Lateral heterogeneity scales in regional and global upper mantle shear velocity models

Matthias Meschede, Barbara Romanowicz

Lateral heterogeneity scales in regional and global upper mantle shear velocity models

Matthias Meschede[1] and Barbara Romanowicz[1,2,3]

[1]*Institut de Physique du Globe de Paris, F-75005 Paris, France. E-mail: matthias.meschede@googlemail.com*
[2]*Collège de France, F-75005 Paris, France*
[3]*Department of Earth and Planetary Science, University of California, Berkeley, CA 94720, USA*

SUMMARY

We analyse the lateral heterogeneity scales of recent upper mantle tomographic shear velocity (*Vs*) global and regional models. Our goal is to constrain the spherical harmonics power spectrum over the largest possible range of scales to get an estimate of the strength and statistical distribution of both long and small-scale structure. We use a spherical multitaper method to obtain high quality power spectral estimates from the regional models. After deconvolution of the employed taper functions, we combine global and regional spectral estimates from scales of 20 000 to around 200 km (degree 100). In contrast to previous studies that focus on linear power spectral densities, we interpret the logarithmic power per harmonic degree l as heterogeneity strength at a particular depth and horizontal scale. Throughout the mantle, we observe in recent global models, that their low degree spectrum is anisotropic with respect to Earth's rotation axis. We then constrain the uppermost mantle spectrum from global and regional models. Their power spectra transfer smoothly into each other in overlapping spectral bands, and model correlation is in general best in the uppermost 250 km (i.e. the 'heterosphere'). In Europe, we see good correlation from the largest scales down to features of about 500 km. Detailed analysis and interpretation of spectral shape in this depth range shows that the heterosphere has several characteristic length scales and varying spectral decay rates. We interpret these as expressions of different physical processes. At larger depths, the correlation between different models drops, and the power spectrum exhibits strong small scale structure whose location and strength is not as well resolved at present. The spectrum also has bands with elevated power that likely correspond to length scales that are enhanced due to the inversion process.

Key words: Fourier analysis; Spatial analysis; Seismic tomography; Statistical seismology.

1 INTRODUCTION

The first global tomographic models of mantle elastic structure developed in the 1980s revealed coherent very long wavelength structures, and in particular, a correlation of the shear velocity structure in the uppermost 200 km of the mantle with surface tectonics (e.g. Woodhouse & Dziewonski 1984), as well as the presence of a 'degree 2' structure in the lower mantle (e.g. Dziewonski 1984; Woodhouse & Dziewonski 1989) which correlates with the corresponding degree 2 structure in the geoid (e.g. Hager 1984).

Since then, owing to the accumulation of high quality broadband data as well as improvements in tomographic techniques, progressively higher resolution has been achieved, and has made it possible to (1) map in more detail the uppermost mantle structure, revealing more focused ridges, back arcs, lateral variations in continental lithosphere structure as well as age dependence of the oceanic lithosphere, (2) identify fast velocity regions corresponding to subducting slabs in the upper mantle transition zone

and (3) more accurately map the two large low shear velocity provinces (LLSVPs) in the lowermost mantle, located antipodally under the central Pacific and under Africa (e.g. Dziewonski *et al.* 2010).

Global shear velocity models now agree remarkably well with each other on the location and shape of long wavelength structure in the uppermost mantle and in the lowermost 500 km of the lower mantle. Moreover, details, for example of the LLSVP boundaries, are consistent among global models to much shorter length scales (Lekic *et al.* 2012).

Because these global models are necessarily limited to long wavelengths, important questions arise: down to which wavelength can we trust the feature strength of a current generation global model? Can we constrain the global higher degree spectrum by making use of information contained in large scale regional models? What is the shape of the spectrum at higher degrees, why does it have this shape, and can it be represented by a simple model spectrum (e.g. power-law decay)?

Documenting the characteristics of the spectrum of lateral heterogeneity in the Earth's mantle across a broad range of wavelengths is important for the understanding of the physical processes that gave rise to this heterogeneity. Spectra that exhibit similar heterogeneities distributed over a large scale range (e.g. according to a power law) are abundant in nature (e.g. mixing, fracturing, etc.). On the other hand, processes with a dominant, maximal or minimal length scale are also plausible (bending of the lithosphere, thin plumes,...) (e.g. Turcotte 1997). At a given depth, the global spherical harmonics spectrum likely reflects a mixture of heterogeneities due to different processes.

The power spectrum has further been used as a method to compare geodynamic and tomographic models in numerous studies (e.g. Bunge & Richards 1996; Mégnin *et al.* 1997; Bunge *et al.* 1998; Becker & Boschi 2002; Phillips & Bunge 2005; Schuberth *et al.* 2009). For example Tackley *et al.* (1993) argued for an important role of the 670 km discontinuity in global mantle circulation based on such a comparison.

Finally, a robust knowledge of the heterogeneity spectrum to short wavelengths is important for predicting the effects of small scale heterogeneity on wave propagation in the context of estimation of scattering effects (e.g. Shearer & Earle 2008; Garcia *et al.* 2009; Sato *et al.* 2012).

For these reasons, the global power spectrum as a function of depth has attracted much attention (e.g. Ritzwoller & Lavely 1995; Becker & Boschi 2002; Dziewonski *et al.* 2010). In addition to the strong 'degree 2' structure that is present in the Earth's mantle, a robust feature of all global models is the relatively 'red' spectrum of heterogeneity in the upper mantle as well as in the D'' region, whereas the spectrum is whiter in the bulk of the lower mantle. This difference correlates with the strength of rms velocity fluctuations, which is largest at the top and bottom of the mantle, and weaker in the mid-mantle. Another consistent feature in most global velocity models is a strong corner between degrees 5 and 8 in the uppermost mantle, which has been compared to the similar looking continent–ocean function, and also to the hotspot distribution (Tanimoto & Anderson 1984; Tanimoto 1990). This corner has also been interpreted as meaning that structure is dominated by features of dimensions larger than 2500–3000 km at these depths (Su & Dziewoski 1992).

Using new, recently published global and regional high-resolution models that employ advanced tomographic techniques based on the inversion of full waveforms, we can analyse the less well constrained higher degree spectrum in more detail. In particular, comparison with regional model spectra, that focus on areas with good data coverage, can help to assess the influence of limited resolution and regularization on the global model spectrum.

In order to compare the spectra from current global and large scale regional tomographic models of the mantle, it is necessary to do so consistently. In particular, in going from the global to the regional scale, one needs to carefully consider the effects of truncation due to windowing on the 2-D power spectral density (Wieczorek & Simons 2005; Dahlen & Simons 2008).

Attempts at global and regional model comparisons have previously been made: Passier & Snieder (1995) used a 2-D Fourier transform approach to compare global and regional spectra. They assumed to be in the Euclidean limit which means they treated a small region of the sphere as a flat plane, and they did not take truncation effects into account. They found a large gap between regional and global power spectra. Chevrot *et al.* (1998) used a similar approach and compared the 2-D Fourier transform of a regional model to an asymptotic expansion of spherical harmonics without

taking window effects into account. They found less of a difference between the models and attributed the previously seen gap to improper comparison between Fourier and spherical harmonics power. However, large differences and strong ringing, by up to an order of magnitude, remained, which prohibits a more detailed quantitative analysis of the shape of the spectrum. Qualitatively, both authors interpreted the spectra in terms of their observed power-law decay rates.

In the first part of this paper, we quickly discuss the meaning of the power spectrum. In contrast to typically used linear power spectral densities we choose to consider the 'logarithmic power per degree l' spectrum that measures the model variance in localized heterogeneities of a certain size. This helps not only to constrain, but also to interpret the power spectrum. We then briefly describe the construction of special band-limited window functions that minimize truncation effects in order to obtain accurate estimates of the spherical power spectrum from recent regional models, following the multi-taper approach of Wieczorek & Simons (2007).

In the second part, we discuss the robust features of the power spectrum by comparing various global models and connecting them at shorter wavelengths with regional power spectral estimates, considering the possible limitations of different models along the way. We show that in the uppermost 200–250 km of the mantle [hereafter 'heterosphere' following Dziewonski *et al.* (2013, AGU abstract)], the most recent global model spectra connect smoothly to the most recent regional ones. From this analysis, a spectral shape that is more complex than a single power-law decay becomes apparent. We explain the spectral shape in terms of heterogeneities with distinct scales and explain in particular the degree 5, but also a higher degree 'corner'. We also show that at the examined depths, the global power spectrum is, to some extent, non-isotropic and non-Gaussian due to specific, localized features. In the deeper model parts (>250 km depth), strong heterogeneity power is present up to the smallest parametrized scales of the models, which is unlikely to be well constrained and which is further inconsistent amongst models. At these depths, we observe further spectral peaks close to the resolution limits of the models. These are likely length scales that are artificially enhanced by the inversion procedure.

2 POWER SPECTRAL ESTIMATION ON THE SPHERE

2.1 Spherical harmonics and their power spectra

We expand models in terms of spherical harmonics, with coefficients c_{lm}. These spherical harmonic functions $Y_{lm}(\mathbf{r})$, where $l \geq 0$ and $|m| \leq l$, form an orthogonal, complete basis for square integrable functions on the surface of the sphere and can be expressed as:

$$Y_{lm}(\mathbf{r}) = \sqrt{(2l+1)\frac{(l-m)!}{(l+m)!}} P_{lm}(\cos\theta)e^{im\phi} \tag{1}$$

where P_{lm} is the associated Legendre polynomial of degree l and order m. The normalization is chosen such that the variance of each $Y_{lm}(\mathbf{r})$ is equal to one:

$$\frac{1}{4\pi}\int_\Omega Y_{lm}^*(\mathbf{r})Y_{l'm'}(\mathbf{r})\,d\Omega = \delta_{ll'}\delta_{mm'}. \tag{2}$$

A seismic velocity model at constant depth, $f(\mathbf{r})$, can then be expanded as:

$$f(\mathbf{r}) = \sum_{lm} c_{lm}Y_{lm}(\mathbf{r}) \tag{3}$$

with coefficients c_{lm}:

$$c_{lm} = \frac{1}{4\pi} \int_\Omega f(\mathbf{r}) Y_{lm}^*(\mathbf{r}) \, d\Omega. \qquad (4)$$

Parseval's identity relates the c_{lm} coefficients to the model $f(\mathbf{r})$ as follows:

$$\frac{1}{4\pi} \int_\Omega f(\mathbf{r})^2 d\Omega = \sum_l S_l = \sum_l \sum_m |c_{lm}|^2, \qquad (5)$$

where we call $S_l = \sum_{m=-l}^{m=l} |c_{lm}|^2$ the 'power per degree l'. It can be seen as the variance of $f(\mathbf{r})$, i.e. the expected square amplitude $f^2(\mathbf{r})$, due to all spherical harmonics with degree l at a randomly chosen point on the sphere. Note that other power spectral densities, notably the '—expected—power per degree l and order m', $[=S_l/(2l+1)]$, are also commonly used. The latter density is comparable to the classical power spectral density of a 2-D Fourier transform, and can be related to the expected variance (under arbitrary rotations) of a single spherical harmonics coefficient when the model is isotropic and Gaussian (e.g. Baldi & Marinucci 2007):

$$\langle c_{lm}^* c_{l'm'} \rangle = \frac{S_l}{2l+1} \delta_{ll'} \delta_{mm'} \qquad (6)$$

On the other hand, it is the power per degree S_l which directly relates the single parameter l to the model variance, and this is what we are interested in.

2.2 Linear versus logarithmic power spectral density

It can be shown that multiplying the spherical harmonics coefficients of a model $f(\mathbf{r})$ with a real valued filter that depends only on degree l is equivalent to convolving the model with an axisymmetric wavelet. Moreover, a set of wavelets—for example a set of Gaussian or Mexican Hat filters—that are scaled (dilated) to different sizes and filter out heterogeneities of different scales—correspond to bandpass filters that are constant in logarithmic, not linear spherical harmonics degree space (e.g. Wiaux *et al.* 2005; Leistedt *et al.* 2013).

It is therefore the logarithmic power per degree which is directly connected to the model variance in a set of dilated, localized features. In contrast, the linear power per degree is related to spherical harmonics that are *not* scaled versions of each other because they always span the whole globe as standing waves. A consequence is that in a linear power per degree spectrum, randomly distributed localized heterogeneities with fixed variance at different scales impose a spectral decay of $1/l$ because they get 'averaged' over l wavelengths. From a different perspective, the scaled heterogeneities distribute their power over a bandwidth that is proportional to their scaling factor. Only the logarithmic power per degree therefore measures the heterogeneity variance that we are interested in (see Appendix A for illustrations). We can transform the power per degree l to a logarithmic density as follows:

$$S_l dl = s(l) d(\log_2 l)$$

$$\Rightarrow s(l) = \log(2) l S_l. \qquad (7)$$

We call $s(l)$ the 'power per octave'. This quantity —sometimes called the power 'per band' or 'per relative frequency'— is well known and has been used for scale analysis in different fields (e.g. Zangvil 1977; Lewalle *et al.* 2007). The degree l represents the horizontal wavelength $2\pi r/l$, where r is the distance from the centre of the Earth. In terms of localized heterogeneities, this corresponds for example to a typical distance between two heterogeneities of

equal sign. The half wavelength $\pi r/l$ is then the typical size of an isolated heterogeneity between two features of opposite sign. We indicate it in brackets behind the degree l in the rest of the paper.

We would like to emphasize that this subtlety can have strong implications for the interpretation of the shape of the power spectrum. The superimposed 'natural' power-law decay of $1/l$ in the presence of localized features hides minima, maxima and also deviations from a power-law decay spectrum. In the logarithmic power per degree, corner frequencies can become maxima and strong power at low degrees can become relatively weak. Furthermore, new maxima and corners can appear. Note that this power-law dependence of $1/l$ depends on the power spectral density, in particular its dimensionality. This spectrum is also often called the 'pink spectrum', and in the context of scattering studies it represents a self-similar medium for the previously mentioned reasons (e.g. Frankel 1989; Klimeš 2002; Sato *et al.* 2012). Nevertheless, deviations from it are usually not or at most vaguely (e.g. Tanimoto 1990) examined.

2.3 Regional power spectral estimates

As in common spectral analysis in Cartesian coordinates, estimating the power spectrum of a function with support in a subregion has to be done with care, especially when examining quickly decaying spectra. Restricting a function to a subregion of the sphere can be described by multiplication of a globally defined function by a window function. Whereas in the ordinary Fourier transform, multiplication in space corresponds to a convolution in the spectral domain, in the spherical harmonics case, multiplication with an axisymmetric window function can be expressed through a coupling matrix that relates the unwindowed power to the windowed power (see Appendix B) (Hivon & Górski 2002; Wieczorek & Simons 2005). Importantly coupling can only occur between degrees closer to each other than the window bandwidth.

A good window function should therefore concentrate its power in the lowest possible degrees to minimize coupling while at the same time optimally covering the region of interest. In other words, the window function should concentrate its power both in the spatial and in the spectral domain. Contrary to the Fourier transform, apart from being smoothed, a windowed spherical harmonics spectrum can also be biased, especially around the window wavelengths, which can be seen as a consequence of the representation of 2-D space by the single parameter l (see Appendix B). In our particular case, this bias needs to be removed as much as possible, because we want to get information from the regional models at the largest possible scales, which are naturally close to the window wavelength, in order to connect them to the global ones. For example, the larger available regional studies correspond to areas of at most 4000–5000 km in extent, and we can only fit two wavelengths of heterogeneities of size 1000 km in such an area, which is around the maximum resolution that we expect in global models.

In order to deal with these issues, the approach chosen in this paper is that developed in (Wieczorek & Simons 2007) which we summarize only briefly here. It is similar to well-known multitaper methods in Cartesian coordinates. These authors provide a method (and the Fortran library SHTOOLS) to calculate sets of band-limited windows that are optimally concentrated in space. We use the first few, most concentrated, of these tapers until we have a reasonable coverage of the region and compute the global spherical harmonics power spectra of the windowed functions. We select the region of interest for each model through the information given by the authors

Table 1. Overview of the models examined in this study. Data period ranges are taken from the publications and are half-widths in case of a smooth bandpass filter.

Name	Publication	Region	Data[a]	Period range	Technique[b]
S362ANI	Kustowski *et al.* (2008)	Whole mantle	sw/bw	35–150s/20s	TT + PA waveforms
S40RTS	Ritsema *et al.* (2011)	Whole mantle	nm/msw/bw	>333s/40–275s/–	TT + splitting functions + PA
SEMUM2	French *et al.* (2013)	Upper mantle	sw	70–325s	SEM + NACT kernels
	Mégnin & Romanowicz (2000)	Lower mantle	sw/bw	80–400s/31–300s	NACT waveforms + kernels
DR2012	Debayle & Ricard (2012)	Upper mantle	sw	50–250s	PA waveforms + kernels
SL2013	Schaeffer & Lebedev (2013)	Upper mantle	sw	11–450s	JWKB waveforms + kernels
EU.52	Fichtner *et al.* (2012)	Europe	sw/bw	30(8)–200s	SEM + adjoint
	Rickers *et al.* (2013)	North Atlantic	sw/bw	25–200s	SEM + adjoint
EU30	Zhu *et al.* (2012)	Europe	sw/bw	25–150s/15–50s	SEM + adjoint
SaCO13	Colli *et al.* (2013)	South Atlantic	sw/bw	>120s	SEM + adjoint
AMSAN.19	Fichtner *et al.* (2009)	Australia	sw/bw	50–200s	SEM + adjoint
AusREM	Kennett *et al.* (2012)	Australia	–	–	Three models combined
Na07	Bedle & van der Lee (2009)	North America	sw	–	PA waveforms
Na11	Yuan *et al.* (2011)	North America	sw/bw	>70s/10–30s	NACT
DNA10	Obrebski *et al.* (2011)	Western US	fsw/bw	18–125s/10–50s	TT + PA
DNA13	Porritt *et al.* (2013)	Western US	fsw/bw/nc	30–125s/10–50s/8–40s	TT + PA
WusSH	Schmandt & Humphreys (2010)	Western US	bw	2.5–20s	TT + FF kernels

[a]nm, normal modes; sw, multimode surface waves; fsw, fundamental mode surface waves; bw, body waves; nc, noise cross correlation.
[b]FF, finite frequency; TT, travel times/ray theory; PA, path-average surface wave theory; NACT, non-asymptotic mode coupling theory; JWKB, semi-classical mode phase perturbations.

but also visually, trying to stay away from edges that are not well resolved. Typically more than 99 per cent of the power of each taper is in the selected region. Performed tests (not shown) demonstrated that edge effects are indeed very small and the power spectrum is also not very sensitive to the exact choice of the region.

For each tapered model, which is non-zero only within the selected area, we can then compute the global spherical harmonics power spectrum. The mean power of the model windowed with the selected tapers yields an estimate of the global power spectrum. The mean of the coupling matrices corresponding to each of the tapers, connect the windowed estimate to the estimate of the global spectrum. The windowed spectral estimate can finally be debiased using a least squares inversion under certain assumptions: We search for a global isotropic and Gaussian distributed medium with a smooth spectrum that has the same windowed spectrum as the regional model. To stabilize the inversion, we reparametrize the regional spectrum using a quadratic B-spline basis with approximately one window bandwidth as knot distance. The smooth debiased spectra can finally be compared to the global power spectra (Appendix B illustrates this process for a synthetic model).

3 OVERVIEW OF SELECTED REGIONAL AND GLOBAL MODELS

Instead of introducing the considered models in every detail, as is necessary to fully understand how they were constructed, we try to emphasize characteristics that show the diversity of the employed modelling strategies. Further details can be found in the respective publications. Model names, publications and a few characteristics including their data and period bands are summarized in Table 1. Models do not always include lateral variations of radial anisotropy. If a model provides separate SH and SV velocities, we compute Voigt-averages. Otherwise we use only the available component that could be biased due to the data used. We expect, however, and tested, that the differences between models due to unaccounted for anisotropy are smaller than the differences that we observe and not important for this study.

3.1 Global whole mantle models

S362ANI (Kustowski *et al.* 2008) is a radially anisotropic model that is based upon surface and body wave phase and traveltime anomalies combined with long period waveform modelling under the 'path-average' approximation. It is parametrized horizontally on a spherical spline grid that has a resolution of approximately 1000 km. Radially it is parametrized with 16 splines that are separated at the 660 km discontinuity.

S40RTS (Ritsema *et al.* 2011) is based on Rayleigh wave phase velocity, body wave traveltimes and normal mode splitting functions that constrain the even degree structure of the Earth. It is parametrized horizontally in spherical harmonics up to degree $l = 40$ and vertically with 21 radial splines. The model is regularized by damping the model norm and different versions are available that correspond to different strengths of damping. If not indicated otherwise, we use the 'official' version (identified by its effective number of free parameters: $N = 8112$), in which, however, the amplitudes of even the lowest degrees are significantly weaker than in the other examined models.

SEMUM2 (French *et al.* 2013) is a radially anisotropic model of the upper mantle (depths <800 km) constructed by waveform modelling. It is parametrized on a spherical and radial spline basis and a depth dependent *a priori* smoothness regularization is employed. It combines 3-D waveform modelling using the spectral element method (SEM) with kernels from non-linear asymptotic mode coupling theory (NACT, Li & Romanowicz 1995). The nominal resolution is determined by the *a priori* smoothness and is around 600 km horizontally and 50 km radially. Below 800 km depth, the model connects smoothly to model SAW2416B (Mégnin & Romanowicz 2000) which was constructed from waveforms of surface and body waves using NACT and is parametrized horizontally in spherical harmonics up to degree $l = 24$, and vertically using B-splines.

3.2 Global upper mantle models

DR2012 (Debayle & Ricard 2012) was constructed by waveform modeling, superimposing surface wave modes in a given 1-D model

representing the average 1-D structure along each path. An *a priori* vertical covariance function with a length scale of 50 km is imposed on the 1-D models and the tomographic model with an a priori horizontal length scale of 400 km is then constructed from the path-average 1-D models.

SL2013 (Schaeffer & Lebedev 2013) is obtained from waveforms by summing normal modes whose phase is perturbed due to the velocity model in their approximate 3-D sensitivity region (JWKB approximation). The kernels are thus more accurate than path average kernels but stay in the limits of sensitivity around the initial 1-D reference model. The model is parametrized with horizontal and radial splines including a superimposed 3-D crust.

3.3 Regional upper mantle models

Models of the South Atlantic region (SaCO13) (Colli *et al.* 2013), Europe (EU.52, EU30) (Fichtner *et al.* 2012; Zhu *et al.* 2012), Iceland (included in EU.52) (Rickers *et al.* 2013), Australia (AMSAN.19) (Fichtner *et al.* 2009) were constructed using accurate full waveform modelling (SEM) and (smoothed) Born kernels that take into account the full 3-D structure (adjoint technique). They are parametrized on a 3-D grid whose resolution can vary with depth, and employ different smoothing procedures in the inversion process. This modelling strategy is computationally demanding, especially at shorter periods, which limits the number of events that can be used, but it can accurately treat even strong heterogeneities with sharp transitions, as appear in a complex crust. We have a second model of Australia (Kennett *et al.* 2012) which is a weighted average of three Australian models, including the previously mentioned one. We also considered two models of North America: NA07 (Bedle & van der Lee 2009) which is based upon fitting multimode Rayleigh wave waveforms. NA11 (Yuan *et al.* 2011) combines multimode surface wave waveform modelling based on a coupled mode approach (NACT, Li & Romanowicz 1995) with SKS splitting measurements to obtain a radially and azimuthally anisotropic model. We finally consider a few smaller scale models of the western United States: the DNA10 (Obrebski *et al.* 2011), DNA13 (Porritt *et al.* 2013) and WUSSH (Schmandt & Humphreys 2010) as well as NWUS (James *et al.* 2011), based on teleseismic traveltimes and surface wave dispersion data.

To summarize, we see several trends that indicate improvements in tomographic imaging in recent years: increasingly, instead of fitting traveltimes of isolated phases, full waveforms are considered. Note however that waveforms have been used for a long time in global tomography: Woodhouse & Dziewonski (1984) already included time domain waveforms and Li & Romanowicz (1996) introduced mantle models entirely based on waveforms. Increasingly, the forward computation and the sensitivities of these waveforms to 3-D structure are computed more accurately, transitioning from asymptotic normal mode based methods to numerical SEM-based methods. Many of these improvements come at the expense of computation time and also the inverse problem can become more difficult to tackle (Panning *et al.* 2009). Different compromises have to be made, and it is not easy to judge what is gained with each combination of the different techniques, because a rigorous investigation would itself require very heavy computations.

Our model selection covers a large variety of modelling techniques for the forward as well as for the inverse problem. Also the scale at which they are employed is diverse and ranges from the largest scales (global) and frequencies (normal modes) to the smallest (regional studies and higher frequency surface/body waves). A common property of all models is their uneven but to some degree correlated station/event distribution which cannot be changed. Still, the individual path distributions of each model differ, especially when comparing regional and global models. Comparing these two types of models, we therefore hope to obtain an estimate of feature strength and shape that is reasonably independent of modelling choices.

4 WHOLE MANTLE SPECTRA

4.1 Power and correlation

In Fig. 1, we show a comparison of the power spectrum of several recent global whole mantle models defined in the standard linear (S_l) versus the chosen logarithmic (s_l) power per degree.

In the standard representation (Fig. 1a), a strong degree 2 is the clearest feature throughout the mantle, a well-known feature, that can be seen directly in the data from mode frequency shifts (Masters *et al.* 1982), or differential traveltimes of S–ScS (e.g. Manners 2008; Lekic *et al.* 2012). Whereas low degrees are strongest everywhere, one observes the strongly 'red' spectrum at the top and the bottom of the mantle. As we have argued, a maximum in linear power at degree two does not necessarily reflect that the model features have this dominant size but certainly that they are to some degree globally distributed according to this wavelength. The particular global distribution of model features according to long wavelength

(a) linear power spectral density

(b) logarithmic power spectral density
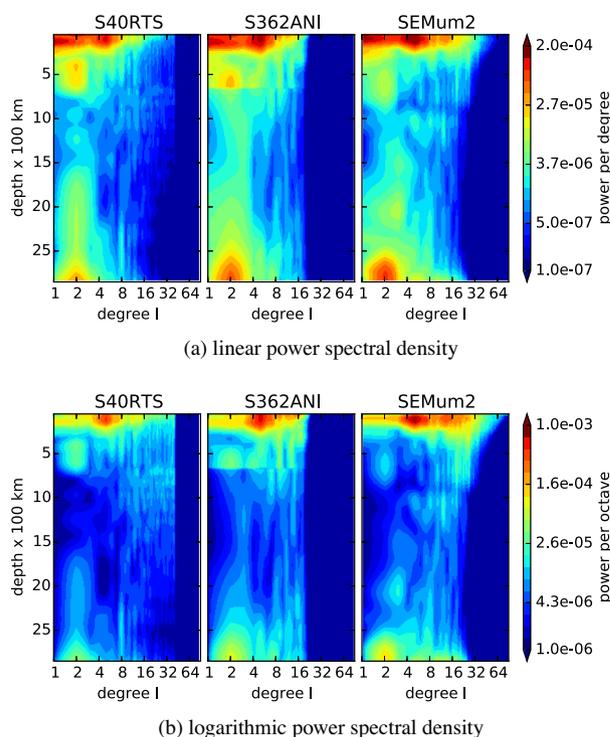
Figure 1. Comparison of whole mantle spectra for the three recent whole mantle models, S362ANI (Kustowski *et al.* 2008), S40RTS (Ritsema *et al.* 2011), SEMum2 (French *et al.* 2013). (a) power per degree spectra: this linear spectral density can be associated with the global distribution of features according to spherical harmonics with a certain wavelength. (b) In contrast, the logarithmic power per octave density can be associated with localized features of a certain size. It is, in general, no longer dominated by low degrees and some clear features appear, such as peaks at degree 5 in the upper 250 km, degree 2 in the transition zone and D″, as well as bands with enhanced small-scale features that differ between models.
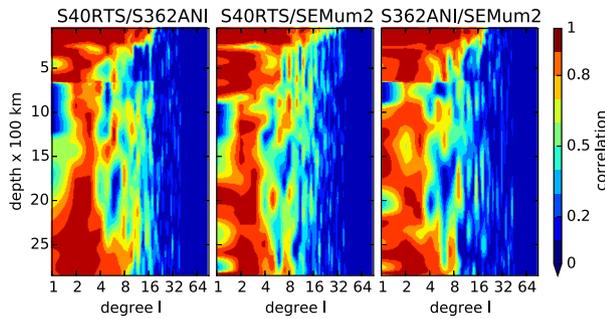
Figure 2. Global mantle model correlations: models correlate by far best in the heterosphere. Deeper, good correlation drops to the lowest degrees, until it increases again towards the lowermost part of the mantle. In general, low degree bands with strong power favour high correlation coefficients, whereas weak or high degree bands do not. Model power, on the other hand, does have strong power in high degree bands, for example in the mid-mantle, that do not correlate well (compare with Fig. 1).

Figure 3. The horizontal distribution function of seismic velocities. We have normalized the maximum of the distribution function to one at each depth. Strong departures from Gaussian behaviour can be observed close to the boundaries.

Figure 4. Antipodal symmetry in the global power spectrum. Even (blue) and odd (red) model power for degrees 1–12. There are possible regions with strong antipodal symmetry in the Earth's mantle in models SEMum2 and S40RTS: the large low shear velocity provinces (LLSVP's) at the bottom of the mantle and the subduction zones (Japan and South America) in the transition zone. At the bottom of the mantle, power is concentrated in degree 2. The subduction zones in the transition zone are small and their power is spread also to higher even degrees (sometimes called a 'zebra' spectrum). S40RTS has enhanced even degree power throughout the lower mantle which could be a consequence of the splitting function constraints that were used to construct the model. Splitting functions are only sensitive to even degree power.
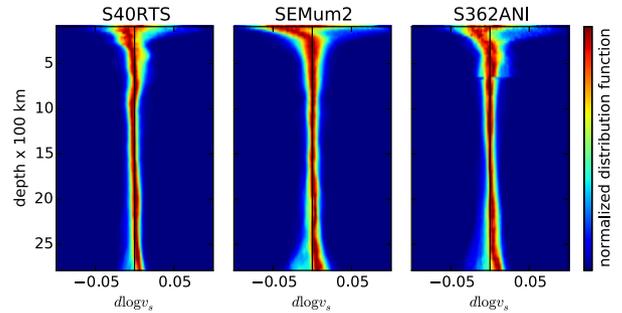
structures reflected by the linear power per degree spectrum, is not the focus of this paper but its significance for mantle dynamics is well recognized (e.g. Dziewonski *et al.* 2010).

The transformation to the local logarithmic (Fig. 1b) power spectral density results in the following changes: In the 'heterosphere' (upper ∼250 km), the strong degrees 1 and 2 become much weaker and the degree 5 corner becomes the isolated dominant feature size. There are some clear differences between the models at higher degrees and the spectrum drops at some point sharply towards the parametrization limits of the models. Degree 2 is now only a distinguishable peak in the transition zone above 660 km and in D″. The logarithmic power per degree confirms that smaller scale features ($l > 8$) dominate the models at intermediate depths.
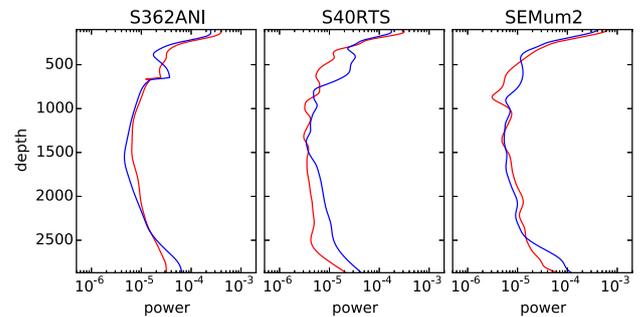
In Fig. 2, we show model correlations for the whole mantle at each degree l. We see that even higher degrees ($l \sim 16$–32) correlate reasonably well in the heterosphere as opposed to the mid-mantle, where the correlation is only high for the lowest degrees $l < 4 \sim 8$. Towards D″, correlation increases again, even more so considering the absolute, not angular wavelength of degree 8 at the core–mantle boundary (CMB) which is equivalent to about degree 16 at the surface. Away from the top and bottom boundaries, we thus see large regions in the mid-mantle in which small scale heterogeneities that do not correlate well across models dominate the total model power. It is only the general long wavelength distribution of these small features for which models agree. We will discuss this further when we examine regional tomographic models.

4.2 Anisotropy and symmetry axes in the global spectrum

In this section, we analyse departures of the global models from Gaussian and isotropic properties. To this end, we show in Fig. 3 the horizontal distribution function of relative velocity perturbations as a function of depth. The maxima of the distribution functions have been normalized to one at each depth. All three global models clearly show that there are regions, especially at the top and bottom boundaries of the models, that are non-Gaussian. Note that a non-Gaussian model implies a phase relation between spherical harmonic coefficients of different degrees l that should therefore not be interpreted independently. The model can nevertheless be isotropic, with coefficients of different order m being independent. The horizontal distribution function of the velocity perturbations has been examined more closely in deep and upper mantle by previous authors (e.g. Hernlund & Houser 2008; Lekic & Romanowicz 2011; Lekic *et al.* 2012).

Fig. 4 shows summed spectral power for degrees 1–12 separated into even and odd degrees. We can further see, that at certain depths, Earth's mantle has preferred antipodal alignment of features near the CMB, and in some models also in the transition zone (SEMum2, S40RTS). Whereas this antipodal alignment is mostly due to degree 2 at the CMB, it extends to higher degrees in the transition zone (zebra spectrum) and can be identified with the strong, antipodal subduction zones in Japan and South America. Model S40RTS, which is amongst other data based on normal mode splitting functions that can only resolve the even structure of the Earth, shows enhanced antipodal features throughout the whole lower mantle.

Apart from departures from Gaussian behaviour and antipodal alignment, heterogeneities could also have special preferred orientations on the Earth. Spherical harmonics can be divided into zonal, that is those with low order m and 'sectoral', that is those with high order m. The zonal harmonics oscillate in latitude and have longer wavelengths in the east–west direction whereas the sectoral harmonics oscillate in the longitudinal direction and are elongated in the north–south direction. We examine whether power is preferentially
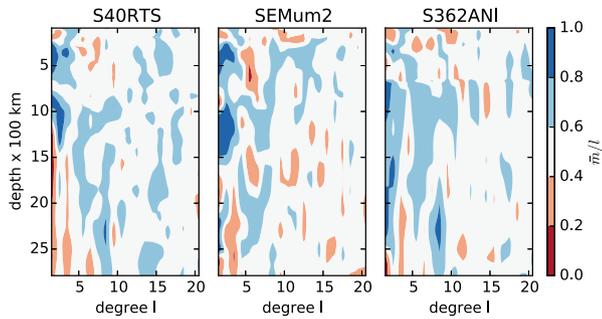
Figure 5. Spherical harmonics anisotropy. The power-weighted average of order m divided by the angular order ($\bar{m}/l$) is preferentially greater than 0.5 (blue colour). This indicates a preferred north–south orientation of heterogeneities.

in one or the other, by computing the quantity A_l:

$$A_l = \sum_{m=-l}^{l} (1 + \delta_{m0}) \frac{c_{lm}^* c_{lm}}{S_l} \frac{|m|}{l}. \tag{8}$$

Note that the zero order needs to be counted twice because it is represented by a single coefficient with similar distribution as the two coefficients for $|m| > 0$ (see eq. 6). In an isotropic model, we expect this quantity to approach $A_l \approx 0.5$. Fig. 5 shows that it oscillates indeed around 0.5 on average with a trend towards north–south orientation of heterogeneities (blue colours), especially in the mid-mantle where it reflects the orientation of the subduction zones.

We can generalize this measurement by searching for an axis of maximal anisotropy in the Earth's mantle: to this end, we rotate the model to points on a grid that cover the whole Earth, compute A_l and save the axis where it becomes maximal (sectoral harmonics) or minimal (zonal harmonics). With this procedure, we can identify anisotropy independent of Earth's axis of rotation. Fig. 6 shows the results for the degree 2 component of the Earth. We show the maximum deviations of A_2 from 0.5 achieved in (a) for sectoral and in (b) for zonal harmonics throughout the mantle. In general, a zonal degree 2 structure is preferred, but there are some depth ranges at which it becomes more sectoral. Although the degree 2 structure correlates relatively well amongst models, S362ANI sometimes prefers a sectoral shape whereas SEMum2 and S40RTS prefer a zonal one. At the bottom of the mantle as well as in the transition zone, where degree 2 dominates total power, zonal harmonics are clearly preferred in all models.

We plot the axes of symmetry where maximum anisotropy is achieved in Fig. 6(c) (sectoral) and (d) (zonal) for evenly distributed depth samples from 90 to 2850 km, if the deviation of A_2 from the isotropic value of 0.5 is larger than 0.35. Symmetry axes cluster very clearly around the equator for the zonal (d) and around the poles (c) for sectoral harmonics. At large depths, the zonal harmonics are centred on the LLSVP's (red colour). At shallower depths, they are mostly centred in the subduction zones of east Asia and South America (blue colour). The symmetry analysis shows that indeed, Earth's rotation axis seems to be (slightly) preferred in the long wavelength structure in most of today's mantle, supporting a deeper connection between both, as has also been proposed in Dziewonski *et al.* (2010).

Because of this anisotropy and the non-Gaussian distribution function, we have to interpret the power spectrum carefully. We can, in any case, interpret power in a spectral band in terms of a convolution with an axisymmetric wavelet.
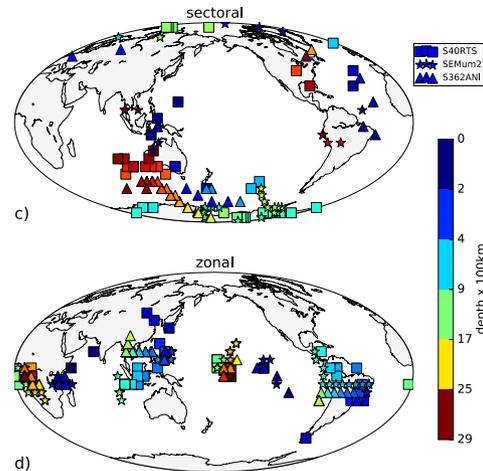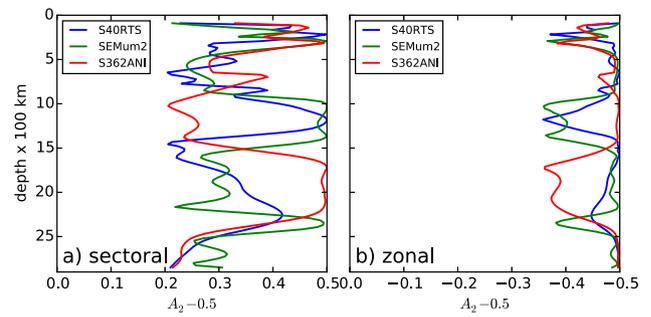

Figure 6. Orientation of symmetry axes of degree 2. We search for an axis that maximizes (a) or minimizes (b) A_2 (see definition in main text), thus indicating an axis of symmetry. In this axis, models are preferably sectoral (a) or zonal (b). In general, the zonal structure (b) is stronger than the sectoral (a). If the zonal structure becomes weaker, the sectoral one, which can otherwise be quite weak, takes over. Panels (c) and (d) show the corresponding symmetry axis of sectoral and zonal degree 2 structure (plotted only if $|A_2 - 0.5| > 0.35$ with S40RTS = squares, SEMum2 = stars, S362ANI = triangles): the sectoral harmonics cluster around the poles and the zonal harmonics cluster around the equator at all depths. Close to the core–mantle boundary (red colour), the symmetry axis is determined by the LLSVPs, but at depths of 900–1700 km (green) it shifts by about 90° in longitude, towards the subduction zones in East Asia and South America.

5 UPPER MANTLE SPECTRA

5.1 Comparison of models in map view

For the qualitative investigation of the different models in the uppermost mantle and the regions from which we obtain spectral estimates, we consider maps at 100 km depth in the upper mantle (Fig. 7). For a good visual comparison, we have adjusted the average shear velocity of the models. Whereas in the later spectral analysis this will play only a minor role because we will be interested in local perturbations, here we find that perturbations from a common average are not a good way to present models because differences in the average model velocities can hide the resolved heterogeneities. This underlines the need for the construction of a reference long wavelength global 3-D model that can also be used in regional model construction. We have ordered the models from left to right roughly by the amount of small-scale structure that they are showing. When we speak of 'apparent resolution' in the following, we refer to the amount of small-scale structure that is visible in a model, which can be biased by artifacts due to the inversion process.
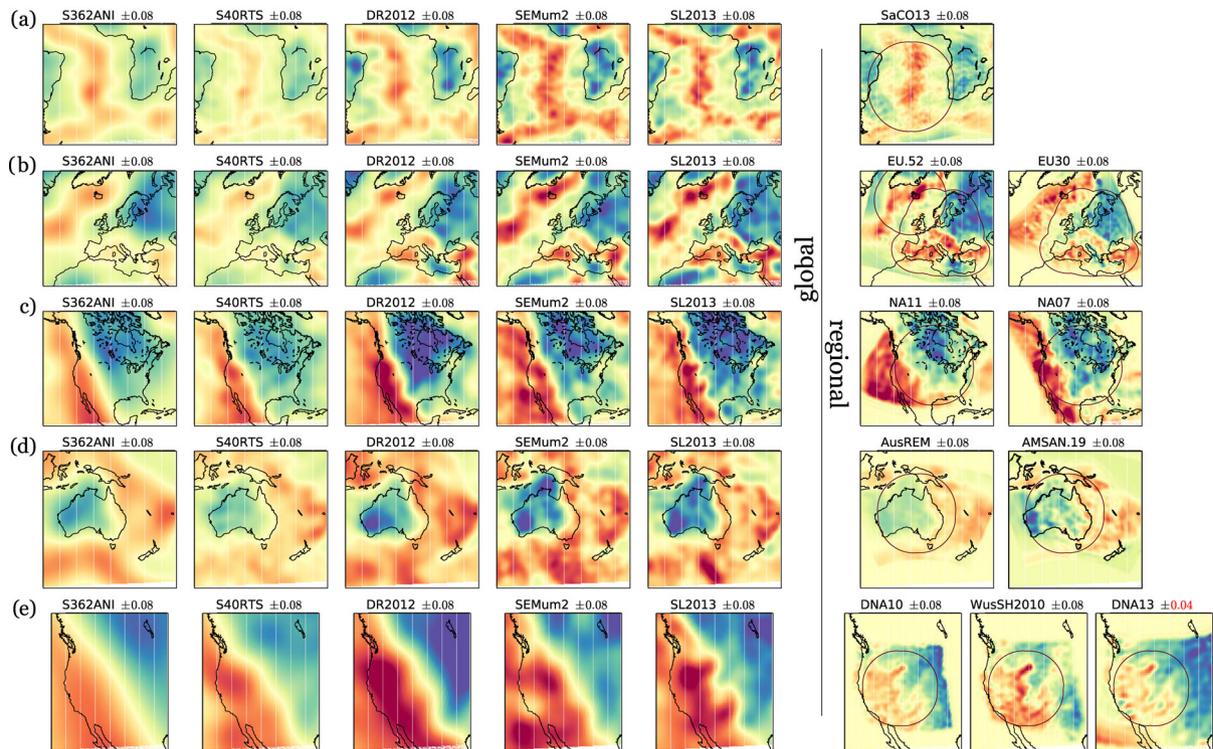
Figure 7. Maps of the different global and regional models in different regions at 100 km depth. Models are approximately sorted by the small-scale power that they have (left to right) and the area that they cover (top to bottom). A contour at 10 per cent of the mean taper power is also shown for the regional models. We use between 3 and 6 taper windows, depending on the region. Each taper has less than 1 per cent of its power outside of the intended area. For model EU.52 (Europe), two contours are shown, the circular one corresponding to the region around Iceland, which is included in EU.52 but not well resolved in EU30. Note that model DNA13 is shown on a different colourscale because it is significantly weaker than the other models.

(1) The South Atlantic (Fig. 7a) is the largest examined region. All models show the general ridge-craton dichotomy and the ridges are well centred on the plate boundary (red line). Additionally, some consistent smaller scale structure can be seen in both global and regional models. For example, a slow velocity anomaly east of South America, in the vicinity of the Trinidade hotspot, is consistently present, as well as some wave speed variations along the ridge. The regional model has a further characteristic: it shows many small-scale fluctuations which are likely due to the inversion process. In this case, the resolved larger features and this inversion noise can be easily separated because they have clearly different scales and we are going to see this in the power spectrum.

(2) Europe (Fig. 7b) is an interesting region because it is tectonically active with dense station coverage. The two considered regional models of Europe claim to reach very high resolutions. It is remarkable to see how similar the global and regional models are, even in their smaller scale features. Starting from the lowest apparent resolution models that show mostly the contrast between the large craton and the North Atlantic ridge, many strong smaller scale features appear as the apparent model resolution increases (e.g. Adriatic sea, variations along the ridges at hotspot locations). Furthermore, amplitudes increase with apparent resolution, especially for smaller scale features such as ridges and hotspots, less so for larger scale features such as cratons. In the regional models, the amplitudes of features of similar size (e.g. Pannonian basin in middle Europe), seem quite similar to the ones in the most recent global models. On the other hand, a closer look reveals that the resolution of the regional models appears higher in some areas than that of the global ones: the Massif Central in France is present in the regional

models, the ridge and hotspot features are also smaller in spatial extent. At this scale, however, some other small scale features become gradually inconsistent amongst the regional models.

(3) In the North American continent (Fig. 7c), as expected, the large scale distribution of slow and fast velocities is consistent across all models. Additionally, some consistent smaller scale features appear in all models, such as the shape of the southern edge of the craton. The craton sometimes has a smoother, sometimes a sharper border and other smaller scale features are somewhat inconsistent between regional and global models. The overall agreement is therefore high only at long wavelengths.

(4) In Australia (Fig. 7d), all models resolve the craton, its edge and the surrounding ridges. There is also a consistently faster western part within the craton. However, at smaller scales, the models are quite different and even the two regional models do not agree very well. It should also be noted that amplitudes, at least for the larger scale features, are much higher in the global models than in the regional ones, which is surprising. Furthermore, enhanced small scale features are visible in the regional model AMSAN.19.
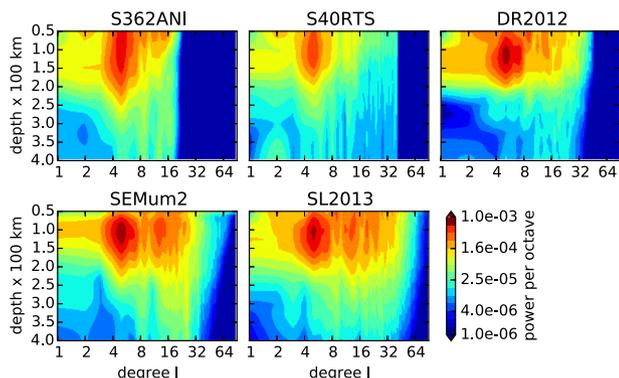
(5) The western United States models (Fig. 7e) cover the smallest of all examined regions and reach the highest nominal resolutions (see Becker 2012, for a regional model comparison). This is an area with high seismic station coverage that makes such a resolution possible. Whereas the model S362ANI shows only a straight tectonic transition from slow to fast velocities, this boundary becomes rugged and complex with increasing apparent resolution. The smaller scale features are consistent between the models but the small-scale regional models have weaker amplitudes, in general, than the global ones. This could be due to the fact that the regional

models are under-representing the longest wavelength component of heterogeneity. Interestingly, even though we are here at the resolution limit of the global models, we can see that they do seem to consistently resolve some of the very small features that are visible in the local studies.
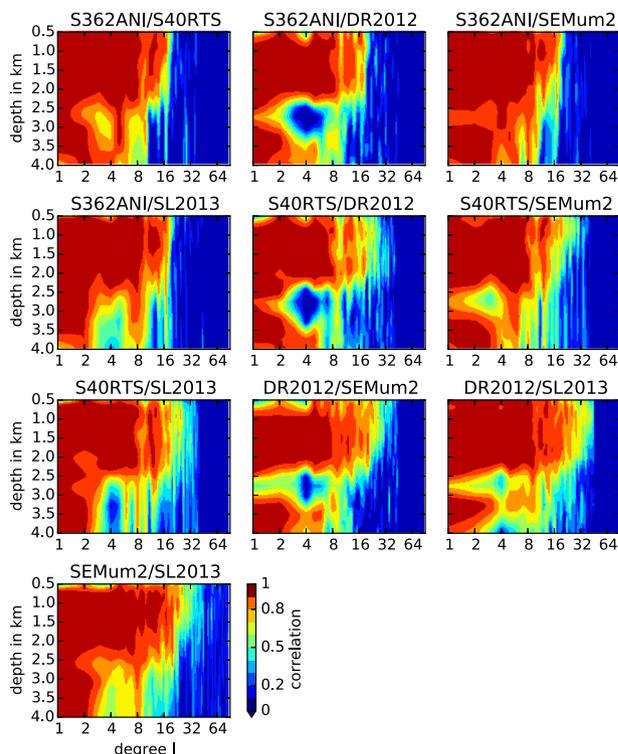
In summary, in the heterosphere, the most recent global and regional models seem to agree on heterogeneities that are smaller than the main typical ridge-craton features. These can be either sharpened larger scale structures or completely 'new' features that do not appear at all in lower resolution models. Small scale structure that does not appear in the smooth global models can have significant amplitude as seen in the European and western United States cases. Of all examined regions, only in Europe do all models agree in shape and amplitude of the features at their respective resolution levels, which is remarkable because they were constructed using, in general, very different modelling techniques.

5.2 Global upper mantle spectra and correlations

We now examine the upper mantle spectrum quantitatively. Before we go to the regional models, we focus on the power per octave spectra of the five global upper mantle models (Fig. 8). The spectra are ordered as before from lowest apparent resolution (upper left-hand panel) to highest apparent resolution (lower right-hand panel). All models show a very clear peak at degree 5 in the heterosphere around 100 km depth. We can also see gradually increasing power from model to model, especially in degrees greater than 5. SEMum2 and SL2013 and to some extent DR2012 show a second spectral peak appearing at around degree 16. Between 150 and 200 km (S362ANI: 200–250 km), the spectra change and the degree 5 peak vanishes until, between about 200 and 400 km depth, models consistently show stronger power in the high than in the low degrees. Fig. 9 shows the corresponding model correlations. As we have seen before, usually models correlate to quite high degrees in the heterosphere, but especially well up to about degree 8. The best correlating models are SEMum2, SL2013 and DR2012, which have correlation coefficients >0.7 up to about degree 20–30 (~1000–650 km). Correlations with models S40RTS and S362ANI drop between degrees 8–16 (~2500–1250 km). Overall, total model correlation is high

Figure 9. Upper mantle global model correlations. In the heterosphere, the correlation coefficient for models SEMum2, SL2013 and DR2012 is >0.7 up to $l = 16$–32 (2500–1250 km). Correlation with S362ANI and S40RTS drops towards degree 16 (500–2500 km). This is consistent with the approximate distribution of model power at these depths. Below 250 km depth, this changes and correlation drops to low degrees even though much of the model power is in the high degrees.

which is reflected in the visual similarity of the models. Immediately below the heterosphere, model correlation drops in general sharply, whereas power stays strong in the higher degrees (especially for DR2012, SEMum2 and SL2013), which could be due to quickly radially varying structure that is differently resolved.

Because heterogeneities are strongest in the heterosphere where models correlate best, we compare the global models with the regional ones in two steps: In Section 5.3, we focus on the heterosphere, examining models, averaged around 100 km depth (that is why we showed map views at this depth). We try to average over a range that is far enough away from the crust as well as from the lower boundary of the heterosphere at around 200 km. In Section 5.4, we interpret the shape of the spectrum in terms of physical processes and we finally focus in Section 5.5 on a few regional models that we examine over the whole upper mantle.

5.3 Combined power spectra of the heterosphere

We now compute the horizontal heterosphere spectra of global and regional models and compare both. To this end, the models have been averaged over depths from 90 to 140 km. Depth averaging changes the spectrum only slightly compared to a single depth slice at 100 km depth indicating that models are smooth enough in radial direction for a meaningful comparison. We recall that the absolute value of the power per octave at degree l can be interpreted as the variance of the model filtered around half wavelength $\pi r_{Earth}/l$. Heterogeneities of even strength and even coverage on all scales result in a flat spectrum. A spectral drop indicates that heterogeneities

Figure 8. Comparison of upper mantle spectra for the global models considered, down to 400 km depth. Models are approximately arranged by their apparent resolution. In the first 200 km, most differences are in the higher degrees ($l = 8$–32). The recent models DR2012, SEMum2 and SL2013 include stronger small scale structure that adds power in this band. If this structure is not artificial due to the inversion process models could therefore have advanced in resolution mostly in this range. Below 200 km, model power is consistently in high degrees.
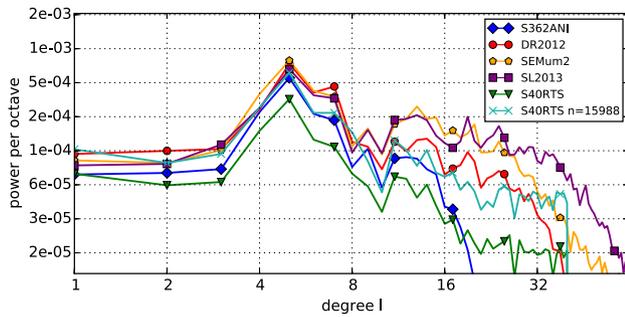
Figure 10. Comparison of 'power per octave' spectra for global models. Models have been averaged between 90 and 140 km depth before computing the spectra, to smooth out rapidly varying structure in the radial direction. Apart of the clear degree 5 peak, we can see a consistent minimum or corner around degrees 8–10 with a subsequent maximum (or at least saddle point). Spectra then drop smoothly (except S40RTS).

with decreasing size become weaker (smaller amplitude) or sparser such that their total power decreases.

Fig. 10 shows the power per octave spectra of the five global models considered and additionally a much less damped version of S40RTS (number of free parameters 15 988). Features of degree $l = 4$–7 (5000–2850 km) clearly dominate the whole spectrum, whereas larger structures ($l = 1$–3) and smaller ones (>8) are weaker. In the most recent models, a minimum around degree $l = 8$–11 (2500–1800 km) is also visible, followed by a broad maximum or at least saddle point, between degrees $l = 11$–16 (1800–1250 km). All models drop clearly in power towards their parametrization limit, except the S40RTS models. Spectral power is most similar at degrees $l \lesssim 7$ (a factor of 2 in relative differences). The largest relative differences can be seen at higher degrees (factor 3 at $l = 15$ and factor 4–8 at $l = 20$). The relative factors in amplitude are 1.4 at $l = 5$ and 2–2.8 at $l = 20$. By far the weakest model is the 'standard' S40RTS version. A less damped version of S40RTS, shown in Fig. 10, yields similar amplitudes as the other models at low degrees (factor 1.4 at $l = 5$, 1.18 in amplitude) but also much stronger higher degree power which is presumably not very well constrained. Although the low degree amplitudes of the less damped version of S40RTS agree better with other models, we have decided to keep the standard version in what follows. Excluding the standard S40RTS, we can further see that models agree best in amplitude around degree 5, where heterogeneities are strongest.

Whereas it is unclear, where and how strongly the model spectra are damped due to regularization, we can see drops in heterogeneity power that can be associated at least with an apparent resolution limit at the following scales: S362ANI: $l = 15$–16 (1250 km), S40RTS: (unclear), DR2012: $l = 28$ (700 km?), SEMum2: $l = 25$ (800 km), SL2013: $l = 36$ (550 km). In the case of S362ANI, SEMum2 and SL2013, these length-scales agree very roughly with the nominal *a priori* model resolution and imposed correlation lengths given by the authors. S40RTS shows yet another particular behaviour: whereas most models show a gradual spectral decay that becomes stronger and stronger with higher l, the power of S40RTS levels off for $l > 20$ up to its apparent resolution limit. A similar behaviour can also be observed in its less damped version. This is suspicious because, in its spherical harmonics parametrization, degrees close to $l = 40$ become more and more delocalized and harder to resolve. We would therefore expect the spectrum to decay gradually. On the other hand, its power at these scales is weaker than in degrees $l = 16$, so that we do not expect these degrees to have strong influence on model features, at least in the heterosphere.
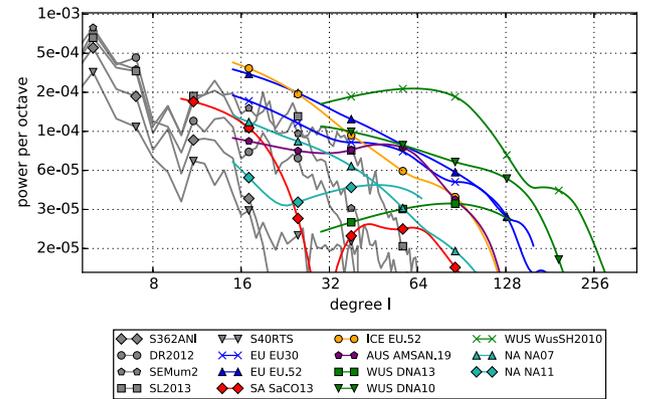
Figure 11. Power per octave spectra of all global (grey) models and debiased regional (colours) models in regions: EU, Europe; SA, South Atlantic; ICE, Iceland; AUS, Australia; WUS, Western USA; NA, North America. Model spectra drop sharply at different degrees at their resolution limit. Such transition bands can not be interpreted as heterogeneity strength. Although there is much variance amongst the models, their ensemble indicates a clear trend that is consistent with the stronger spectra of the recent models DR2012, SEMum2, SL2013. Some spectra (SaCO13, AMSAN.19, NA11 and DNA13) show 'bumps' close to their apparent resolution limit that can indicate artificially strengthened structure due to the inversion process.

From the spectral differences, we can see that models resolve heterogeneity amplitudes differently, independent of their nominal resolutions. The relative differences are larger when features are smaller or weaker as we would expect. Differences depend on the employed inversion techniques, including data coverage, forward and inverse modelling and we cannot definitely conclude which amplitude is more 'correct'.

From the subdivision of the global model spectrum into three regions: degree 5 peak, intermediate peak or saddle-point at degrees $l \gtrsim 13$ with a final spectral drop, we can already see that the uppermost mantle spectrum is more complex than a power-law decay spectrum with a single length scale.

We now turn to the regional power spectral estimates. Fig. 11 shows the power per octave spectra of our regional model selection with the global model spectra shown in the background (grey). The regional spectra have been debiased and are smooth on scales on the order of their taper bandwidth. We show the power only for those degrees that can be reliably estimated, that is degrees higher than about the taper bandwidth. Even though the regional models focus on smaller regions, they do not necessarily have higher resolution than the global models. For a reasonable comparison, we need to select the spectral bands of a model that it resolves at least apparently. A clear indicator for a models apparent resolution limit are transition bands with a corner and strong spectral decay rates.

Clear evidence for such a transition band can be seen in the South Atlantic model SaCO13, which is based on long period data only. It shows a sharp drop in power around degree $l = 16$ (1250 km) where it reaches its apparent resolution limit. In other models, a spectral drop can ultimately be observed at their very high-end spectrum.

There are a few models (SaCO13, NA11, AMSAN.19, WusSH and DNA13) that have an isolated peak in heterogeneity power at high degrees. This can be reflected in the maps (Fig. 7) as 'spotty' small-scale structure that is likely an artifact of the inversion process. In the case of the SaCO13 model, this 'secondary' scale length is clearly visible as a spectral peak at degree $l \lesssim 50$ (400 km). Here, it is well separated from the rest of the model power (i.e. degrees $l < 20$), and model features can easily be separated into 'true' and 'artificial' features. In other models, a secondary peak is more subtle
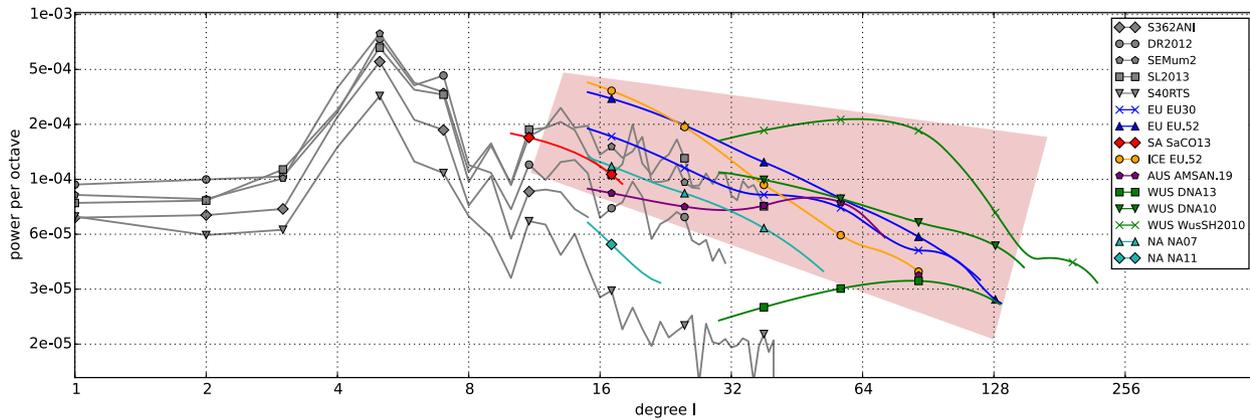
Figure 12. Global and cleaned regional model spectra. We have removed spectral bands that appear to surpass the models resolution limit, such as strong spectral drops (compare with Fig. 11) or bands that surpass nominal resolution, given by the authors or the data that they use. Although there is much variance, the models indicate a clear spectral trend that coincides with the spectra of the recent global models DR2012,SEMum2,SL2013. Only the horizontally anisotropic model NA11, in the examined band, and the largest scales in model DNA13 have weaker heterogeneities and are slightly off the trend.

(AMSAN.19, WusSH, NA11 and DNA13) but still visible. Interestingly, model DNA10 does not show this behaviour while its successor DNA13 does: Whereas the former contains large features with small scale structure at their boundaries, the latter predominantly consists of small scale features (see Fig. 7). Most of these peaks are close to the parametrization limit of the respective models, supporting the inference that these features are artifacts of the inversion process. This shows that smaller model features can be 'polluted', that is, artificially strengthened compared to larger scales.

Fig. 12 shows the combined global and regional spectra with identified unresolved bands taken out. In addition, we have removed the secondary peak of model NA11 around degrees 40–50 which also seems to be an artifact. The general model trend, which spans all models except NA11 and S40RTS, is indicated by a red box. This trend overlaps with the spectra of the global models DR2012, SEMum2 and SL2013 around degree $l = 20$–30. Its average is slightly higher than DR2012, closest to SEMum2 and SL2013 and it spans a range of about 4 ($l = 20$) to 8 ($l = 100$) in relative power which corresponds to a factor of 2 ($l = 20$) to 3 ($l = 100$) in amplitude. This confirms the visual impression from Fig. 7 that amplitudes of DR2012, SEMum2 and SL2013 are closest to those of the regional models. Within the indicated trend, the EU.52 model has the highest amplitudes whereas EU30 matches SEMum2 and SL2013. The models SaCO13, NA07 and AMSAN.19 agree well with the model trend. SaCO13 and NA07 are between the power spectra of SEMum2,SL2013 and DR2012 and do not have higher apparent resolution than the global models. AMSAN.19 is closest to DR2012 in overlapping bands.

Assuming that the European and Icelandic spectra are well constrained at these from their perspective larger scales ($l = 20$, 1000 km) and overestimate the global average because they are obtained for specific, tectonically active regions, at least SEMum2 and SL2013 seem not to have significantly reduced amplitudes at this scale ($l = 20$, 1000 km). This does of course not mean that their structure at these scales is unaffected by regularization but its strength appears consistent with, and not weaker than the regional model trend. At smaller scales, the DNA10, DNA13 and WusSH models also generally agree with the considered trend but span a very wide range of spectral power (factor 8).

Model NA11 seems to be weaker than the other models at these scales ($l = 20$, 1000 km) which confirms the visual impression from the map view. A possible but not further examined explanation is that

it includes azimuthal anisotropy and that the inclusion of additional parameters results in a smoother model.

The variations in the regional spectra show that their spectral shape is not very well constrained. The regional models that agree best in spectral amplitude and shape, but also visually with the global ones, especially with SEMum2 and SL2013, are the Europe and Iceland models (EU30,EU.52). These regional models benefit from both a good distribution of stations and events and a fairly advanced waveform modelling technique. As we previously discussed, the spectral shape is well resolved in all global models up to about degree 8, including the degree 5 peak. The spectral minimum around degree 8–10, and the second broad peak around degrees 11–16, which is present in the recent tomographic models (SEMum2 SL2013, less pronounced in DR2012), also seem to be robust, because they are supported by the regional models.

5.4 Interpretation of the heterosphere spectrum

In this section, we propose an explanation for the particular shape of the uppermost mantle spectrum and especially its two spectral peaks around degrees 5 and degrees 11–16. For this purpose, we again consider a global model average between 90 and 140 km depth, and separate it into faster and slower than average parts, which generates two new 'artificial' models.

These models are physically meaningful because, at these depths, regions that are significantly faster than average correspond primarily to continental lithosphere, and in particular, old cratonic regions. Regions that are slower than average, on the other hand, correspond mostly to upwelling regions at plate boundaries, such as mid-ocean ridges and backarc basins. This can be objectively assessed, for example, from model cluster analysis (e.g. Lekić & Romanowicz 2011). The precise threshold value for separating fast and slow regions is somewhat arbitrary. However, a value near the global average velocity at the corresponding depth works well for this purpose (e.g. fig. 3d in Lekić & Romanowicz 2011). We choose a threshold value of $|dvs/vs| = 0.01$. This separation creates artificial edges with a discontinuous gradient. We have evaluated the influence of these edges on the resulting spectra by testing different threshold values, and have found that it is very small and only important at the very high end of the spectrum.

Fig. 13 shows the global power spectra of the fast and slow parts of the three most recent global models and illustrates how
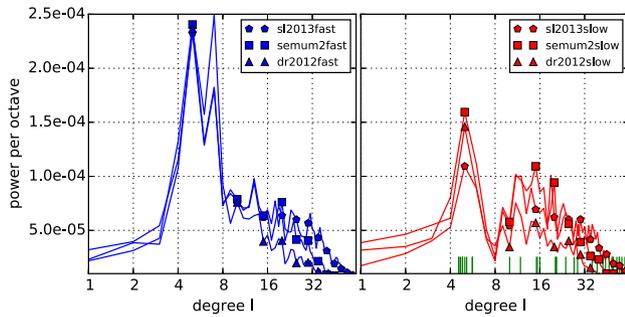
Figure 13. Spectra of fast (blue) and slow (red) heterogeneities in semi-logarithmic scale. Models have been averaged from 90 to 140 km depth. Due to the logarithmic power per octave density, the area under the graph can be interpreted as the variance contribution of this spectral band. Whereas fast velocities are clearly dominated by large scale structure around degree 5–7, slow velocities have one sharp peak at degree 5 and another broad peak around degree 16. The latter feature is most pronounced in SEMum2, DR2012 and SL2013, which have strong, narrow ridges, but can also be seen to a lesser degree in S362ANI and S40RTS (not shown). Fast heterogeneities are stronger than the slow ones around degree 5 but weaker around degree 16. Green lines show the sizes of tectonic plates (see text).

this approach might help determine the origin of the two peaks in the spectrum. Both fast and slow spectra show a sharp peak at degree 5 in all three models. Only the slow part shows a minimum around degree 8 with a secondary peak at higher degrees. The slow heterogeneities have more energy in the higher degree spectrum, indicating that they are in general of smaller scale. Note that if this secondary peak was due to artificial edges created during the separation procedure, it would also appear in the fast spectrum.

Fig. 14 shows the slow and fast parts of model SEMum2. Similar results are obtained for other global models. We analyse the model by using two filters that focus on two spectral bands: one band includes the degree 5 peak. The other includes all higher degrees with significant power. We can localize spectral power in the model by interpreting the filtering operation as a convolution with an axisymmetrical wavelet (see Appendix A). Model spectra, filter bands and the associated wavelets are shown in Figs 14.3(a)–(c). Due to the shape of the filters, the wavelets are indeed localized and have one strong main lobe and one strong side lobe. Figs 14.1(a)–(c) (slow part) and Figs 14.2(a)–(c) (fast part) show the unfiltered models (a) and the effect of the two filters (b) and (c).

The slow model is separated in two different scales: the degree 5 filter (Fig. 14.1b) shows all major plate boundaries with strong slow velocities (red). Because the main lobe of the filter is positive, this means these come from slow features that have approximately this size or smaller. Additionally, however, there are large blue regions. These regions come from the negative sidelobes of the wavelet and these features thus reflect the presence of a specific distance between the slow features. The slow model convolved with the high degree filter (Fig. 14.1c) emphasizes the narrow structures that we also see in the original 'slow' model and that form elongated, interconnected 'lines'.

The fast model filtered around degree 5 (Fig. 14.2b), shows the distribution of the main cratons, while the 'slow' regions from the sidelobes of the filter are not as strong (except perhaps in the North Atlantic), indicating a less regular distribution. In the 'fast' model filtered at small scales (Fig. 14.2c), shorter wavelength features are mostly isolated and unevenly distributed.

We can support this visual assessment by looking at the velocity distribution histograms of these filtered models that are shown in Figs 14.4(a)–(c), respectively for the fast model part, the slow model

part, and the original model SEMum2. The histograms confirm the previous description: large scale and small scale histograms of the fast model parts (Fig. 14.4a) are very similar. Both have a long tail in the positive direction, showing that there are indeed fast heterogeneities that respond strongly to the respective filter sizes in the model and no strong slow heterogeneities that respond strongly to the filter sidelobes.

The slow model histograms are very different (Fig. 14.4b) and more symmetric around zero. The histogram of the model filtered around degree 5 shows two distinct peaks that reflect the slow heterogeneities themselves and a dominant distance between them. In contrast to the fast model part, a tail in the negative velocity direction appears only in the histogram corresponding to the model filtered at short scales.

The original model (Fig. 14.4c) reflects the combined behaviours.

Slow velocities therefore have a characteristic distance corresponding to the sidelobes of the degree 5 filter: ∼8000 km. As we have mentioned before, these slow regions delineate the tectonic plates, that thus seem to have a characteristic size themselves. Interestingly, plates that are larger, the Pacific, Africa and Indo-Australia plates, all show slow velocity subdivisions that split them roughly into regions of degree $l = 5$ size (see Figs 14.1a and 2a).

Indeed, direct measurement of the area of tectonic plates at the surface gives a similar result: In Fig. 13, plate sizes taken from Bird (2003) are displayed as green bars in units of angular degrees assuming that plates are approximately circular caps. We have combined the South America and Nazca plates as well as the Africa and Somalia plates that are not divided by (strong) slow velocities at these depths. Then we have split the plates that show subdivisions in slow velocities in two equal parts: these are the Pacific Plate and the Africa–Somalia Plate. India and Australia are also treated as separate plates. This is of course only a rough comparison but it shows how the largest plates cluster very accurately around degree $l = 5$. Without the extra subdivision, some plates are somewhat larger but the peak is still pronounced. However, subdivisions of the Pacific and of the African plates have been noted in a different context before [Pacific subdivision: Montagner (2002), African subdivision at west and central African rift zones: Sebai *et al.* (2006), Australia-India subdivision at 90-east ridge: Delescluse & Chamot-Rooke (2007)]. We would like to point out that the characteristic degree 5 peak is most pronounced at these depths. At shallower depths, for example the combined Africa–Somalia as well as the South America–Nazca plates are separated by slow velocity anomalies that slightly decrease the characteristic size.

The distinct plate size has been analysed from the surface expressions of tectonic plates (Bird 2003; Morra *et al.* 2013) and preferred plate sizes and arrangements have also been theoretically examined (Anderson 2002). As we have seen, their dominant scale becomes very clear in the heterosphere where additional and missing plate subdivisions cluster them even stronger around the degree 5 peak. This peak is furthermore isolated, indicating that the plates are even to some extent ordered according to this distance over multiple plate distances.

Our analysis therefore indicates that the characteristic size of tectonic plates is at the origin of the strong degree $l = 5$ peak. Cratonic regions have similar but also smaller size as they are embedded within these plates. They add constructively to the pattern of slow plate boundary heterogeneities. Naturally, Earth's topography, which is dominated by the variation between continental and oceanic elevations also reflects this characteristic length-scale (e.g. Dziewonski & Romanowicz 2014). Slow velocities further exhibit
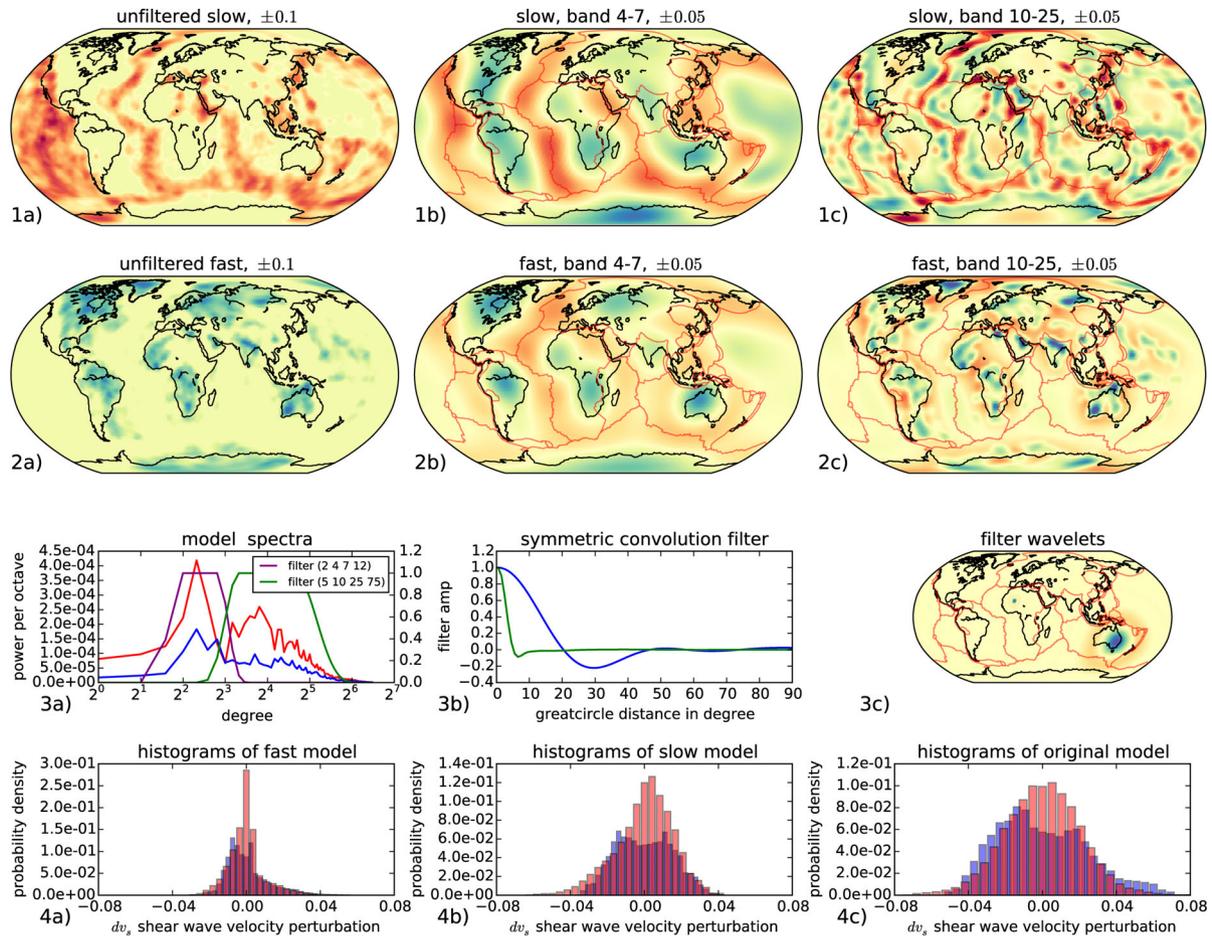
Figure 14. Maps and spectra of the slow and fast model parts. Maps of slow and fast model parts are shown in rows 1 and 2. Models have been clipped at a threshold value of $\frac{dvs}{vs} = +0.01$ and convolved with a large (column 2) and small (column 3) wavelet. Row 3, column 1 shows the respective power per octave spectra (red/blue), the low degree filter that belongs to the large wavelet (purple) and the high degree filter that belongs to the small wavelet (green). Both wavelets are shown in row 3, column 2 and in map view in column 3 (note that the small wavelet is very small and located in the Sahara). Row 4 shows histograms in both bands (low degree filter in blue, high degree filter in red) of the fast model (column 1), the slow model (column 2) and the original SEMum2 model (column 3). See text for a detailed explanation.

enhanced power at higher degrees (degree 11–16), which reflect the actual characteristic size, not distance, of ridges, backarc-basins or of hotspots, as seen in the tomographic models.

A similar separation of slow and fast model spectra, as the one presented in this section, can be done at different depths (not shown). While both spectra have a peak at degree 5 in the depth range that we have examined above, they transfer differently to deeper model parts. The fast model's degree 5 peak extends deeper than the degree 5 peak of the slow model. The higher degree peak at $l = 16$ of the slow model becomes weaker but persists at deeper depths. This indicates that slow velocities keep their characteristic size but lose their long range tectonic signature with depth.

We have identified two structures, fast embedded cratons and slow upwellings, both organized by the tectonic plates. All these three form heterogeneities at different scales in the heterosphere. They leave their signature in the spectrum which has, as a consequence, a more complex shape than a simple power-law decay.

5.5 Combined upper mantle spectra and correlations

The high resolution European models reach very small scales, cover a relatively large area and agree well with the global models at larger scales. They give us the possibility to study the higher degree spectrum and model correlations in detail over a larger depth range than just the heterosphere. The only global models that have a large-enough overlapping spectral band with them are SEMum2 and SL2013.

Fig. 15 shows the power spectra (upper row) and model correlations (lower row) from 50 to 800 km depth for degrees $l = 20 - 150$ (lower degrees are too large for the region and strongly biased). All models have strong power in the heterosphere that decreases rapidly between 200 and 300 km depth. Below this depth, the spectra look rather different. The global models have only weak heterogeneities, the regional models are stronger with significant small scale power. The EU.52 model, which is defined on a grid that coarsens with depth, is dominated by larger scale heterogeneities and shows a cutoff at its grid resolution limit, at least for depths below 400 km ($0.1°, 0.25°, 0.5°, 1°, 2°$ for depths larger than $50, 200, 400, 670, 1420$, corresponding respectively to degrees: $900, 360, 180, 90, 45$). Also in EU30, strong heterogeneities can be seen up to about degree $l = 150$ which is close to its grid resolution limit ($0.5°$ corresponding to degree 180). Because it includes boundary topography, the 660 km discontinuity is visible as a sharp spike.

We see that the apparent horizontal model resolution appears to depend to a large extent on the model's *a priori* constraints
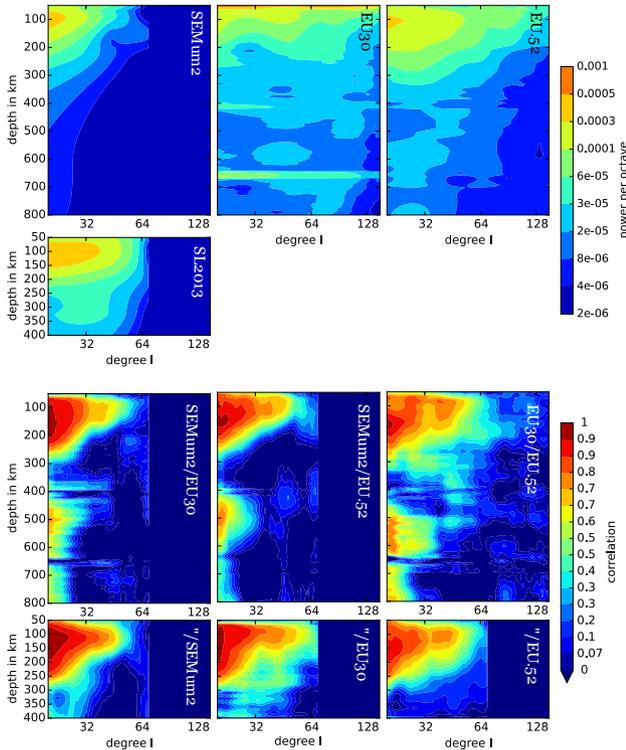
Figure 15. Spectra and model correlations of the upper mantle in Europe for degrees $l > 20$: upper rows: power per octave spectra ($l > 20$) from 50 to 800 km depth. Model SL2013 is shown only down to 400 km, its lower depth limit. As we expect from a regional model, EU30 and EU.52 have much more power in smaller scale structure. Especially deeper than 200–300 km, they have significant power up to their parametrization limit. Lower rows: model correlations for degrees $l > 20$ in the same depth range. All models correlate well to very high degrees in the upper 200 km. Below, as in the global models, only lower degrees correlate well, which means that much of the strong high degree power at these depths is inconsistent between the models.

and significant power is present everywhere within this range. We would, however, expect that resolution also decreases with depth due to weaker heterogeneities that are more difficult to image. If not in the model itself, this reduced resolution is visible in the model correlations in the lower rows of Fig. 15: models correlate very well in the heterosphere (to degree 30–40 ∼ 700–500 km), but only poorly below 300 km depth (for $l > 20$) although much of the model power is still in the higher degrees.

Fig. 16 displays in more detail the global and regional spectra around a depth of 500 km. Models agree in the order of magnitude of heterogeneity strength but there are no common trends or common dominant scales visible as in the heterosphere, except for the distinct flat power per octave spectrum up to at least degree 30.

The lack of correlation in spectral bands with strong heterogeneity means that horizontal model features at these depths may, at least in some models, be quite strongly contaminated by errors due to the inversion process, most likely due to a combination of poor sampling and theoretical approximations. Weaker velocity variations in general, as well as the proximity of strong heterogeneities at shallower depths, make these regions difficult to resolve. On the other hand, in some areas of the Earth, stronger and meaningful model features in the depth range 200–400 km have been identified, such as subduction zones or, in model SEMum2, quasi-periodic (width ∼ 1000 km, degrees 16–32), horizontally elongated low velocity 'fingers' (French *et al.* 2013) in all ocean basins, whose
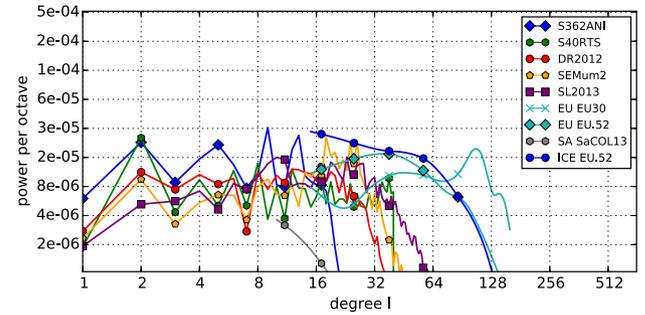
Figure 16. Combined global and european spectra for models that have been averaged from 460 to 540 km depth. SEMum2 is the only model with somewhat increased power at scales around 800 km but in general model power per octave is flat towards higher degrees. Note, however, that power at such short scales is uncorrelated between the models.

robustness has been verified by careful resolution tests, and whose expression has also been reported in regional models (e.g. Katzman *et al.* 1998; Colli *et al.* 2013).

6 CONCLUSIONS

We have obtained spectral estimates of the heterosphere around 100 km depth from global and regional models up to about degree $l = 100$ (∼200 km). The latest regional and global upper mantle models transfer reasonably well into each other at scales of around 1000 km. We therefore do not expect model amplitudes to be strongly reduced by regularization at these scales. The combined power spectrum can be used to constrain the influence of intermediate and smaller scale heterogeneities on seismic waves on the global scale.

The interpretation of the spectrum in terms of localized heterogeneities, as expressed by the logarithmic power per degree l, reveals the presence of different length scales in the heterosphere: the strong peak around degree $l = 5$ reflects the dominant plate and craton sizes. Slow narrow linear features at divergent plate boundaries exhibit a distinct shorter scale length. Such distinct localized features indicate that the spectrum is to some extent non-Gaussian up to high degrees which is confirmed by the horizontal distribution function of velocities. Moreover, throughout the whole mantle, the low degree spectrum is slightly non-isotropic and oriented along Earth's rotation axis.

Plate-scale features dominate the overall heterogeneity power of the heterosphere, and smaller structures between 3000 and 700 km size have relatively even strength before they weaken towards the model resolution limits. Regional models that constrain even smaller scales tend to have weaker power around degree 100 and show a trend of decreasing strength. Since these models represent tectonically active regions that might have stronger than global average heterogeneity power, we have some indication that the global average decreases towards higher degrees and smaller scale features become weaker or sparser.

Instead of isotropically distributed heterogeneities described by a single scale-length and power-law decay process, the heterosphere spectrum can be subdivided in bands that reflect different physical processes. Describing the power spectrum as an isotropic random medium, for example as a von-Karman or exponential medium with a single scale-length and power-law decay rate over all scales, can oversimplifies the heterosphere and is therefore not appropriate for a statistical description of this (important) region. In particular, it is difficult to infer the strength of smaller scales from the strength

of the large scales due to the different physical processes that are involved. On the other hand, the regional spectral estimates, that are consistent with the global ones at intermediate scales and reach higher resolutions, are more suited for this purpose. The different types of heterogeneity dominating at different scales lead to a distinct spectral shape that current tomographic models consistently resolve, at least at the larger scales.

Below the heterosphere, between 200 and ~600 km depth, models correlate only at larger scales. In contrast, dominant heterogeneity power shifts to small scales, in some models up to the smallest scales allowed for by the model parametrization. Even though individual meaningful features can be identified in some models, strong inconsistent small-scale features are present, indicating contamination by noise at these scales and depths which limits our ability to robustly derive peak amplitudes, model roughness or model gradients at these depths at the present time.

ACKNOWLEDGEMENTS

REFERENCES

Anderson, D.L., 2002. How many plates?, *Geology,* **30**(5), 411–414.

Baddour, N., 2010. Operational and convolution properties of three-dimensional Fourier transforms in spherical polar coordinates, *J. Opt. Soc. Am. A,* **27**(10), 2144–2155.

Baldi, P. & Marinucci, D., 2007. Some characterizations of the spherical harmonics coefficients for isotropic random fields, *Stat. Probab. Lett.,* **77**(5), 490–496.

Becker, T.W., 2012. On recent seismic tomography for the western United States, *Geochem. Geophys. Geosyst.,* **13**(1), doi:10.1029/2011GC003977.

Becker, T.W. & Boschi, L., 2002. A comparison of tomographic and geodynamic mantle models, *Geochem. Geophys. Geosyst.,* **3**(1), doi:10.1029/2001GC000168.

Bedle, H. & van der Lee, S., 2009. S velocity variations beneath North America, *J. geophys. Res.,* **114**(B7), B07308, doi:10.1029/2008JB005949.

Bird, P., 2003. An updated digital model of plate boundaries, *Geochem. Geophys. Geosyst.,* **4**(3), doi:10.1029/2001GC000252.

Bunge, H.-P. & Richards, M.A., 1996. The origin of large scale structure in mantle convection: effects of plate motions and viscosity stratification, *Geophys. Res. Lett.,* **23**(21), 2987–2990.

Bunge, H.-P., Richards, M.A., Lithgow-Bertelloni, C., Baumgardner, J.R., Grand, S.P. & Romanowicz, B., 1998. Time scales and heterogeneous structure in geodynamic earth models, *Science,* **280**(5360), 91–95.

Chevrot, S., Montagner, J.-P. & Snieder, R., 1998. The spectrum of tomographic earth models, *Geophys. J. Int.,* **133**(3), 783–788.

Colli, L., Fichtner, A. & Bunge, H.-P., 2013. Full waveform tomography of the upper mantle in the South Atlantic region: imaging a westward fluxing shallow asthenosphere?, *Tectonophysics,* **604,** 26–40.

Dahlen, F. & Tromp, J., 1998. *Theoretical Global Seismology,* Princeton Univ. Press.

Dahlen, F.A. & Simons, F.J., 2008. Spectral estimation on a sphere in geophysics and cosmology, *Geophys. J. Int.,* **174**(3), 774–807.

Debayle, E. & Ricard, Y., 2012. A global shear velocity model of the upper mantle from fundamental and higher Rayleigh mode measurements, *J. geophys. Res.,* **117**(B10), B10308, doi:10.1029/2012JB009288.

Delescluse, M. & Chamot-Rooke, N., 2007. Instantaneous deformation and kinematics of the India–Australia Plate, *Geophys. J. Int.,* **168**(2), 818–842.

Dziewonski, A., 1984. Mapping the lower mantle: Determination of lateral heterogeneity in P velocity up to degree and order 6, *J. geophys. Res.: Solid Earth (1978–2012),* **89**(B7), 5929–5952.

Dziewonski, A. & Romanowicz, B., 2014. Overview, in *Treatise of Geophysics,* Vol. 1, Elsevier, in press.

Dziewonski, A., Lekic, V. & Romanowicz, B., 2013. Planet within a planet: implications of principal component analysis of global tomographic models, in *AGU Fall Meeting,* Abstracts #DI32A-01.

Dziewonski, A.M., Lekic, V. & Romanowicz, B.a., 2010. Mantle anchor structure: an argument for bottom up tectonics, *Earth planet. Sci. Lett.,* **299**(1–2), 69–79.

Fichtner, A., Kennett, B.L.N., Igel, H. & Bunge, H.-P., 2009. Full seismic waveform tomography for upper-mantle structure in the Australasian region using adjoint methods, *Geophys. J. Int.,* **179**(3), 1703–1725.

Fichtner, A., Trampert, J., Cupillard, P., Saygin, E., Taymaz, T., Capdeville, Y. & Villase, A., 2012. Multi-scale full waveform inversion, *Geophys. J. Int.,* pp. 1–25.

Frankel, A., 1989. A review of numerical experiments on seismic wave scattering, *Pure appl. Geophys.,* **131**(4), 639–685.

French, S., Lekic, V. & Romanowicz, B., 2013. Waveform tomography reveals channeled flow at the base of the oceanic asthenosphere, *Science,* **342**(6155), 227–230.

Garcia, R.F., Chevrot, S. & Calvet, M., 2009. Statistical study of seismic heterogeneities at the base of the mantle from PKP differential traveltimes, *Geophys. J. Int.,* **179**(3), 1607–1616.

Hager, B.H., 1984. Subducted slabs and the geoid: constraints on mantle rheology and flow, *J. geophys. Res.: Solid Earth (1978–2012),* **89**(B7), 6003–6015.

Hernlund, J.W. & Houser, C., 2008. On the statistical distribution of seismic velocities in Earth's deep mantle, *Earth planet. Sci. Lett.,* **265**(3-4), 423–437.

Hivon, E. & Górski, K., 2002. MASTER of the cosmic microwave background anisotropy power spectrum: a fast method for statistical analysis of large and complex cosmic microwave background, *Astrophys. J.,* **1,** 2–17.

James, D.E., Fouch, M.J., Carlson, R.W. & Roth, J.B., 2011. Slab fragmentation, edge flow and the origin of the yellowstone hotspot track, *Earth planet. Sci. Lett.,* **311**(1), 124–135.

Katzman, R., Zhao, L. & Jordan, T.H., 1998. High-resolution, two-dimensional vertical tomography of the central pacific mantle using ScS reverberations and frequency-dependent travel times, *J. geophys. Res.: Solid Earth (1978–2012),* **103**(B8), 17 933–17 971.

Kennett, B. L.N., Fichtner, a., Fishwick, S. & Yoshizawa, K., 2012. Australian Seismological Reference Model (AuSREM): mantle component, *Geophys. J. Int.,* **192**(2), 871–887.

Klimeš, L., 2002. Correlation functions of random media, *Pure appl. Geophys.,* **159**(7–8), 1811–1831.

Kustowski, B., Ekström, G. & Dziewoski, A.M., 2008. Anisotropic shear-wave velocity structure of the Earth's mantle: a global model, *J. geophys. Res.,* **113**(B6), 1–23.

Leistedt, B., McEwen, J.D., Vandergheynst, P. & Wiaux, Y., 2013. S2LET: a code to perform fast wavelet analysis on the sphere, *Astron. & Astrophys.,* **558,** doi:10.1051/0004-6361/201220729.

Lekic, V. & Romanowicz, B., 2011. Tectonic regionalization without a priori information: a cluster analysis of upper mantle tomography, *Earth planet. Sci. Lett.,* **308**(1–2), 151–160.

Lekic, V., Cottaar, S., Dziewonski, A. & Romanowicz, B., 2012. Cluster analysis of global lower mantle tomography: a new class of structure and implications for chemical heterogeneity, *Earth planet. Sci. Lett.,* **357,** 68–77.

Lewalle, J., Farge, M. & Schneider, K., 2007. D.6.1 Wavelet transforms, in *Handbook of Experimental Fluid Mechanics,* pp. 1378–1395, eds Tropea, C., Yarin, A. & Foss, J.F., Springer-Verlag.

Li, X.-D. & Romanowicz, B., 1995. Comparison of global waveform inversions with and without considering cross-branch modal coupling, *Geophys. J. Int.,* **121**(3), 695–709.

Li, X.-D. & Romanowicz, B., 1996. Global mantle shear velocity model developed using nonlinear asymptotic coupling theory, *J. geophys. Res.,* **101**(B10), 22 245–22 272.

Manners, U., 2008. Investigating the core-mantle boundary region with S and P diffracted waves, *PhD thesis,* University of California, San Diego.

Masters, G., Jordan, T.H., Silver, P.G. & Gilbert, F., 1982. Aspherical Earth structure from fundamental spheroidal-mode data, *Nature,* **298,** 609–613.

Mégnin, C. & Romanowicz, B.a., 2000. The three-dimensional shear velocity structure of the mantle from the inversion of body, surface and higher-mode waveforms, *Geophys. J. Int.,* **143,** 709–728.

Mégnin, C., Bunge, H.-P., Romanowicz, B. & Richards, M.a., 1997. Imaging 3-D spherical convection models: what can seismic tomography tell us about mantle dynamics?, *Geophys. Res. Lett.,* **24**(11), 1299–1302.

Montagner, J.-P., 2002. Upper mantle low anisotropy channels below the Pacific Plate, *Earth planet. Sci. Lett.,* **202**(2), 263–274.

Morra, G., Seton, M., Quevedo, L. & Müller, R.D., 2013. Organization of the tectonic plates in the last 200 Myr, *Earth planet. Sci. Lett.,* **373,** 93–101.

Obrebski, M., Allen, R.M., Pollitz, F. & Hung, S.-H., 2011. Lithosphere-asthenosphere interaction beneath the western United States from the joint inversion of body-wave traveltimes and surface-wave phase velocities, *Geophys. J. Int.,* **185**(2), 1003–1021.

Panning, M.P., Capdeville, Y. & Romanowicz, B.a., 2009. Seismic waveform modelling in a 3-D Earth using the Born approximation: potential shortcomings and a remedy, *Geophys. J. Int.,* **177**(1), 161–178.

Passier, M.L. & Snieder, R.K., 1995. On the presence of intermediate-scale heterogeneity in the upper mantle, *Geophys. J. Int.,* **123**(3), 817–837.

Phillips, B. & Bunge, H., 2005. Heterogeneity and time dependence in 3D spherical mantle convection models with continental drift, *Earth planet. Sci. Lett.,* **233**(1–2), 121–135.

Porritt, R.W., Allen, R.M. & Pollitz, F.F., 2013. Seismic imaging east of the Rocky Mountains with USArray, *Earth planet. Sci. Lett.,* **1,** 1–10.

Rickers, F., Fichtner, A. & Trampert, J., 2013. The IcelandJan Mayen plume system and its impact on mantle dynamics in the North Atlantic region: evidence from full-waveform inversion, *Earth planet. Sci. Lett.,* **367,** 39–51.

Ritsema, J., Deuss, A., van Heijst, H.J. & Woodhouse, J.H., 2011. S40RTS: a degree-40 shear-velocity model for the mantle from new Rayleigh wave dispersion, teleseismic traveltime and normal-mode splitting function measurements, *Geophys. J. Int.,* **184**(3), 1223–1236.

Ritzwoller, M. & Lavely, E., 1995. Three-dimensional seismic models of the Earth's mantle, *Rev. Geophys.,* **33**(1), 1–66.

Sato, H., Fehler, M.C. & Maeda, T., 2012. *Seismic Wave Propagation and Scattering in the Heterogeneous Earth,* Springer-Verlag.

Schaeffer, A.J. & Lebedev, S., 2013. Global shear speed structure of the upper mantle and transition zone, *Geophys. J. Int.,* **194**(1), 417–449.

Schmandt, B. & Humphreys, E., 2010. Complex subduction and small-scale convection revealed by body-wave tomography of the western United States upper mantle, *Earth planet. Sci. Lett.,* **297**(3–4), 435–445.

Schuberth, B. S.a., Bunge, H.P. & Ritsema, J., 2009. Tomographic filtering of high-resolution mantle circulation models: can seismic heterogeneity be explained by temperature alone?, *Geochem. Geophys. Geosyst.,* **10**(5), doi:10.1029/2009GC002401.

Sebai, A., Stutzmann, E., Montagner, J.-P., Sicilia, D. & Beucler, E., 2006. Anisotropic structure of the African upper mantle from Rayleigh and Love wave tomography, *Phys. Earth planet. Inter.,* **155**(1–2), 48–62.

Shearer, P.M. & Earle, P.S., 2008. Observing and modeling elastic scattering in the deep Earth, *Adv. Geophys.,* **50,** 167–193.

Su, W.-j. & Dziewoski, A.M., 1992. On the scale of mantle heterogeneity, *Phys. Earth planet. Inter.,* **74**(1–2), 29–54.

Tackley, P., Stevenson, D., Glatzmaier, G. & Schubert, G., 1993. Effects of an endothermic phase transition at 670 km depth in a spherical model of convection in the Earth's mantle, *Nature,* **361,** 699–704.

Tanimoto, T., 1990. Predominance of large-scale heterogeneity and the shift of velocity anomalies between the upper and lower mantle, *J. Phys. Earth,* **38,** 493–509.

Tanimoto, T. & Anderson, D., 1984. Mapping convection in the mantle, *Geophys. Res. Lett.,* **11**(4), 287–290.

Turcotte, D.L., 1997. *Fractals and Chaos in Geology and Geophysics,* Cambridge Univ. Press.

Wiaux, Y., Jacques, L. & Vandergheynst, P., 2005. Correspondence principle between spherical and euclidean wavelets, *Astrophys. J.,* **632**(1), 15–28.

Wieczorek, M.A. & Simons, F.J., 2005. Localized spectral analysis on the sphere, *Geophys. J. Int.,* **162**(3), 655–675.

Wieczorek, M.A. & Simons, F.J., 2007. Minimum-variance multitaper spectral estimation on the sphere, *J. Fourier Anal. Applicat.,* **13**(6), 665–692.

Woodhouse, J. & Dziewonski, A., 1989. Seismic modelling of the Earth's large-scale three-dimensional structure, *Phil. Trans. R. Soc. Lond., A: Math. Phys. Sci.,* **328**(1599), 291–308.

Woodhouse, J.H. & Dziewonski, A.M., 1984. Mapping the upper mantle: three-dimensional modeling of Earth structure by inversion of seismic waveforms, *J. geophys. Res.,* **89**(B7), 5953–5986.

Yuan, H., Romanowicz, B., Fischer, K.M. & Abt, D., 2011. 3-D shear wave radially and azimuthally anisotropic velocity model of the North American upper mantle, *Geophys. J. Int.,* **184**(3), 1237–1260.

Zangvil, A., 1977. On the presentation and interpretation of spectra of large-scale disturbances, *Mon. Wea. Rev.,* **106**(8), doi:10.1175/1520-0493(1978).

Zhu, H., Bozda, E., Peter, D. & Tromp, J., 2012. Structure of the European upper mantle revealed by adjoint tomography, *Nat. Geosci.,* **5**(7), 493–498.

# APPENDIX A: SCALED FILTERS AND THE POWER PER OCTAVE SPECTRUM

As in the 1-D Fourier transform, convolving a function $f(\mathbf{r})$ with coefficients f_{lm} on the sphere with an axisymmetric filter h with coefficients h_{l0} can be written as a filtering operation (e.g. Baddour 2010):

$$h \star f = \frac{c_{lm} h_{l0}}{\sqrt{2l+1}} \qquad (A1)$$

we can write this equation in terms of a simple spectral filter $H(l)$ with which each coefficient is multiplied and that depends only on the degree l.

$$h \star f = H(l) c_{lm}. \qquad (A2)$$

The spherical harmonics coefficients that are associated with a given filter $|H(l)|^2$ that is applied to the power spectrum are therefore $h_{l0} = \sqrt{2l+1} H(l)$. The filter response function is then the spherical harmonics transform of h_{l0} and can be examined in space. Fig. A1 shows an example of a set of simple octave width bandpass filters and corresponding (normalized) wavelets. The logarithmic tiling creates a set of scaled zero-mean wavelets that have similar shape with a main lobe of width $360°/(sl)$ and side lobes at $\pm 360°/(sl)$ and multiples. If we decrease the bandwidth or sharpen the edges of these filters, we get more energy in the side lobes and a slower spectral decay. The dilated wavelets are further amplitude (i.e. magnitude not square magnitude) normalized and the application of such a filter can be interpreted as as a local axisymmetric
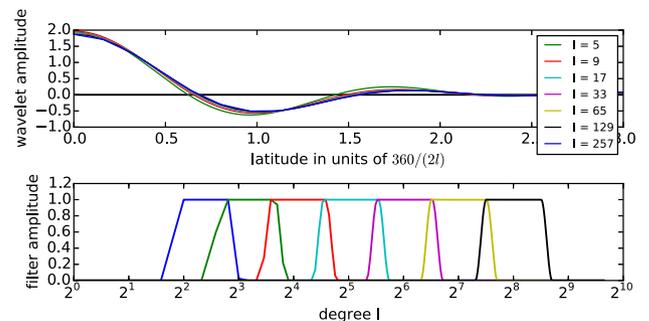
Figure A1. Octave width filters in space (upper row) and frequency (lower row). The octave band allows for reasonable localization and the filter main lobe extends approximately over $360°/2l$, where l is the logarithmic centre of the bandpass. Side lobes span a distance of $360°/l$. Power in spectral bands can be understood as the model variance after convolving with such a filter.
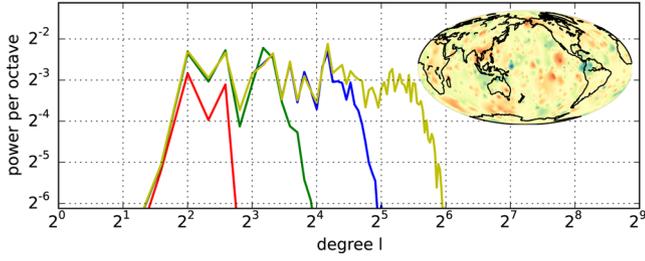
Figure A2. Synthetic model which has been constructed by adding heterogeneities at four different scales l_i at random locations on the sphere. The number of heterogeneities is proportional to the square of l_i such that they have constant power at each scale. As we can see, the constant total model power at different scales causes the logarithmic power per octave density to be constant as well.

'averaging' operation. The power in the associated spectral band can then be interpreted as the model variance after such an averaging operation.

Fig. A2 shows the map and power per octave spectrum of a synthetic model with known heterogeneities at different scales. It was constructed by adding 'Mexican hat' heterogeneities of equal amplitude, stretched to different scales at random locations to the model. The number of heterogeneities is chosen such that the total added power at each scale is constant and we can indeed recover this directly via the particular definition of the spectrum. This illustrates, that the power per octave can indeed be used as a measure of the total variance due to heterogeneities of a particular scale.

APPENDIX B: REGIONAL ESTIMATION OF GLOBAL POWER SPECTRA

Here, we examine the effect of windowing on power spectral estimations on the sphere. The important conclusion that a windowed spectral estimate is not only smoothed but might also be biased, is unrelated to the geometry of the sphere and can also be derived in the plane, using a 2-D Fourier transform (e.g. Hivon & Górski 2002). It can be seen as a consequence of the representation of a 2-D frequency space by a single radial magnitude component (e.g. the degree l). Nevertheless, we focus our derivation completely on the sphere. The first part is based upon Hivon & Górski (2002) and Wieczorek & Simons (2005, 2007) and in the second we illustrate the procedure using a synthetic model spectrum. In this section the upper index corresponds to the spherical harmonics order *m* to save space but without any further meaning.

The spherical harmonics coefficients c_l^m of a function f with power per degree F_l, multiplied with a window function w with power per degree W_l can be written as:

$$c_l^m = \int_\Omega f(\mathbf{r}) w(\mathbf{r}) Y_l^{m*}(\mathbf{r}) \, d\Omega. \tag{B1}$$

Expanding f and w with their spherical harmonics coefficients f_l^m and w_l^m, as well as power per degree spectra F_l and W_l, leads to:

$$c_{l_3}^{m_3} = \sum_{l_1 m_1} f_{l_1}^{m_1} \sum_{l_2 m_2} w_{l_2}^{m_2} \int_\Omega Y_{l_1}^{m_1} Y_{l_2}^{m_2} Y_{l_3}^{m_3*} \, d\Omega. \tag{B2}$$

We rewrite this equation in terms of a matrix K that relates input coefficients to output coefficients.

$$c_{l_3}^{m_3} = \sum_{l_1 m_1} f_{l_1}^{m_1} K_{l_1 l_3}^{m_1 m_3}, \tag{B3}$$

where K depends only on the window function and can be written as:

$$K_{l_1 l_3}^{m_1 m_3} = \sum_{l_2 m_2} w_{l_2}^{m_2} \int_\Omega Y_{l_1}^{m_1} Y_{l_2}^{m_2} Y_{l_3}^{m_3*} \, d\Omega. \tag{B4}$$

The triple spherical harmonic products can be rewritten in terms of *Wigner-3j* symbols that satisfy a useful orthogonality relation (e.g. Dahlen & Tromp 1998, eqs C197 and C204):

$$\int_\Omega Y_{l_1}^{m_1} Y_{l_2}^{m_2} Y_{l_3}^{m_3*} \, d\Omega$$

$$= \frac{((2l_1+1)(2l_2+1)(2l_3+1))^{\frac{1}{2}}}{(-1)^{m_1+m_2+m_3}} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \tag{B5}$$

$$\sum_{m_1 m_2} \begin{pmatrix} l_1 & l_2 & l \\ m_1 & m_2 & m \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l' \\ m_1 & m_2 & m' \end{pmatrix} = \frac{\delta_{ll'} \delta_{mm'}}{2l+1}. \tag{B6}$$

With these expressions, the expected power $\langle S_l \rangle$ of an isotropic Gaussian model (see eq. 6) after windowing can finally be written as:

$$\langle S_{l_3} \rangle = \sum_{m_3} \langle c_{l_3}^{m_3*} c_{l_3}^{m_3} \rangle \tag{B7}$$

$$= \sum_{m_3} \sum_{l_1 m_1 l_2 m_2} \langle f_{l_1}^{m_1} f_{l_2}^{m_2*} \rangle K_{l_3 l_1}^{m_3 m_1} K_{l_3 l_2}^{m_3 m_2*} \tag{B8}$$

$$= \sum_{l_1} \frac{F_{l_1}}{2l_1+1} \sum_{m_1 m_3} |K_{l_3 l_1}^{m_3 m_1}|^2. \tag{B9}$$

With relations (B5) and (B6) the sum over K can be greatly simplified:

$$\sum_{m_3 m_1} |K_{l_3 l_1}^{m_3 m_1}|^2 = \sum_{lm} w_l^m w_l^{m*} (2l_1+1)(2l_3+1) \begin{pmatrix} l_1 & l_3 & l \\ 0 & 0 & 0 \end{pmatrix}^2. \tag{B10}$$

The expected power can therefore be written:

$$\langle S_{l_3} \rangle = (2l_3+1) \sum_{l_1} F_{l_1} \sum_{l_2} W_{l_2} \begin{pmatrix} l_1 & l_3 & l_2 \\ 0 & 0 & 0 \end{pmatrix}^2 \tag{B11}$$

$$\langle S_l \rangle = \sum_{l'} F_{l'} M_{ll'}. \tag{B12}$$

The *coupling matrix* M relates the input power in degree l' of a global model to the output power l of the windowed model:

$$M_{ll'} = (2l+1) \sum_{l_w} W_{l_w} \begin{pmatrix} l & l' & l_w \\ 0 & 0 & 0 \end{pmatrix}^2. \tag{B13}$$

If the window power is limited by an upper degree l_{max}, selection rule properties of the Wigner-3j symbols (e.g. Dahlen & Tromp 1998; Hivon & Górski 2002; Wieczorek & Simons 2005) ensure that coupling occurs only between degrees that are within range of the window bandwidth. Limiting window power to the lowest possible degrees is therefore important to get accurate spectral measurements, especially in the presence of a fast spectral decay.
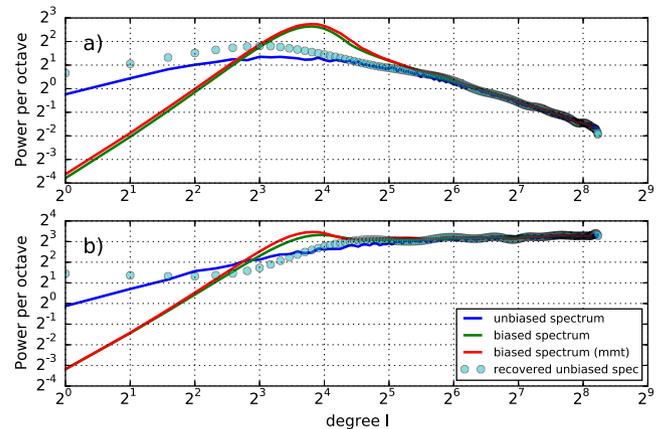
For a particular windowed power spectrum, we can invert this matrix and search for an associated global isotropic and Gaussian model for which we expect to get this windowed spectrum. Note that
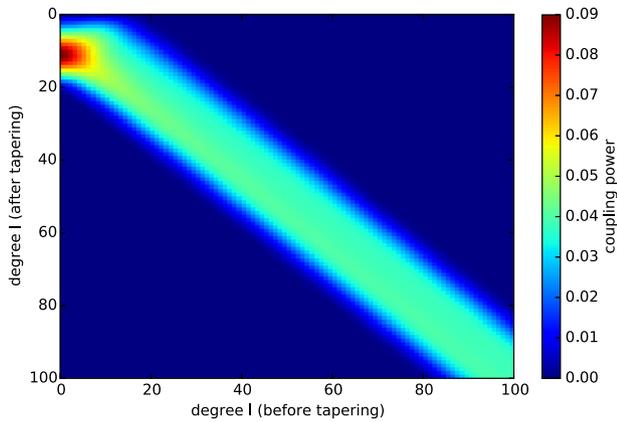
Figure B1. Spherical harmonics coupling matrix corresponding to the 5 best concentrated windows with power up to degree $l_w = 20$ in a spherical cap of $20°$ radius as commonly used in the multitaper spectral estimation process. Colour indicates how the power of an input function is transformed into output power due to windowing. At degrees with the strong taper power, the matrix becomes strongly asymmetric and therefore spectral power has to be interpreted with care. Low output degrees become under-represented, whereas especially the window degrees become over-represented.

Figure B2. Examples of the regional to global inversion process in the case of (a) an exponential and (b) an asymptotically scale-free (b) spectrum, perturbed by 30 per cent noise (blue). The spectrum is biased (green) using the coupling matrix, or equivalently measured (red) in a few realizations and then unbiased by inverting the coupling matrix (turquoise dots). The inversion searches for an isotropic Gaussian medium whose expected regional spectrum equals the measured one.

this procedure is not limited to axisymmetric windows if the model has independent Gaussian coefficients. Also, because the expected model is linearly related to the window power, we can first sum several windowed spectra and the power of the associated windows and then invert for the global spectrum.

Fig. B1 shows the coupling matrix for the five tapers with bandwidth $l = 20$ that are best concentrated in a spherical cap with radius $20°$ as we commonly use them in this paper. We can see that it is asymmetric, especially at degrees with strong window power and input power becomes preferentially mapped to higher output power. At high degrees, the matrix becomes symmetric and can be seen as a convolution. Compared to Fourier analysis in a single dimension, the asymmetry of this matrix adds a new aspect to the windowing procedure. Unfortunately the spectral band of strong asymmetry is important for the comparison of the global and regional models, because it is at the scales that both can resolve.

Fig. B2 illustrates the inversion process for a synthetic model under the influence of this asymmetric coupling matrix and its implications on the measurements. Fig. B2(a) shows the example of a noisy (10 per cent noise) synthetic spectrum with decreasing power per octave (blue). The multi-taper measurement of 10 realizations (red), the spectrum biased with the coupling matrix of Fig. B1 (green) are also shown. We can see that this procedure leads to increased power especially at degrees with strong window power, and decreased power at lower degrees. Inversion of the coupling matrix and the multitaper spectrum approximately yields the original global spectrum (turquoise dots). Errors due to the inversion are not negligible, but smaller than the differences in the tomographic models that we analyse in this paper. We can resolve the spectral decay rate relatively well, as well as spectral power at degrees larger than about the window bandwidth. Fig. B2(b) shows a similar example for an asymptotically flat power per octave spectrum.