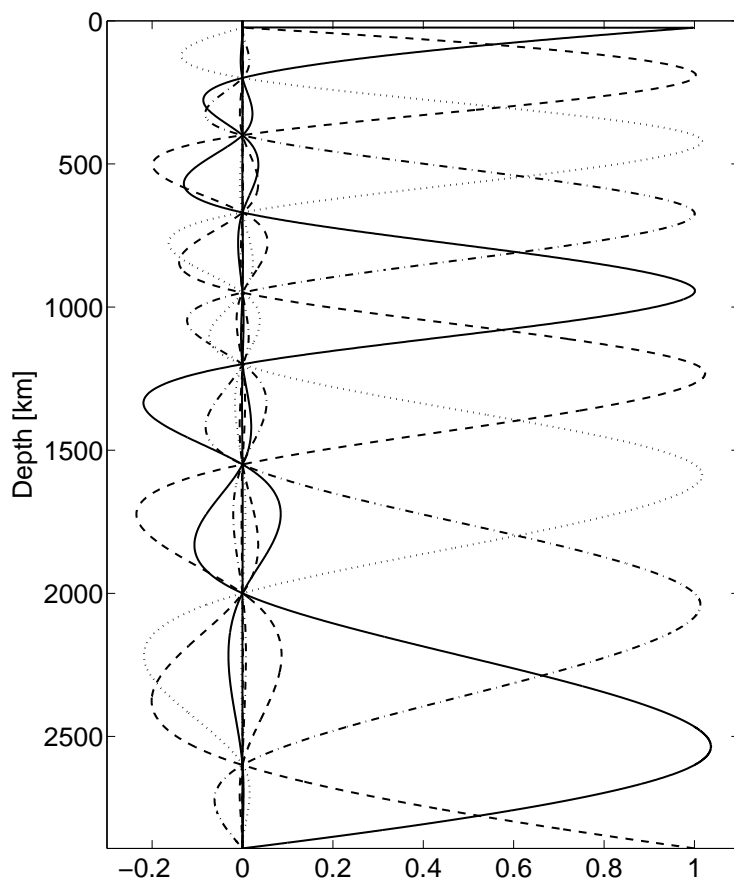


Figure 1.

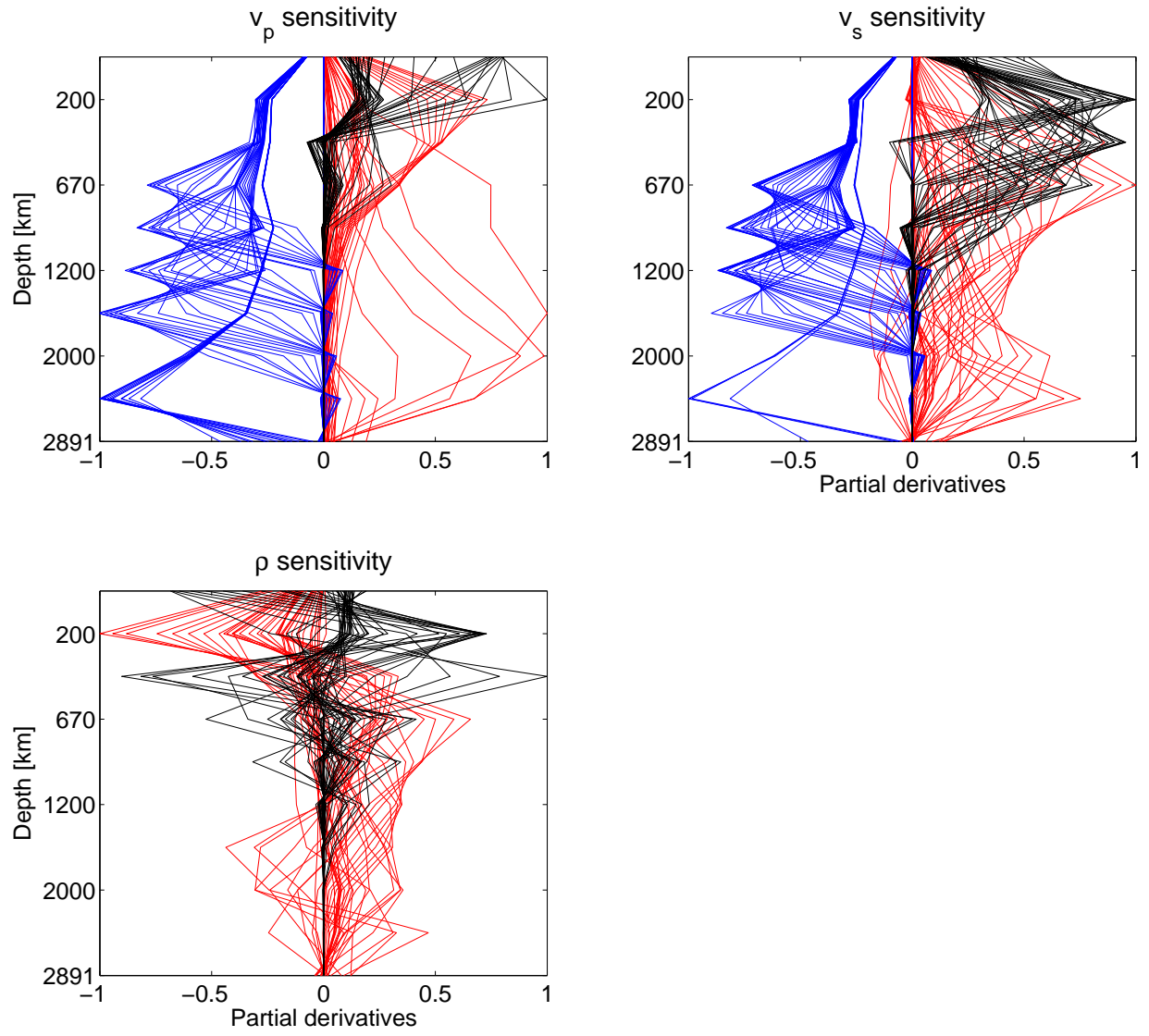


Figure 2.

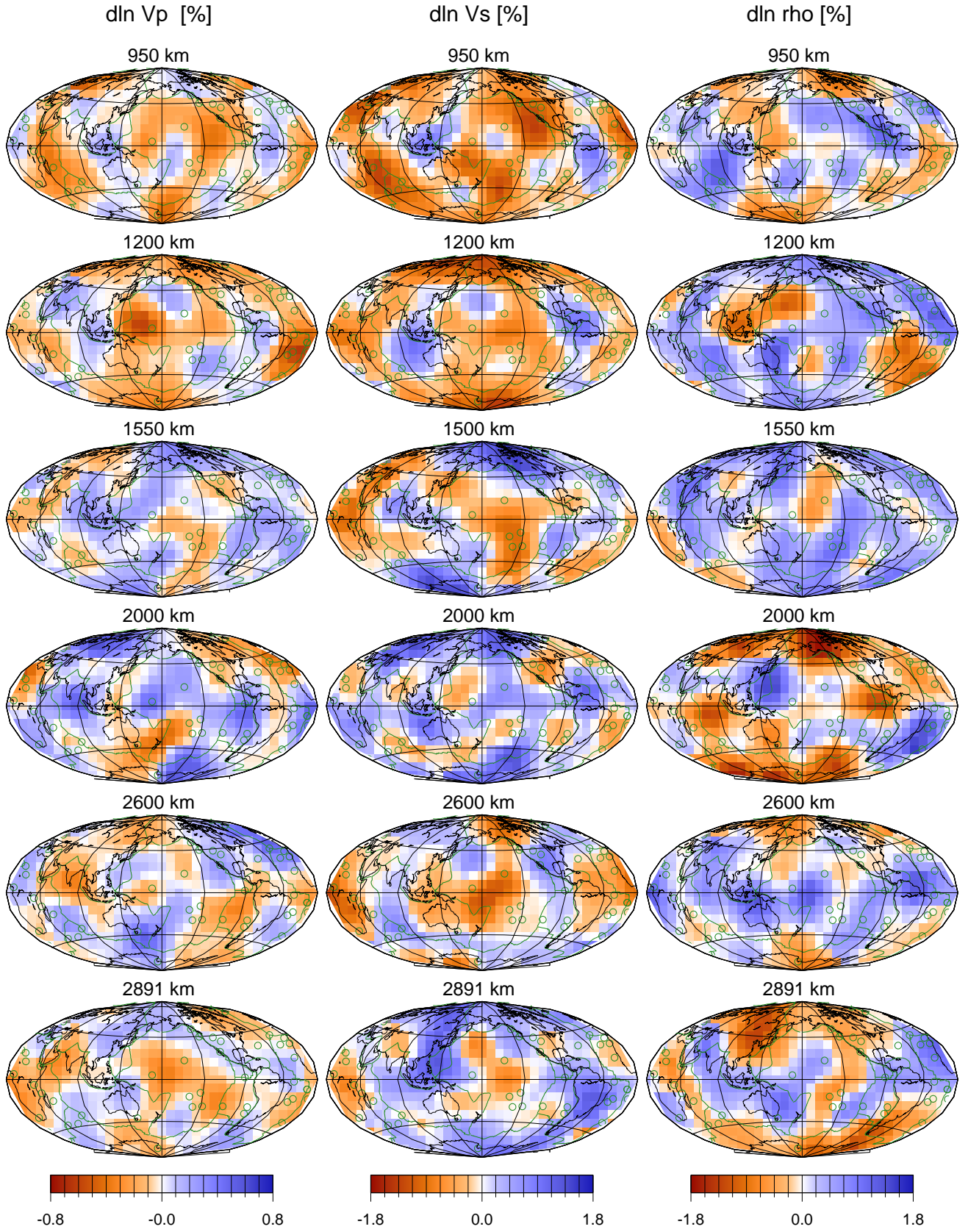


Figure 3.

February 10, 2012, 7:14pm

D R A F T

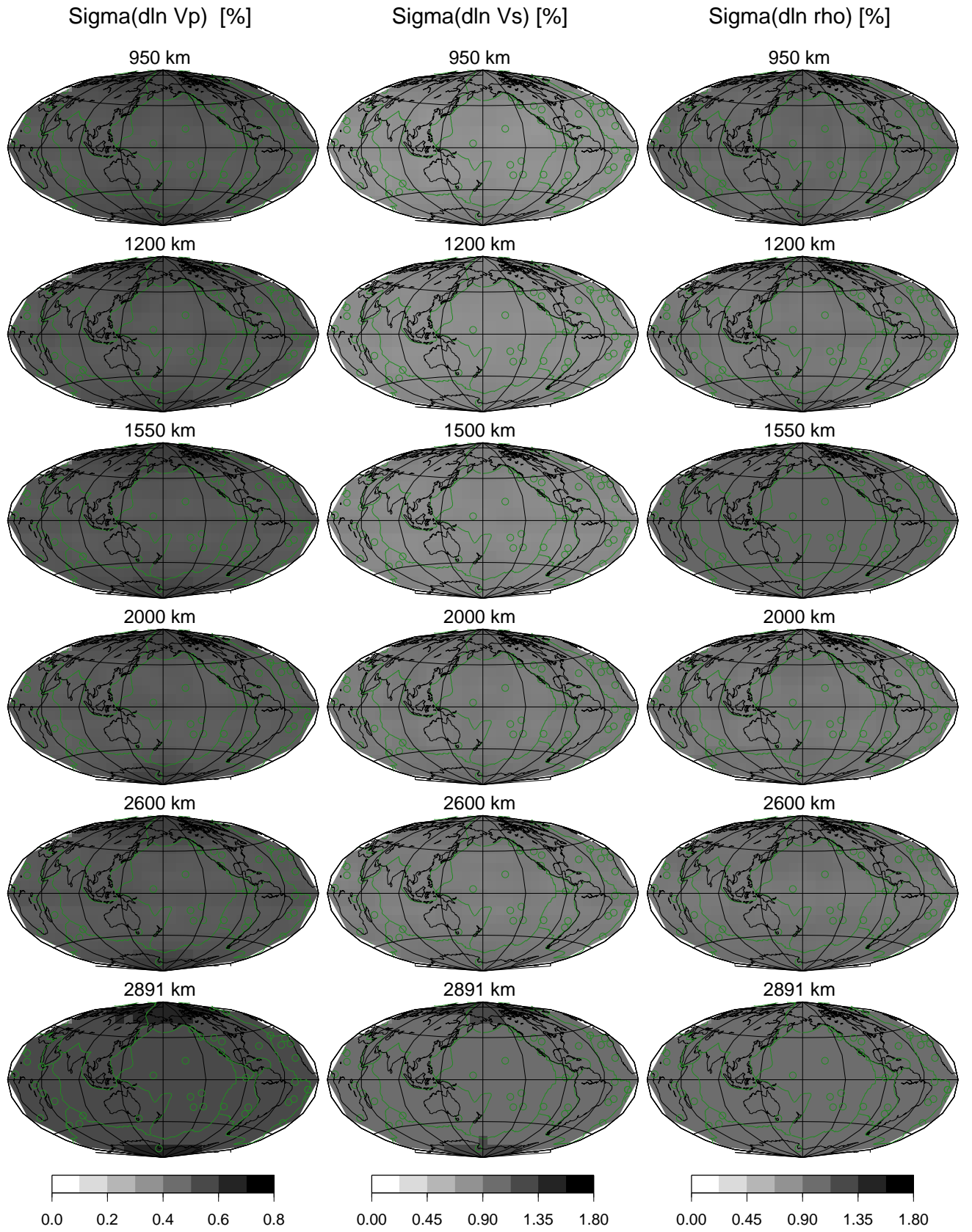


Figure 4.

February 10, 2012, 7:14pm

D R A F T

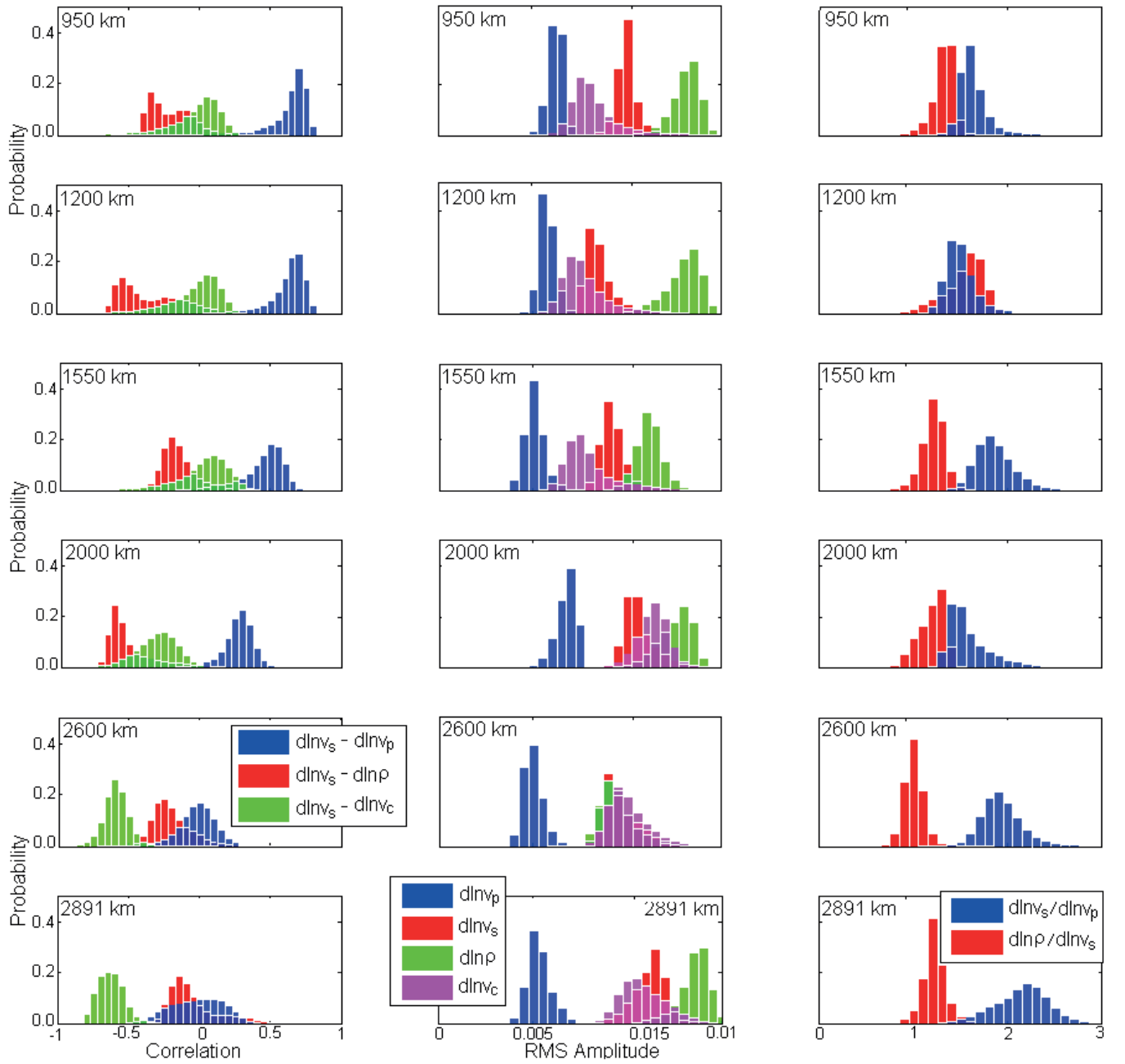


Figure 5.

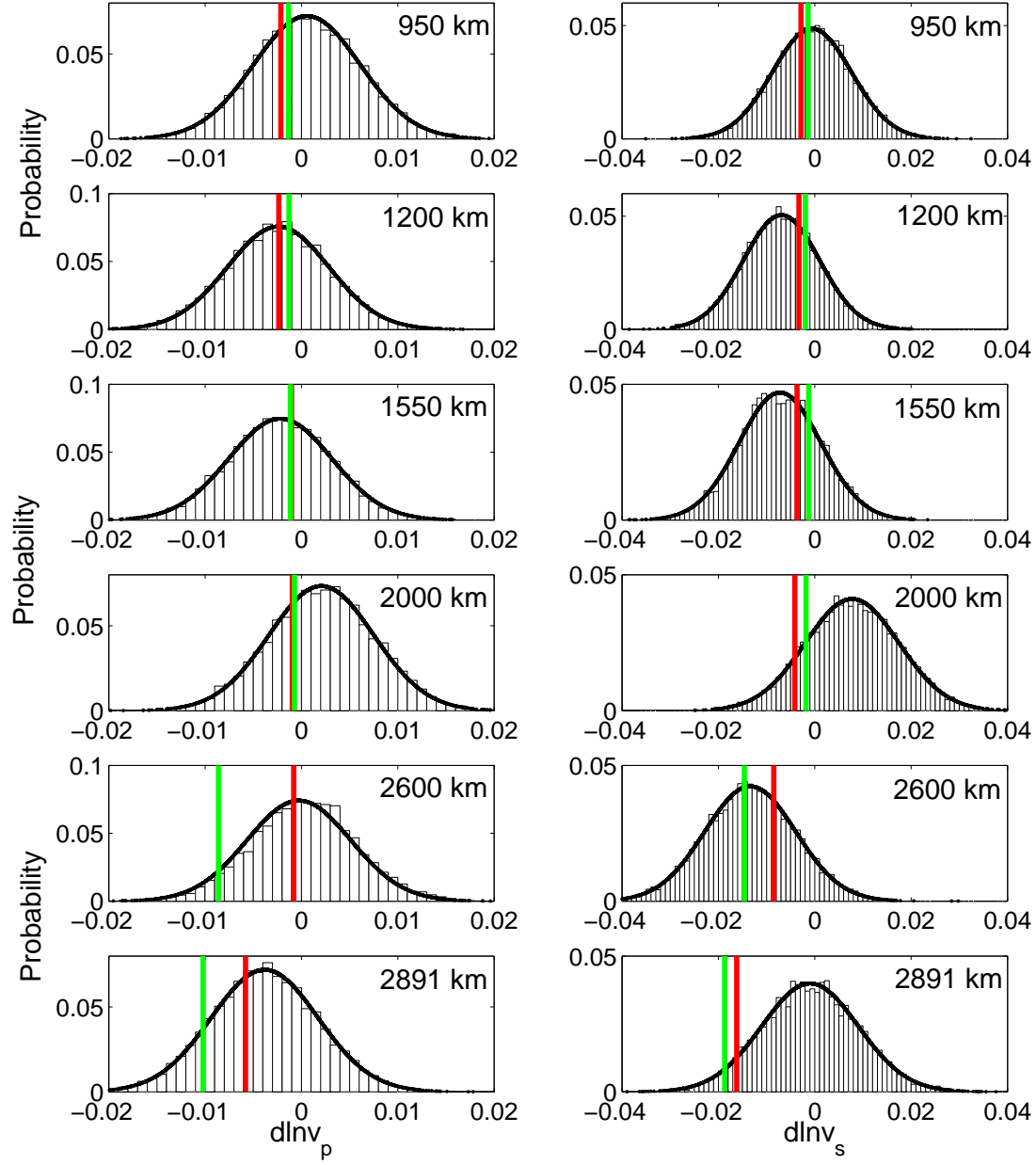


Figure 6.

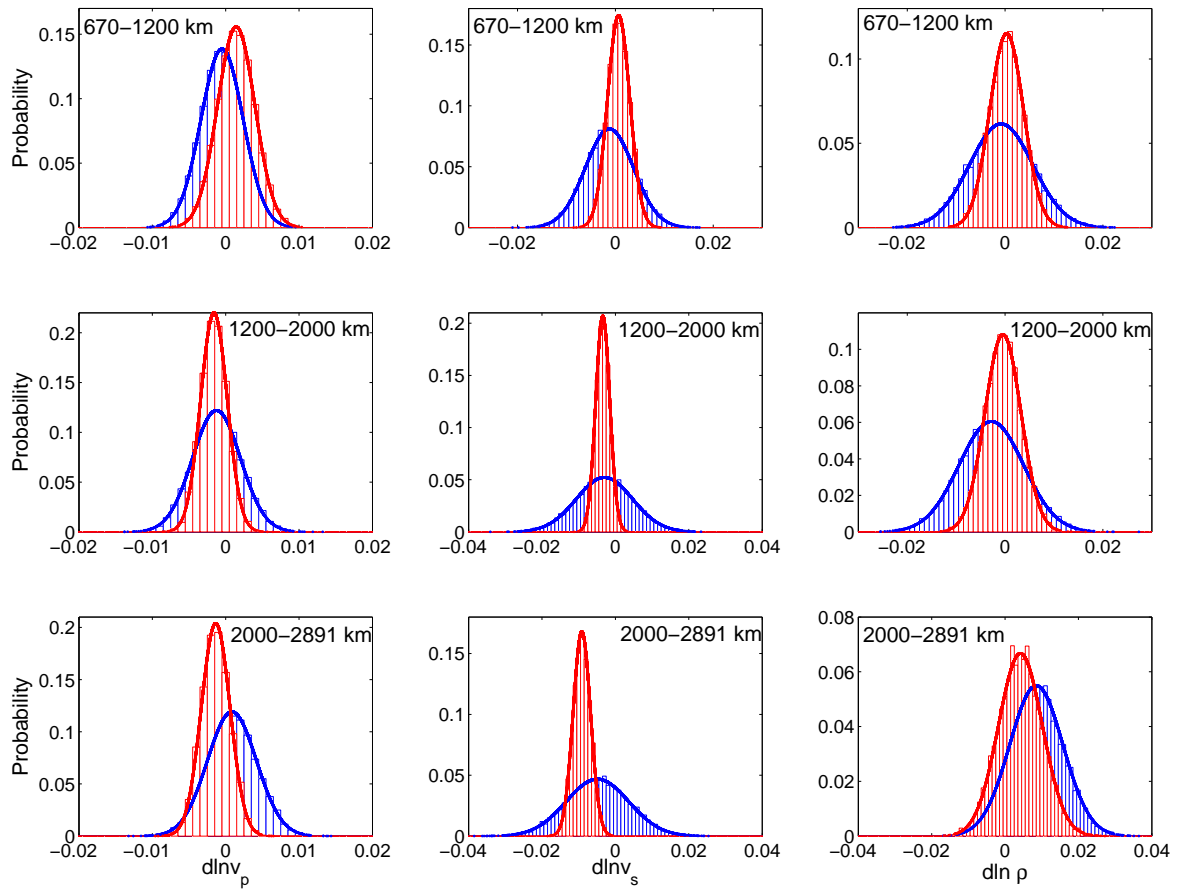


Figure 7.

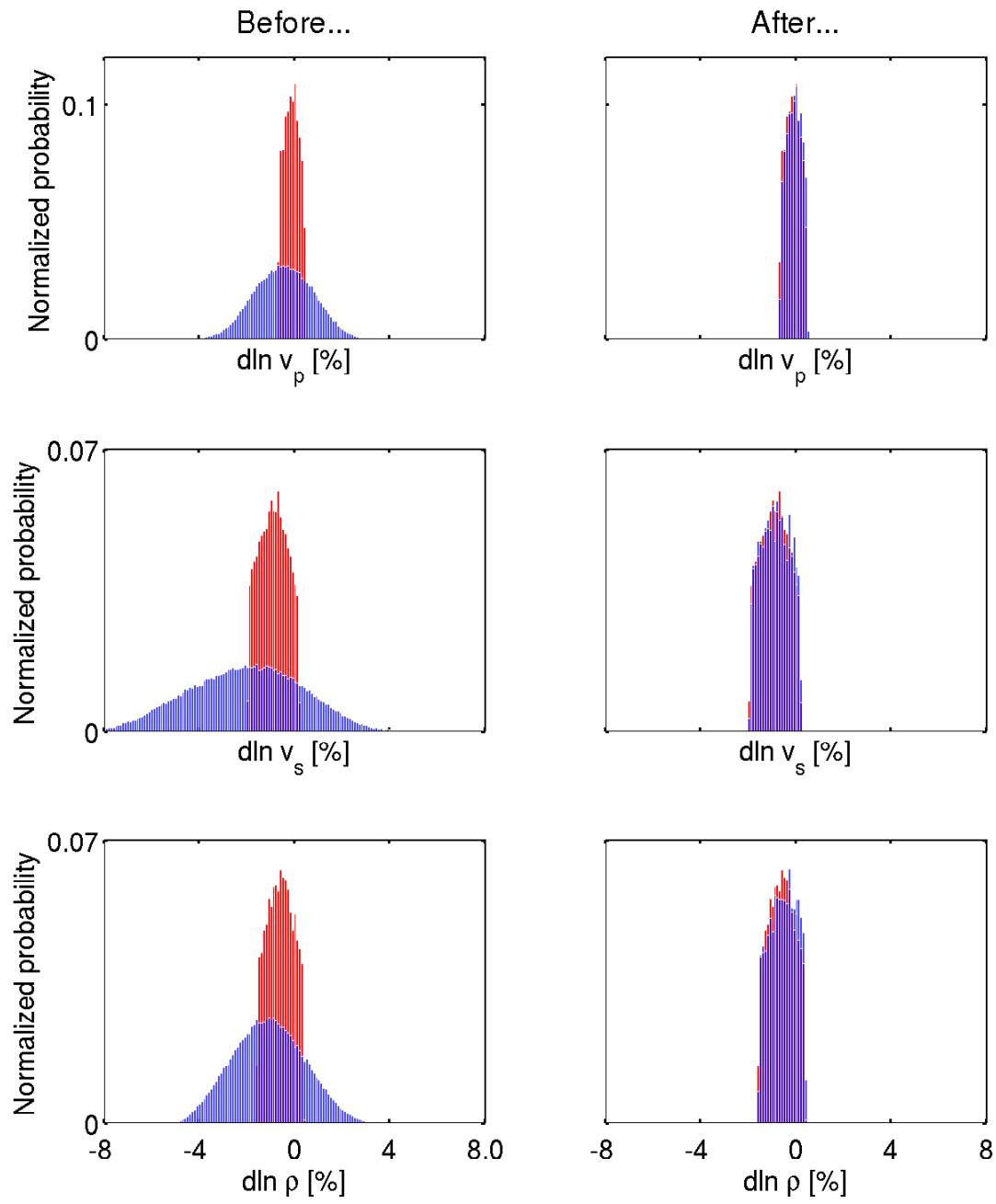


Figure 8.

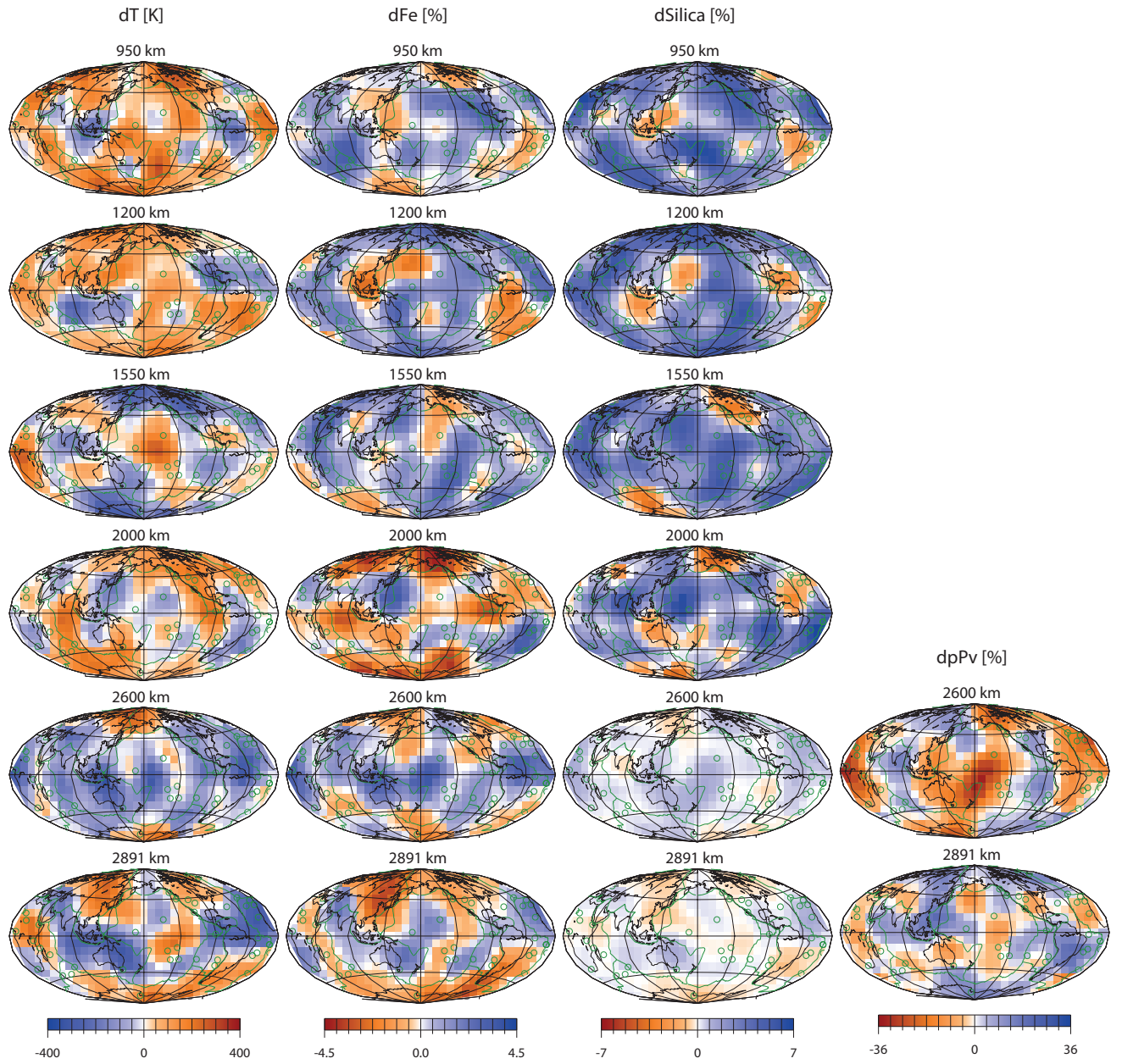


Figure 9.

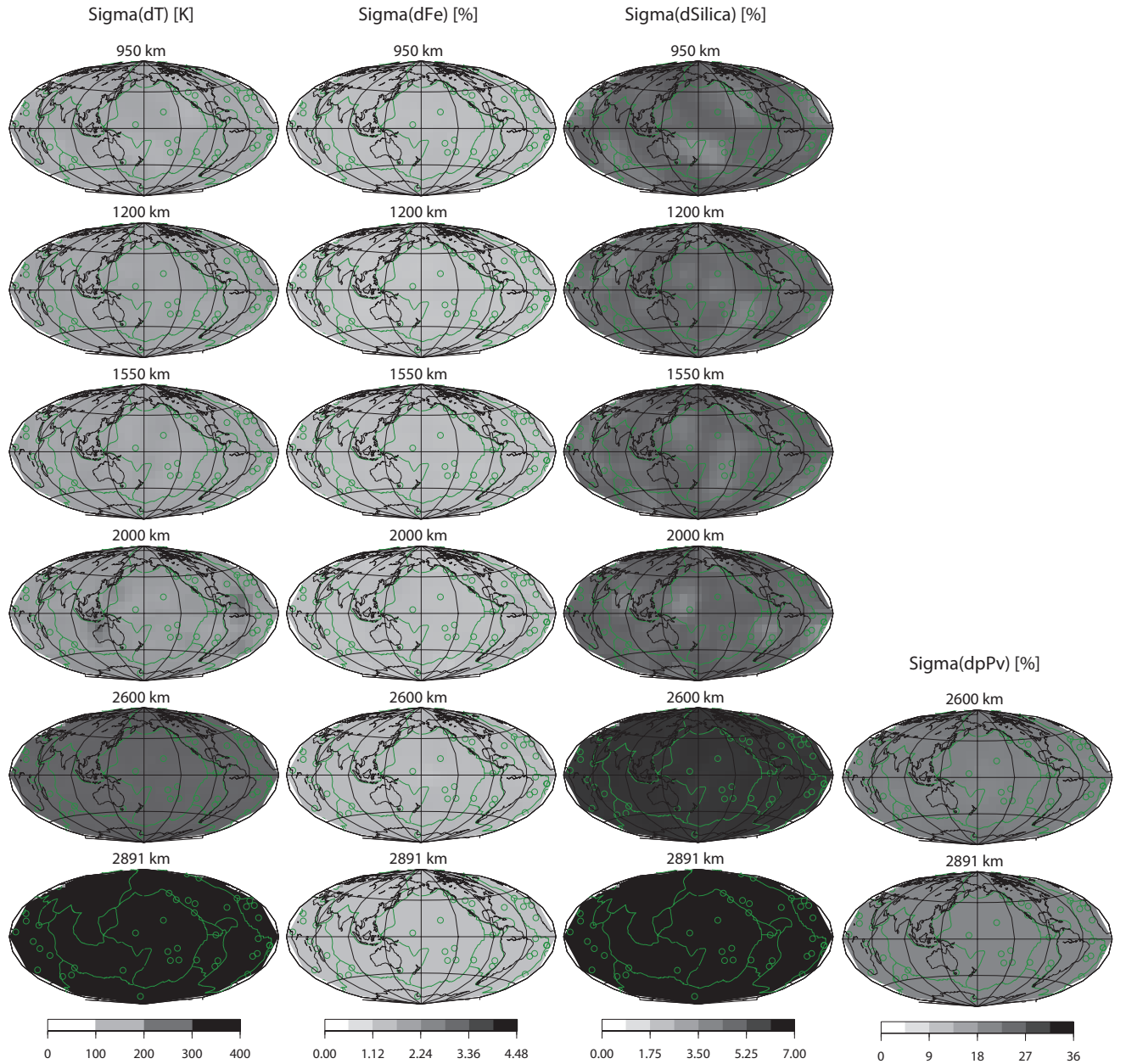


Figure 10.

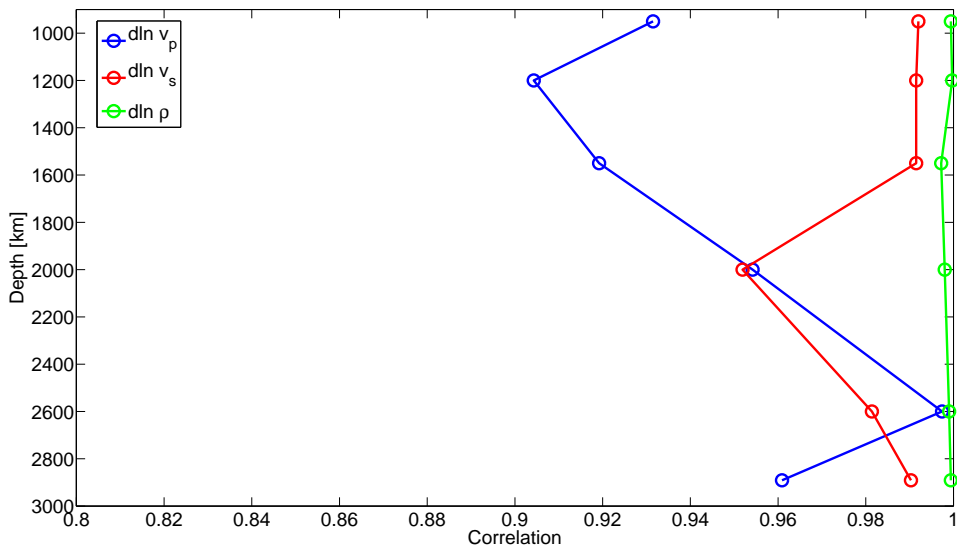


Figure 11.

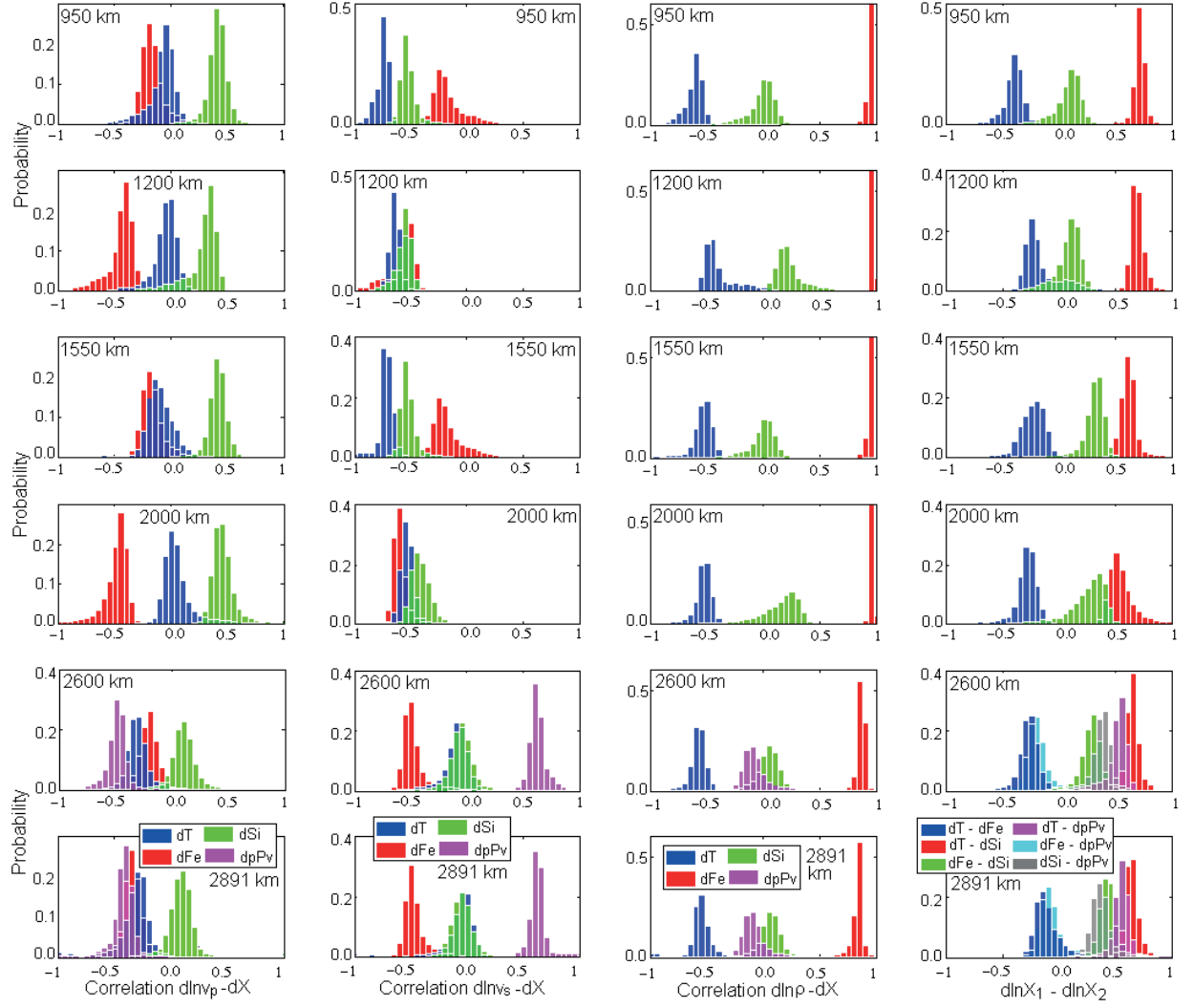


Figure 12.

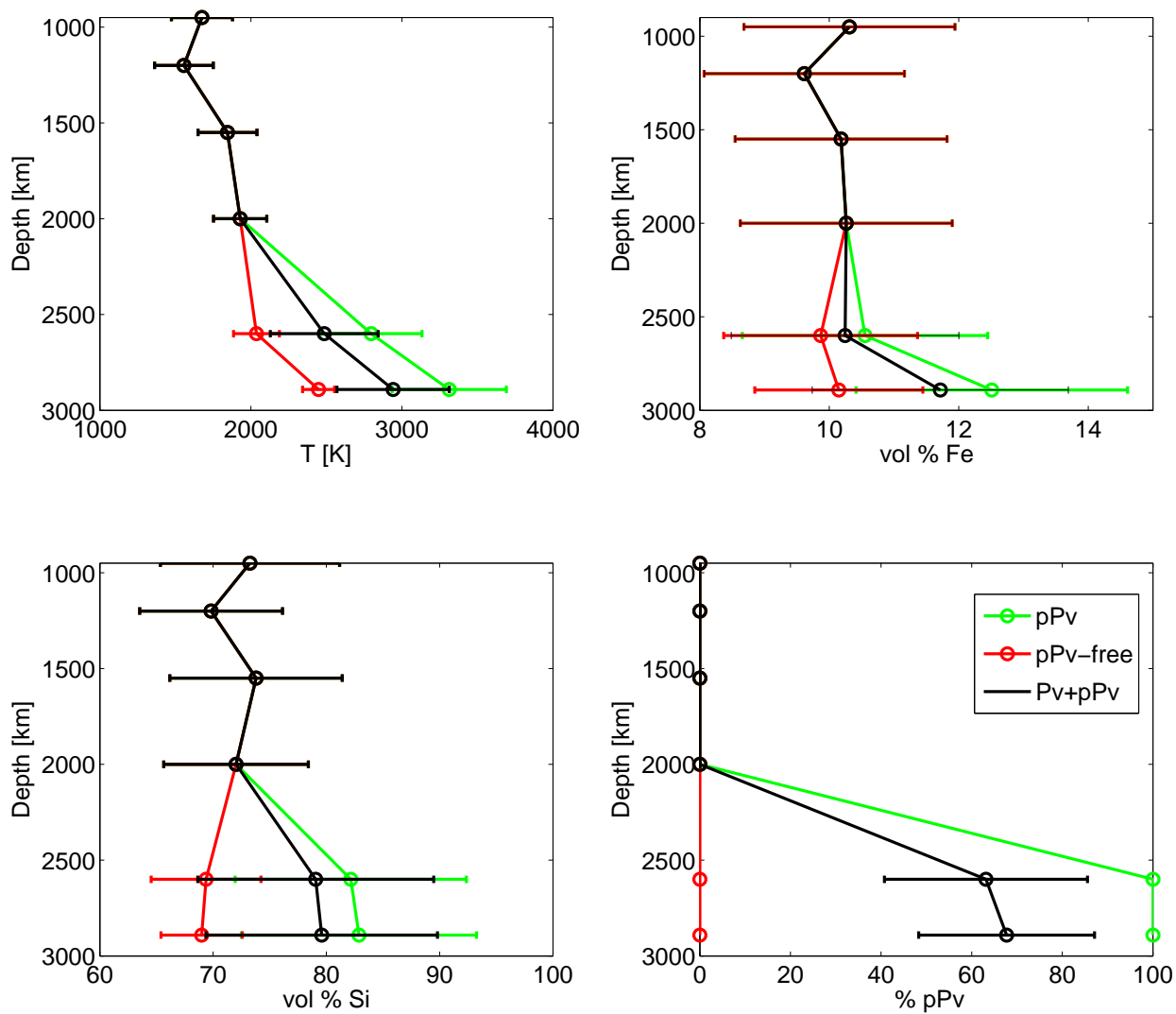


Figure 13.

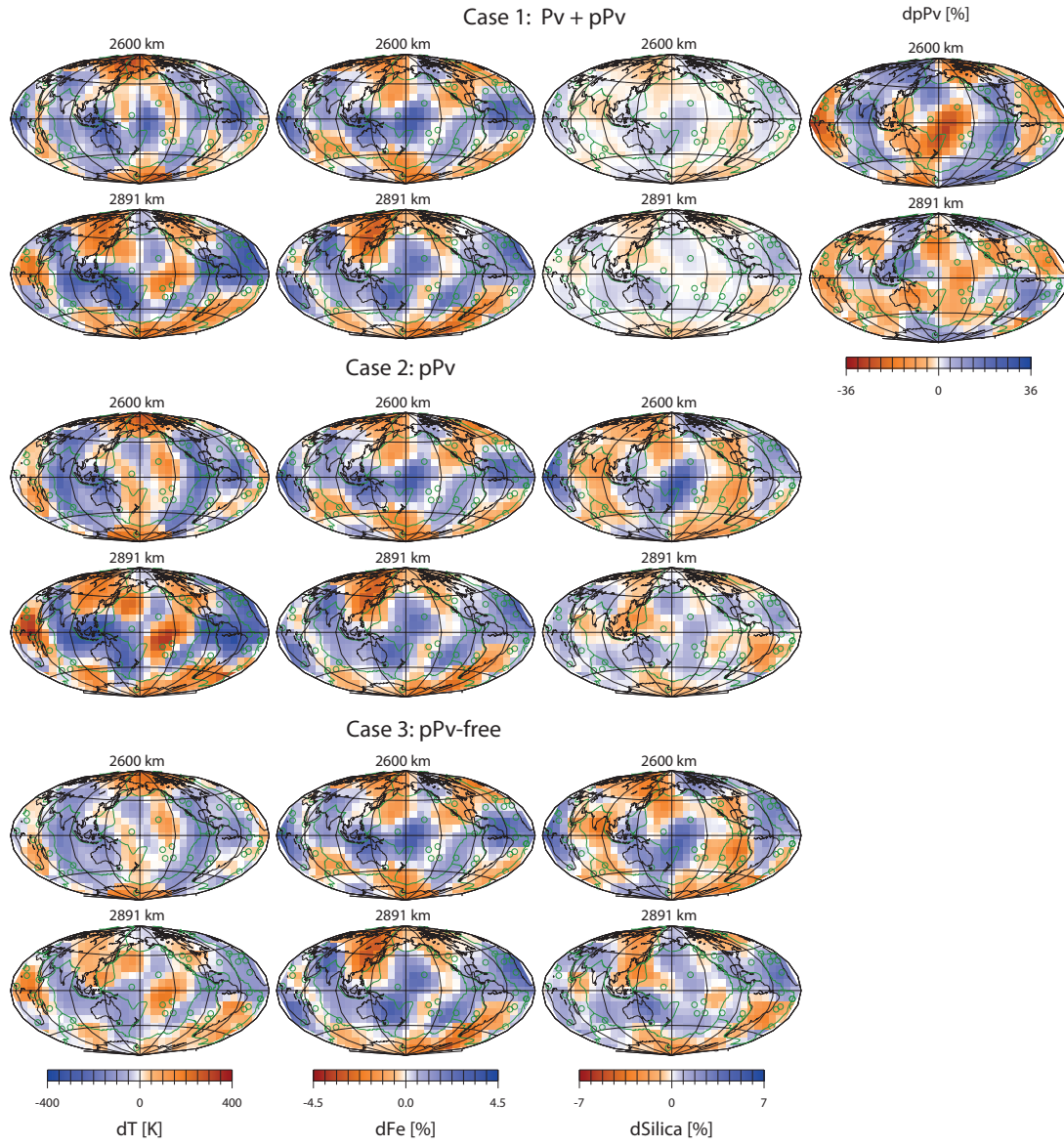


Figure 14.

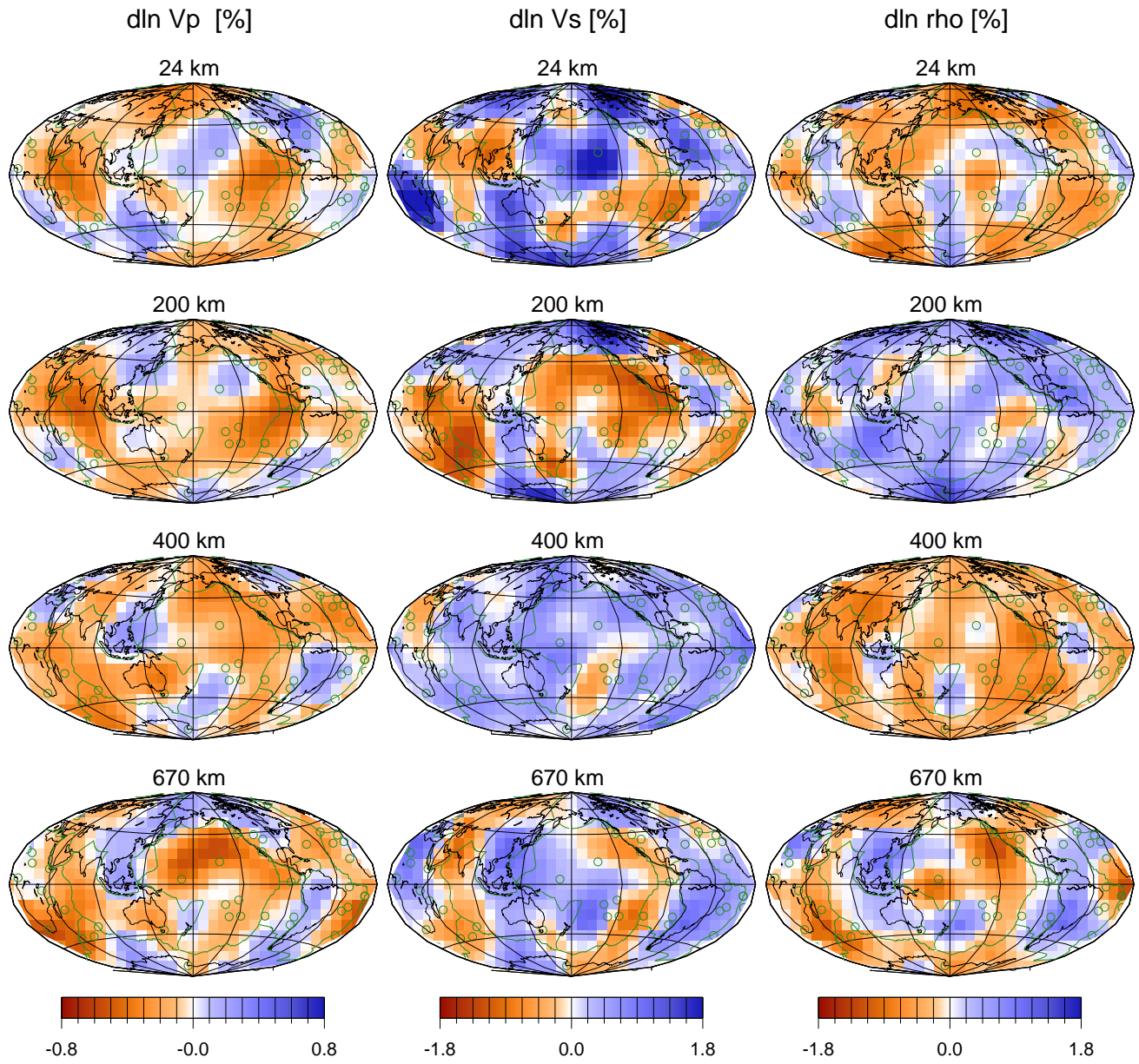


Figure 15.

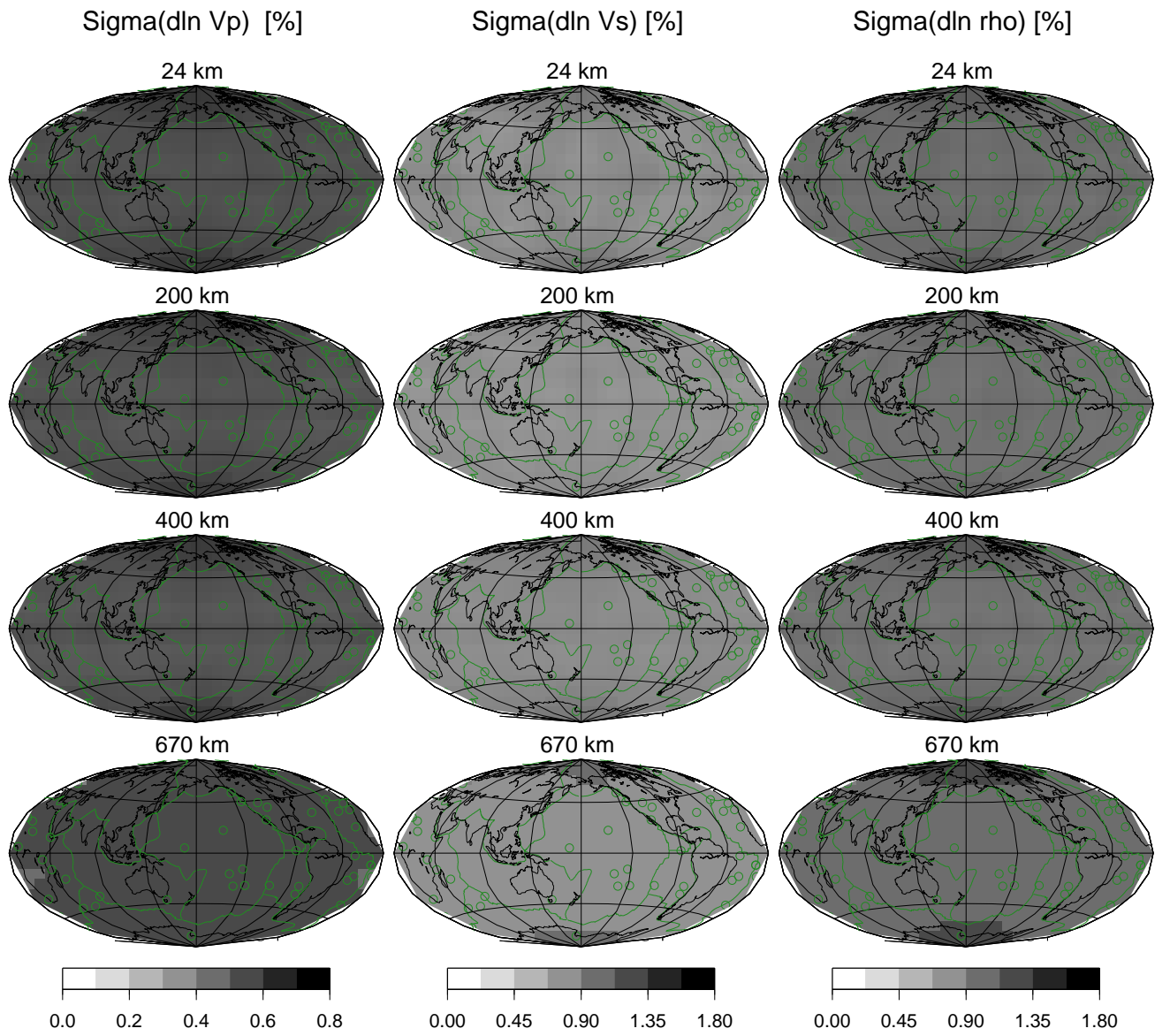


Figure 16.

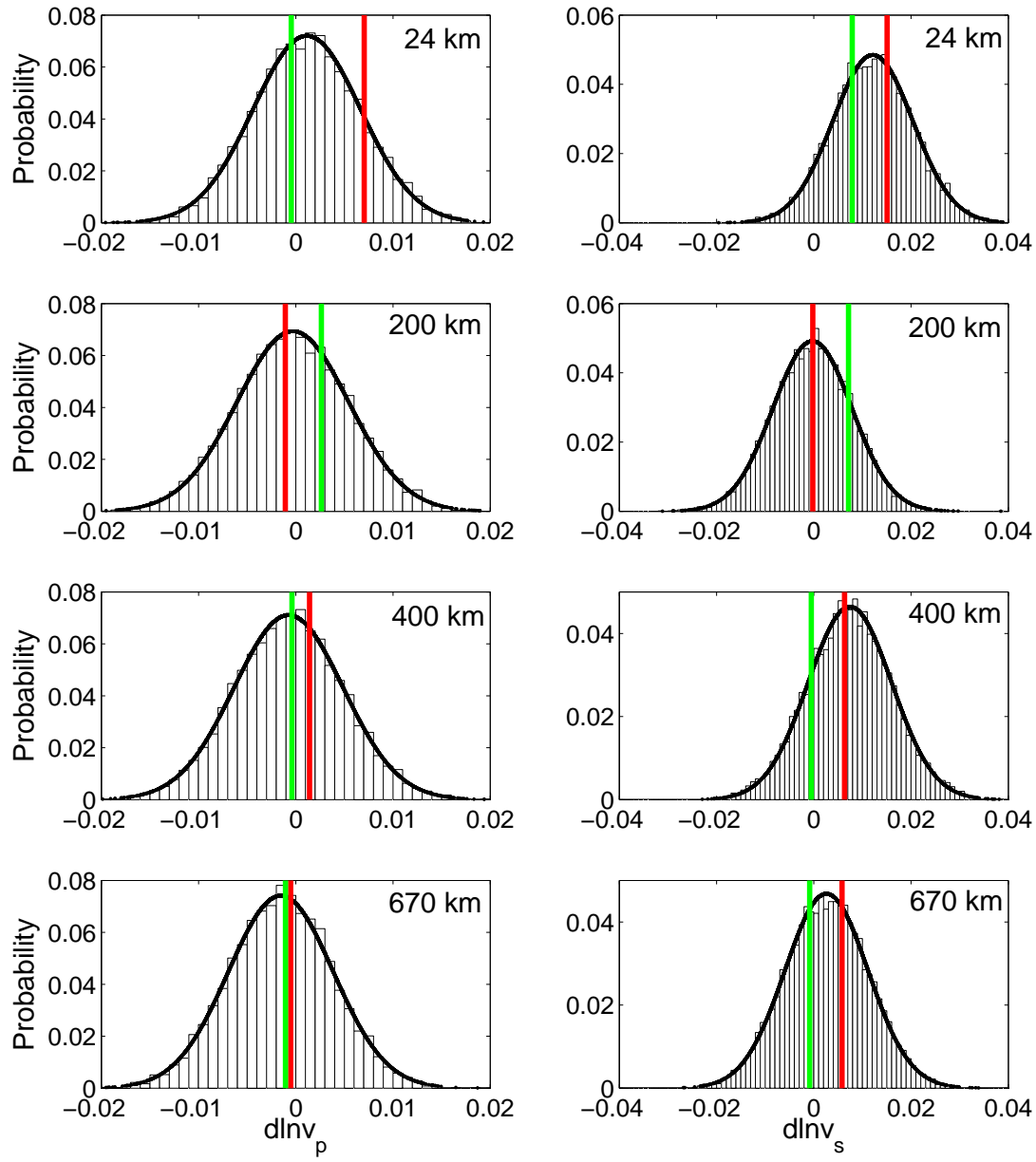


Figure 17.

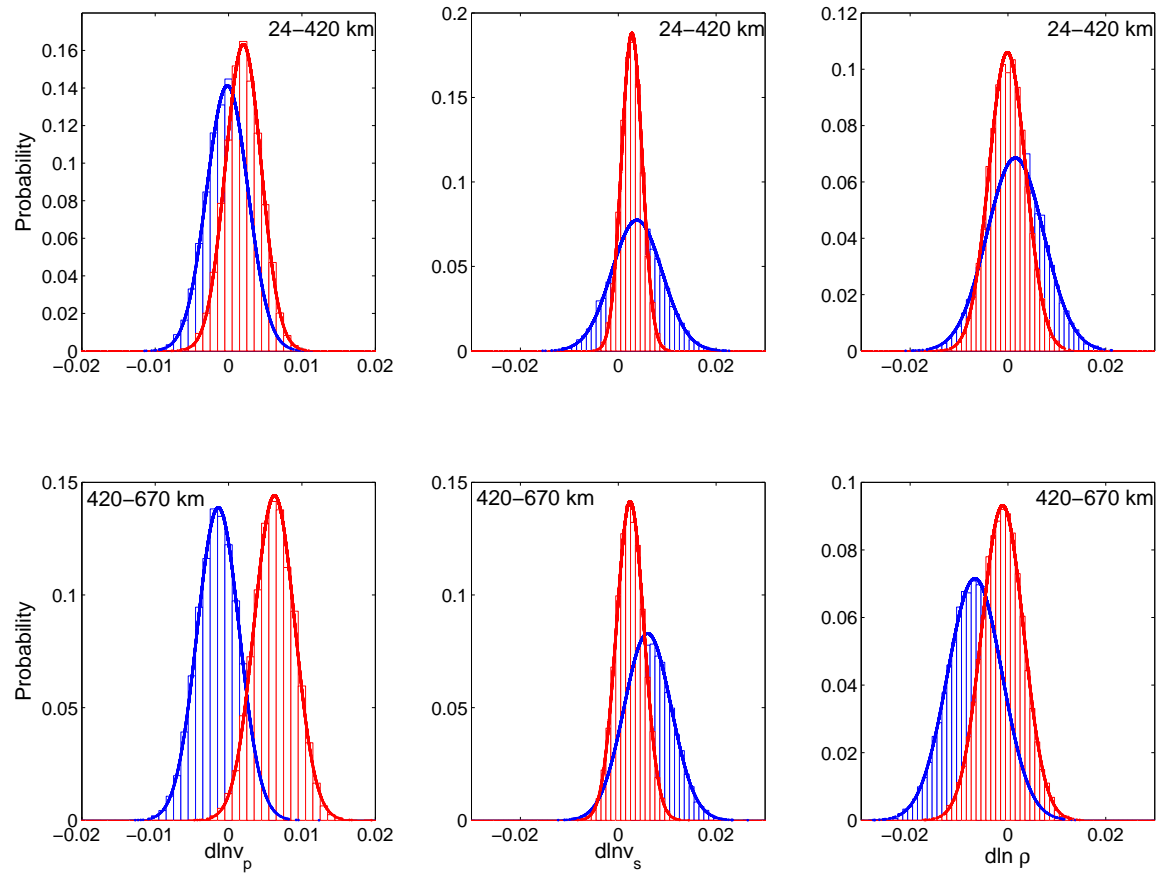


Figure 18.