Available online at www.sciencedirect.com
ScienceDirect
Earth and Planetary Science Letters 259 (2007) 526–540
EPSL
www.elsevier.com/locate/epsl

Stochastic analysis of shear-wave splitting length scales ☆

Thorsten W. Becker *, Jules T. Browaeys [1], Thomas H. Jordan

Department of Earth Sciences, University of Southern California, MC 0740, Los Angeles, CA 90089-0740, USA

Received 13 December 2006; received in revised form 6 April 2007; accepted 3 May 2007
Available online 10 May 2007
Editor: C.P. Jaupart

Abstract

The coherence of azimuthal seismic anisotropy, as inferred from shear-wave splitting measurements, decreases with the relative distance between stations. Stochastic models of a two-dimensional vector field defined by a von Karma'n [T. von Karma'n, Progress in the statistical theory of turbulence, J. Mar. Res., 7 (1948) 252–264.] autocorrelation function with horizontal correlation length L provide a useful means to evaluate this heterogeneity and coherence lengths. We use the compilation of SKS splitting measurements by Fouch [M. Fouch, Upper mantle anisotropy database, accessed in 06/2006, http://geophysics.asu.edu/anisotropy/upper/] and supplement it with additional studies, including automated measurements by Evans et al. [Evans, M.S., Kendall, J.-M., Willemann, R.J., 2006. Automated SKS splitting and upper-mantle anisotropy beneath Canadian seismic stations, Geophys. J. Int. 165, 931–942, Evans, M.S., Kendall, J.-M., Willemann, R.J. Automated splitting project database, Online at http://www.isc.ac.uk/SKS/, accessed 02/2006]. The correlation lengths of this dataset depend on the geologic setting in the continental regions: in young Phanerozoic orogens and magmatic zones $L \sim 600$ km, smaller than the smooth $L \sim 1600$ km patterns in tectonically more stable regions such as Phanerozoic platforms. Our interpretation is that the relatively large coherence underneath older crust reflects large-scale tectonic processes (e.g. continent–continent collisions) that are frozen into the tectosphere. In younger continental regions, smaller scale flow (e.g. slab anomaly induced) may predominantly affect anisotropy. In this view, remnant anisotropy is dominant in the old continents and deformation-induced anisotropy caused by recent asthenospheric flow is dominant in active continental regions and underneath oceanic plates. Auxiliary analysis of surface-wave anisotropy and combined mantle flow and anisotropic texture modeling is consistent with this suggestion. In continental regions, the further exploration of a stochastic description of seismic anisotropy may form a useful counterpart to deterministic forward modeling, particularly if we wish to understand the origin of discrepancies in heterogeneity estimates based on different seismological data sets.

Keywords: shear-wave splitting; seismic anisotropy; correlation length; Stochastic model

1. Introduction

Global seismology records elastic anisotropy in the lithosphere and upper mantle (e.g. Anderson, 1989; Montagner and Guillot, 2000). One of the most direct observations is from shear-wave splitting (Vinnik et al.,

☆ Submitted to Earth and Planetary Science Letters in original form on December 13, 2006. Revised version as of April 6, 2007.

* Corresponding author. Tel.: +1 213 740 8365; fax: +1 213 740 8801.
E-mail address: twb@usc.edu (T.W. Becker).
[1] Now at: The University of Texas at Austin, University Station, Austin, Texas, USA.

0012-821X/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.epsl.2007.05.010

1984), and numerous measurements for SKS phases have now led to substantial coverage in some of the continental plates (Silver, 1996; Savage, 1999). However, resolution in other regions, particularly the oceanic plates, is still severely limited by station–receiver geometry and logistics (Fig. 1). From the few measurements in oceanic regions, the orientations of inferred fast propagation from shear-wave splitting tend to align both with azimuthal anisotropy from fundamental-mode surface waves and predictions from mantle flow models (Montagner et al., 2000; Gaboret et al., 2003; Becker et al., 2003; Behn et al., 2004). This implies a common origin of anisotropy: deformation of intrinsically anisotropic olivine in convective flow in the upper ~300 km of the mantle (McKenzie, 1979; Vinnik et al., 1995).

In continental regions, however, marked variations in the spatial coherence of patterns and discrepancies between different seismological datasets occur. While some regions such as the eastern United States (US) show relatively smooth patterns (Fouch et al., 2000), continental splitting data can also be highly spatially variable, even at the smallest scales (Savage, 1999; Huang et al., 2000; Fouch et al., 2004; Heintz and Kennett, 2006), and surface wave anisotropy is less well fit by large-scale geodynamic models (Becker et al., 2003, 2007). Possible reasons for this mismatch include lack of resolution in input models for the flow computations, or an inadequate rheological description of mantle convection. Yet, incorporation of regional flow patterns around stiff continental keels or plume upwellings only partially explain observations of continental anisotropy (Fouch et al., 2000; Savage and Sheehan, 2000; Walker et al., 2005). Where there is sufficient coverage, depth dependent anisotropy may be invoked (e.g. Simons et al., 2002; Simons and van der Hilst, 2003; Debayle et al., 2005; Heintz and Kennett, 2006), and particularly an integrated approach using data from different disciplines may allow unraveling local settings (Fouch and Rondenay, 2006). However, for global datasets improvements in mantle circulation models such as the introduction of lateral viscosity variations (LVVs) has not yet led to substantially better model fits for surface waves or SKS splitting (Becker et al., 2006a,b; Conrad et al., in press).

This difficulty of matching stable-continent anisotropy with geodynamic models contrasts with geologically active regions, for example the western US. Anisotropy in these regions may be fit successfully by invoking asthenospheric flow (Silver and Holt, 2002; Savage et al., 2004; Becker et al., 2006a,b), presumably because the lithosphere is thinner than in older continental regions. An explanation for these differences between oceanic or young continental regions on the one hand, and old continental regions on the other hand, is that frozen-in structure dominates anisotropy in

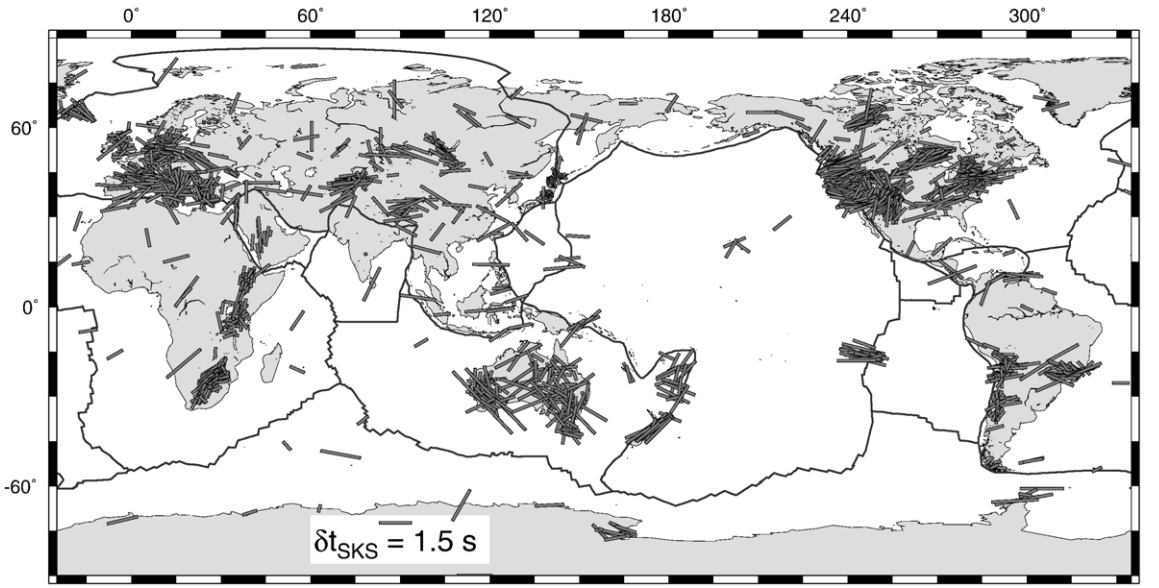

Fig. 1. Global distribution of SKS and SKKS splitting measurements as used for this study. Stick orientation and length denote fast azimuth ϕ and delay time δt, respectively. The original data (7705 measurements, mainly from the compilation by Fouch (2006) and automated measurements by Evans et al. (2006a,b) have been averaged on a 0.5° × 0.5° grid.

the old continents (Silver and Chan, 1988). In the thick lithosphere (tectosphere) a long ($\gtrsim$ 500 Ma) sequence of tectonic events may be recorded in rock fabrics. These remnant structures and the corresponding anisotropy are likely only moderately affected by asthenospheric flow which we may hope to reconstruct for the last $\sim$ 60 Ma, at best (Bunge et al., 2003). We wish to establish a framework to test suggestions of a variable contribution of the asthenosphere and lithosphere to anisotropy (cf. Fouch and Rondenay, 2006), and to provide quantitative measures of spatial coherence and smoothness.

A statistical description of anisotropy may be a useful, general approach to characterize the global nature of anisotropic heterogeneity. Particularly continental domains may be adequately described by a stochastic model due to their complex tectonic history. While it may not be possible to predict long-term continental deformation histories by means of circulation models, we can still extract information from the existing patterns in anisotropy that can guide our understanding of plate tectonic processes. Stochastic descriptions are standard in exploration seismology, but such models have not been explored in much detail for global anisotropy. This study intends to extract a subset of the relevant parameters of the heterogeneity of anisotropy, the horizontal coherence of the stochastic model. We focus on fast propagation orientations from *SKS* splitting and use a two-dimensional (2-D) von Karma'n (1948) model. This stochastic medium is characterized by the horizontal wavelength of the heterogeneity λ and the Hurst (smoothness) exponent ν. Such an approach has been shown to be relevant for geophysical data (e.g. Goff and Jordan, 1988; Saito, 2006).

We proceed to describe the stochastic model and illustrate it with a test application to mapped crustal structure in Australia. Then, we apply the analysis to the global SKS splitting dataset, and close by discussing our findings for different tectonic subsets in the context of surface-wave observations of anisotropy and mantle flow models.

2. Stochastic model of azimuthal anisotropy

We are concerned with azimuthal anisotropy where seismic wave velocities vary as a function of propagation direction within the horizontal plane. Such anisotropy can be inferred from shear-wave splitting; measurements at a seismic station at location $\mathbf{x}$ typically consist of a time delay δt between the arrivals of the split shear-wave trains and the azimuthal orientation ϕ of the fast shear wave polarization (e.g. Silver, 1996). The splitting measurement represents a nonlinear projection of the underlying 3-D anisotropic structure to the surface (Rümpker and Silver, 1998; Levin et al., 1999; Saltzer et al., 2000; Chevrot, 2006). Interpretations of the measurements are, however, commonly based on a single, equivalent layer with a horizontal hexagonal symmetry axis (Silver and Chan, 1991). The anisotropy could theoretically be accumulated anywhere along the ray path, between the P to S conversion at the core–mantle boundary and the station for SKS arrivals. However, most surface wave studies find that anisotropy is concentrated in the uppermost $\sim$300 km of the mantle, consistent with comparisons of local S with teleseismic SKS splitting (Meade et al., 1995; Fischer and Wiens, 1996). The physical explanation is that dislocation creep, required for lattice preferred alignment of olivine, can be expected to be dominant at these depths (McNamara et al., 2003; Podolefsky et al., 2004; Becker, 2006). We shall thus proceed to interpret our results in terms of structure in the uppermost mantle and lithosphere, recognizing that this is a simplification.

Shear-wave splitting measurements collected at a given station should ideally have broad back-azimuth coverage in order to better characterize anisotropy, e.g. to detect a tilted axis or layering (Silver and Savage, 1994; Levin et al., 1999; Schulte-Pelkum and Blackman, 2003; Chevrot and van der Hilst, 2003). We expect that the presence of layered anisotropy, or strong lateral variations, will affect delay times more than fast orientations (Schulte-Pelkum and Blackman, 2003; Chevrot, 2006). In practice, back-azimuth coverage is often limited, particularly for temporary deployments. Moreover, the global SKS splitting compilations that are available to us were derived from different measurement methods and only include event information and uncertainty estimates for ϕ and δt for some of the studies. A formal analysis of systematic and random errors is therefore not possible at this point. The sensitivity of the shear wave to azimuthal anisotropy is also a function of depth and wavelength. This introduces a low-pass filtering effect on lateral coherence, and the width of the sensitivity kernels at shallow depth constitutes the lower bound for the correlation lengths we can detect (Chevrot, 2006). We proceed to average measurements on a $0.5°\times0.5°$ grid (roughly corresponding to the Fresnel-zone width), taking into account the orientational (180°-periodic) nature of the data. A further assumption we make for simplicity is that the resulting fast orientations are indicative of an anisotropic medium that can be effectively described by purely horizontally varying heterogeneity. We will derive a general error estimate that lumps these different uncertainties together.

2.1. Correlation function

A stochastic medium is governed by the correlation function which describes the spectral character and coherence of the random heterogeneity. We choose the parameterization that was introduced by von Karma'n (1948) for turbulent fluids. This function has been used to model the seafloor topography and has found applications in seismic wave propagation studies (e.g. Goff and Jordan, 1988; Saito, 2006). The von Karma'n correlation function depends on the relative distance r and a characteristic length scale λ as:

$$\rho_{\lambda,\nu}(r) = 2^{(1-\nu)}(r/\lambda)^{\nu}\frac{K_{\nu}(r/\lambda)}{\Gamma(\nu)} \quad \text{with} \quad r \geq 0, \qquad (1)$$

where K_ν is the modified Bessel function of the second kind, ν is the Bessel function order, $\nu \in [0, 1]$, and Γ is the gamma function (Abramowitz and Stegun, 1972). The function $\rho_{1,\nu}(r)$ is shown in Fig. 2a for different values of ν; at fixed λ, larger ν implies spatially smoother fields. Eq. (1) includes, as a special case for $\nu = 0.5$, the decreasing exponential function commonly used to model stochastic media for wave scattering studies (e.g. Chernov, 1960; Roth and Korn, 1993):

$$\rho_{\lambda,0.5}(r) = \exp(-r/\lambda). \qquad (2)$$

In our stochastic model, the ϕ orientations are generated from a two-dimensional, random vector field. We present a general model that includes Cartesian vector components for completeness but, given the limitations of the splitting data, we will later focus on ϕ orientations only. Our model is based on the assumption that Cartesian components of the vectors $\mathbf{s}(\mathbf{x}) = \{s_X(\mathbf{x}), s_Y(\mathbf{x})\}$ are stationary Gaussian variables with mean

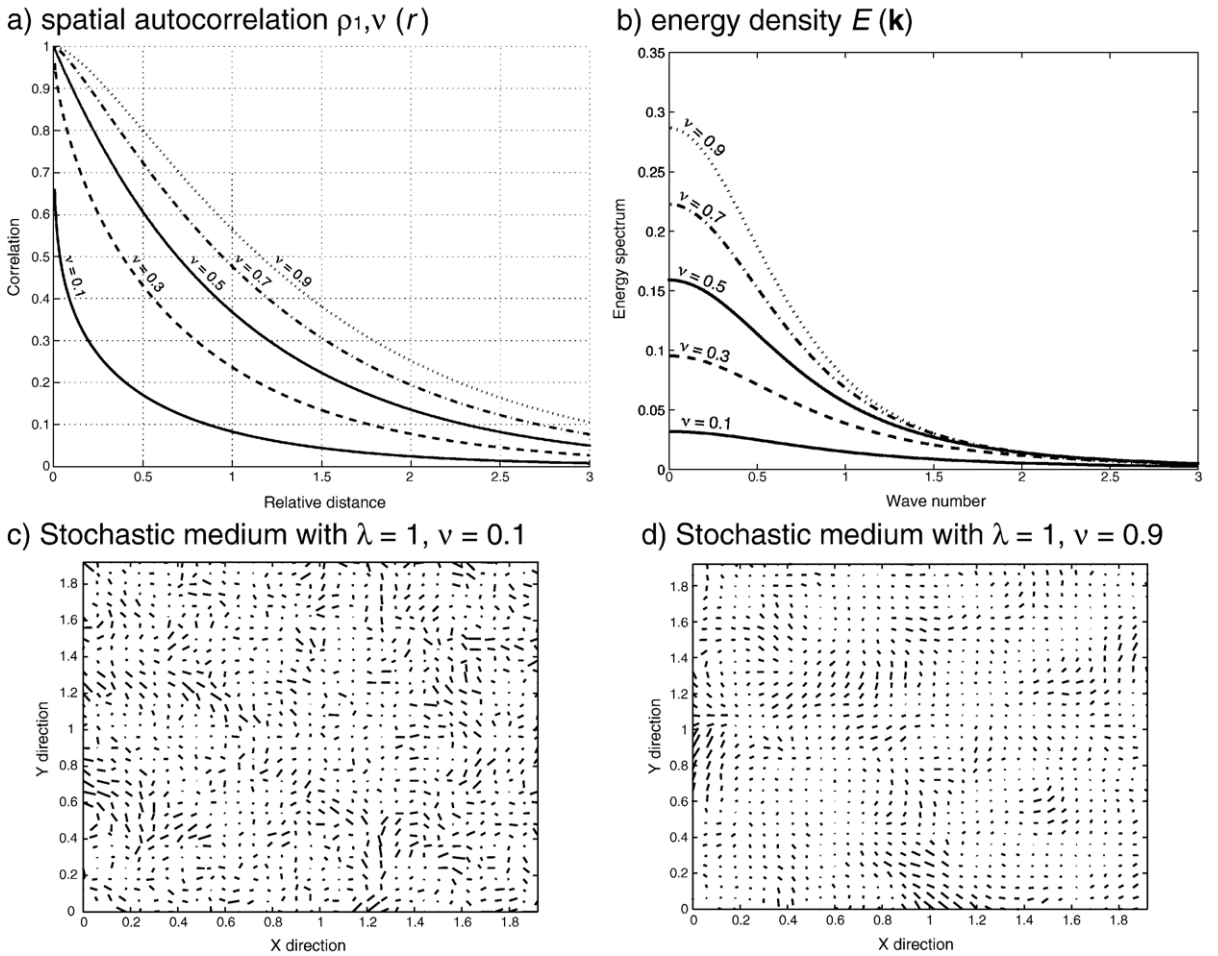

Fig. 2. (a): Spatial autocorrelation function $\rho_{1,\nu}(r)$ at $\lambda = 1$ for different values of the Hurst exponent ν. (b): The corresponding energy spectra of the stochastic process, the case $\nu = 0.5$ corresponds to a decreasing exponential. (c and d): Examples of a 2-D stochastic vector field using the correlation function $\rho_{1,\nu}(r)$ for Cartesian components with $\nu = 0.1$ (c) and $\nu = 0.9$ (d).

{0, 0} and standard deviation σ. The random field is characterized by a joint, correlated Gaussian probability on each of the vector components and the auto-covariance of the relative distance $r=|\mathbf{r}|$ between two locations, where vectors are $\mathbf{s}=\mathbf{s}(\mathbf{x})$ and $\mathbf{s}'=\mathbf{s}(\mathbf{x}+\mathbf{r})$:

$$\mathbf{C}(r) = \begin{pmatrix} \langle s_X s'_X \rangle & \langle s_X s'_Y \rangle \\ \langle s_X s'_X \rangle & \langle s_X s'_Y \rangle \end{pmatrix} = \begin{pmatrix} \sigma^2\rho(r) & 0 \\ 0 & \sigma^2\rho(r) \end{pmatrix}. \quad (3)$$

Here, $\rho(r)$ is the spatial autocorrelation, Eq. (1), and $\langle . \rangle$ is the expectation operator. The two Cartesian components s_X and s_Y are independent random variables with variance σ^2. This vector field has the joint probability density:

$$p(s_X, s_Y) = \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}} \exp\left(-\frac{s_X^2 - 2\rho s_X s_Y + s_Y^2}{2\sigma^2(1-\rho^2)}\right), \quad (4)$$

where ρ is the correlation coefficient. The spatial Fourier transform of the autocovariance function is by definition the energy density spectrum $E(\mathbf{k})$ of the stochastic process:

$$E(\mathbf{k}) = \sigma^2 \frac{\nu\lambda}{\pi}(1 + (\lambda k)^2)^{-(\nu+1)}, \quad (5)$$

where the spectral coordinate vector is $\mathbf{k}=\{k_X, k_Y\}$ and $k=|\mathbf{k}|$. The radial dependence of the energy density spectrum $E(\mathbf{k})$ is plotted in Fig. 2b, for $\sigma=1$, $\lambda=1$ and different values of ν. The curves exhibit an asymptotic power-law dependence in $k^{-2(\nu+1)}$. The medium has a smoother (redder) spectral heterogeneity character for larger values of ν, but spatial smoothness at constant ν can also be realized by a larger correlation length λ.

To create realizations of such random media in the Fourier domain, we compute the spectrum $\Phi(\mathbf{k})$ from the energy density:

$$\Phi(\mathbf{k}) = \sqrt{E(\mathbf{k})}\exp(i\varphi(\mathbf{k})), \quad (6)$$

where the phase $\varphi(\mathbf{k})$ is obtained from white Gaussian noise with zero mean and standard deviation σ. The inverse Fourier transform of one realization $\Phi(k_X, k_Y)$ then produces one of the components $s_X(\mathbf{x})$ or $s_Y(\mathbf{x})$. Using this approach, rather than simply assigning a random phase in the Fourier domain, ensures that the vector components are Gaussian and have the desired spectral character. Fig. 2c and d shows two stochastic vector field realizations using a discrete inverse fast Fourier transform for $\lambda=1$, $\sigma=1$ and $\nu=0.1$ as well as $\nu=0.9$. An increase of the Hurst exponent ν increases the smoothness of the medium at small scales (cf. Chemingui, 2001; Saito, 2006).

2.2. Statistics of azimuthal orientations

We are interested in the statistical distribution of data that can be described by an azimuth, ϕ. If the azimuth denotes an orientational quantity with 180°-periodicity, such as azimuthal anisotropy, an appropriate measure of the difference between two azimuths ϕ and ϕ' is given by

$$\cos 2\Psi \quad \text{where} \quad \Psi = \sigma - \sigma'. \quad (7)$$

There is a large body of literature on the statistics of orientational data, and statistics on a sphere (e.g. Watson, 1983). Accordingly, we explored several alternative metrics of the alignment between azimuths, including a complex representation of splitting vectors. While more comprehensive treatments may exist, we found that the $\cos 2\Psi$ measure is convenient and appropriate for our purpose. Importantly, we are able to provide an analytical solution for our stochastic model which is developed below. The angular deviation between azimuths, Ψ, is computed accounting for the fact that data are given on a sphere following Bird and Li (1996), although this correction is small for distances $r \lesssim 3000$ km. The mean difference for azimuthal anisotropy can be computed from $\langle \cos 2\Psi \rangle(r)$ for any data pair separated by a relative distance r. The limiting cases for a set of azimuths are:

$\langle \cos 2\Psi \rangle \rightarrow 1$ for correlated orientations, and
$\langle \cos 2\Psi \rangle \rightarrow 0$ for uncorrelated orientations.

The statistical orientational difference is then defined as:

$$\hat{\Psi} = \frac{1}{2}\arccos(\langle \cos 2\Psi \rangle), \quad (8)$$

where $\hat{\Psi} \in [0, 90°]$ and completely random, uncorrelated orientations yield $\hat{\Psi}=45°$. The theoretical mean $q_m(\rho)=\langle \cos m\Psi \rangle$ for the vector field $\mathbf{s}(\mathbf{x})$ (Eq. (3)), for a fixed relative distance defining ρ, is the quadruple integral:

$$q_m(\rho) = \int_{-\infty}^{+\infty} ds_X \int_{-\infty}^{+\infty} ds'_X \int_{-\infty}^{+\infty} ds_Y \int_{-\infty}^{+\infty} ds'_Y \times \cos(m\Psi) p(s_X, s'_X) p(s_Y, s'_Y).$$

Because the vector components are Gaussian variables with zero mean and standard deviation σ, the joint

probability density function p used is the one in Eq. (4). This integral can be reduced to:

$$q_m = \frac{1}{2\pi\sigma^4(1-\rho^2)} \times \int_0^\infty \mathrm{d}s s \int_0^\infty \mathrm{d}s' s' \int_0^{2\pi} \mathrm{d}\Psi \times \cos(m\Psi) G(\Psi, s, s', \rho, \sigma),$$

$$G(\Psi, s, s', \rho, \sigma) = \exp\left(-\frac{s^2 + s'^2 - 2\rho s s' \cos\Psi}{2\sigma^2(1-\rho^2)}\right),$$

where s and s' are the respective vector lengths and the correlation ρ satisfies $0 \le \rho \le 1$. The statistical orientational quantity $\langle \cos 2\Psi \rangle$ is related to the particular case $m=2$ with limiting values $q_2(0)=0$ and $q_2(1)=1$ for completely uncorrelated or correlated media, respectively. To complete our stochastic model, we need to evaluate this integral; an analytical solution exists for $m=2$:

$$q_2(\rho) = \langle \cos 2\Psi \rangle = \frac{1-\rho^2}{\rho^2} \ln(1-\rho^2) + 1. \quad (9)$$

We have numerically verified Eq. (9) and will use this expression along with the auto-correlation function Eq. (1) to model orientational data.

We found that the $\langle \cos 2\Psi \rangle$ measure is most convenient for the analysis of orientational data. However, the mapping of Eq. (9) introduces a stretching of the correlation lengths of $\hat{\Psi}$ and the random vector field to that of q_2. The special case of the exponential ρ_2 decay for $\nu=0.5$ will be used most often below. In this case, the drop-off of correlation functions can be matched by rescaling the correlation lengths. The $\lambda_{\nu=0.5}$ characteristic length scale of $\langle \cos 2\Psi \rangle$ is then related to a correlation length for $\hat{\Psi}$, L, by

$$L \approx 0.242 \lambda_{\nu=0.5}, \quad (10)$$

i.e. the angular fast azimuth difference in Cartesian space, Ψ, decays spatially over length scales ~4 times shorter than the decay for the $\langle \cos 2\Psi \rangle$ measure. When we discuss L correlation lengths below, those are based on Eq. (10).

2.3. Error estimates

To estimate the stochastic model parameters we use N fast orientations. From those, we compute $N(N-1)/2$ data pairs which are sorted into N_b bins according to their relative distance. This yields mean angular deviation estimates $\langle \cos 2\Psi \rangle_i$ where $i=1, \ldots, N_b$. We chose a constant distance bin width of 100 km for all datasets; the lower bound for this parameter is determined by the number of data in the tectonic subsets we consider below. This width limits the spatial scales that are accessible by the stochastic analysis and so affects best-fit parameter values for λ and ν moderately. However, the different azimuthal anisotropy datasets can be compared to each other in a consistent framework as long as the width remains the same.

Standard errors σ_i of observed values $\langle \cos 2\Psi \rangle_i$ at each relative distance bin i can be estimated by the delete-d jackknife method (Efron and Stein, 1981) where error bounds are computed by removing n subsets of the data. For spatially coherent data, we choose as these subsets actual, regularly bounded geographic boxes that taken together make up the original spatial region of interest. All data points of each geographic box are removed for the jackknife step, while the other data is binned into the distance bins as before (without spatial averaging). If the data are given on irregular, spatially incoherent domains (as for our tectonic regionalization below), we create subsets by randomly selecting only a fraction of the data; all other data is again regularly binned with distance. For n total subsets of data that can possibly be removed, delete-d jackknife then removes $d = \sqrt{n}$ from the sample, a process with $n_s = \binom{n}{d}$ possible selections. Using all of these possible combinations, one obtains the jackknife estimate of best value and standard error:

$$\langle \cos 2\Psi \rangle_i = \frac{1}{n_s} \sum_{k=1}^{n_s} \langle \cos 2\Psi \rangle_i^{(k)}, \quad (11)$$

$$\sigma_i = \sqrt{\frac{n-d}{n_s d} \sum_{k=1}^{n_s} \left(\langle \cos 2\Psi \rangle_i^{(k)} - \langle \cos 2\Psi \rangle_i^{(\cdot)} \right)^2}, \quad (12)$$

where $\langle \cos 2\Psi \rangle_i^{(k)}$ is the k-th jackknife replication of $\langle \cos 2\Psi \rangle_i$.

Besides the uncertainties in estimating $\langle \cos 2\Psi \rangle$ at every distance bin, a correlation function estimated from imperfect data can exhibit a discontinuity at zero relative distance where $\langle \cos 2\Psi \rangle$ is expected to be unity. In variograms used in geostatistics, this discontinuity is called the nugget effect (Isaaks and Srivastava, 1989). This experimental phenomenon is caused both by the measurement errors and the fact that very small scales are not accessible, due to the finite size of bins. While dealing with shear wave splitting data, the horizontal resolution is also limited by the Fresnel zone width, and the offset in the data at $r=0$ is thus most likely caused by a combination of binning artifacts and measurement error.

The discontinuity at zero relative distance can be translated into an effective standard error on the azimuthal measurements. The error $e(\mathbf{x})$ in the measurement is assumed to be an uncorrelated spatial process, with zero mean. The tilde superscript $\tilde{\cdot}$ denotes data with errors and azimuthal measurements at two locations, expressed as:

$$\tilde{\varphi} = \varphi + e,\ \tilde{\varphi}' = \varphi' + e',\ \tilde{\Psi} = \Psi + \gamma, \tag{13}$$

where $\langle\gamma\rangle = e - e'$. The assumption of a zero mean and uncorrelated errors implies:

$$\langle\gamma\rangle = 0, \quad \text{and} \quad \langle\gamma^2\rangle = 2\langle e^2\rangle = 2\varepsilon^2. \tag{14}$$

For a small error, one can approximate (using Eq. (9))

$$\langle\cos 2\gamma\rangle \approx 1 - 2\langle\gamma^2\rangle, \tag{15}$$

$$\langle\cos 2\,\tilde{\Psi}\rangle \approx \cos\left(2\sqrt{2\varepsilon}\right)\left[\frac{1-\rho^2}{\rho^2}\ln(1-\rho^2)+1\right]. \tag{16}$$

At zero relative distance, the orientational difference is explained by an effective measurement error, $\cos 2\sqrt{2\varepsilon}$, which we estimate by extrapolating the binned $\cos 2\tilde{\Psi}$ estimates for each data subset to zero distance. Below, we drop the tilde superscripts; all $\langle\cos 2\Psi\rangle$ estimates refer to data with errors that are corrected following Eq. (16).

3. Results

We first apply our statistical description of azimuthal anisotropy to a test case of tectonic features mapped in the Australian crust and then proceed to estimate correlation lengths from shear-wave splitting.

3.1. Crustal structure orientations in Australia

The Australian continent constitutes an instructive case to test our statistical model. This cratonic region shows crustal features as mapped, e.g., by Wellman (1976). Fig. 3a shows our digitized version of his map of tectonic lineations; we computed 951 individual orientations for each straight segment whose locations were assigned to the segment's mid point. For relative distance bins of width 100 km, each bin contains and $\gtrsim$1000 data pairs up to 2000 km (Fig. 3), from which we compute the spatial decay of orientational coherence. Fig. 3b shows the binned data with error bars from the delete-d jackknife technique which uses a $10°\times10°$ box grid covering the Australian region. Only boxes with more than 28 data points are used such that the total number of boxes is $n=10$. The number of boxes removed in the jackknife process is $d=3$, leading to $\binom{n}{d} = 120$ replications. The χ^2 misfit function is calculated for model parameters λ, ν as:

$$\chi^2(\lambda,\nu) = \sum_{i=1}^{N_b}\left(\frac{\langle\cos 2\Psi\rangle_i - q_2(\rho_{\lambda,\nu}(r_i))}{\sigma_i}\right)^2, \tag{17}$$

where N_b is the number of relative distance bins centered at r_i and σ_i is from Eq. (12). We plot $\Delta\chi^2 = \chi^2 - \chi^2_{\min}$, where $\chi^2_{\min}$ is the best misfit, at intervals that would correspond to 69%, 90%, 95%, and 99% significance levels, respectively, if errors were normally distributed for a linear, two degree of freedom fit (e.g. Press et al., 1993, p. 695ff).

The $\langle\cos 2\Psi\rangle$ values as inferred from Fig. 3a show an exponential-like decay (Fig. 3b). After approaching $\langle\cos 2\Psi\rangle \approx 0$ at $r\sim1350$ km, there are some fluctuations in the signal for larger r. Those are related to structural features not well represented by our stochastic model and we have excluded the data points shown in open squares from the χ^2 computation. The optimal correlation function, $q_2(\rho_{\lambda^*,\nu^*})$, along with the misfit surface is shown in the inset of Fig. 3b. The minimum χ^2 occurs for a correlation length of $\lambda^* \sim 900$ km and a Hurst exponent $\nu^* \sim 0.7$; uncertainties are inferred from the width of the $\Delta\chi^2 = 6.2$ ($\sim$95% confidence) contour and given in the figure legend. The standard deviation of the measurement error as calculated for the first bins is $\varepsilon = 18°$. The exact cutoff choices for the fitting procedure do not affect numerical results for best-fit values significantly. Similarly, if we averaged the tectonic orientations on a $0.5°\times0.5°$ grid for consistency with the SKS splitting analysis below, best-fit parameters are only changed slightly, i.e. $\nu^* = 0.7$ and $\lambda^* = 850$ km vs. the 900 km from above. A Hurst exponent of $\nu \approx 0.7$–0.8, as preferred by the first bins of the data (compare solid and dashed lines in Fig. 3b), indicates a relatively smooth structure, compatible with visual comparison of Figs. 2c and 3a. The simple exponential correlation function with $\nu = 0.5$ leads to a misfit that is just outside the $\Delta\chi^2 = 2.3$ ($\sim$68%) contour. Clearly, the actual geological structure may display rougher, possibly fractal, structure than the map representation by Wellman (1976) which is deliberately simplified and so introduces "mapping smoothness".

Importantly, the $\Delta\chi^2$ contours in Fig. 3b show that there is a trade-off between ν and λ. As anticipated in the discussion of Fig. 2 and Eqs. (1) and (5), one may achieve similar model fits by increasing λ while decreasing ν, and vice versa. While this dataset is just good enough to make inferences about the preferred Hurst exponent, the situation is worse for the SKS data.

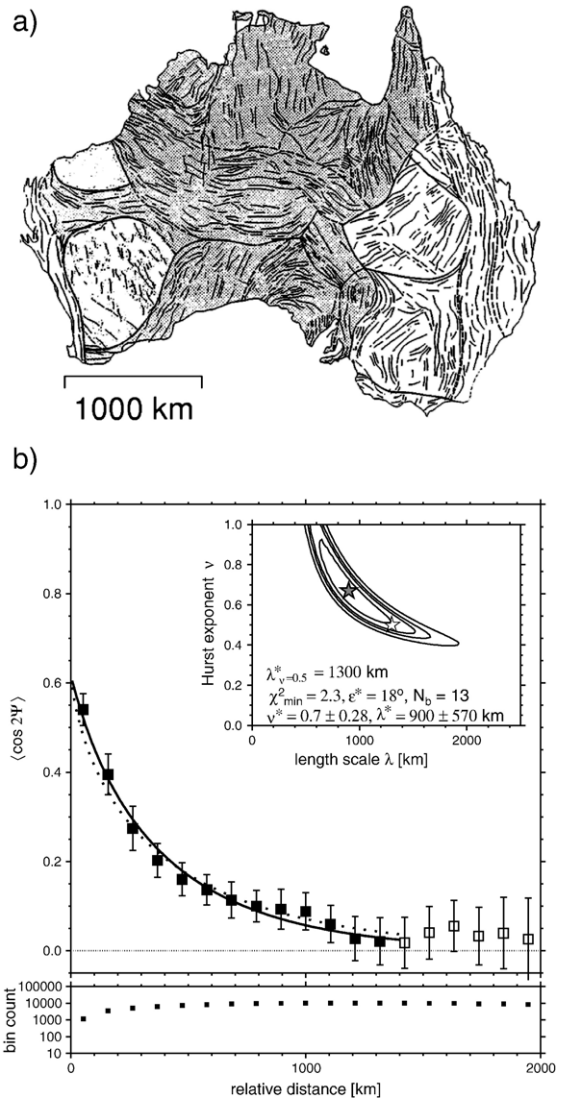

Fig. 3. (a): Tectonic features in Australia (a, modified after a map by Wellman (1976)); we measured 951 structure orientations. (b): Statistical estimate of the axial correlation $\langle \cos 2\Psi \rangle$ as a function of the relative distance for tectonic lineaments. Open squares show results from all bins, filled squares are those data points which were used for the model fit (χ^2 computation); error bars are provided by delete-d jackknife. Inset shows constant $\Delta\chi^2=\chi^2-\chi^2_{\min}$ misfit contours (Eq. (17)) at $\Delta\chi^2=2.3$, 4.6, 6.2, and 9.2 levels. Here, $\chi^2_{\min}$ is the smallest misfit, and for normally distributed errors and two degrees of freedom, these $\Delta\chi^2$ contours would correspond to 68%, 90%, 95%, and 99% significance levels, respectively. The solid curve, solid gray star, and values in the legend specify the best-fit λ–ν values (ν^*, ν^*) for $\chi^2_{\min}$, where parameters ranges are read off the $\Delta\chi^2=6.2$ (95%) contour, and ε is inferred from the y-axis intersect. The dashed curve and open star correspond to the best-fit range for $\nu=0.5$, $\lambda^*_{\nu=0.5}$. Lower plot shows the number of pairs in each distance bin.

We therefore also report the corresponding wavelength for $\nu=0.5$, $\lambda^*_{\nu=0.5}$, which is ~1300 km. From Eq. (10), we find that the variations of $\hat{\Psi}$ obey length scales of $L\sim330$ km. Such relatively short wavelength variations are probably related to the Proterozoic and Archean structures in Australia (cf. Fig. 3a). They are also broadly consistent with the finding that azimuthal anisotropy from surface waves shows a highly heterogeneous pattern at shallowest depths in the Australian continent (Simons and van der Hilst, 2003; Heintz and Kennett, 2006).

3.2. Shear-wave splitting

We now turn to azimuthal anisotropy as imaged by splitting of SKS and SKKS phases (Fig. 1). Most of our data are from the Arizona State University (ASU) database (Fouch, 2006), which is a continuously updated compilation of published splitting measurements based on earlier work (Silver, 1996; Schutt and Kubo, 2001). These data consist of 3630 splits from 1377 different stations and were augmented by measurements from the ISC Automated Splitting Project (Evans et al., 2006a,b). We accessed the ISC web site for splits from the following networks, with time period and numbers of stations given in parentheses: Canadian National Seismic Network (1989–2001; 33), China Digital Seismic Network (1986–2001; 10), GEOFON (1993–2001; 43), GEOSCOPE (1987–2001; 24), IRIS/IDA (1986–2001; 41), and IRIS/USGS (1988–2001; 60). Measurements were performed with the eigenvalue method, and all data that passed the following quality criteria were used: maximum eigenvalue ratio of the corrected traces smaller than 0.07, maximum δt error of 1.25 s, and maximum δt of 3.5 s (for definitions, see Evans et al., 2006a,b). The ISC dataset then contributes 2060 splits at 170 stations. Together with a few added regional studies, our complete SKS and SKKS dataset has 7705 entries, from 1474 stations. We found that our general conclusions could have been arrived at with the ASU database alone, but clustering effects were less severe for the expanded dataset. We only use non-null delay time estimates and spatially average the splitting data on a $0.5°\times0.5°$ grid. The simplest way this can be done for orientational data is by averaging the components of vectors that have twice the original azimuth, and then dividing the mean azimuth by two. The SKS grid averaging procedure yields 1302 data points as used for further analysis.

We have tried different approaches of analyzing subsets of the data, based on geographic regions and other criteria. Here, we report results from the *GTR*-1 tectonic regionalization (Jordan, 1981). *GTR*-1 is a fairly coarse, $5°\times5°$, parameterization of crustal geology; we consider this scheme, however, sufficient for our purposes. The globe is subdivided into three different

oceanic plate regions (codes A–C; 9% of the averaged splits), and three continental types based on their broad tectonic history during the Phanerozoic: Precambrian shields and platforms (S, 15% of the data, "shields" in the remainder), Phanerozoic platforms (P, 15%, "platforms"), and Phanerozoic orogenic zones and magmatic belts (Q, 62%, "orogenic zones").

Considering the delay times of the splitting data, our averaged splits have mean $\langle\delta t\rangle \pm$ standard deviation of $\langle\delta t\rangle = 0.9 \pm 0.5$ s (original data: $\langle\delta t\rangle = 1.2 \pm 0.9$ s). Data from platforms are more tightly clustered around smaller delay times than data from Precambrian shields or orogenic zones. However, the differences are small (mean delay times differ by ~0.2 s) and probably not significant in a tectonic sense. Given the systematic and procedural difficulties in interpreting particularly the delay times of shear-wave splitting (e.g. Chevrot, 2006), we do not think that further analysis is warranted at this point. We therefore choose to limit this study to inferred fast azimuths only, and defer a joint analysis of orientation and amplitude of splitting vectors until denser and consistently measured datasets are available.

3.3. Spatial correlation of azimuthal anisotropy from splitting

All 1302 fast azimuths from splitting are analyzed in Fig. 4, each bin contains ~5000 data pairs up to 4000 km. As for the Australia dataset, we use a geographic box for the jackknife errors. With 25° × 25° size and a minimum number requirement of 35, we find $n = 9$. Fig. 4 shows the axial correlation $\langle\cos 2\Psi\rangle$ as a function of relative distance for data with error bars. The preferred Hurst exponent is $\nu \sim 0.4$, and the estimated effective measurement error is $\varepsilon = 18°$. This ε value is compatible with cited standard errors from the shear-wave splitting literature, typically reported to be of order $\varepsilon = 15° \pm 5°$ (e.g. Silver and Chan, 1991; Sandvol and Hearn, 1994). As for the tectonic data in Fig. 3, we only used data plotted as solid squares in Fig. 4 for the χ^2 computation. Undulations in $\langle\cos 2\Psi\rangle$ estimates for $r \gtrsim 3000$ km are caused by clustering of data that leads to second order correlation wavelengths. We also find less severe deviations from the exponential-like decay of $\langle\cos 2\Psi\rangle$ at $r \sim 1900$ km, where the jackknife technique indicates large uncertainties.

In terms of correlation lengths, there is a clear preference for larger values than for the Australia tectonic dataset. If we restrict the χ^2 fit to $\langle\cos 2\Psi\rangle$ estimates for distances $r < 3000$ km, the formal best-fit is achieved at $\lambda^* \sim 4850$ km for $\nu \sim 0.4$. Both values have large uncertainties based on the extent of permissible ranges

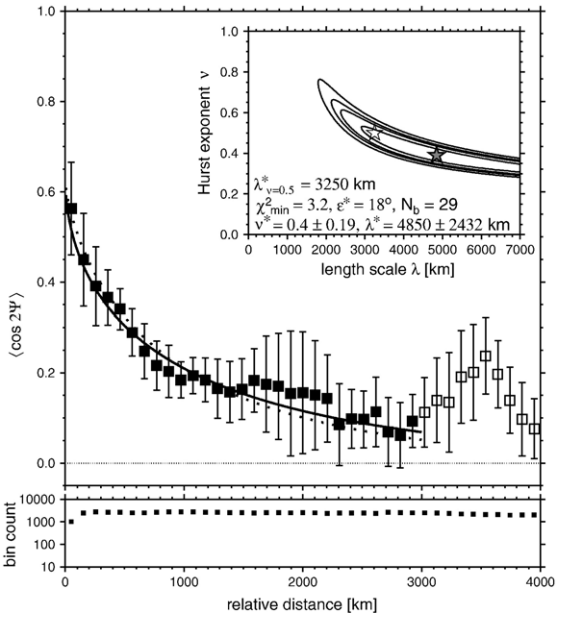

Fig. 4. Statistical estimate of the axial correlation $\langle\cos 2\Psi\rangle$ as a function of the relative distance for all shear-wave splitting data, averaged on a 0.5° × 0.5° grid (cf. Fig. 1). For description of plot, see Fig. 3b, but note different x-axes (distance range).

from the $\Delta\chi^2 = 6.2$ (~95%) contour (inset of Fig. 4). Lower Hurst exponents are preferred by the data (solid best-fit curve and gray star) compared to the crustal structure in Australia which has $\nu = 0.7$–0.8, probably because of map smoothing. However, the exponential auto-correlation function model leads to comparable misfits (dashed curve and open star), at correlation length of $\lambda^*_{\nu=0.5} \sim 3250$ km. This corresponds to a $L \sim 790$ km correlation length in $\hat{\Psi}$(Eq. (10)), well within the spatial coverage of our data.

Only more complete datasets will allow a more robust inference if one stochastic heterogeneity model is better than another in a statistical sense. In the following, we will therefore focus on the simpler, exponential type auto-correlation function with $\nu = 0.5$ when comparing different stochastic models to avoid complications due to the trade-off between λ or L and ν. Comparing L for the mapped crustal signal in Australia (~310 km) with the splitting data (~790 km) that sense the mantle, we find that the anisotropy signal is ~2.5 times more spatially coherent laterally on global scales. Shear-wave splitting in Australia may locally obey a shorter wavelength correlation than the global SKS data, and there may be large lateral and vertical anisotropy variations that complicate the signal. However, this needs to be addressed elsewhere given the relatively small number of SKS observations in that region (Heintz and Kennett, 2006).

3.4. Correlation length of tectonic regions

Subsets of the splitting data are difficult to reliably analyze in terms of the $\langle \cos 2\Psi \rangle$ correlations. However, it is still instructive to consider different tectonic regimes. If we exclude oceanic regions from the complete SKS dataset as shown in Fig. 4, our results are not affected substantially, as oceanic regions make up less than 10% of the data. The oceanic results themselves are not sufficient to estimate any individual length scale there. Subdividing the continental data into tectonic settings, Fig. 5a shows the axial correlation function for tectonically active areas as assigned from the Phanerozoic orogenic zones and magmatic belts of *GTR*-1 (Q), and Table 1 compares the best-fit parameters for different datasets. The orogenic subset is sufficient to reliably estimate a best-fit correlation length of $\lambda^*_{\nu=0.5} \sim 2400$ km, or $L \sim 580$ km. This implies shorter correlation lengths for the tectonically active areas than for the global dataset.

Fig. 5b shows the corresponding axial correlation function of the Phanerozoic platforms (P). The estimates are clearly more scattered than for orogenic regions or the complete dataset, as we now have only ~15% of the data (~190 samples) and the number of pairs in each distance bin is only ~50. This is reflected in the χ^2 curve for the exponential $\nu = 0.5$ decay; it is not possible to confidently assign a specific correlation length to this dataset as the $\Delta\chi^2$ curves for platforms $\nu = 0.5$ show a wide range of permissible values. Fixing ν^* at 0.5, the best-fitting $\lambda^*_{\nu=0.5} \sim 6550$ km, or $L \sim 1590$ km, is larger than for orogenic regions by a factor of ~2.7. Data from Precambrian shields (Fig. 5c) show even more scatter than the platforms plot, as there is now a very uneven distribution of data pairs in the distance bins. Moreover, there are second order fluctuations in the $\langle \cos 2\Psi \rangle$ curve that are related to insufficient sampling. If we accept a single autocorrelation function fit, the trend toward longer correlations lengths from the platform subset is confirmed, with $L \gtrsim 1500$ km. If we only consider the decay of $\langle \cos 2\Psi \rangle$ for shields up to 1000 km (the first wiggle of the $\langle \cos 2\Psi \rangle$ curve in Fig. 5c), we would determine $L \sim 800$ km. This value is still larger than the correlation length we found for orogenic regions, but not as distinct as the large correlations that can be determined for platforms.

As is evident from the scatter and trade-off curves in Fig. 5, exact values for λ^* and ν^* are not well constrained for the presently available SKS data. One may estimate overall smoothness, but often this may either be achieved by lowering the Hurst exponent, or by increasing the correlation length. The statistical model parameters are affected by the clustering of data and are also somewhat

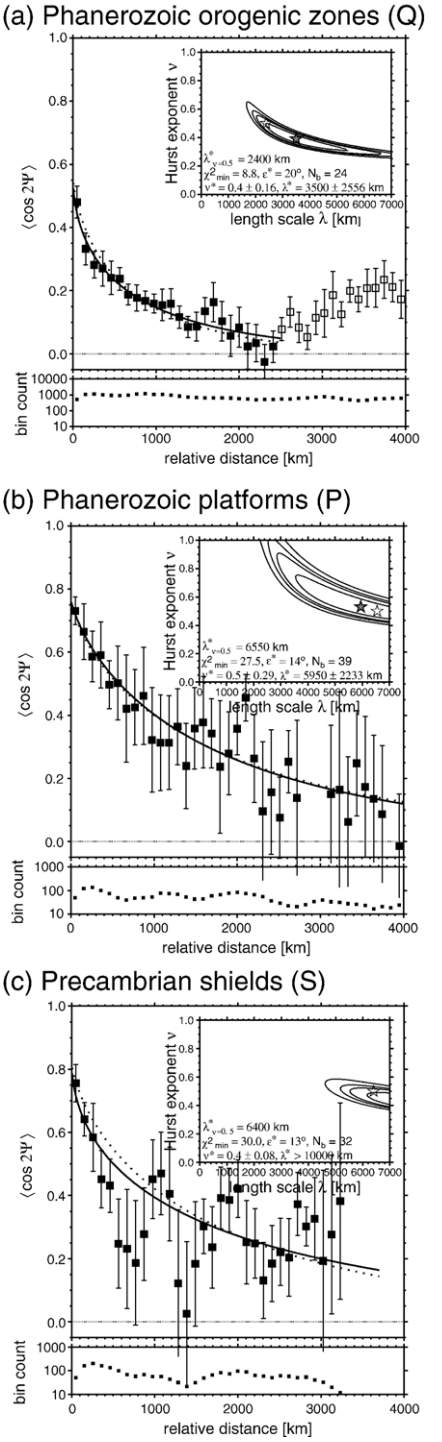

Fig. 5. Statistical estimate of the axial correlation $\langle \cos 2\Psi \rangle$ as a function of the relative distance for subsets of splitting data: Phanerozoic orogenic zones and magmatic belts (a, Q regions of *GTR*-1), Phanerozoic platforms (b, P), and Precambrian shields (c, S). For description of plots, see Fig. 4.

Table 1
Best-fitting stochastic model parameters for the Hurst exponent, ν^*, correlation length for fixed $\nu=0.5$, $\lambda^*_{\nu=0.5}$, corresponding correlation length for $\hat{\Psi}$ (from Eq. (8)), L, and estimated measurement uncertainty ε as shown in Figs. 3, 4 and 5

Type of data	Figure	$\langle\cos 2\Psi\rangle$		$\hat{\Psi}$	
		ν^*	$\lambda^*_{\nu=0.5}$ (km)	L (km)	ε (°)
Australia crustal map	3	0.7	1300+	310	18
SKS: global	4	0.4	3250	790	18
SKS: orogens (Q)	5a	0.4	2400	580	20
SKS: platforms (P)	5b	0.5	6550	1590	14
SKS: shields (S)	5c	0.4?	≳6400?	≳1550?	13

Q: Phanerozoic orogens and magmatic regions; P: Phanerozoic platforms; S: Precambrian shields, all from *GTR*-1. Compare Fig. 6.

dependent on the choices of the inversion strategy with regard to binning, error estimates for $\langle\cos 2\Psi\rangle$, and choices or estimates of ε. For the test case of the Australian crustal structure, ν is easier to determine, and we found a preference of a higher Hurst exponent ($\nu\sim0.7$) there. This compares to the SKS estimate for orogenic regions of $\nu\sim0.4$, and indicates that splitting data, and so mantle structure by inference, in regions of active tectonic may have a rougher spectrum than the crustal structure map of Australia (which may very well reflect the smoothed nature of Wellman's (1976) map).

A more interesting result is that the correlation lengths of tectonically more active regions within the continents are smaller than the correlation lengths of more stable regions. We experimented with a range of fitting and binning schemes and always found that geologically older regions such as platforms are inferred to have longer correlation lengths, or to exhibit smoother structure, than the more recently active, orogenic regions. Consistent differences in correlation lengths are also found for geographic subsets, e.g. when the North American data is subdivided into western US (geologically more active, East of the Rocky Mountains) and eastern US (eastern seaboard, Canadian cratons). For those two subsets, we again find longer correlation lengths for the geologically less active eastern US.

The tectonic subsets have irregular geographic bounds, and this may affect correlation lengths. However, further analysis shows that the data span similar distance ranges that go well beyond the best-fit λ^* correlation lengths. Histograms of pair distances for orogenic and platform subsets show bimodal distributions with peaks at $r\sim1000$ km and ~9000 km. Intermediate distance ranges are similarly well sampled for orogens and platforms. The shield subset is more irregular, with a third distance cluster at $r\sim15000$ km and poor sampling at ~4000 km; the problems with the low-numbers of shields were discussed above. However, we expect that the general differences in correlation lengths for orogens and platforms are not caused by the different distance distributions.

4. Discussion

Detection of smoother, longer wavelength patterns in older compared to younger crust may seem counter-intuitive at first. Given the complex nature of the Proterozoic assemblages around Archean cratons, old continental crust and lithosphere may be expected to appear very heterogeneous in shear-wave splitting. Strong local gradients in delay times are indeed observed in continental lithosphere splitting (Fouch and Rondenay, 2006), and the mapped crustal structure from Australia indicates relatively shorter correlation lengths. However, the opposite is the case for our analysis of global shear wave splitting statistics where older continental regions appear more laterally coherent. One interpretation of these results is that the tectonically active regions are dominated by smaller scale convective processes in the mantle (small-scale convection or regional slab anomalies, e.g. Farallon sinker under the western US), whereas old continents display larger scale, inherited structure (old collisions, e.g. Greenville front). If anisotropy in continents is mainly due to frozen-in past tectonic episodes, the long-wavelength correlations in splitting may record continent–continent collisions throughout the super-continental cycle. Such a model would allow for the observed sharp, local gradients between geologic domains that are internally relatively smooth (Huang et al., 2000; Fouch et al., 2004) and would also explain why the inferred correlations lengths for old continents (~1600 km) are relatively large compared to individual craton dimensions.

Alternatively, our estimates of correlation lengths may reflect lateral variations in the present-day mantle. Underneath shields and platforms, higher viscosities in the tectosphere may lead to less vigorous, or differently shaped flow than underneath younger continents, where the mantle may be relatively hot and weak (Fouch et al.,

2000; Becker et al., 2006a,b; Conrad et al., in press). Such variations may also affect the partitioning of anisotropy between lithosphere and asthenosphere which appear to both contribute to anisotropy in continental settings (Fouch and Rondenay, 2006). Longer correlation lengths may then indicate that present-day asthenospheric flow is relatively more important than lithospheric anomalies. This suggestion of a relatively stronger lithospheric signal for shorter correlation length regions is, however, somewhat at odds with the expectation of a relatively thin lithosphere in the western US. There, mantle flow-derived models also explain SKS splitting well (Becker et al., 2006a,b).

To answer if the SKS patterns reflect mainly present-day convection or recorded convective history, it would be desirable to compare L estimates for the continental regions with those from oceanic plates. For the latter, we would expect current plate tectonic processes as evidenced by plate motions to be the dominant cause for anisotropy. Unfortunately, the SKS dataset is not sufficient to determine correlation lengths in the oceans. We can, however, turn to surface wave inversions for azimuthal anisotropy (e.g. Montagner and Tanimoto, 1991; Ekström, 2001; Trampert and Woodhouse, 2003), or to geodynamic models that incorporate estimates of plate-motion related mantle flow (Becker et al., 2006a,b; Becker, 2006). Surface wave models provide a less direct measure of anisotropy than SKS as such models involve inversions where damping and parameterization impose data-independent length-scales. However, coverage is basically global, albeit at low effective resolution of ϕ-type structure (up to spherical harmonic degree $\lesssim 20$) because of uneven ray density and other complications (e.g. Tanimoto and Anderson, 1985; Trampert and Woodhouse, 2003; Becker et al., 2007).

With these caveats, we computed correlation curves in analogy to Fig. 5 for surface waves. Those were based on associating ϕ with the azimuthal anisotropy fast azimuths of Ekström's (2001) Rayleigh-wave phase velocity maps at 50 s (peak sensitivity at ~ 75 km depth), using only the 2Φ component of the Smith and Dahlen (1973) type anisotropy expansion. We use an even-area sampling of the seismological model at $\sim 2° \times 2°$ and find that $\langle \cos 2\Psi \rangle$ from surface waves prefers smooth Hurst exponents of $\nu^* \sim 1$ but falls off more rapidly than the $\nu = 1$ predictions at distances $r \gtrsim 2000$ km. This indicates that the damping choices of the seismological inversions are mapped into the correlation functions and makes our stochastic model inadequate to determine correlation lengths. We therefore pick the distance at which $\langle \cos 2\Psi \rangle$ drops to zero, r_0, as a measure of relative coherence regardless of the functional shape of $\langle \cos 2\Psi \rangle$. The fact that $\langle \cos 2\Psi \rangle$ crosses zero, and does not remain at small but positive values as for the SKS data, implies that there is a dominant wavelength in the surface wave models which is related to the spherical harmonics parameterization.

To test lower bounds on resolution, we applied our correlation analysis to random spherical harmonic fields that were spatially parameterized similarly to the phase velocity maps by Ekström (2001). The shortest fall-off in coherence we expect to be detectable in such models is $r_0 \sim 1000$ km. For the actual Rayleigh wave structure, we find larger values: $r_0 \sim 1600$, 2050, and 2300 km for orogens, platforms, and shields, respectively. This progression toward longer correlation in older continental regions is consistent with our SKS splitting results. Fig. 6 shows relative correlation lengths with respect to orogenic regions for both surface waves and our SKS results. The drop-off distance for oceanic regions for surface waves are a factor of ~ 2 larger than the continental values, $r_0 \sim 4500$ km. While uneven ray coverage may explain part of the relative de-coherence within the continents, we expect the general result of longer coherence in oceans to be robust. It may indicate that asthenospheric flow is the main cause of anisotropy in the oceanic regions and imprints a longer wavelength signal there, compared to the continental regions which are overall more affected by frozen-in texture.

To put our findings on observed anisotropic upper mantle structure into context, we can turn to geodynamic models. Mantle flow may be estimated from plate motions and internal density anomalies as inferred from isotropic seismic tomography (Gaboret et al., 2003; Becker et al., 2003; Behn et al., 2004). Such models involve strong assumptions; importantly, mantle flow is often assumed to be steady-state and so only representative of the last few 10 s of Ma of tectonic history. However, geodynamic models allow insights into how asthenospheric flow may affect anisotropy correlation lengths. We consider both inferred mantle currents themselves, and the anisotropy that arises through progressive deformation of advected rocks in that flow field (lattice preferred orientation of olivine). To this end, we sampled mantle flow velocities at asthenospheric depths (~ 250 km), and the predicted surface wave anisotropy. The latter is based on vertical kernel-averages of elastic anisotropy from texture computations; details of the flow models and anisotropy estimates are described in Becker et al. (2003), Becker (2006) and Becker et al. (2006a,b). Both spatial estimates of fast azimuths were evenly sampled on the same spatial grid that was used for the surface wave data. For geodynamic models, additional smoothness constraints are imposed by the

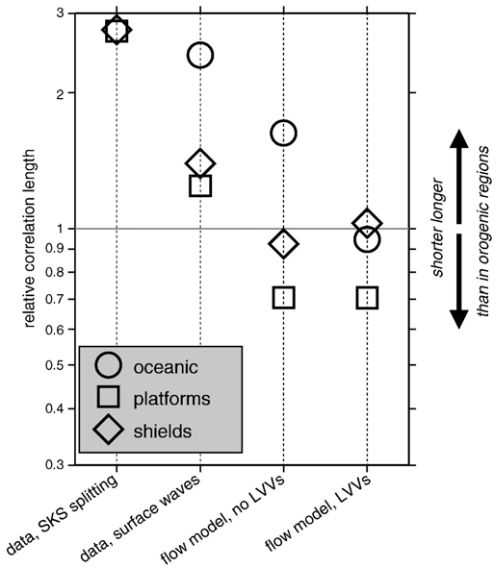

Fig. 6. Ratio of correlation lengths for azimuthal anisotropy from SKS splitting, Rayleigh waves, and flow models in different tectonic settings (oceanic: ABC, platforms: P, shields: S) relative to orogenic and magmatic regions (Q of *GTR*-1). SKS splitting estimates are computed by dividing L values as shown in Table 1 by the L value for orogens (no oceanic plate estimate is possible for SKS). Surface wave ratios are for 50 s Rayleigh waves from Ekström (2001), we use r_0 values as discussed in main text. Flow model anisotropy is based on computing mantle circulation, estimating elastic anisotropy, and computing kernel-averaged, 50 s Rayleigh wave anisotropy (for details, see Becker et al., 2003, 2006a,b). Correlation length ratios are computed from $\lambda^*_{\nu=1}$ for computations without and with strong lateral viscosity variations (LVVs). Analysis uncertainties in all ratios are about twice the symbol size.

input models (resolution of tomography) and results are most applicable for oceanic plates (Becker et al., 2003, 2007).

We test a simple rheological mantle structure with purely radially varying viscosity that is similar to the best-fit model in Becker et al. (2003) and also consider a flow model with strong LVVs, where viscosity obeys a joint diffusion–dislocation creep law as based on laboratory results for olivine [$\eta_{\rm eff}$ model of Becker (2006)]. In general, we find that smooth anisotropic structure with $\nu^* \sim 0.9$–1 is preferred by the geodynamic models, and our stochastic description (Eq. (1)) captures the $\langle \cos 2\Psi \rangle$ behavior well. The global correlation lengths for the model without LVVs are $\lambda^*_{\nu=1} \sim 3200$ km for velocities, and $\lambda^*_{\nu=1} \sim 1850$ km for predicted 50 s Rayleigh wave azimuthal anisotropy. Such a shortening of coherence is expected, as anisotropy is formed by shearing, and strain-rates are the spatial derivative of the velocities. Focusing on correlation lengths for anisotropy synthetics, we predict $\lambda^*_{\nu=1} \sim 2200$ km for oceanic regions. If we isolate continents, we find $\lambda^*_{\nu=1} \sim 1350$, 950, and 1250 km for orogenic, platform, and shield regions, respectively (Fig. 6). This means that for flow models without LVVs, oceanic regions are estimated to be longer wavelength than continental plate anisotropy, similar to our surface wave observation. Within the continents, tectonically young orogenic regions are found to be laterally more coherent than older ones (particularly platforms), unlike what we observed for SKS splitting and surface waves.

If we allow for strong LVVs, including continental keels and relatively weak sub-oceanic asthenosphere (Becker, 2006), the situation is reversed for the continental/oceanic plate distinction (Fig. 6). In the oceanic regions, we find $\lambda^*_{\nu=1} \sim 1750$ km, and within continental lithosphere $\lambda^*_{\nu=1} \sim 1850$, 1300, and 1900 km for orogenic regions, platforms, and shields, respectively. Oceanic correlation length estimates are thus slightly smaller than those for orogenic regions for LVV flow. This decrease compared to the model without LVVs is related to the relatively more complex flow underneath the oceanic plates for LVVs (Becker, 2006). Shields are slightly longer wavelength than orogenic regions for LVV models, as the flow hypothesis for tectonic differences would lead us to expect. However, the increase in correlation relative to orogens is minor (~within uncertainty of correlation length ratio estimates), and much smaller than the observed seismological values. Also, the old platforms remain relatively unchanged for LVV flow and are still estimated to have shorter correlation lengths than the orogenic zones. Becker's (2006) continental keel model is somewhat different from the *GTR*-1 parameterization, and the relationship between LVVs, the planform of convective flow, and resulting seismic anisotropy certainly needs to be evaluated further (Becker et al., 2006a,b; Conrad and Lithgow-Bertelloni, 2006; Becker et al., 2007; Conrad et al., in press). However, we take these findings as indications that the longer correlation lengths in older crust that are observed both in SKS splitting and surface wave azimuthal anisotropy are more likely due to frozen-in past tectonic episodes than variations in present-day convective flow.

5. Conclusions

Stochastic models of azimuthal seismic anisotropy provide a new way of characterizing the continental lithosphere and upper mantle. Defining such stochastic parameters is useful to better understand the present discrepancies in heterogeneity estimates from different kinds of seismological data. We are able to detect longer correlation lengths, or a smoother signal, in the

Phanerozoic platforms ($L \gtrsim 1500$ km) than in Phanerozoic orogenic zones ($L \sim 600$ km). A comparison of the SKS results with surface wave models and geodynamic estimates provides evidence that a frozen-in record of large-scale processes, such as continent–continent collisions, is archived in the thick tectosphere. Anisotropy underneath old continents may be dominated by long-term tectonics, while anisotropy in younger continental regions and the oceanic plates is predominantly caused by asthenospheric flow over the last few 10 s of Ma.

Acknowledgments

This project was initiated with the help of Olav van Genabeek and James B. Gaherty and the manuscript benefited from the comments of two anonymous reviewers. We thank Matt Fouch, the ISC, and Göran Ekström for making their respective SKS splitting databases and surface wave models available. We also thank Boris Kaus, Clare Steedman and Katrin Plenkers for helping with completing the dataset. Most figures were prepared with the Generic Mapping Tools (Wessel and Smith, 1991). This material is based upon work supported by the National Science Foundation under Grant No. 0509722. This research was also supported by the Southern California Earthquake Center. SCEC is funded by NSF Cooperative Agreement EAR-0106924 and USGS Cooperative Agreement 02HQAG0008. The SCEC contribution number for this paper is 1052.

References

Abramowitz, M., Stegun, I.A., 1972. Handbook of Mathematical Functions. Dover Publications, Inc.

Anderson, D.L., 1989. Theory of the Earth. Blackwell Scientific Publications, Boston.

Becker, T.W., 2006. On the effect of temperature and strain-rate dependent viscosity on global mantle flow, net rotation, and plate-driving forces. Geophys. J. Int. 167, 943–957.

Becker, T.W., Kellogg, J.B., Ekström, G., O'Connell, R.J., 2003. Comparison of azimuthal seismic anisotropy from surface waves and finite-strain from global mantle-circulation models. Geophys. J. Int. 155, 696–714.

Becker, T.W., Chevrot, S., Schulte-Pelkum, V., Blackman, D.K., 2006a. Statistical properties of seismic anisotropy predicted by upper mantle geodynamic models. J. Geophys. Res. 111, B08309. doi:10.1029/2005JB004095.

Becker, T.W., Schulte-Pelkum, V., Blackman, D.K., Kellogg, J.B., O'Connell, R.J., 2006b. Mantle flow under the western United States from shear wave splitting. Earth Planet. Sci. Lett. 247, 235–251.

Becker, T.W., Ekström, G., Boschi, L., Woodhouse, J.W., submitted for publication. Length-scales, patterns, and origin of azimuthal seismic anisotropy in the upper mantle as mapped by Rayleigh waves, Geophys. J. Int. Online at http://geodynamics.usc.edu/(becker/preprints/bebw07.pdf.

Behn, M.D., Conrad, C.P., Silver, P.G., 2004. Detection of upper mantle flow associated with the African Superplume. Earth Planet. Sci. Lett. 224, 259–274.

Bird, P., Li, Y., 1996. Interpolation of principal stress directions by nonparametric statistics: global maps with confidence limits. J. Geophys. Res. 101, 5435–5443.

Bunge, H.-P., Hagelberg, C.R., Travis, B.J., 2003. Mantle circulation models with variational data assimilation: inferring past mantle flow and structure from plate motion histories and seismic tomography. Geophys. J. Int. 152, 280–301.

Chemingui, N., 2001. Modeling 3-D anisotropic fractal media. Technical Report. Stanford Exploration Project, Report, vol. 80. Stanford University.

Chernov, L.A., 1960. Wave Propagation in a Random Medium. McGraw-Hill, New York.

Chevrot, S., 2006. Finite-frequency vectorial tomography: a new method for high-resolution imaging of upper mantle anisotropy. Geophys. J. Int. 165, 641–657.

Chevrot, S., van der Hilst, R.D., 2003. On the effects of a dipping axis of symmetry on shear wave splitting measurements. Geophys. J. Int. 152, 497–505.

Conrad, C.P., Lithgow-Bertelloni, C., 2006. Influence of continental roots and asthenosphere on plate–mantle coupling. Geophys. Res. Lett. 33. doi:10.1029/2005GL02562.

Conrad, C.P., Behn, M.D., Silver, P.G., in press. Global mantle flow and the development of seismic anisotropy: differences between the oceanic and continental upper mantle, J. Geophys. Res.

Debayle, E., Kennett, B.L.N., Priestley, K., 2005. Global azimuthal seismic anisotropy and the unique plate-motion deformation of Australia. Nature 433, 509–512.

Efron, B., Stein, C., 1981. The jackknife estimate of variance. Ann. Stat. 9, 586–596.

Ekström, G., 2001. Mapping azimuthal anisotropy of intermediate-period surface waves (abstract). EOS, Trans. AGU 82 (47), S51E-06.

Evans, M.S., Kendall, J.-M., Willemann, R.J., 2006a. Automated SKS splitting and upper-mantle anisotropy beneath Canadian seismic stations. Geophys. J. Int. 165, 931–942.

Evans, M.S., Kendall, J.-M., Willemann, R.J., 2006. Automated splitting project database, Online at http://www.isc.ac.uk/SKS/, accessed 02/2006.

Fischer, K.M., Wiens, D.A., 1996. The depth distribution of mantle anisotropy beneath the Tonga subduction zone. Earth Planet. Sci. Lett. 142, 253–260.

Fouch, M. Upper mantle anisotropy database, Online, 2006. Accessed in 06/2006, http://geophysics.asu.edu/anisotropy/upper/.

Fouch, M.J., Rondenay, S., 2006. Seismic anisotropy beneath stable continental interiors. Phys. Earth Planet. Inter. 158, 292–320.

Fouch, M.J., Fischer, K.M., Parmentier, E.M., Wysession, M.E., Clarke, T.J., 2000. Shear wave splitting, continental keels, and patterns of mantle flow. J. Geophys. Res. 105, 6255–6275.

Fouch, M.J., Silver, P.G., Bell, D.R., Lee, J.N., 2004. Small-scale variations in seismic anisotropy near Kimberley, South Africa. Geophys. J. Int. 157, 764–774.

Gaboret, C., Forte, A.M., Montagner, J.-P., 2003. The unique dynamics of the Pacific Hemisphere mantle and its signature on seismic anisotropy. Earth Planet. Sci. Lett. 208, 219–233.

Goff, J.A., Jordan, T.H., 1988. Stochastic modeling of seafloor morphology: inversion of sea beam data for second-order statistics. J. Geophys. Res. 93, 13589–13608.

Heintz, M., Kennett, B.L.N., 2006. The apparently isotropic Australian upper mantle. Geophys. Res. Lett. 33, L15319. doi:10.1029/2006GL026401.

Huang, W.C., Ni, J.F., Tilmann, F., Nelson, D., Guo, J., Zhao, W., Mechie, J., Kind, R., Saul, J., Rapine, R., Hearn, T.M., 2000. Seismic polarization anisotropy beneath the central Tibetan plateau. J. Geophys. Res. 105, 27979–27989.

Isaaks, E.H., Srivastava, R.M., 1989. An Introduction to Applied Geostatistics. Oxford University Press, New York.

Jordan, T.H., 1981. Global tectonic regionalization for seismological data analysis. Bull. Seismol. Soc. Am. 71, 1131–1141.

Levin, V., Menke, W., Park, J., 1999. Shear wave splitting in the Appalachians and the Urals: a case for multilayered anisotropy. J. Geophys. Res. 104, 17975–17993.

Meade, C., Silver, P.G., Kaneshima, S., 1995. Laboratory and seismological observations of lower mantle isotropy. Geophys. Res. Lett. 22, 1293–1296.

McKenzie, D.P., 1979. Finite deformation during fluid flow. Geophys. J. R. Astron. Soc. 58, 689–715.

McNamara, A.K., van Keken, P.E., Karato, S.-I., 2003. Development of finite strain in the convecting lower mantle and its implications for seismic anisotropy. J. Geophys. Res. 108, 2230. doi:10.1029/2002JB001970.

Montagner, J.-P., Guillot, L., 2000. Seismic anisotropy in the Earth's mantle. In: Boschi, E., Ekström, G., Morelli, A. (editors). Problems in Geophysics for the New Millenium. Istituto Nazionale di Geofisica e Vulcanologia, Editrice Compositori, Bologna, Italy, pp. 217-253.

Montagner, J.-P., Tanimoto, T., 1991. Global upper mantle tomography of seismic velocities and anisotropies. J. Geophys. Res. 96, 20337–20351.

Montagner, J.-P., Griot-Pommera, D.-A., Lavee', J., 2000. How to relate body wave and surface wave anisotropy? J. Geophys. Res. 105, 19015–19027.

Podolefsky, N.S., Zhong, S., McNamara, A.K., 2004. The anisotropic and rheological structure of the oceanic upper mantle from a simple model of plate shear. Geophys. J. Int. 158, 287–296.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1993. Numerical Recipes in C: The Art of Scientific Computing, 2 edition. Cambridge University Press, Cambridge.

Roth, M., Korn, M., 1993. Single scattering theory versus numerical modeling in 2-D random media. Geophys. J. Int. 112, 124–140.

Rümpker, G., Silver, P.G., 1998. Apparent shear-wave splitting parameters in the presence of vertically varying anisotropy. Geophys. J. Int. 135, 790–800.

Saito, T., 2006. Synthesis of scalar-wave envelopes in two-dimensional weakly anisotropic random media by using the Markov approximation. Geophys. J. Int. 165, 501–515.

Saltzer, R.L., Gaherty, J.B., Jordan, T.H., 2000. How are vertical shear wave splitting measurements affected by variations in the orientation of azimuthal anisotropy with depth? Geophys. J. Int. 141, 374–390.

Sandvol, E., Hearn, T., 1994. Bootstrapping shear-wave splitting errors. Bull. Seismol. Soc. Am. 84, 1971–1977.

Savage, M.K., 1999. Seismic anisotropy and mantle deformation: what have we learned from shear wave splitting? Rev. Geophys. 37, 65–106.

Savage, M.K., Sheehan, A.F., 2000. Seismic anisotropy and mantle flow from the Great Basin to the Great Plains, western United States. J. Geophys. Res. 105, 13715–13734.

Savage, M.K., Fischer, K.M., Hall, C.E., 2004. Strain modelling, seismic anisotropy and coupling at strike-slip boundaries: applications in New Zealand and the San Andreas fault. In: Grocott, J., Tikoff, B., McCaffrey, K.J.W., Taylor, G. (Eds.), Vertical Coupling and Decoupling in the Lithosphere. Geol. Soc. Lond., Spec. Pubs, vol. 227. Geological Society of London, London, pp. 9–40.

Simons, F.J., van der Hilst, R.D., 2003. Seismic and mechanical anisotropy and the past and present deformation of the Australian lithosphere. Earth Planet. Sci. Lett. 211, 271–286.

Simons, F.J., van der Hilst, R.D., Montagner, J.-P., Zielhuis, A., 2002. Multimode Rayleigh wave inversion for heterogeneity and azimuthal anisotropy of the Australian upper mantle. Geophys. J. Int. 151 (3), 738–754.

Silver, P.G., 1996. Seismic anisotropy beneath the continents: probing the depths of geology. Annu. Rev. Earth Planet. Sci. 24, 385–432.

Silver, P.G., Chan, H.H., 1988. Implications for continental structure and evolution from seismic anisotropy. Nature 335, 34–39.

Silver, P.G., Chan, H.H., 1991. Shear wave splitting and subcontinental mantle deformation. J. Geophys. Res. 96, 16429–16454.

Silver, P.G., Holt, W.E., 2002. The mantle flow field beneath Western North America. Science 295, 1054–1057.

Silver, P.G., Savage, M.K., 1994. The interpretation of shear wave splitting parameters in the presence of two anisotropic layers. Geophys. J. Int. 119, 949–963.

Schulte-Pelkum, V., Blackman, D.K., 2003. A synthesis of seismic P and S anisotropy. Geophys. J. Int. 154, 166–178.

Schutt, D., Kubo, A., 2001. Anisotropy Resource Page. Carnegie Institution of Washington. http://www.ciw.edu/schutt/anisotropy/aniso_source.html, access 04/2001.

Smith, M.L., Dahlen, F.A., 1973. The azimuthal dependence of Love and Rayleigh wave propagation in a slightly anisotropic medium. J. Geophys. Res. 78, 3321–3333.

Trampert, J., Woodhouse, J.H., 2003. Global anisotropic phase velocity maps for fundamental mode surface waves between 40 and 150 s. Geophys. J. Int. 154, 154–165.

Tanimoto, T., Anderson, D.L., 1985. Lateral heterogeneity and azimuthal anisotropy of the upper mantle: Love and Rayleigh waves 100–250 s. J. Geophys. Res. 90, 1842–1858.

Vinnik, L., Kosarev, G.L., Makeyeva, L.I., 1984. Anisotropy of the lithosphere from the observations of SKS and SKKS phases. Proc. Acad. Sci. USSR 278, 1335–1339.

Vinnik, L.P., Green, R.W.E., Nicolaysen, L.O., 1995. Recent deformations of the deep continental root beneath southern Africa. Nature 375, 50–52.

von Karma'n, T., 1948. Progress in the statistical theory of turbulence. J. Mar. Res. 7, 252–264.

Walker, K.T., Bokelmann, G.H.R., Klemperer, S.L., Nyblade, A., 2005. Shear-wave splitting around hotspots: evidence for upwelling-related mantle flow? In: Foulger, G., Nat-land, J.H., Presnall, D.C., Anderson, D.L. (Eds.), Plates, Plumes, and Paradigms. . Geological Society of America Special Paper, vol. 388. Geological Society of America, pp. 171–192.

Watson, G.S., 1983. Statistics on Spheres. John Wiley & Sons, New York.

Wellman, P., 1976. Gravity trends and the growth of Australia: a tentative correlation. J. Geol. Soc. Aust. 23, 11–14.

Wessel, P., Smith, W.H.F., 1991. Free software helps map and display data. EOS, Trans. AGU 72, 445–446.