Structure and Evolution of the Australian Continent: Insights from Seismic and Mechanical Heterogeneity and Anisotropy

by

Frederik Jozef Maurits Simons

Kandidaat in de Geologie (1994), Katholieke Universiteit Leuven
Licentiaat in de Geologie (1996), Katholieke Universiteit Leuven

Submitted to the Department of Earth, Atmospheric and Planetary Sciences
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2002

Author ....
Department of Earth, Atmospheric and Planetary Sciences
March 12, 2002

Certified by ....
Robert D. van der Hilst
Professor
Thesis Supervisor

Accepted by ....
Ronald G. Prinn
Head, Department of Earth, Atmospheric and Planetary Sciences

Abstract

In this thesis, I explore the geophysical structure and evolution of the Australian continental lithosphere. I combine insights from isotropic and anisotropic seismic surface-wave tomography with an analysis of the anisotropy in the mechanical properties of the lithosphere, inferred from the coherence between gravity anomalies and topography. With a new high-resolution waveform tomographic model of Australia, I demonstrate that the depth of continental high wave speed anomalies does not universally increase with age, but is dependent on the scale and the tectonic history of the region under consideration. I construct an azimuthally anisotropic three-dimensional model of the Australian upper mantle from Rayleigh-wave waveforms. I compare Bayesian inverse methods with discretely parameterized regularization methods, and explore the use of regular, tectonic and resolution-dependent tomographic grids. I advocate the use of multitaper spectral estimation techniques for coherence analysis of gravity and topography, applied to Australian isostasy. I investigate the importance of internal loading, the directional anisotropy of the gravitational response to loading, and the estimation bias affecting the long wavelengths of the coherence function. I develop a method for non-stationary coherence analysis which enables a complete characterization of continental strength by the dependency of gravity-topography coherence on wavelength, direction and geologic age. Combining high-resolution, depth-dependent anisotropy measurements from surface-wave tomography with the mechanical anisotropy from gravity/topography coherence, I assess the validity of two competing theories regarding the cause of continental anisotropy (vertically coherent deformation or simple asthenospheric flow) quantitatively for the very first time.

Acknowledgments

I dedicate this work to my parents, in the memory of my grandfathers, Jozef Simons (1888-1948) and Maurits Vandemeulebroecke (1908-1990), after whom I was named.

As our very first meeting in a café in Leuven progressed from coffee (several) to beer (tantamount) to aspirin (for me), Professor Rob van der Hilst set the tone for the evolution from tutelage to advising to collaboration which I have had the privilege to enjoy over the course of the last six years. The same day, I met Rob's wife, Dr. Alet Zielhuis (she had fries), who, introducing me to the intricacies of seismic waveform inversion, has remained a collaborator and co-author of my papers.

Professor Maria Zuber stimulated my pursuit of problems relating to coherence analysis. I am grateful for her support and inspiration, and continue to admire her multi-planetary vision and her efficiency at getting things done.

Hrafnkell Kárason not only taught me how to write my first computer program (and many more after that) but was also always first in line to discuss science, real and perceived, unique and non-unique. The friendship with his extended family of *sons* and *dottirs* made the greater Boston area a place to *live* rather than *work*.

I thank Sam Bowring, Don Forsyth, Brad Hager, Tom Jordan, Jun Korenaga, Jean-Paul Montagner, and many more people who, through formal or casual interactions, helped shape, refine, or refute my ideas. Oded Aharonson, Mark Behn. Margaret Boettcher, Sergei Lebedev, Maureen Long, Laurent Montesi, Eliza Richardson, Rebecca Saltzer, Phil Tracadas, and the Great Unnamed, were all *semper parati*, and much appreciated for it.

I also thank Pallavi Nuka: together, we will write Chapter 7.

Contents

3 Anisotropic Isostatic Response 115

List of Figures

Introduction

The geological diversity of the Australian continent is manifest in its physical properties. Factors controlling the evolution (creation, survival and destruction) of continental lithosphere include its temperature and chemical composition, density and elastic properties, and its seismic and mechanical thickness. The Australian lithosphere is both heterogeneous and anisotropic with respect to the geophysical variables that characterize it, and this variation is present on multiple spatial and temporal scales. For instance, the three-dimensional variation of shear wave speeds in the upper mantle, inferred from their effect on the seismic waveforms recorded at the surface, clearly delineates the Australian continental "keel", but on a smaller, regional scale, shear wave speed deviations reveal a continent of a highly complex structure.

The Australian continent consists of three major geological age domains. The easternmost two-thirds contain several Paleozoic fold belts which have undergone Cenozoic igneous activity. The center of Australia is composed of Proterozoic orogens with an additional Archean craton (the Gawler craton) in the South, and an amalgamate of Proterozoic-Archean units in the North (grouped together as the North Australian craton). The Kimberley region is a Proterozoic block. The age of its basement is disputed but is probably Archean. The westernmost third of Australia is the site of the Archean Pilbara and Yilgarn cratons, which are bordered by Proterozoic orogens.

One of the goals of our study of the Australian continent is to determine and interpret the seismic, hence "instantaneous" elastic response and to unravel the evolution of the continent from how these parameters vary with its age. On an entirely different time scale, the expression of topographic features in the regional gravity field is indicative of the long-term mechanical properties of the continental plate. The character of isostatic compensation is apparent from the elastic response to loading. This represents the strength or effective elastic thickness of the lithosphere on long time scales. Strength or thickness variations can be inferred from the correlation of surface topography to gravity anomalies.

The development of texture through the alignment of anisotropic mantle minerals by deformation is expressed in the dependence of seismic wave speeds on propagation direction or on the polarization of particle motion. How seismic anisotropy relates to the deformation of continents has remained enigmatic; for instance, it is not known if and how seismic anisotropy is related to the large-scale and long-term mechanical properties of the continent. As directions that are mechanically weak will accumulate more gravity anomaly for a given amount of topography than the isotropic average, the time- and depth integrated history of strain accumulation will be recorded in the anisotropy of the isostatic compensation mechanism.

Australia is ideally suited for seismic tomographic studies because of its central location with regard to the surrounding plate boundaries, which form an almost circular belt of high seismicity around it. The seismic waveforms collected by the SKIPPY experiment form the basis of the three-dimensional model of shear-wave speed variations presented in **Chapter I** (published in *Lithos* (**48**), 17–43, 1999). The main contributions of Chapter I to the study of wave speeds in the Australian upper mantle are the addition of data from permanent seismic stations in areas where the SKIPPY data had less than ideal resolution, an improved parameterization aimed at enhancing the depth resolution of the model, and the reduction of the dependence on earthquake location errors. Resolution tests and an analysis of the spectral power of the wave speed anomalies show that our regional model of Australia resolves structures at smaller spatial length scales than models based on global data sets, while at the same time being compatible with such global models. A comparison of wave speed anomalies to the age of the Australian crust shows the diminishing correlation of lithospheric structure with depth and decreasing length scale.

In **Chapter II** we present an azimuthally anisotropic wave speed model of the Australian upper mantle (*Geophysical Journal International*, in press). Much thought is given to the parameterization of azimuthally anisotropic wave speed inversions, and resolution, trade-off, and variance reduction are carefully explored. From theoretical as well as practical considerations, our discrete, regularized inversion with (ir)regular surface cells is shown to be compatible with Bayesian inversions which assume an *a priori* Gaussian model covariance structure. The dense data coverage allows the characterization of continental anisotropy in unprecedented detail. At first glance, the top 200 km of the upper mantle presents a regime of strong and rapidly varying anisotropic directions, whereas between 200 and 400 km depth, a smoother pattern of lower-magnitude anisotropy is observed. The increasing correlation of anisotropic directions with the direction of absolute plate motion is a first indication that a top region of frozen lithospheric anisotropy is overlying a more actively deforming lower part, an issue which is explored further in subsequent chapters.

The degree to which surface and subsurface loads are isostatically compensated and the pattern of resulting gravity anomalies are dependent on the effective elastic thickness of the lithosphere, a measure of its mechanical strength. Treating isostatic compensation as a linear system, the coherence, which relates input (topography) to output (gravity anomalies), is an estimate of the isostatic transfer function, which can be compared to forward models of elastic layers overlying fluid half-spaces. The development of quadratic spectrum estimators for stationary fields has greatly improved our ability to estimate (cross-)spectral properties such as coherence with great accuracy and resolution. In **Chapter III** (published in *Journal of Geophysical Research*, (**105**) 19,163–19184, 2000), we explore the use of the Thomson multitaper method, which makes use of Slepian eigenwindows, for an optimal determination of power and coherence of gravity and topography in the frequency domain.

The Thomson method presents a marked improvement over other, single-window methods used in previous studies. We illustrate the effect of the use of this method on the determination of isotropic elastic thickness estimates. Perhaps the greatest gain from multitaper methods is derived from their ability to bypass wavenumber averaging and characterize the directionally anisotropic isostatic response of gravity to loading. The measured mechanical anisotropy of the Australian continent is proposed to be tectonically controlled. In particular, for central Australia, our observations are consistent with the suggestion that the parallel faults in that area act to make the lithosphere weaker in the direction perpendicular to them. The central Australian mechanical anisotropy can also be related to the N–S direction of maximum stress and possibly the presence of E–W running zones weakened due to differential sediment burial rates.

An accurate characterization of the spatial variations of coherence anisotropy of the Australian continent is required in order to make quantitative comparisons between mechanical and seismic anisotropy. In simple models and deformation experiments, seismic anisotropy develops by an alignment of the fast axis of anisotropic minerals, perpendicular to the direction of maximum compression. As we are able to identify regions on the Australian continent which preferentially accumulate strain, expressed by a directional predominance of gravity anomalies that show a significant coherent with the surface topography, we can assess the predictive value of the directions of inferred fossil strain on the direction of seismic anisotropy varying with depth. While Slepian functions are by design ideally suited for the accurate analysis of frequency dependent stochastic properties of data having a finite spatial extent, characterizing the spatial variations of spectral properties requires a joint localization in the space-frequency plane. In **Chapter IV** (*Journal of Geophysical Research*, under review), we explore the use of a class of orthogonal Hermite polynomial functions for the characterization of spatially and azimuthally varying spectral coherence functions. Such Hermite functions have completely isotropic concentration domains. As a consequence, the space of space-frequency stochastic properties can be explored without sacrificing spatial to frequency resolution, and without introducing an artificial anisotropic bias in the coherence function. We use the Hermite multitaper method to map out areas with detectable mechanical anisotropy on the Australian continent, defined as areas with a direction in which the isostatic compensation is more Airy-like than the isotropic average. The comparison between seismic and mechanical anisotropy is addressed in more detail in the last chapter.

The last two chapters use the knowledge of the seismic and mechanical structure of the Australian lithosphere gained from Chapters I—IV to test the validity of existing paradigms in continental structure and evolution. In particular, in **Chapter V** (submitted to *Geophysical Research Letters*) we test the notion that the older parts of the Australian continent are present as deep penetrating keels that are seismically fast and mechanically strong. The covariation of seismic and mechanical thickness with the age of the crust appears more complex than often assumed. The departure from the global trend of the correlation of seismic and mechanical properties with each other and with age shows the danger of making oversimplified generalizations. The detailed complexity of the Australian lithosphere is further evidence for the secular nature of lithospheric properties and in line with recent observations of the small scale variability of the geophysical, geochemical and geodynamical behavior of continents.

Chapter VI applies our new knowledge of the depth-variation of seismic anisotropy (from surface-wave tomography) and the time- and depth integrated deformation history of the continent (from the directional dependence of gravity anomalies on topography) to the issue of continental deformation. Combining both types of data we discriminate between two popular models of continental deformation traditionally based on surface observations with limited vertical resolution. We delineate an upper continental regime in which seismic anisotropy and fossil deformation are primarily related to each other (above 200 km), separated from a deeper regime in which the presently active mantle flow may be responsible for the generation of seismic anisotropy (below 200 km).

With the exception of the appendices, the material contained in this thesis has been published (Chapters I & III), accepted (Chapter II), or submitted for publication (Chapters IV & V). Chapter VI, which highlights the important conclusions of our research is in preparation for a journal with a broad scientific readership. As the titles of our chapters vary from the published form, we list our publications below, in the same order:

Frederik J. Simons, Alet Zielhuis, and Rob D. van der Hilst,
The deep structure of the Australian continent from surface-wave tomography,
Lithos **48**, 17–43, 1999.

Frederik J. Simons, Rob D. van der Hilst, Jean-Paul Montagner, and Alet Zielhuis,
Multimode Rayleigh wave inversion for heterogeneity and azimuthal anisotropy of the Australian upper mantle,
Geoph. J. Int., 2002, in press.

Frederik J. Simons, Maria T. Zuber and Jun Korenaga,
Isostatic response of the Australian lithosphere: Estimation of effective elastic thickness and anisotropy using multitaper spectral analysis,
J. Geoph. Res. **105** (B8), 19,163–19,184, 2000.

Frederik J. Simons, Rob D. van der Hilst, and Maria T. Zuber,
Isostatic and seismic anisotropy of Australia from non-stationary gravity-topography coherence analysis and surface wave tomography,
J. Geoph. Res., 2002, under review.

Frederik J. Simons, Rob D. van der Hilst,
Age-Dependent Seismic Thickness and Mechanical Strength of the Australian Lithosphere,
Geoph. Res. Let, 2002, under review.

Frederik J. Simons and Rob D. van der Hilst,
Structure and deformation of the Australian continent,
In preparation, 2002.

Chapter 1

Isotropic Surface Wave Tomography[1]

Abstract

We present a new model of three-dimensional variations of shear wave speed in the Australian upper mantle, obtained from the dispersion of fundamental and higher-mode surface waves. We used Rayleigh wave data from the portable arrays of the SKIPPY project and from permanent stations (from AGSO, IRIS and GEOSCOPE), amounting to about 1600 source-receiver combinations. AGSO data have not been used before and provide better data coverage of the Archean cratons in western Australia. Compared to previous studies we also improved the vertical parameterization and the weighting scheme that accounts for variations in data quality and we reduced the influence of epicenter mislocation on velocity structure. The dense sampling by seismic waves provides for unprecedented resolution of continental structure, but the wave speed beneath westernmost Australia is not well constrained owing to insufficient station coverage. Global compilations of geological and seismological data (using regionalizations based on tectonic behavior or crustal age) suggest a correlation between crustal age and the thickness and composition of the continental lithosphere. However, the age and the tectonic history of crustal elements vary on wavelengths much smaller than have been resolved with global seismological studies. Using our detailed regional upper mantle model we investigate how the seismic signature of tectonic units changes with increasing depth. At large wavelengths, and to a depth of about 200 km, the inferred velocity anomalies corroborate the global pattern and display a progression of wave speed with crustal age: slow wave propagation prevails beneath the

---

[1] Published as: **The Deep Structure of the Australian Continent from Surface Wave Tomography**, *Lithos* (**48**), 17–43, 1999.

Paleozoic fold belts in eastern Australia and wave speeds increase westward across the Proterozoic and reach a maximum in the Archean cratons. The high wave speeds that we associate with Precambrian shields extend beyond the so-called Tasman Line, which marks the eastern limit of Proterozoic outcrop. This suggests that parts of the Paleozoic fold belts are underlain by Proterozoic lithosphere. We also infer that the North Australia craton extends off-shore into southwestern Papua New Guinea and beneath the Indian Ocean. For depths in excess of 200 km a regionalization with smaller units reveals a more complex pattern. Some tectonic subregions of Proterozoic age are marked by pronounced velocity highs to depths exceeding 300 km, but others do not and, surprisingly, the Archean units do not seem to be marked by such a thick high wave speed structure either. The Precambrian cratons that lack a thick high wave speed "keel" are located near passive margins, suggesting that convective processes associated with continental break-up may have destroyed a once present tectosphere. Our study suggests that deep lithospheric structure can vary as much within domains of similar crustal age as between units of different ages, which hampers attempts to find a unifying relationship between seismological units and crustal age domains.

1.1 Introduction

The inference that continental geotherms in the Archean were not much different from the averages seen today, despite pervasively (200–300°) hotter mantle temperatures, has been supported (but also refuted, by, e.g., *Strong and Stevens* [1974]) by both observations [*Burke and Kind*, 1978; *Boyd and Gurney*, 1986; *Boyd*, 1989] and numerical simulations [*Richter*, 1984, 1985; *Lenardic*, 1998] (see also *Jaupart and Mareschal* [1999] and *Nyblade* [1999]). Combined with experimental data on diamond stability in the mantle, the age of the continental lithospheric mantle (CLM) inferred from diamond inclusions [*Richardson et al.*, 1984] and Re-OS dating of mantle xenoliths [*Pearson et al.*, 1995; *Pearson*, 1999], and the correlation between composition of continental mantle and the age of the overlying crust [*Griffin et al.*, 1998], this suggests the presence of a thick ($\geq$175 km) Archean and Early-Proterozoic subcontinental lithosphere that has stabilized shortly after being formed and has remained coupled to the crust ever since. This "cold" lithosphere, with geothermal gradients generally less than 23°km^{-1} [*Burke and Kind*, 1978], must have been stabilized against convective disruption by its composition and its rheological properties [*Jordan*, 1988; *Richter*, 1988; *Shapiro et al.*, 1999a; *De Smet et al.*, 1999].

Cooling of the lithosphere explains the geophysical signature (e.g., gravity, bathymetry, seismic properties) of oceanic plates generally well [*Sclater et al.*, 1981], but the large CLM thickness inferred from seismic imaging cannot be explained by cooling alone. If a cold, negatively buoyant continental lithosphere is stabilized by rheology (that is, the strength of the plate) the thermally induced density contrasts would produce anomalies in the long-wavelength gravity field and geoid, which are not observed [*Kaula*, 1967; *Jordan*, 1975b, 1978; *Shapiro et al.*, 1999b]. Considering geological, geochemical, and geophysical evidence, Jordan [1975b,1978] postulated a mechanism for producing a petrologically distinct chemical boundary layer under the cratonic parts of the continents. In his isopycnic model the negative buoyancy of cold lithosphere is compensated by positive compositional buoyancy. Neutrally buoyant and probably more viscous than the surrounding mantle, this "keel" of depleted, refractory continental mantle, or tectosphere, is thought to remain coupled to the crust for billions of years and thus plays a crucial role in regulating continental development and stability [*Jordan*, 1981a, 1988]. The presence of such "keels" has been corroborated by seismological studies that reveal a (300–500 km) thick, anomalously fast lithosphere under some of the oldest parts of the continents [*Sipkin and Jordan*, 1975; *Jordan*, 1975a; *Sipkin and Jordan*, 1976; *Woodhouse and Dziewonski*, 1984; *Su et al.*, 1994; *Masters et al.*, 1996; *Ekström et al.*, 1997; *van der Lee and Nolet*, 1997b]. To first order this relationship seems to hold globally [*Okal*, 1977; *Jordan*, 1981b; *Polet and Anderson*, 1995], but, as we show in this study, significant departures may occur on smaller length scales.

The lateral variation in upper mantle structure, such as the differences between oceanic and continental regions and between stable or tectonically active domains, can be studied with a variety of seismic imaging techniques. The type of data used depend on the specific research objectives. *Bostock* [1999] reviews techniques used for high-resolution imaging of the crust and lithospheric mantle with phase conversions and scattering of seismic body waves. Here we consider variations in elastic properties on a larger scale than considered by Bostock; these can be constrained by a variety of methods and data, for example dispersion curves of fundamental-mode surface waves [*Toksöz and Anderson*, 1966; *Priestley*, 1999], the dispersion of singlets in the eigenfrequency multiplets of fundamental spheroidal and toroidal modes [*Dahlen*, 1976], vertical delay times of (multiply) reflected body wave phases [*Sipkin and Jordan*, 1975, 1976], and higher mode phase and group velocities [*Cara*, 1979]. With the advent of different methods of surface-wave tomography [*Woodhouse and Dziewonski*, 1984; *Nolet*, 1990; *Trampert and Woodhouse*, 1995, 1996; *Ekström et al.*, 1997] direct images have been made of the aspherical structure of the upper mantle. Surface waves, propagating horizontally, provide better radial resolution than vertically incident body waves, especially when higher modes are included in the analysis. Body waves potentially provide better lateral resolution of lithospheric structure. In general, global data sets resolve structure on the scale of sub-continental age provinces. For example, the two-fold division of Australia into a Phanerozoic and Precambrian part is resolved in the global models [*Ekström et al.*, 1997], but these do not constrain variations in structure over distances less than about 1000 km [*Laske and Masters*, 1996]. Regional studies can provide better horizontal resolution, in particular if additional data from regional seismometer networks are available.

Australia is well suited to investigate deep continental structure. The makeup of the continent is extremely varied (Fig. 1-1). Based on outcrop, continental Australia can be divided in a western, central, and eastern domain (see Fig. 1-1c) but smaller-scale units have been identified as well (see Figs. 1-1b, 1-1a). The western third of the continent comprises granites and greenstones of the Pilbara and Yilgarn blocks, which formed in the Archean (3,500 and 3,100 Ma, respectively) and have been stable since at least 2,300 Ma [*Plumb*,

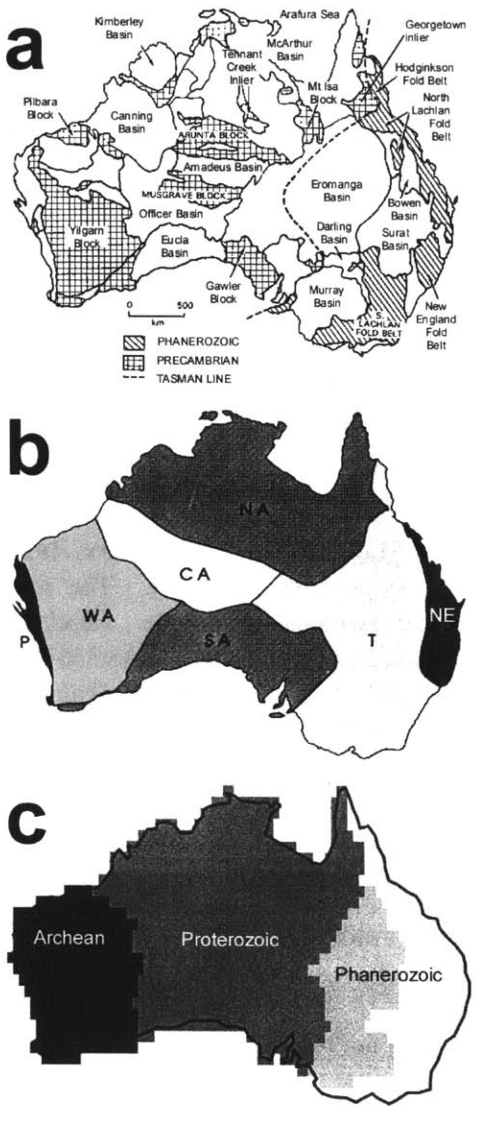


Figure 1-1: The geology of Australia at different scales and with different definitions of the units. (*a*) Most detailed representation considered in this chapter. The dashed line is the Tasman Line, which divides the Precambrian western Australia from Phanerozoic eastern Australia (from *Zuber et al.* [1989]). (*b*) Australian crustal elements, representing continent-scale groups of geophysical domains (from *Wellman* [1998]). CA, Central Australia; NA, Northern Australia; P: Pinjarra; WA, Western Australia; SA, South Australia; T, Tasman; NE, New England. (*c*) A coarse, fourfold regionalization of the Australian continent based on crustal age.The crustal age decreases from Archean in the westernmost part, to predominantly Proterozoic in the central part, and Phanerozoic to the right of the Tasman line.

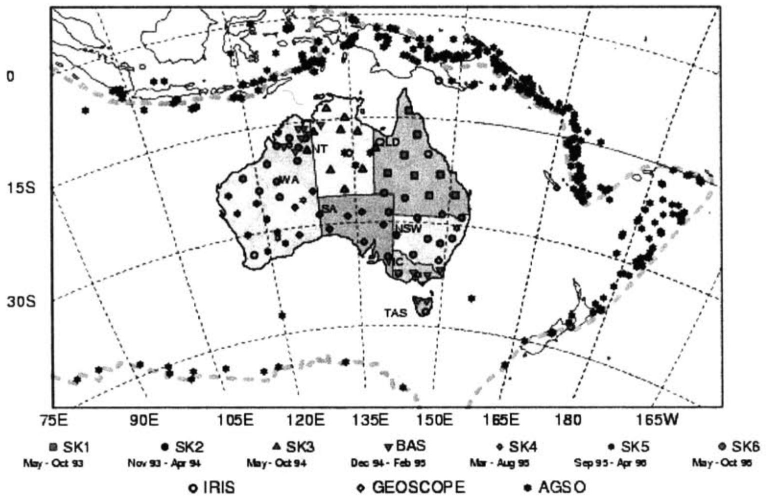


Figure 1-2: Locations of the SKIPPY, IRIS, GEOSCOPE, and AGSO stations used in this study. The solid stars depict epicenters of all earthquakes (source: *Engdahl et al.* [1998]) for which data were used in this study (N=336). (Abbreviations used: *QLD*: Queensland, *NSW*: New South Wales, *VIC*= Victoria, *SA*= South Australia,*TAS*= Tasmania, *NT*= Northern Territory, *WA*= West Australia.) Gray dashed lines depict plate boundaries.

1979]. There is little or no Phanerozoic cover of these cratonic units. Central Australia consists of a series of intracratonic Late-Proterozoic–Early-Paleozoic basins separated by fault-bounded blocks exposing Mid-Proterozoic basement rocks [*Lambeck*, 1983]. The Tasman Line — first drawn by Hill in 1951 [*Veevers*, 1984] but different interpretations exist — separates this Precambrian outcrop from exposed Phanerozoic formations in the east. Its definition is based largely on surface geology and, in regions of limited exposure (such as across the Eromanga basin), on lineations in gravity and magnetic anomaly maps [*Murray et al.*, 1989; *Wellman*, 1998]. Moreover, the continent is favorably located with respect to zones of active seismicity (Fig. 1-2), which provide ample sources for seismic tomographic imaging, and the SKIPPY seismometry project [*van der Hilst et al.*, 1994] has provided data coverage that allows surface wave tomography with unprecedented resolution.

In this chapter we briefly describe the tomographic technique that we have used to determine lateral variations in shear wave speed in the Australian mantle and discuss the relationship between the thickness of the high wave speed lithosphere and variations in crustal age. We present our model both in terms of the inversion results and as wave speed averages over well-defined geotectonic regions. Using regionalizations at different spatial scales we show that the deep structure of the Australian continent varies significantly, not only across the large-scale tectonic units but also within domains of roughly the same age. These observations are important for our understanding of the Australian continent but also have ramifications for studies of deep continental structure and evolution on the basis of global regionalizations defined on crustal age and tectonic history.

1.2 Imaging With Seismic Waves: Seismic Tomography

Tomography is a technique for reconstructing a function (”the unknowns”, or ”the model”) from projections (”the data”) along a set of curves. This relationship is often expressed as an integral over a certain volume V, $\int_V g_i(\mathbf{r})\delta x(\mathbf{r})\, d\mathbf{r} = b_i$, with $g_i(\mathbf{r})$ the Fréchet derivative (or sensitivity kernel) describing the functional dependence of the measurements b_i on the model perturbations $\delta x(\mathbf{r})$ (or in linearized form as a system of normal equations $\mathbf{Ax} = \mathbf{b}$, with $\mathbf{A}$ the sensitivity matrix containing the containing the appropriate Fréchet derivatives). If the medium under study is the human body, then the function might be the density of organ tissue, and the data used to constrain it might be the intensity of X-rays after a beam has passed through (the principle behind many kinds of medical tomography [*Herman*, 1979]). For imaging the Earth, one uses seismic waves, which are affected by anomalous structure, resulting in a phase arrival time, amplitude, or entire waveform that differs from the one expected in a spherical reference Earth model. Such differences are then interpreted in terms of velocity and attenuation variations of seismic waves within the Earth. A major complication is that the sources, earthquakes or man-made explosions, and receivers (seismometers) are distributed very unevenly over the surface of the Earth so that some regions are constrained by many data whereas others are not sampled at all. This renders the tomographic problem underdetermined (that is, not all unknowns can be determined independently) and the solution is non-unique. Out of a large number of solutions we choose a solution by minimizing a penalty function, which typically includes regularization terms (also known as damping) and accounts for *a priori* information.

Different kinds of seismic tomography exist (see, for example, the overview by *Nolet* [1987]). The data that can be used differ in frequency content and in the way they sample Earth's interior (see Fig. 1-3). Body-wave tomography often makes use of travel-time perturbations from the reference model. With surface waves, which typically have lower frequencies than body waves, the principle is the same, but they sample the Earth in a different manner. Rather than having their sensitivity to Earth structure concentrated along a ray path (such as in Fig. 1-3b) this sensitivity is given by a frequency-dependent kernel (Fig. 1-4). Frequency is a proxy for depth; the higher the frequency of the waves the more their sensitivity shifts to shallower depths. For a given frequency, the fundamental mode surface waves (Fig. 1-4a) sample shallower structure than the higher modes (Fig. 1-4b), and inclusion of the latter thus provides for increased depth resolution.

1.3 Partitioned Waveform Inversion (PWI)

The interpretation of waveforms in terms of aspherical variations in Earth's structure is often partitioned into (1) a non-linear waveform inversion that, for each record, seeks to determine the 1D structure between source and receiver (see Sec. 1.3.1) and (2) a geometrical, linear tomographic inversion which combines the individual path constraints into a 3D model for Earth structure (see Sec. 1.3.2). Partitioned Waveform Inversion (PWI), developed by *Nolet* [1990], has previously been used to study the upper mantle beneath Europe [*Nolet*, 1990; *Zielhuis and Nolet*, 1994a, b], South Africa [*Cichowicz and Green*, 1992],

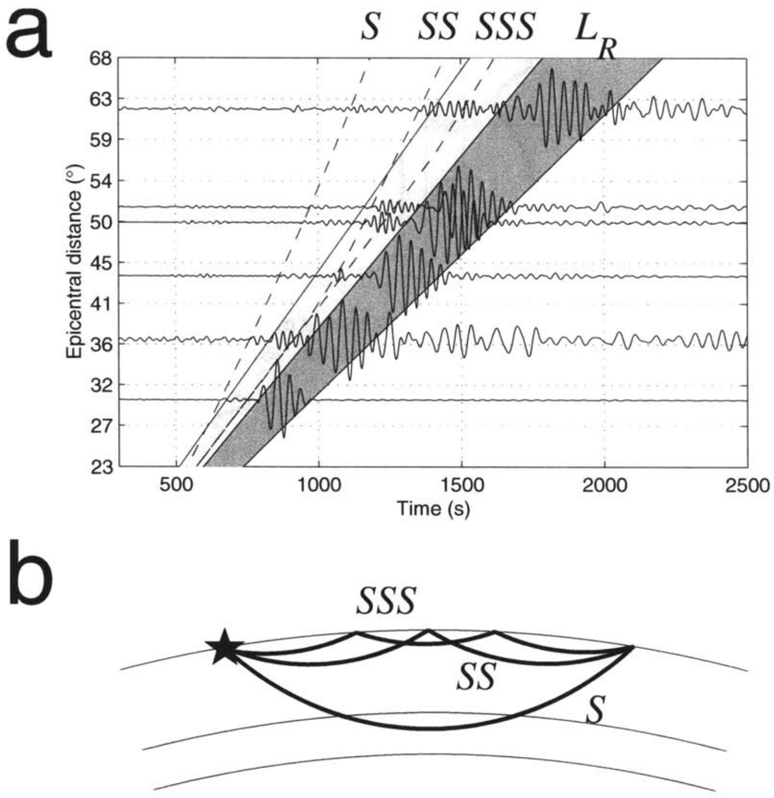


Figure 1-3: Body and surface-wave phases. (*a*) Composite record section of vertical-component seismograms arranged in order of increasing epicentral distance to two different earthquakes. The dashed lines represent the arrival times of the body-wave phases *S*, *SS* and *SSS* calculated from the *ak135* wave speed model [*Kennett et al.*, 1995]. The Rayleigh-wave surface wave L_R can be thought of as a limiting case of multiple reflections propagating across the surface. The solid lines and shaded areas indicate two group velocity windows. The first window, defined between 4.9 and 4.2 km s^{-1} selects phases sampling the upper-mantle (the “’higher-mode window’). The second window, between 4.2 and 3.4 km s^{-1} selects fundamental-mode surface waves. We note that for each record the precise windows are set manually upon visual inspection of the waveforms. Seismograms have been filtered between 10 and 45 mHz for the higher modes, and between 10 and 25 mHz for the fundamental modes. (*b*) Ray geometry in the upper mantle of the body-wave phase shown in (*a*).

North America [*van der Lee and Nolet*, 1997a, b], and Australia [*Zielhuis and van der Hilst*, 1996]. Here we review some basic aspects of the technique; for more complete descriptions we refer to *Nolet et al.* [1986], *Nolet* [1990], and *Zielhuis and Nolet* [1994a].

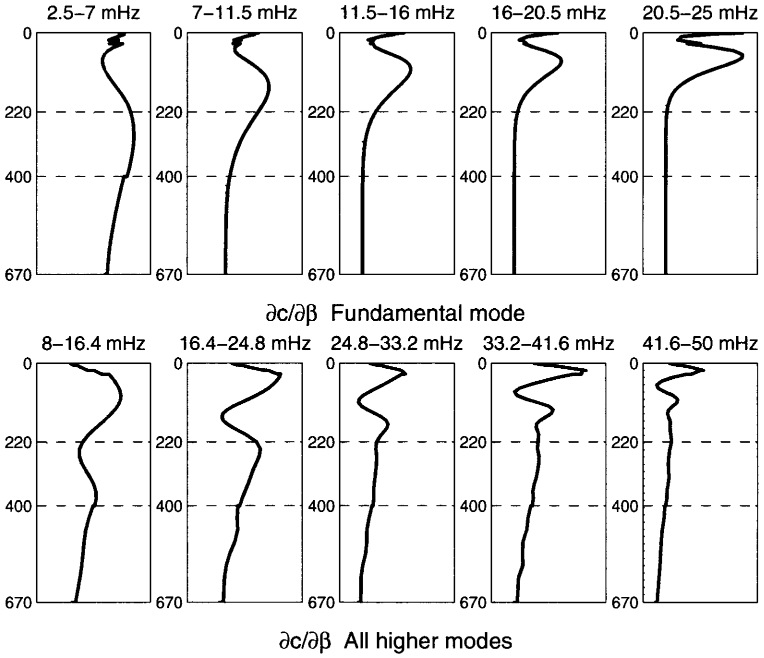


Figure 1-4: Sensitivity Fréchet kernels for surface wave propagation. The kernels represent $\partial c/\partial\beta(r)$, i.e. the sensitivity of the phase velocity c of a particular set of surface wave modes to a perturbation of the shear-wave speed β at a particular depth. (*Top*) Fundamental mode. (*Bottom*) Higher modes. Surface-wave studies done with fundamental modes are mostly sensitive to shallow structure, while most of the sensitivity at depth is due to the higher modes.

1.3.1 Waveform inversion for path-averaged structure

In the first step of PWI individual seismograms are analyzed and, within the restrictions discussed below, inverted for shear velocity variations with depth, $\delta\beta(r)$ (with r radius), averaged along the source-receiver path. This involves the matching of observed waveforms with theoretical (synthetic) seismograms, which are computed as a sum of surface-wave modes using the JWKB approximation [*Woodhouse*, 1974]. The waveform synthesis requires information about the earthquake focal mechanisms, which is obtained from the

Harvard Centroid Moment Tensor (CMT) catalog [*Dziewonski et al.*, 1981] and the National Earthquake Information Center (NEIC) [*Sipkin*, 1994].

The JWKB formalism assumes that lateral heterogeneity is sufficiently smooth compared to the wavelength of the seismic waves used, which imposes a lower limit for frequency. For our application, with perturbation scale lengths typically greater than $\sim$400 km, epicentral distances ranging from 1000 to 4000 km, and a typical phase speed of about 4.5 km s^{-1}, a lower frequency limit of not much less than 10 mHz is predicted on theoretical grounds [*Kennett*, 1995; *Wang and Dahlen*, 1995; *Dahlen and Tromp*, 1998]. In practice this criterion can be relaxed a bit and for the fundamental modes we consider a lower limit of 5 mHz (although most data are for frequencies higher than 8 mHz). The path-average approximation also implies an upper-frequency limit, both for the fundamental mode and the overtones, because cross-branch coupling between modes, as would arise from lateral heterogeneity [*Kennett*, 1984; *Li and Tanimoto*, 1993; *Marquering and Snieder*, 1995] and which brings out the ray character of the higher modes, such as *SS*, is not accounted for. The detrimental effects of ignoring mode coupling are aggravated with increasing frequency. However, *Marquering et al.* [1996] and *Marquering and Snieder* [1996] have demonstrated that if data coverage is dense the results will be little changed by the adoption of mode coupling techniques in the inversion. Moreover, the largest differences would occur outside the depth range that is also constrained by the horizontally propagating fundamental modes; therefore, we only discuss structure to a depth of 400 km. (We note that structure beneath stations or events that is not well sampled by rays from different directions, such as on the edge of the model, may exhibit spurious vertical structure resulting from the one-dimensionality of the kernels. This may be the case for the deep fast anomaly beneath the NWAO station at Narrogin, West Australia.) At high frequencies the fundamental-mode Rayleigh waves become sensitive to steep gradients in shallow structure (Fig. 1-4), such as at the transition from oceanic to continental crust, which can severely distort the waveforms. Along with low-passing the data, we minimized such effects by accounted for variations in crustal thickness within the region under study on the basis of crustal thickness information from converted phases recorded at the SKIPPY stations [*Shibutani et al.*, 1996; *Clitheroe and van der Hilst*, 1998].

Kennett and Nolet [1990] and *Kennett* [1995] conclude that with an upper frequency limit of 20 mHz for the fundamental mode and 50 mHz for the higher modes the surface-mode summation used in PWI provides a representation of the seismic wave field that is adequate for our purposes. Because the admissible frequency limits differ we use group velocity windows to isolate the fundamental and higher mode part of the records so that we can analyze them within the frequency bands discussed above (see also Fig. 1-3).

An example of waveform fits is given in Fig. 1-5. The wave trains for the fundamental and higher modes have been normalized to unit amplitude. Both sections of the seismogram are fitted separately within the applicable frequency band. The initial fits depict the difference between the observed data (thick solid lines) and the synthetic records (thin lines) produced from a reference model. For this event in Southern Sumatra the fundamental-mode Rayleigh waves arrive earlier than predicted from the average Earth for all stations considered, suggesting relatively fast wave propagation across the western and central part of the continent. But at some stations the mismatch is larger than at others; compare, for instance, the records of ARMA (Armidale, New South Wales) and CTA (or CTAO, Charters Towers, Queensland). These differences are indicative of lateral variations in wave speed (for these earthquakes the data indicate that average wave speed along the path to ARMA is faster than along the one to CTA). For the same stations, the difference between the observed and synthesized overtones are smaller than the fundamental mode, suggesting that the wave speed variations in the deep part of the model are smaller than in the shallow part. The bottom panel of Fig. 1-5 shows the excellent (final) fit of the waveforms, both for the higher frequency overtones as well as for the lower-frequency fundamental modes.

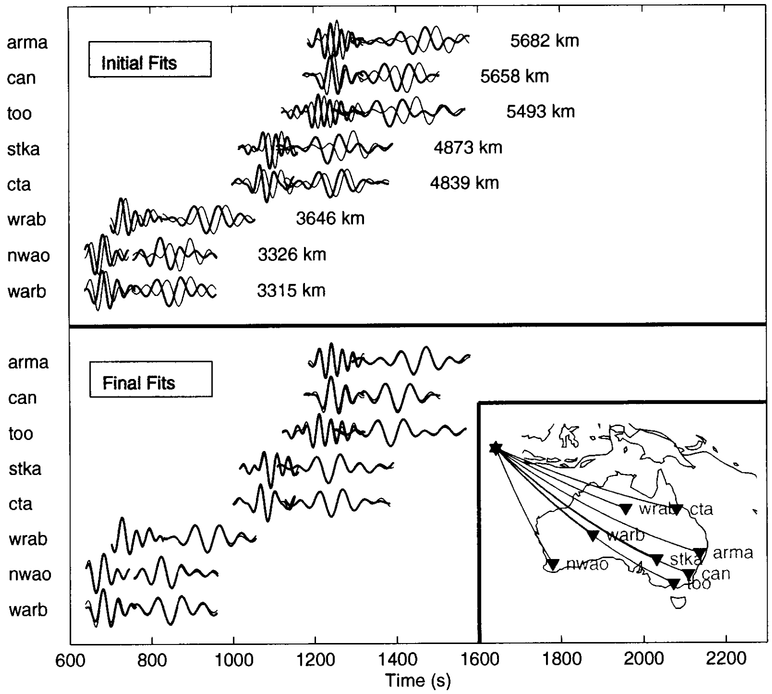


Figure 1-5: Waveform fitting. Observed data are plotted as thick solid lines, predictions are depicted with thin lines. Seismograms are for a Mb 5.8 event in Southern Sumatra, at (103.9°E, 5.7°S), located at 56 km depth. Great circle paths to the stations are plotted in the inset. "Initial fits" are predictions made by surface-wave summation, using an assumed reference models. "Final fits" are obtained by the nonlinear inversion procedure described in *Nolet et al.* [1986] (see text).

1.3.2 Tomographic inversion for 3D structure

Once the path-average of the variation of wave speed with depth ($\delta\beta(r)$) is determined for each source-receiver combination these 1D profiles are used as observations in a tomographic inversion for 3D variations in shear wave speed ($\delta\beta(r, \theta, \varphi)$, with r radius, θ colatitude, and φ longitude). This linear tomographic inversion may be performed by a variety of methods, such as the ones described by *Nolet* [1990], *Zielhuis* [1992], *Zielhuis and Nolet* [1994b], or *van der Lee and Nolet* [1997b].

The tomographic problem is parameterized by means of local basis functions (equal-area blocks in latitude and longitude direction; and box-car and triangular basis functions for radius). The system of equations is solved in a generalized least square sense, using the LSQR iterative algorithm [*Paige and Saunders*, 1982; *Nolet*, 1985]. The damping applied is a combination of a first-order gradient damping, which minimizes the differences in structure between adjacent cells, and a norm damping, which produces a bias toward the reference model used. The model presented in this chapter results from careful experimentation with all variables involved (damping, weights, cell size), and parameters such as variance reduction and retrieval of synthetic input models have been used as guidance. After 200 iterations of the LSQR inversion a variance reduction of about 90% was obtained. The reference model used for the 3D inversion has a crustal thickness of 30 km, which is a reasonable average for the region under study. For the mantle we used a modified PREM model, smoothly interpolated over the 220 km discontinuity [*Dziewonski and Anderson*, 1981; *Zielhuis and Nolet*, 1994b].

1.3.3 Differences with respect to previous studies of Australia

We have made several modifications to the methods by *Zielhuis and Nolet* [1994b], including a different form of weighting of the individual data fits, a parameterization with an increased number of basis functions in radial direction, and we added a parameter that can absorb effects of epicenter mislocation. The effects of these improvements are, however, subtle, and most of the differences with previous models [*Zielhuis and van der Hilst*, 1996; *van der Hilst et al.*, 1998] can be attributed to the use of an expanded data set.

Firstly, we assign an uncertainty to the individual fits based on a weighted combination of the signal bandwidth, the length of the group velocity windows, the χ^2–norm, and the zero-lag cross-correlation value of the synthetic and observed waveforms, as well as the ability to fit both the fundamental and higher mode data. The reciprocals of the uncertainties obtained were used to weight the data in the inversion. Secondly, like *Zielhuis and Nolet* [1994b] we used a combination of boxcar and triangular basis functions, but we have added additional node points in order to extract more information on deep structure from the large number of higher modes in our data set. Thirdly, we have used the hypocenter locations from the global data file by *Engdahl et al.* [1998], but in this region source mislocations can be substantial owing to sparse station coverage of the southern hemisphere. A mislocation of 20 km, on an epicentral distance of 2000 km can cause a spurious wave speed anomaly of 1%. Without attempting a formal earthquake relocation we aimed to absorb such effects of source mislocation in a *denuisancing* parameter in the linear inversion for 3D structure. This has an effect similar to damping, and varying the degree of source relocation allows us to assess the structural features in our model that are required by the waveform data only.

1.4 Data used in this study

Between May 1993 and October 1996 the Australian National University operated the SKIPPY seismometry project [*van der Hilst et al.*, 1994]. This project involved 6 arrays of up to 12 portable broad-band seismometers that together synthesized a nationwide array (Fig. 1-2) and was intended to exploit Australia's location with respect to regional seismicity and to provide dense data coverage for a range of tomographic imaging techniques. The individual arrays were deployed for about 6 months at a time. Parts of the large set of surface wave data were used in previous studies [*Zielhuis and van der Hilst*, 1996; *van der Hilst et al.*, 1998], which focused on central and eastern Australia because data coverage in the west was not satisfactory at the time. In addition to the data from the SKIPPY experiment we have used data from broadband permanent stations from the IRIS (Incorporated Research Institutions for Seismology), GEOSCOPE, and AGSO (Australian Geological Survey Organization). The ~1600 vertical-component seismograms from about 340 seismic

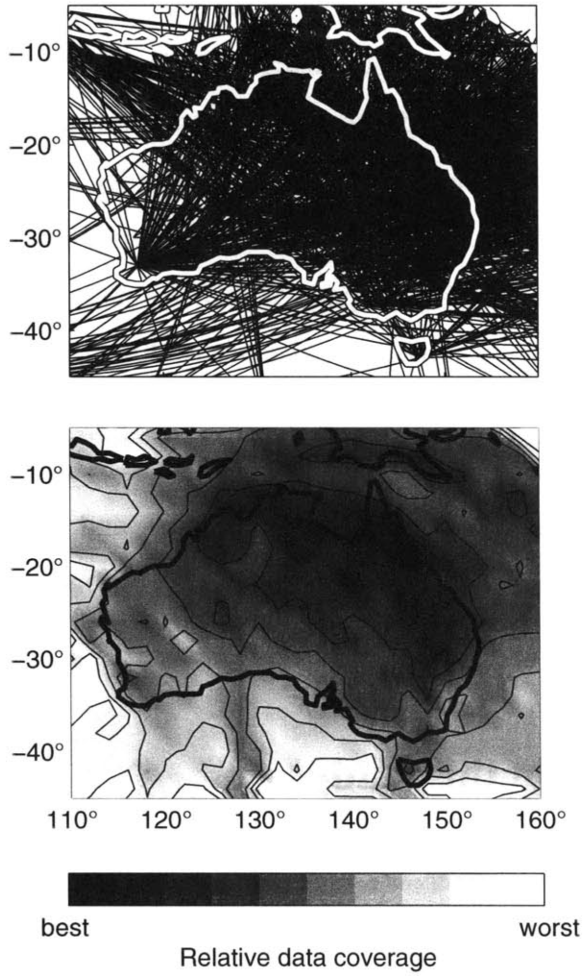


Figure 1-6: (*Left*) Great-circle paths of the 1596 event-receiver combinations used in this study. (*Right*) Path coverage. Path lengths and variance of the directions of the rays crossing in 2deg × 2deg cells, two indicators of tomographic quality, are expressed on a relative scale (1 indicates well sampled; 0 indicates no sampling).

events provide excellent data coverage (Fig. 1-6). Further improvements in the resolution, in particular of the western part of the continent, are still expected, given the continued monitoring of earthquake activity by AGSO.

We used the portion of the vertical-component seismogram from the arrival of direct *S* up to, and including, the arrival of the fundamental mode of the Rayleigh wave. This time window includes multiple body wave reflections at the free surface, such as *SS* and *SSS* (Fig. 1-3). The group velocity windows used for the isolation of the fundamental and higher modes are approximately 3.4–4.2 and 4.0–5.0 km s^{-1}, respectively; for each individual record the exact limits are set after visual inspection. With this selection, the fundamental mode is truncated before the scattered and multipathed surface waves arrive, and body waves with turning points in the deep mantle are excluded. For example, direct *S* is the first body wave phase considered at a distance of 2000 km, but at 4000 km this is *SS*, and at 6000 km *SSS*, etc. (Fig. 1-3). In recognition of the frequency limitations imposed by the approximations implicit in the method (see above), fundamental mode data were fit in the 5–25 mHz range and the higher modes were modeled within 8–50 mHz. About 40% of all seismograms contained higher-mode windows for which good fits could be obtained.

1.5 Results

Here we discuss briefly the spatial resolution and the general aspects of our model, compare our results to variations in shear wave speed as inferred from a global inversion, and describe in detail how the shear wave speed varies within several well-defined tectonic units that constitute the Australian continent.

1.5.1 Spatial resolution

Image quality does not only depend on the sheer number of data but also on the azimuthal distribution of the (crossing) paths [*Aki and Richards*, 1980; *Menke*, 1989; *Lay and Wallace*, 1995]. For each cell, we added the variance of the directions (between 0 and π) of the rays (normalized from 0 to 1) to the normalized sum of the path lengths to provide a better measure of the quality of data coverage (Fig. 1-6). The data coverage is good throughout the Australian continent but degrades towards the southwest. Since all AGSO data have not yet been used we expect further improvements in this part of the continent.

In order to assess the reliability of the images we have performed test inversions with synthetic data calculated from different input models; the ability to reconstruct an input model from the synthetic data is then used to assess how well real structure can be constrained by the available data. We have used different synthetic input models, with harmonic wave speed variations as well as spike tests (for a discussion of such resolution tests, see, e.g., *Spakman and Nolet* [1988] or *Humphreys and Clayton* [1988]), and others in which we tested the robustness of a specific structure in the model. Based on these tests and on theoretical considerations [*Zielhuis*, 1992; *Zielhuis and van der Hilst*, 1996; *Kennett and Nolet*, 1990] we conclude that the horizontal resolution in the best resolved parts of the continent approaches 250 km, and the vertical resolution ranges from 50 km (at 100 km depth) to between 100 and 150 km (around 300 km depth). As expected from the data coverage (Fig. 1-6), the resolution in the central and east is still superior to that in western Australia.

For the purpose of this chapter we evaluated whether the data can constrain the wave speed variations on the length scales considered in the finest regionalization described below (see Sec. 1.5.4). In the experiment, different wave speeds were assigned to various tectonic regions of the Australian continent. Fig. 1-7 displays the results of two such tests. In the first, the input anomalies of Fig. 1-7c were put at 80 km depth, with zero perturbations elsewhere in the model. We calculated synthetic data for all 1600 paths (both for fundamental modes and the overtones) and repeated the linearized inversion for 3D structure. The result for that layer is given in Fig. 1-7a. In the second experiment the input pattern (Fig. 1-7c) was placed at 210 km depth, with the response shown in Fig. 1-7b. These tests show that the wave speed variations are well resolved on the length scales we

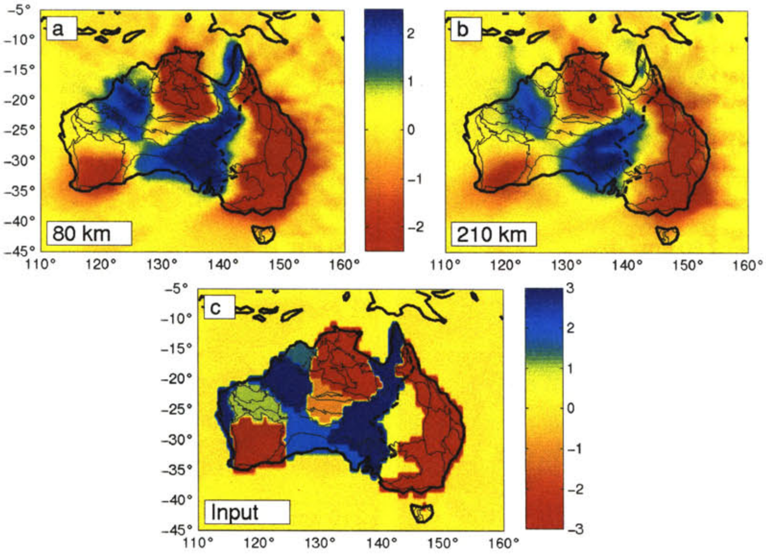


Figure 1-7: Results of resolution experiment. Input models were constructed by assigning constant wave speed anomalies to different geotectonic regions and at different depths. Anomalies are in percent from a spherical reference model. Synthetic data were generated on the basis of all wave paths used in actual inversion. (*a*) Recovery of anomaly placed at 80 km. (*b*) Recovery at 210 km. (*c*) Input anomaly used in both tests. Note the difference between both color schemes. Thick dashed lines give approximate location of Tasman Line.

are interested in, but at shallow depth the image quality is better than at larger depth. In general wave speed contrasts are well resolved, but the amplitude of the wave speed variations is less well determined. We note that in the regionalized presentation of our results (see Sec. 1.5.4) the wave speeds were averaged over the regions identified, but in the test inversions shown here such averaging was not done.

1.5.2 Shear wave speed variations in the Australian upper mantle

Figs. 1-8 and 1-9 display some of the tomographic results. The shear wave speed anomalies $(\beta(r,\theta,\varphi) - \beta_0(r))$ are plotted as percentages of the reference model $(\beta_0(r))$ which is a modified version of PREM (without the discontinuity at 220 km). Wave propagation is

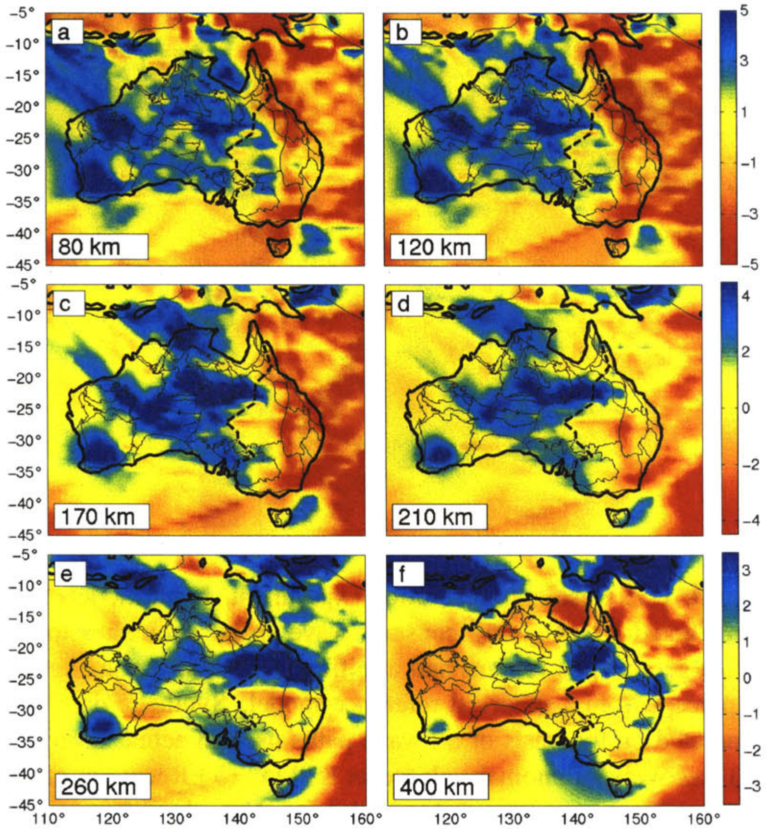


Figure 1-8: Depth slices through our preferred velocity model. Anomalies are in percentage from a spherical reference model (see text). The reference velocities used are 4500 m s^{-1} at 80, 120 and 170 km depth (*a-c*); 4513.5 m s^{-1} at 210 km (*d*); 4581 m s^{-1} at 260 km (*e*) and 4851.5 m s^{-1} at 400 km depth (*e*). The color scales used are for *a* and *b*, *c* and *d* and *e* and *f*, with diminishing values of saturation which reflects the decreasing magnitude of the anomalies with depth. Thick dashed lines give approximate location of Tasman Line.

slow beneath the Phanerozoic eastern part of Australia and fast beneath the Proterozoic and Archean domains, at least to a depth of 200 km (e.g., Figs. 1-8a-c and 1-9). The Phanerozoic high wave speed lithosphere is relatively thin, generally less than 80 km (note the cross-sections have been cropped to show the depth range between 15 and 450 km), and overlies a pronounced low-velocity zone that extends to approximately 200 km depth

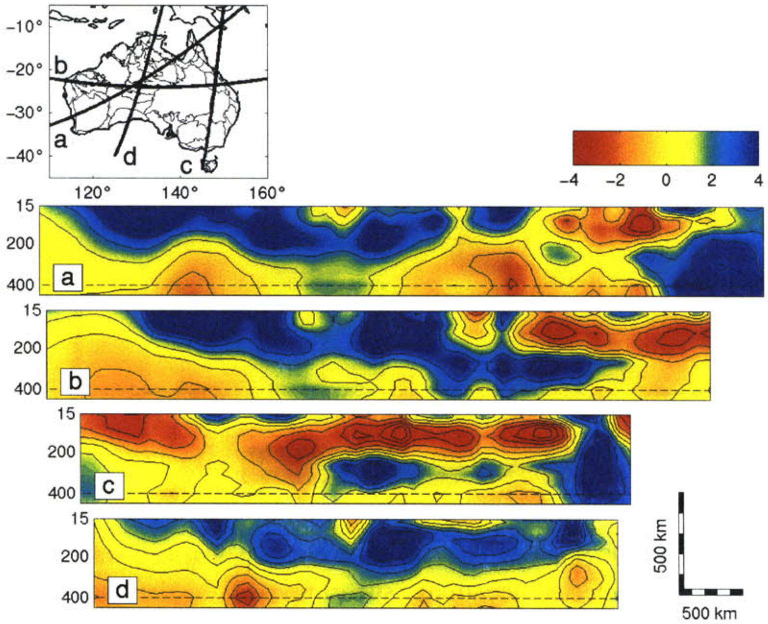


Figure 1-9: Profiles through the model. Anomalies in percentage from a spherical reference model (see text). The high wave speeds in the northeastern corner of the map (inset), which show at the right hand side of cross-sections (*a*) and (*c*), reflect the recent subduction of the Pacific beneath the Indo-Australian plate. The high wave speed feature to the southwest of Tasmania, (*c*), may be related to the Australia–Antarctic discordance [*Gurnis et al.*, 1998] but this part of our model is not well sampled (see Fig. 1-6). The scale bar in (*d*) represents a 10° interval at Earth's surface.

(Fig. 1-9) (see also *Goncz et al.* [1975], *Goncz and Cleary* [1976], and *Zielhuis and van der Hilst* [1996]).

In agreement with *Zielhuis and van der Hilst* [1996] our model reveals that at depths shallower than 150 km the wave speed gradient from the eastern to the central domains occurs east of where it would be expected on the basis of the exposure at the surface of Proterozoic rock (limited by the Tasman Line). The presence of high wave speed lithosphere east of the Tasman line was confirmed by the analysis of fundamental mode dispersion between stations along the same great-circle path [*Passier et al.*, 1997] (Fig. 1-10). In several regions, the lateral wave speed contrast coincides with surface outlines of sedimentary

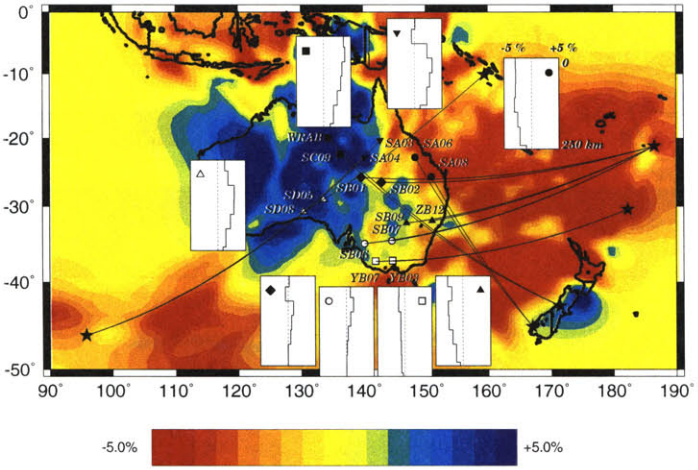


Figure 1-10: Stations, events, great circle paths, and wave speed profiles at selected locations, superimposed on the shear wave speed variations at a depth of 140 km depth [*Zielhuis and van der Hilst*, 1996]. The panels display wave speed variation to a depth of 250 km in eastern and central Australia as inferred from the differential dispersion of fundamental model Rayleigh waves between stations along the same great circle path [*Passier et al.*, 1997]. In all velocity panels, perturbations range from -5% to 5% and the depth ranges from 0 to 250 km. (Modified after *Passier et al.* [1997]). Note that our current models are based on more data than were available to *Zielhuis and van der Hilst* [1996].

basins in easternmost Australia, in particular the western margin of the Bowen and Surat basins. At greater depths the wave speed divide shifts westward, and at 170 and 210 km depth it appears to parallel the Tasman line and the western margin of the Eromanga basin. In northeastern Australia, the wave speed contrast is located to the west of the Georgetown Inlier, and this Proterozoic unit is part of a rather thin high wave speed lid with a pronounced low velocity zone underneath (see Fig. 1-9). In contrast, the Proterozoic shields of central Australia are delineated by high wave speeds down to at least 250 km. The area of the Late-Paleozoic Alice Springs orogeny (Amadeus basin and Musgrave and Arunta blocks) stands out from the adjacent shields by lower wave speeds. This may reflect the thick layer of sediments in this region. Interestingly, also the Kimberley block, which is often interpreted as the westward continuation of the North Australia craton (see, e.g., *Shaw et al.* [1995] and Fig. 1-1b) has a seismic signature that differs significantly from that of the Proterozoic craton. Likewise, the deep structure beneath the Canning basin differs from that beneath central Australia and from the Archean cratons further South (Fig. 1-8b-d).

Below 200 km depth the central cratons continue to be marked by fast anomalies, but the Archean Yilgarn, Pilbara, and Gawler cratons do not seem to be marked by wave speeds that are significantly higher than the reference model, with the exception of the localized high wave speeds under station NWAO in the southern Yilgarn craton, which might be an artifact of using one-dimensional sensitivity kernels (see Sec. 1.3.1). With the exception of the region encompassing the Alice Springs orogeny, the mantle beneath 300 km depth beneath the Proterozoic and Archean appears to be rather homogeneous. In this depth range, the most pronounced fast anomalies are associated with the subduction zones to the North and Northeast of Australia and with the intriguing structure east of the Mt. Isa block (at 23°S). Resolution tests and inversions with subsets of the data removed demonstrate that the latter anomaly is well resolved and required by the data used.

1.5.3 Global vs. regional models

Our regional model reveals variations in shear wave speed on smaller scales and with larger amplitude than indicated by global tomography models. To better understand such differences we compared our maps to the global images both in the spatial and spectral domains. Figure Fig. 1-11a displays the wave speed variations according to the global model by *Ekström et al.* [1997], which contains information from about 1700, globally distributed earthquakes, and only a small number of permanent stations on or near the Australian continent. We obtain a similar result if we apply our inversion technique to fundamental mode data from the 6 regional IRIS and GEOSCOPE stations that were also used by *Ekström*

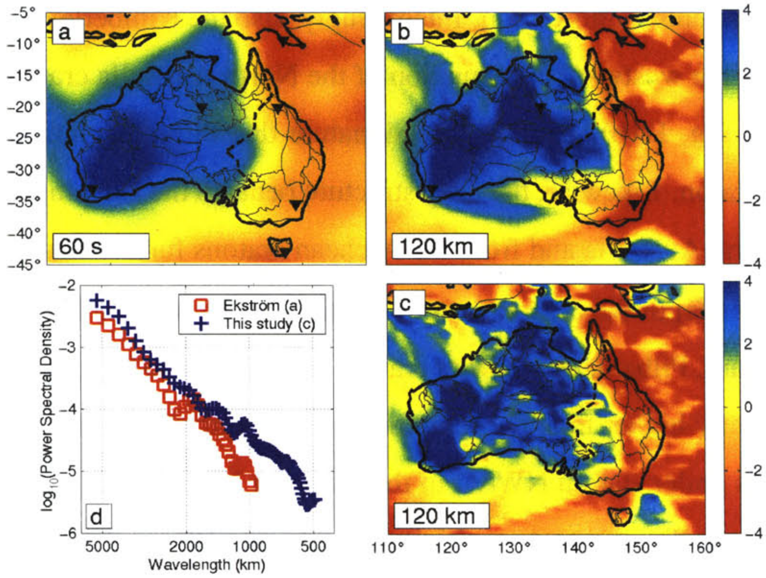


Figure 1-11: Comparison with global model in spatial domain. (*a*) Rayleigh-wave inversion result by *Ekström et al.* [1997], at 60 s. This phase-velocity map records the anomalies of fundamental modes of some (narrow) frequency band around 60 s, or 17 mHz. From inspection of the kernels of Fig. 1-4, such map represents structural wave speed deviations with a maximum sensitivity around 120 km depth. We therefore compare it to our inversion results at those depths. (*b*) Wave speed anomalies at 120 km, using only fundamental modes from permanent seismometer stations (from IRIS and GEOSCOPE). The resolution of the result is roughly equivalent to the one in (*a*), with certain differences attributable to differences in path geometry. In (*a*) and (*b*), the location of the 6 permanent stations is indicated by the black triangles. (*c*) Our preferred model at 120 km depth. Thick dashed lines give approximate location of Tasman Line. (*d*) Power spectrum of the shear wave speed model at 80 km depth. The results are compared with the model by *Ekström et al.* [1997].

et al. [1997] (Fig. 1-11b), although we retrieve somewhat lower wave speeds beneath the marginal basins east of Australia. Following *Chevrot et al.* [1998] we also compared the power spectrum of our model with that from *Ekström et al.* [1997], which was parameterized in spherical harmonics up to degree 40, corresponding to a wavelength of about 1000 km. Across the entire spectrum our regional model does have somewhat higher amplitudes than the global model by [*Ekström et al.*, 1997], but particularly so at wavelengths smaller than 2000 km (Fig. 1-11d). This is primarily a consequence of degrading resolution at those wavelengths in the global model (see Fig. 13 in *Ekström et al.* [1997]). From this comparison we conclude that when evaluated at similar wavelengths the spatial and spectral characteristics of global and our regional models are generally in good agreement with each other, and the superior resolution at high wave numbers (comparing, e.g., Figs. 1-11a and c) is not dictated by the parameterization but results from the dense data coverage provided by the SKIPPY project and the inclusion of higher-mode Rayleigh waves.

1.5.4 Seismic signature of distinct tectonic regions

"Tectonic" regionalizations, whereby seismic properties are averaged over well-defined geological domains, have formed the basis of first-order testing of the tectosphere hypothesis [*Jordan*, 1975a; *Okal*, 1977; *Jordan*, 1981b; *Polet and Anderson*, 1995], the synthesis of global seismic data [*Nolet et al.*, 1994], and the calculation of travel-time corrections [*Guðmundsson and Sambridge*, 1998]. We have used regionalizations to investigate the relationship between the thickness of the high velocity lid and crustal age at different length scales.

In our first experiment we used a three-fold tectonic regionalization based roughly on crustal age (Fig. 1-1c), with a further subdivision of the eastern domain in two parts motivated by wave speed maps (see *Zielhuis and van der Hilst* [1996] and Fig. 1-8). We averaged the wave speed perturbations over these domains. Fig. 1-12a-b shows the rather dramatic westward progression of seismic wave speed, which correlates well with the general increase in crustal age (Fig. 1-1c). However, deeper in the upper mantle this relationship breaks down and beneath 200 km depth the central Proterozoic is characterized by fast wave propagation while the average wave speeds in the Archean units are comparable to those under the Phanerozoic at large (Fig. 1-12c-d). Radial wave speed profiles, calculated for each domain, further illustrate that the recently tectonized eastern seaboard of Australia is generally slow, with a pronounced low-velocity zone centered around 150 km (Fig. 1-13). The Precambrian central and western parts do not show this reduction in wave speed;

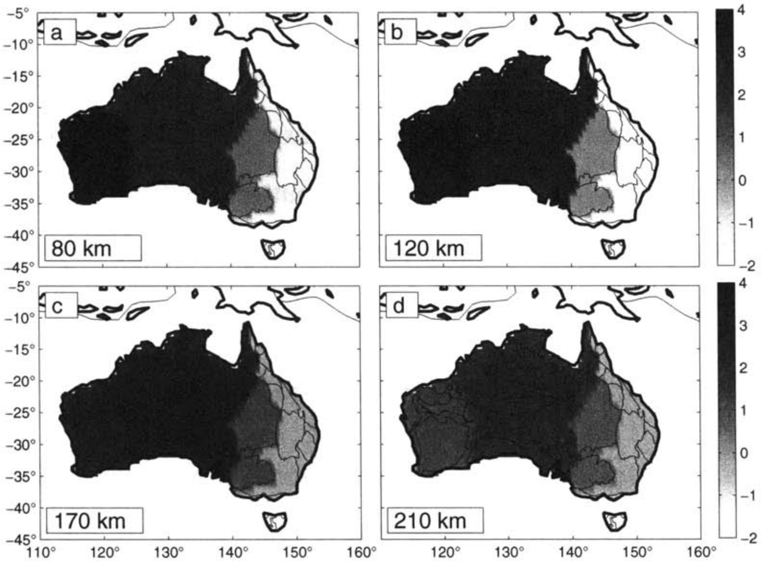


Figure 1-12: Four-part regionalized representation of the shear wave speed model. This regionalization is based on large-scale variations in crustal age (see Fig. 1-1c).

instead, high-velocity lithosphere beneath these provinces is generally 200–250 km thick. Note that for z<150 km the Archean is faster than the Proterozoic but that the high wave speed signature does not extend to as large a depth (Fig. 1-13). Our results are in general agreement with the results by *Gaherty and Jordan* [1995] and *Gaherty et al.* [1996] who argued for a difference of similar magnitude between fast continental and slow oceanic mantle to a depth exceeding 250 km.

It is obvious from geological maps and results of tomography that significant variations in crustal structure and seismic properties (and our ability to constrain them) occur on much smaller length scales than represented by the large scale regions discussed above. Therefore we designed another regionalization based on geological units that better represent variations in tectonic age. This tectonic subdivision (Fig. 1-14) is based on the crustal elements map by *Shaw et al.* [1995] and refinements by *Wellman* [1998]; it consists of a northern, central, western, and southern Australian domain, as well as the Tasman group and the

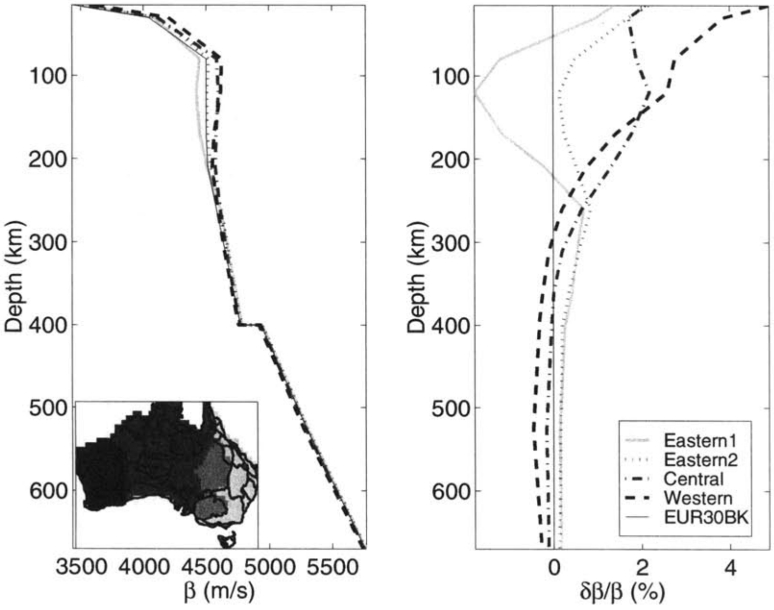


Figure 1-13: Age-dependent wave speed variations with depth. (*Left*) In absolute wave speeds. (*Right*) As wave speed perturbations from a continental reference model. (*Inset*) Geographic location of the regions used (see Fig. 1-1). Compare with Fig. 1-12 for a plan view. Below ~250 km, the differences of average wave speed between those regions have become an order of magnitude smaller than in the shallow mantle.

New England fold belt (see Fig. 1-1b), and the Pinjarra (the southwestern-most part). On the basis of our wave speed maps (Fig. 1-8) we treat the Kimberley block separately from the northern Australian domain, the Canning basin apart from Central Australia, and the Pilbara apart from the Yilgarn craton. In Fig. 1-14a-b we maintain our earlier division of the Tasman domain into a unit comprising the (western parts of the) Eromanga and Murray basins and a unit with the eastern Lachland and New England fold belts (see Fig. 1-1a). In Fig. 1-14c-f we have subdivided the Eastern and Western Tasman domain in a northern and southern part as to prevent the anomalous fast structure centered around (145°E,23°S) from biasing a larger region.

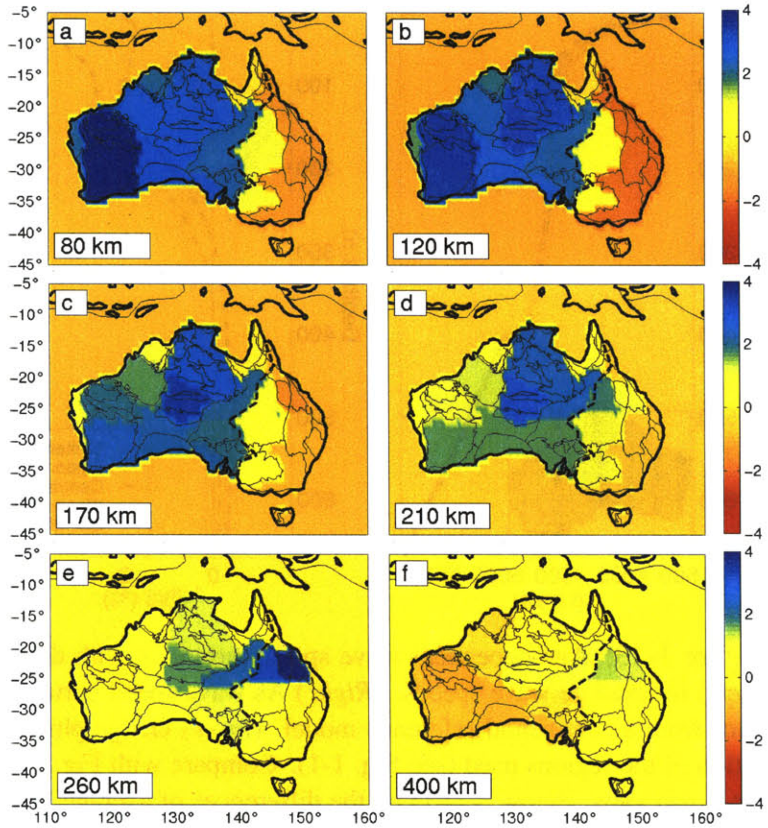


Figure 1-14: Detailed regionalized representation of the shear wave speed model. In (*a*) and (*c*), a 12-part subdivision was chosen, whereas in (*c*)-(*f*), the domain centered over the Eromanga basin was subdivided to bring out the anomalous character of the fast structure centered at 23°S. For more discussion, see text. Thick dashed lines give approximate location of Tasman Line.

This regionalization (Fig. 1-14) largely confirms the westward progression of seismic signature at depths less than 200 km but also reveals significant departures of the general trends. The high wave speed in the Archean western Australian domain can largely be attributed to the Yilgarn block; the Pilbara block is not marked by higher-than-average wave speeds but resolution in this particular area may be lacking. Recall that the high wave speeds beneath the station NWAO may be an artifact of using 1D kernels, and the signature of the Yilgarn region may thus be biased. The high wave speed under central Australia is mostly associated with the central block-and-basin structure and the North Australia craton. Beneath the Kimberley block and the Canning basin the pronounced positive wave speed anomalies vanish beneath 150 km depth, and at a slightly larger depth also the Archean Yilgarn, Pilbara, and Gawler units lose their deep seismic expression. However, the Proterozoic North Australia craton is still faster than average at depths in excess of 250 km. Surprisingly, the wave speed anomaly beneath the region encompassed by the Alice Springs orogeny seems to persist to depths in excess of 300 km. The upper mantle and transition zone beneath Eastern Australia seems more complex than its Precambrian counterpart (see also Fig. 1-9c and Sec. 1.5.2) and on average the transition beneath geological domains in eastern Australia are faster than in the west (Fig. 1-9b,c).

1.6 Discussion

The tectosphere concept as formulated by *Jordan* [1988] implies that the negative thermal buoyancy of the CLM is largely offset by changes in composition. Here we do not image composition and, therefore, use a less restricted definition. High wave speed lids that are more than ~200 km thick would be dynamically unstable if the negative buoyancy owing to the low temperature is not somehow compensated by changes in composition [*Jordan*, 1975a, b; *Anderson and Bass*, 1984; *Polet and Anderson*, 1995; *De Smet et al.*, 1999; *Shapiro et al.*, 1999a]. In the following we loosely refer to “continental keels” or ”tectosphere” when the thickness of the high wave speed lid as inferred from our results exceeds this value.

1.6.1 Regional deviations from a global pattern

On the basis of global wave speed models, *Jordan* [1981b], *Nolet et al.* [1994], *Polet and Anderson* [1995], and others, concluded that there is a good agreement between the velocity anomalies and the surface pattern of crustal formation ages. The Precambrian domains are fast, whereas most Phanerozoic regions are underlain by a seismically slow upper mantle. Such observations have often been used to corroborate the correlation between crustal age and the composition of the CLM inferred from geological data. Global wave speed models have suggested that the high wave speeds persist to up to 500 km depth [*Su et al.*, 1994; *Masters et al.*, 1996] but the depth extent of the high-velocity anomalies under stable continental regions has not been well constrained.

The fundamental and higher model data used in our study are not consistent with these earlier results and, instead, constrain the thickness of the high wave speed CLM to less than 250–300 km. For Australia a strong correlation between surface age and seismic structure does exists, but only at long wavelengths ($\lambda > 1500$ km) and for depths shallower than $\sim$200 km. At smaller length scales there are parts of Precambrian shields that are marked by high wave speeds to even larger depth but there are also many such regions without a deep seismic expression. At the length scales considered in our study, the "tectosphere" beneath the Proterozoic domains is as thick as or thicker than beneath Archean units, but the wave speed in the latter appears to be higher (Fig. 1-13). However, some Proterozoic domains, for instance the Georgetown Inlier in northern Queensland, are not marked by a deep, seismologically fast keel. On the other hand, some regions devoid of Proterozoic or older outcrop appear to have a tectosphere-like CLM. Significant variations in seismic properties thus occur beneath geological provinces of comparable crustal age or tectonic setting and similar complexity may exist for other continents as well. Hence, for Australia there does not seem to be a simple relationship between tectosphere structure and crustal age.

1.6.2 Eastern Australia

Beneath easternmost Australia the high wave speed lid is thinner than 80 km in most places, entirely absent in some (for instance beneath the Queensland Volcanic Province), and mostly underlain by a pronounced low velocity zone at around 150 km depth (see *Zielhuis and van der Hilst* [1996] and references therein). See also Fig. 1-9. Tectonically this area is relatively young, and in the southeast volcanism may have ceased only a few thousand years ago [*Johnson*, 1989]. The geothermal gradient is steep, the present-day surface heat-flow is high (70–100 $\mathrm{mWm^{-2}}$) [*Cull and Denham*, 1979], and studies of mantle conductivity [*Lilley et al.*, 1981; *Finlayson*, 1982] suggest that volatile-rich material or partial melt may be present. Therefore, the pervasive lower-than-average wave speeds in this region are most likely the result of thermal processes associated with the Late-Cretaceous ($\sim$80 Ma) opening of the Tasman Sea and accompanying volcanism [*Storey*, 1995], volatile infusion during subduction prior to that event, and the recent volcanism in the region.

To the west of the low wave speed region, but still east of the Tasman line, the wave speed signature suggests the presence of a seismically fast root. Although it is not as thick as the CLM beneath the Proterozoic proper, this could imply that some of the early Paleozoic fold belts (for example the Adelaidean and western part of the Lachland) and basins (e.g., western part of the Murray basin) are underlain by Precambrian basement [*Zielhuis and van der Hilst*, 1996]. This interpretation, which was recently corroborated by Re-Os model ages of up to 1.96 Ga for upper mantle xenoliths in western Victoria [*Handler et al.*, 1997], would have implications not only for understanding the structural development of these fold belts and basins but also for the paleogeographic reconstruction of the Rodinia supercontinent [*Powell*, 1998, *K. Karlstrom*, pers. comm., 1999].

It is unlikely that the anomalously fast structure at a depth greater than 250 km, i.e., beneath the pronounced low velocity zone (see Figs. 1-8c-f and 1-9b), is related to stable lithospheric structure formed *in situ*. Instead it may represent allochthonous material.

Perhaps it represents one or more accreted terranes of continental tectosphere that was originally formed elsewhere. However, given the north-north-eastward motion of the Australian plate relative to the lower mantle [*DeMets et al.*, 1990], a perhaps more plausible explanation is that fragments of formerly oceanic lithosphere that were subducted beneath the island arcs to the north and northeast of Australia have subsequently been overridden by the continent.

1.6.3 Central and western Australia

As mentioned above, the seismic signature of the tectosphere in central and western Australia changes significantly within geological provinces of comparable age and tectonic setting, and the thickest tectosphere appears to be associated with the Proterozoic shields, not with the Archean cratons. Moreover, some of the fast anomalies beneath central and eastern Australia do not correlate with surface geology in an obvious fashion. A detailed interpretation of these observations is beyond the scope of this chapter, but we speculate on a few scenarios that might explain the observed complexity.

Assuming that the interpretation of the thick high wave speed structure in terms of stable continental keels (tectosphere) is basically correct, two fundamentally different explanations can be given. One possibility is that the differences within and between the Archean–Proterozoic provinces are entirely due to how the structures were originally formed, and have persisted ever since. Conditions for melt depletion associated with tectosphere formation could have varied from region to region. However, this complexity is not readily consistent with the correlation between crustal age and mantle composition inferred from geological data (e.g., *Griffin et al.* [1998]), which implies a certain uniformity of formation processes. Another explanation is that a thick ($z>250$ km) tectosphere formed everywhere but only survived beneath central Australia. *Shapiro et al.* [1999a] showed that a cold tectosphere could be supported by high viscosity alone but that compositional buoyancy is required to annihilate part of the gravity signal and to stabilize the “keel” during formation when the temperature was higher and the viscosity lower. The CLM can be eroded by later thermal processes, which would reduce its viscosity, in particular when it is already weakened by hydratation, for example by volatile infusion during subduction [*Nolet and Zielhuis*, 1994]. The removal of CLM by plumes, subduction, or small-scale convection associated with continental break-up has been invoked to explain the deep structure beneath western part of the Russian Platform [*Nolet and Zielhuis*, 1994], the anomalously low subcrustal P_n and S_n velocities beneath the southwestern margin of the Proterozoic Baltic Shield [*Bannister et al.*, 1991], and the absence of a seismically fast CLM beneath the Rocky Mountain front in North America [*Eggler et al.*, 1988], the Proterozoic Grenville province in the eastern part of United States [*van der Lee and Nolet*, 1997b], and the Archean Sino-Korean craton (see *Griffin et al.* [1998] and references therein). The position of the ”rootless” Archean Yilgarn, Pilbara, and Gawler blocks near passive margins suggests that an original deep high-velocity lid could have been destroyed by convective processes associated with continental break-up when western and southern Australia parted from India and Gondwana (onset $\sim$130 Ma) and Antarctica (starting $\sim$80 Ma) [*Veevers*, 1984]. Much like the rifting of Australia from New Zealand, the separation of Greater India from Australia was preceded and accompanied by extensive magmatism and related thermal processes [*Storey*, 1995].

Alternatively, the deep ($z>200$ km) high wave speeds represents temperature perturbations associated with subcontinental downwellings instead of stable tectosphere. This may apply to the localized high wave speed structure beneath the central Australian region involved in the Late-Paleozoic Alice Springs orogeny, which may continue to lower mantle depths [*Grand*, 1998, pers. comm.]. In this scenario, the permanent thick part of the continental plate might have an almost uniform thickness of about 200–250 km beneath all domains of Archean or Proterozoic age, and the deep, high-velocity anomalies under central and northern Australia reflect convective instabilities under the continent. Cold convective drips beneath continents have been suggested before [*Gurnis*, 1988; *De Smet et al.*, 1999] and were invoked to explain the deep high wave speeds beneath the Canadian shield [*Pari and Peltier*, 1996] (see, however, *Shapiro et al.* [1999b]). For North America, *Li et al.* [1998]) showed that the undulations of the 410 km discontinuity, which marks the isochemical phase change from olivine to wadsleyite, are small and do not correlate with the presence of a cold CLM. From this they conclude that the cold drips must be small, if they exist at all. For Australia such studies have not yet been completed.

1.7 Conclusions

We have analyzed about 1600 broad-band, vertical-component seismograms provided by the permanent IRIS, GEOSCOPE, and AGSO stations and the portable seismograph stations of the SKIPPY project. The combination of dense wave-path coverage and the use of both the fundamental and higher modes of the Rayleigh (surface) wave allowed us to delineate aspherical wave speed variations of the Australian continent in unprecedented detail. Our current inversions account for isotropic wave speed variations only. In this chapter we have focused on the presentation of the results, with special emphasis on the relation between wave speed variations in the upper mantle and the geological age of the overlying crust. In order to investigate whether that relationship depends on length scale of the wave speed variations and crustal structures considered, we have performed different tectonic regionalizations of our new wave speed model. We reach the following conclusions:

[1] For large scale lengths, and above ~200 km, there is dramatic westward progression of lithospheric wave speed with increasing crustal age, in agreement with global studies. In general, the Phanerozoic lithosphere (<80 km thick) is underlain by low velocity anomalies (with respect to a continental reference model); such a LVZ is absent beneath the Precambrian units, which instead are underlain by fast S-wave anomalies to about 200–300 km.

[2] Even on these large scales (λ >1500 km) the high wave speed lid of the Proterozoic provinces (central Australia) is generally thicker (250–300 km) than that of the Archean cratons (175–250 km), but the wave speed anomalies are higher in the latter. Our estimates of the thickness of the Archean CLM are slightly larger than, but within the level of uncertainty of, those for the Kaapvaal craton [*Priestley*, 1999]

[3] The Early-Paleozoic mountain belts in southeastern Australia (Adelaidean and western Lachland) and the western part of the Phanerozoic Murray basin are underlain by tectosphere-like upper mantle (~150 km thick). The inference that crustal deformation has occurred on a Proterozoic ramp is consistent with mid-Proterozoic Re-Os ages for upper mantle xenoliths from the Mt. Gambier volcanic complex several hundred km east of the outcrop divide (the so-called Tasman line) between the Proterozoic central shields and the Phanerozoic fold belts [*Handler et al.*, 1997]. These observations are important for understanding the extent of Precambrian lithosphere in Australia as well as the reconstruction of the Proterozoic super-continent Rodinia.

[4] The high wave speed structures that we interpret as old continental keels often continue off-shore, for instance northward into New Guinea, and southward into the Indian Ocean. The latter may represent stretched continental lithosphere resulting from rifting between Australia and Antarctica. Likewise, the inferred high wave speed keel just east of the Tasman Line could represent thinned lithosphere due to extension associated with the break-up of Rodinia (post-750 Ma).

[5] At shorter wave lengths, parts of the Precambrian shields are underlain by high wave speeds to depths greater than 200 km but other such domains lack a deep seismic signature altogether.

[6] The large variability of the seismic signature of the deep continent may be explained by (1) as yet poorly understood differences in the original fractionation and depletion of the upper mantle and — more likely — (2) the local disruption and removal of once thick

tectosphere by later tectonic processes. We notice that the "rootless" Archean cratons are located near old passive plate margins, suggesting that destabilization of tectospheric mantle has occurred by small-scale convection associated with continental break-up and subsequent rifting. Alternatively, some of the deep high wave speed structures may reflect cold downwellings (e.g., beneath central Australia) or the relics of allochthonous terranes (either exotic tectospheric blocks accreted onto eastern Australia, or remnants of subducted slab overridden by the plate on its northeastwardly journey).

[7] Variations in deep continental structure within regions of similar crustal age and tectonic setting are significant and often as large as variations between such age units. This implies that the wave speed models do not always support a simple relationship between mantle composition and crustal age and that global correlations based on geological regionalizations on scales smaller than about 1000 km may not be appropriate.

1.8 Appendix: Partitioned Waveform Inversion

This section presents a development of the theory of PWI as published by *Nolet* [1990] and *Zielhuis and Nolet* [1994a] with our changes as discussed in the text.

In a laterally homogeneous Earth, we may obtain the seismic wave as a sum of (higher modes of) surface waves. We denote the overtone number by n, the complex excitation coefficient of a mode by $A^0_{jn}(\omega)$ and the wavenumber at angular frequency ω as $k^0_n(\omega)$. The path index is j. The seismic spectrum $\tilde{s}_j(\omega)$ in a particular station at epicentral distance Δ_j is then written as:

$$\tilde{s}_j(\omega) = \sum_{n=0}^{\infty} A^0_{jn}(\omega) \exp\left(ik^0_n(\omega)\Delta_j\right). \tag{1.1}$$

We treat the wavenumber as a phase integral over the path:

$$k^0_n(\omega)\Delta_j \rightarrow \int_{P_j} k_n[\theta(\Delta), \varphi(\Delta), \omega]\, d\Delta. \tag{1.2}$$

The approximations involved in the use of this representation of the seismic wave field have been discussed above. Now,

$$\tilde{s}_j(\omega) = \sum_{n=0}^{\infty} A^0_{jn}(\omega) \exp\left(i \int_{P_j} k_n[\theta(\Delta), \varphi(\Delta), \omega]\, d\Delta\right). \tag{1.3}$$

The average wavenumber perturbation from the 1D reference model r is given by

$$<\delta k^r_{jn}(\omega)>= \frac{1}{\Delta_j} \int_{P_j} \left(k_n(\theta, \varphi, \omega) - k^r_n(\omega)\right) d\Delta, \tag{1.4}$$

and the average velocity perturbation from this reference model by

$$<\delta \beta^r_j(r)>= \frac{1}{\Delta_j} \int_{P_j} \left(\beta(\theta, \varphi, r) - \beta^r(r)\right) d\Delta. \tag{1.5}$$

The phase integral gets perturbed by the presence of lateral heterogeneity. Neglecting the effects of backscattering and variations in the amplitude factor A_{jn}^0, we may model the seismogram as follows:

$$\begin{aligned}
\int_{P_j} k_n(\theta, \varphi, \omega)\, \Delta &= k_n^0(\omega)\Delta_j + \int_{P_j}\int_0^a \left\{ \frac{\partial k_n(\omega)}{\partial \beta(r)} \right\} \delta\beta^r(\mathbf{r})\, dr\, d\Delta \qquad (1.6)\\
&= k_n^0(\omega)\Delta_j + \int_0^a \left\{ \frac{\partial k_n(\omega)}{\partial \beta(r)} \right\} \int_{P_j} \delta\beta^r(\mathbf{r})\, d\Delta\, dr\\
&= k_n^0(\omega)\Delta_j + \int_0^a \left\{ \frac{\partial k_n(\omega)}{\partial \beta(r)} \right\} dr <\delta\beta_j^r(r)> \Delta_j .
\end{aligned}$$

The sensitivity kernel of Fréchet derivatives may be found in *Nolet* [1990] or *Takeuchi and Saito* [1972]. This approach ignores perturbations in the excitation coefficients A_{jn}^0 due to perturbations in the elastic structure. To a certain extent, this neglect of relevant information is corrected for by allowing for different models (oceanic and continental models with varying crustal thicknesses) for source and receiver structure and a separate starting model for the path structure. We develop the mean velocity perturbation into a set of basis functions $h_i(r)$. The basis functions are effectively a set of averaging kernels that may take the form of triangular or boxcar functions. We write

$$\frac{1}{\Delta_j} \int_{P_j} \delta\beta^r(\mathbf{r})\, d\Delta = \sum_{i=1}^{M} \gamma_{ji} h_i(r), \qquad (1.7)$$

from which, taking

$$h_i(\omega) = \int_0^a \left\{ \frac{\partial k_n(\omega)}{\partial \beta(r)} \right\} h_i(r)\, dr, \qquad (1.8)$$

we can write the spectrum of the synthetic seismogram as:

$$\tilde{s}_j(\omega) = \sum_{n=0}^{\infty} A^0_{jn}(\omega) \exp\left(i\Delta_j \left\{ k^0_n(\omega) + \sum_{i=1}^{N} \gamma_{ji} h_i(\omega) \right\} d\Delta \right). \tag{1.9}$$

The synthetic seismogram is then obtained by Fourier-transforming $\tilde{s}_j(\omega)$ to the time domain. The objective is to obtain a waveform fit which we obtain by minimizing a penalty function Ψ that is of order q - we want to have an accurate prediction s_j of the observed seismogram d_j that minimizes the misfit as follows *Nolet* [1987]:

$$\Psi_q(\gamma) = \frac{1}{q} \sum_{j=1}^{L} \int_0^{t_j} w_j^q \left\| \mathcal{R} s_j(\mathcal{S}\gamma, t) - \mathcal{R} d_j(t) \right\|^q dt. \tag{1.10}$$

Both $\mathcal{R}$ and $\mathcal{S}$ are interactively changed as the iterations progress. $\mathcal{R}$ is a filtering operator and $\mathcal{S}$ specifies a projection of the model parameter vector γ onto a subspace of lesser dimensionality *Nolet et al.* [1986]. The frequency filter avoids convergence to local minima in the earlier iterations, as we are usually first fitting the long-wavelength components (or even the envelope of the signal) of the seismogram, which have considerable energy. The subspace method adaptively preconditions the matrix. Usually, q is 1,2 or ∞. Taking $q = 2$ (the Euclidean or least-squares norm) greatly simplifies the math and treats outliers on the basis of an assumed normal distribution. w_i represents a weighting factor - at its best, this corresponds to the inverse of the data covariance.

To $\Psi_q(\gamma)$ may be added a damping term of the form (in the case $q = 2$)

$$\alpha \gamma^{\mathrm{T}} \mathbf{C} \gamma, \tag{1.11}$$

where $\mathbf{C}$ takes care of smoothing the data and may be given by the *a priori* covariance matrix of the model parameters. α is a trade-off parameter that weights the relative contribution of data misfit and model norm in the minimization procedure.

At the solution, $\nabla\Psi_q(\gamma)$ vanishes, so in its neighborhood we may use a second-order multi-variable Taylor expansion as follows

$$\Psi_q(\gamma^{\mathrm{est}}) \approx \Psi_q(\gamma^{\mathrm{true}}) + \frac{1}{2}\left(\gamma^{\mathrm{est}} - \gamma^{\mathrm{true}}\right) \cdot \mathbf{H}^{\mathrm{est}}(\gamma) \cdot \left(\gamma^{\mathrm{est}} - \gamma^{\mathrm{true}}\right), \tag{1.12}$$

where **H** the Hessian matrix of second derivatives of the misfit criterion with respect to the model parameters, evaluated at the estimate:

$$H_{ij} = \frac{\partial^2 \Psi_q}{\partial \gamma_i \partial \gamma_j}. \tag{1.13}$$

Since **H** is a symmetric matrix, its eigenvectors are orthogonal, and we may determine a basis S by diagonalizing **H** as

$$\mathbf{H} = \mathbf{S}^{\mathrm{T}} \cdot \mathbf{\Lambda} \cdot \mathbf{S}, \tag{1.14}$$

where $\mathbf{\Lambda}$ is a diagonal matrix of eigenvalues. Therefore, S will transform the model parameters γ into a set

$$\boldsymbol{\eta} = \mathbf{S}^{\mathrm{T}} \cdot \boldsymbol{\gamma}, \tag{1.15}$$

with uncorrelated errors such that their covariance matrix, the dyad $<\Delta\eta\Delta\eta>$, is diagonal. If we define ϵ as a region of confidence around the optimum solution, we define an acceptable fit as contained within

$$\frac{1}{2}\Delta\boldsymbol{\gamma} \cdot \mathbf{H} \cdot \Delta\boldsymbol{\gamma} < \epsilon, \tag{1.16}$$

which implies for the transformed parameters that

$$|\Delta\eta_i| \leq \sqrt{\frac{2\epsilon}{\lambda_i}}. \tag{1.17}$$

In the application of the method of PWI by *Zielhuis and Nolet* [1994a] and *Zielhuis and van der Hilst* [1996], the acceptable misfit is an *a priori* noise estimator. As we define an alternative uncertainty based on the actual characteristics of the waveform fits (as discussed briefly in the text), the individual fits get weighted with a more objective criterion.

We therefore rewrite the average wave speed perturbation as

$$\begin{aligned} \frac{1}{\Delta_j} \int_{P_j} \delta\beta(\mathbf{r})\, d\Delta &= \sum_{i=1}^{M} \gamma_{ji} h_i(r) \\ &= \sum_{p=1}^{M} \sum_{q=1}^{M} S_{jpq} \eta_{jq} h_{jp}(r) \\ &= \sum_{q=1}^{M} \eta_{jq} g_{jq}(r), \end{aligned} \tag{1.18}$$

where we defined

$$g_{jq}(r) = \sum_{p=1}^{M} S_{jpq} h_{jp}(r). \tag{1.19}$$

The $\sum_{q=1}^{M} \eta_{jq} g_{jq}(r)$ now define a set of linear constraints on the velocity structure beneath each path j. In a second, linear inversion step we will have to attempt to invert for a velocity model that satisfies those individual constraints. The effects of an inaccurate event location, however, are to overestimate (if the event has been located too far to the receiver, or underestimate positive velocity anomalies (if the real location lies further from the receiver than assumed). Although we will not attempt to relocate earthquakes, it seems defendable and certainly useful to account for possible effects in mislocation by introducing an extra parameter in the inversion.

In the case of an unperturbed (reference) velocity model $v^r(r)$, an event mislocation that places the event at a distance δs from its assumed distance s, which will be interpreted as a mean velocity perturbation $\delta v^r(r)$ according to

$$<\delta v^r(r)> \sim \frac{<v^r(r)>\, s}{s+\delta s} - <v^r(r)>, \tag{1.20}$$

where again, the integrals are over the great circle.

We may thus perceive our "data" $\sum_{q=1}^{M} \eta_{jq} g_{jq}(r)$ as partly caused by true velocity deviations, and a residual part as due to a wrong source-receiver distance. Since the amount of mislocation takes one value for every path but the reference velocity model is depth-dependent, the inferred velocity anomalies at various depths are affected differently. We include one extra inversion parameter ζ per path, as follows.

$$\sum_{q=1}^{M} \eta_{jq} g_{jq}(r) = \frac{1}{\Delta_j} \int_{P_j} \delta\beta^r(\mathbf{r})\, d\Delta + \zeta_j <\delta v_j^r(r)>, \tag{1.21}$$

where $<\delta v^r(r)>$ is the sensitivity of our data to this mislocation factor that is determined by the reference model for the datum and, through equation (1.20), by some *a priori* value of the source mislocation δs. This extra parameter should work as a fudge factor in that the individual velocity constraints should be sufficient to obtain a good model; at the same time this extra factor may absorb unexplained structure.

In order to isolate the individual coefficients η_{jq} that go with the different basis functions $g_{jq}(r)$, a second set of basis functions $\tilde{g}_{jl}(r)$ is constructed (also from knowledge of the Hessian matrix, which is a different matrix for every path) which satisfy the orthogonality condition

$$\int_0^a \tilde{g}_{jl}(r)\, g_{jm}(r)\, dr = \delta_{lm}. \tag{1.22}$$

Taking the inproduct of those basis functions one after the other, we obtain M equations (where we now explicitly write the uncertainties on the parameters $\Delta\eta_{jk}$ as obtained from equation (1.17):

$$\begin{aligned}
\eta_{jk} \pm \Delta\eta_{jk} &= \frac{1}{\Delta_j}\int\limits_{P_j}\int\limits_0^a \delta\beta^r(\mathbf{r})\,\tilde{g}_{jk}(r)\,dr\,d\Delta + \zeta_j \int\limits_0^a \tilde{g}_{jk}(r)\ <\delta v^r(r)>\ dr \\
&= \frac{1}{\Delta_j}\int\limits_{P_j}\int\limits_0^a \delta\beta(\mathbf{r})\,\tilde{g}_{jk}(r)\,dr\,d\Delta \\
&\quad + \zeta_j \left(\frac{s_j}{s_j+\delta s_j} - 1\right) \int\limits_0^a \tilde{g}_{jk}(r)\ <v^r(r)>\ dr
\end{aligned} \tag{1.23}$$

.

Every path's velocity perturbation $\delta\beta(\mathbf{r})$ was expressed with respect to a particular reference Earth model $\beta^r(r)$=$v^r(r)$=$<v^r(r)>$:

$$\delta\beta(\mathbf{r}) = \beta(\mathbf{r}) - \beta_j^r(r), \tag{1.24}$$

so

$$\begin{aligned}
\eta_{jk} \pm \Delta\eta_{jk} &= \frac{1}{\Delta_j}\int\limits_{P_j}\int\limits_0^a \delta\beta(\mathbf{r})\,\tilde{g}_{jk}(r)\,dr\,d\Delta \\
&\quad + \frac{\zeta_j}{\Delta_j}\left(\frac{s_j}{s_j+\delta s_j} - 1\right) \int\limits_{P_j}\int\limits_0^a \tilde{g}_{jk}(r) v^r(r)\,dr\,d\Delta,
\end{aligned} \tag{1.25}$$

and we may add the reference model used back into it

$$\begin{aligned}
\eta_{jk} \pm \Delta\eta_{jk} &+ \frac{1}{\Delta_j}\int\limits_{P_j}\int\limits_0^a \beta_j^r(r)\,\tilde{g}_{jk}(r)\,dr\,d\Delta \\
&= \frac{1}{\Delta_j}\int\limits_{P_j}\int\limits_0^a \beta(\mathbf{r})\,\tilde{g}_{jk}(r)\,dr\,d\Delta + \zeta_j Q_{jk},
\end{aligned} \tag{1.26}$$

where we called, noting of course that $\Delta_j = s_j$

$$\mathcal{Q}_{jk} = \frac{1}{\Delta_j}\left(\frac{s_j}{s_j + \delta s_j} - 1\right) \int\limits_{P_j}\int\limits_0^a \tilde{g}_{jk}(r)\beta^r(r)\, d\Delta\, dr. \tag{1.27}$$

For the full analysis, we take a global reference model $\beta^R(r)$ so that

$$\begin{aligned} \eta_{jk} \pm \Delta\eta_{jk} \quad &+ \quad \frac{1}{\Delta_j}\int\limits_{P_j}\int\limits_0^a [\beta_j^r(r) - \beta^R(r)]\,\tilde{g}_{jk}(r)\, dr\, d\Delta \\ &= \quad \frac{1}{\Delta_j}\int\limits_{P_j}\int\limits_0^a \delta\beta(\mathbf{r})\,\tilde{g}_{jk}(r)\, dr\, d\Delta + \zeta_j \mathcal{Q}_{jk}, \end{aligned} \tag{1.28}$$

where the perturbation is given by

$$\delta\beta(\mathbf{r}) = \beta(\mathbf{r}) - \beta^R(r). \tag{1.29}$$

Normalizing to unit variance, and writing q_{jk} for the scaled and shifted datum and $G_{jk}(r)$ for the scaled kernel:

$$q_{jk} \pm 1 = \int\limits_{P_j}\int\limits_0^a G_{jk}(r)\delta\beta(\mathbf{r})\, dr\, d\Delta + \zeta_j \mathcal{Q}'_{jk}, \tag{1.30}$$

i.e., with

$$G_{jk}(r) = \frac{\tilde{g}_{jk}}{\Delta_j\, \Delta\eta_{jk}} \qquad \mathcal{Q}'_{jk} = \frac{\mathcal{Q}_{jk}}{\Delta\eta_{jk}}, \tag{1.31}$$

$$q_{jk} = \frac{1}{\Delta\eta_{jk}}\left\{\eta_{jk} + \frac{1}{\Delta_j}\int\limits_{P_j}\int\limits_0^a [\beta_j^r(\mathbf{r}) - \beta^R(r)]\,\tilde{g}_{jk}(r)\, dr\, d\Delta\right\}. \tag{1.32}$$

The true velocity anomaly $\delta\beta(\mathbf{r})$ is now expressed in terms of a set of basis functions:

$$\delta\beta(\mathbf{r}) = \sum_i^m b_i s_i(\mathbf{r}). \tag{1.33}$$

The basis functions $s_i(\mathbf{r})$ define all of space. For the horizontal coordinate, the $s_i(\theta, \varphi)$ are define as discrete rectangular cells on the surface; for the vertical coordinate, $s_i(r)$ can be identical to the basis functions $h_i(r)$ we used previously in the nonlinear inversion:

$$s_i(\mathbf{r}) = l_i(\theta, \phi) h_i(r), \tag{1.34}$$

where $l_i(\theta, \phi)$ takes on values of 1 or 0 only so it effectively filters out each individual surface cell. We now have

$$q_{jk} \pm 1 = \int\limits_{P_j}\int\limits_0^a G_{jk}(r) \left(\sum_i^m l_i(\theta, \phi) h_i(r) b_i \right) dr\, d\Delta + \zeta_j \mathcal{Q}'_{jk}. \tag{1.35}$$

The surface-parallel path integral leads to a tensor L_{ij} that contains the segment length of ray j in cell i:

$$\int\limits_{P_j} l_i(\theta, \phi)\, d\Delta = L_{ji}, \tag{1.36}$$

so

$$q_{jk} \pm 1 = \sum_i \left(L_{ji} \int\limits_0^a G_{jk}(r) h_i(r)\, dr \right) b_i + \zeta_j \mathcal{Q}'_{jk}. \tag{1.37}$$

This leads to the following set of linear equations for path j, individual path parameter k and each of the basis functions i of the total model:

$$q_{jk} \pm 1 = \sum_i^m A_{jik} b_i + \zeta_j \mathcal{Q}'_{jk}. \tag{1.38}$$

For n paths, the dimensions of the matrices involved are:

$$|A| = (n\,k) \times m \quad (1.39) \qquad |b| = m \times 1 \quad (1.40) \qquad |q| = (n\,k) \times 1 \quad (1.41)$$

These dimensions are a maximum. In the process of the construction of the matrix, parameters with uncertainties that are of the order of 100 % have been discarded, as well as those that have sensitivity values lower than a cut-off value due to poor path sampling (typically parameters belonging to cells with a ray length smaller than 50m).

In a matrix formulation, we get

$$(\mathbf{A}|\mathcal{Q}')\,(\mathbf{b}|\zeta) = \mathbf{q}, \tag{1.42}$$

which we represent as

$$\mathbf{A}'\,\mathbf{b}' = \mathbf{q}. \tag{1.43}$$

This may be solved by the LSQR algorithm [*Nolet*, 1985]

Our new matrix equation is of the following schematic form:

$$\begin{array}{ccccc}
 & \begin{array}{c}\text{sensitivities to}\\ m\text{ model parameters}\\ \downarrow\;\downarrow\;\downarrow\end{array}\qquad\begin{array}{c}\text{sensitivity to}\\ \text{fudge factor}\\ \downarrow\;\downarrow\end{array} & \begin{array}{c} m\text{ model}\\ \text{parameters}\\ \downarrow\end{array} & & \begin{array}{c} j\times k\text{ data}\\ \downarrow\end{array}\\
\begin{array}{cc} & \rightarrow\ k\text{ pa-}\\ \rightarrow & \rightarrow\ \text{rame-}\\ & \rightarrow\ \text{ters}\\ \rightarrow & \\ \rightarrow & j\text{ events}\end{array} & \left(\begin{array}{c|ccc} \cdots & \cdots & 0 & 0\\ \int_{P_j}\int_0^a G_{jk}(r)\delta\beta(\mathbf{r})\,dr\,d\Delta & 0 & \mathcal{Q}'_{jk} & 0\\ \cdots & 0 & 0 & \cdots\end{array}\right) & \left(\begin{array}{c}\cdots\\ b_i\\ \cdots\\ \hline \cdots\\ \zeta_i\\ \cdots\end{array}\right) & = & \left(\begin{array}{c}\cdots\\ q_{jk}\\ \cdots\end{array}\right).\\
 & & \begin{array}{c}\uparrow\\ j\text{ fudge}\\ \text{factors}\end{array} & &
\end{array}\tag{1.44}$$

Chapter 2

Anisotropic Surface Wave Tomography[1]

Abstract

We present an azimuthally anisotropic three-dimensional shear-wave speed model for the Australian upper mantle from the dispersion of fundamental and higher modes of Rayleigh waves. We compare two tomographic techniques to map path-average Earth models into a 3-D model for heterogeneity and azimuthal anisotropy. Method I uses a rectangular surface cell parameterization and depth basis functions that represent independently constrained estimates of radial Earth structure. It performs an iterative inversion for three-dimensional structure with norm damping and gradient regularization. Method II uses a direct inversion of individual depth layers constrained by Bayesian assumptions about the model covariance. We recall that Bayesian inversions and discrete regularization approaches are theoretically equivalent, and with a synthetic example we show that they can give similar results. With the exception of westernmost Australia, we can retrieve structure on length scales of about 250 km laterally and 50 km in the radial direction, to within 0.8% for the velocity, 20% for the anisotropic magnitude and 20° for its direction. On length scales of 1000 km and longer, down to about 200 km, there is a good correlation between velocity heterogeneity and geologic age. At shorter length scales and at depths below 200 km, however, this relationship breaks down. The observed magnitude and direction of maximum anisotropy do not, in general, appear to be correlated to surface geology. The pattern of anisotropy appears to be rather complex in the upper 150 km, whereas a smoother pattern of fast axes is obtained at larger depth. If some of the deeper directions of anisotropy are

---

[1] In press as: **Multimode Rayleigh Wave Inversion for Heterogeneity and Azimuthal Anisotropy of the Australian Upper Mantle**, *Geophysical Journal International*, 2002.

aligned with the approximately N–S direction of absolute plate motion, this correspondence is not everywhere obvious, despite the fast (7 cm/yr) northward motion of the Australian plate. This suggests that other mechanisms are important in creating the deeper anisotropy. Predictions of *SKS* splitting times and directions, an integrated measure of anisotropy, are poorly matched by observations of shear wave birefringence.

2.1 Introduction

The Earth is heterogeneous and anisotropic on a wide spectrum of length scales. Polarization anisotropy (modeled by transverse isotropy with a vertical symmetry axis) affects fundamental [*Anderson*, 1961] and higher-mode [*Lévêque and Cara*, 1985] surface wave propagation. Body wave [*Hess*, 1964] and surface wave velocities [*Forsyth*, 1975] are dependent on propagation azimuth. Shear wave birefringence reflects transverse isotropy with a horizontal symmetry axis [*Silver*, 1996]. Since anisotropy requires both the presence of anisotropic crystals and their (strain-induced) preferred orientation on seismic length-scales [*Wenk and Tomé*, 1999], anisotropy measurements contain information about the mineralogy and dynamics of the mantle [*Artyushkov*, 1984; *Montagner and Anderson*, 1989; *Silver and Chan*, 1991; *Vinnik et al.*, 1992].

Knowledge about continental velocity heterogeneity and anisotropy may answer a number of outstanding questions. These include the very definition of the base of the lithosphere, the relation between the depth extent of the lithosphere and the age of the overlying crust, and the deformation of the continental lithospheric mantle and the underlying asthenosphere.

It has been argued that continental seismic anisotropy is best explained by a combination of "frozen" (subcrustal) anisotropy, developed during major orogenic events, underlain by an actively deforming asthenosphere which generates anisotropy by the subhorizontal alignment of olivine crystals due to present-day mantle flow [*Silver*, 1996; *Plomerová et al.*, 1998]. However, considerable ambiguity exists regarding the amount and location of azimuthal anisotropy that can be inferred from body waves. Moreover, simple physical models fail to explain the globally observed pattern of splitting [*Barruol and Hoffman*, 1999]. Conflicting results of shear wave birefringence [*Clitheroe and van der Hilst*, 1998; *Özalaybey and Chen*, 1999] indicate that the anisotropy of the Australian continental mantle is complex and not yet well understood.

In general, surface wave waveform tomography allows the delineation of lateral heterogeneities in the upper mantle with better radial resolution than body wave travel time tomography. A variety of methods has been developed to extract the variation of wave speed with depth from surface waves, [e.g *Nolet et al.*, 1986; *Cara and Lévêque*, 1987; *Stutzmann and Montagner*, 1993; *van Heijst and Woodhouse*, 1997], and a number of tomographic methods to invert for aspherical structure exist [*Woodhouse and Dziewonski*, 1984; *Montagner*, 1986; *Nolet*, 1990; *Zielhuis and Nolet*, 1994a]. Inversions of surface waves for heterogeneity have traditionally been restricted to fundamental modes but more recent models have incorporated higher mode measurements, thereby enhancing the resolution down to depths of about 400 km. Inversions for both heterogeneity and azimuthal anisotropy have been made using fundamental modes [e.g. *Montagner and Nataf*, 1988; *Montagner and Jobert*, 1988; *Nishimura and Forsyth*, 1988, 1989] and multi-mode waves [*Stutzmann and Montagner*, 1994; *Lévêque et al.*, 1998; *Debayle and Kennett*, 2000a].

The SKIPPY experiment [*van der Hilst et al.*, 1994] has produced a unique broadband data set of seismic waves propagating across the Australian lithosphere (see Fig. 2-1). The partitioned waveform inversion of vertical component fundamental and higher mode Rayleigh wave data has yielded a detailed picture of the 3-D isotropic S-wave speed variations of the continent and the adjacent oceanic areas [*Zielhuis and van der Hilst*, 1996; *Simons et al.*, 1999]. With similar data, *Debayle* [1999] and *Debayle and Kennett* [2000a] have published azimuthally anisotropic models. *Debayle and Kennett* [2000b] have also considered polarization anisotropy by the inclusion of Love waves into the analysis.

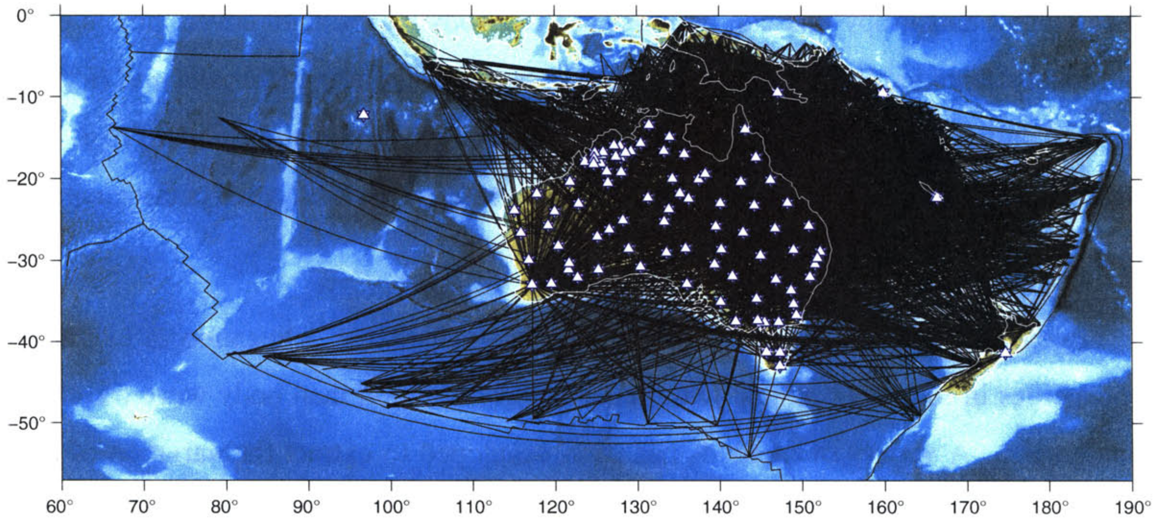


Figure 2-1: Path coverage. About 2250 vertical component seismograms provided waveforms which were fitted. The stations belong to the SKIPPY, KIMBA, AGSO, IRIS and GEOSCOPE arrays. We used earthquake locations from *Engdahl et al.* [1998].

Here, we present a model for the Australian upper mantle that takes into account both wave speed heterogeneity and azimuthal anisotropy, and we quantify resolution and uncertainty by means of synthetic tests, trade-off estimates, and an exact calculation of the horizontal resolution matrix. We compare our results with global models from other seismological observables such as body-wave shear wave splitting. While there is a general agreement between the long-wavelength structures imaged in our present model and the one by *Debayle and Kennett* [2000a] (hereafter: DK2000), there are substantial differences in heterogeneity and anisotropic structure at wavelengths smaller than about 500 km even in regions for which both groups claim good resolution. Both groups attempt to invert for 3-D structure starting from path-average models. However, DK2000 use the regionalized inversion method due to *Montagner* [1986], while we use the same discretely parameterized, regularized inversion method as *Zielhuis and Nolet* [1994a]. We compare these tomographic techniques in theory and in their application to data.

2.2 Inversion partitioning and data

Our method and the one of DK2000 both involve an inversion for aspherical Earth structure that is partitioned into an initial waveform inversion that seeks to determine the average structure under the great-circle path between source and receiver and a subsequent tomographic inversion for 3-D Earth structure (which will be discussed in Sec. 2.4).

Our waveform inversion is due to *Nolet et al.* [1986]. It produces a minimal set of independently constrained estimates of radial Earth structure [*Zielhuis*, 1992] that are used as data in the 3-D tomographic inversion. The independence of the linear constraints and the decorrelation of their errors is achieved by matrix diagonalization. This is the central distinguishing feature of Nolet's method. Furthermore, the waveform analysis can be applied to different analysis windows in the group velocity versus frequency space [see *Simons et al.*, 1999, their Fig. 3], which can be used, for instance, to separate fundamental from higher modes. For their waveform inversion, DK2000 rely on the method by *Cara and Lévêque* [1987], which linearizes the waveform inversion by using the envelope of modal cross-correlation functions and fits the fundamental mode and the overtones simultaneously [*Debayle*, 1999].

Both approaches rely on a path-average approximation and neglect mode coupling. As a result, some effects of complex source mechanisms and mantle heterogeneity are not adequately modeled. However, when applied within the frequency bounds derived for both fundamental and higher modes [*Kennett and Nolet*, 1990; *Kennett*, 1995] and for a large number of crossing paths, the path average method leads to an acceptable approximation of the 3-D upper mantle heterogeneity [*Marquering et al.*, 1996; *Marquering and Snieder*, 1996].

For the present study we have expanded the data set of 1600 paths used by *Simons et al.* [1999] with additional data from the SKIPPY project and from permanent observations operated by IRIS, GEOSCOPE and AGSO, and we applied the waveform inversion to a total of ~2250 paths. In selecting the events we have avoided regions with complex source and mantle structure by limiting the data to events that originate mostly along or within the boundaries of the Indo-Australian plate (see Fig. 2-1).

Where data quality allows we fit the fundamental mode between 5–25 mHz (40–200 s period) or a narrower frequency band within this range, while the higher modes are fitted between 8–50 mHz (20–125 s) or less. With these frequency bounds the fundamental modes can provide sensitivity to structure down to at least 400 km, and the higher modes provided additional detail. The possibility of using two data windows can be exploited to prevent the highly energetic fundamental modes from dominating the waveform inversion. For more details about the application to SKIPPY data, see *Zielhuis and van der Hilst* [1996].

Montagner and Jobert [1988], and many others, have pointed out the importance of using the correct crustal velocity model. At high frequencies, Rayleigh modes are sensitive to variations in crustal thickness and velocity. We use crustal thickness determinations from broadband receiver functions at each of the SKIPPY and permanent network stations [*Shibutani et al.*, 1996; *Clitheroe et al.*, 2000a, b]. This is arguably the most accurate crustal model to date of the Australian continent.

2.3 Azimuthal anisotropy

Following *Montagner and Nataf* [1986], the perturbation of the phase speed of Rayleigh waves (δc_R) in a slightly anisotropic Earth can be modeled by a transversely isotropic perturbation (δL) and two combinations of elastic constants (G_c and G_s) modulated by, respectively, the cosine and sine of twice the azimuth of the ray (Ψ):

$$\delta c_R = \frac{\partial c_R}{\partial L}(\delta L + G_c \cos 2\Psi + G_s \sin 2\Psi). \tag{2.1}$$

In simplified index notation (C_{ij}) for the elastic tensor Γ_{ijkl} [*Montagner*, 1996] the parameters of Eq. 2.1 are:

$$L = (C_{44} + C_{55})/2, \tag{2.2}$$

$$G_c = (C_{55} - C_{44})/2, \tag{2.3}$$

$$G_s = C_{54}. \tag{2.4}$$

Similar to *Lévêque et al.* [1998] we define an effective shear wave speed

$$\hat{\beta}_V = \left(\frac{L + G_c \cos 2\Psi + G_s \sin 2\Psi}{\rho}\right)^{\frac{1}{2}}. \tag{2.5}$$

If no azimuthal anisotropy is present or if it is effectively averaged out over all azimuths, $\rho\beta_V^2 = L$ and $G_c = G_s = 0$ so that Eq. 2.5 reduces to β_V, i.e. the wave-speed of a vertically polarized shear wave in a transversely isotropic medium with a vertical symmetry axis [*Takeuchi and Saito*, 1972]. Assuming $G_c/L \ll 1$ and $G_s/L \ll 1$ we can write:

$$\hat{\beta}_V = \beta_V + \frac{G_c}{2\rho\beta_V}\cos 2\Psi + \frac{G_s}{2\rho\beta_V}\sin 2\Psi. \tag{2.6}$$

In the tomographic inversion we will take as a starting point the radial velocity averages that result from the waveform inversion (see Sec. 2.2), and solve for heterogeneity (β_V) and the anisotropic parameters ($A_1 = G_c/(2\rho\beta_V)$ and $A_2 = G_s/(2\rho\beta_V)$). We calculate the anisotropic magnitude as $(G_c^2 + G_s^2)^{1/2}$ and the direction of maximum velocity as $\tan^{-1}(A_2/A_1)/2$.

The ubiquitous use of the parameterization represented by Eqs. 2.1 and 2.6 is based on the dominance of the $\partial c_R/\partial L$ partial for fundamental mode surface wave propagation [*Montagner and Nataf*, 1986]. Since $\partial c_R/\partial L$ is also dominant for higher modes, both for isotropic and realistically anisotropic media (see the Appendix), a parameterization in terms of β_V, G_c and G_s is adequate when higher modes are included.

To fit seismograms, we used the $\partial c_R/\partial L$ partial derivative with an isotropic reference model. The resulting 1-D path model of β_V is an adequate representation of the Earth when no further azimuthally anisotropic information is available, because the sensitivity kernels of Rayleigh wave phase velocities to perturbations in *S* velocity in isotropic and transversely isotropic Earth models are very similar [*Anderson and Dziewonski*, 1982], although a small error is probably made [*Lévêque and Cara*, 1985].

2.4 The tomographic problem

2.4.1 Construction of a linear matrix problem

With the exception of a few computational details, our method of *S* wave velocity tomography differs little from the procedure described by *Zielhuis and Nolet* [1994a]. The model vector $\mathbf{m}$ containing 3-D shear-wave speed perturbations $\delta\beta(\mathbf{r})$ from a reference model $\beta(r)$ is related to the data vector $\mathbf{d}$ of independent linear constraints by a sensitivity matrix $\mathbf{G}$. The data errors are uncorrelated by construction. We seek the model vector $\mathbf{m}$ that solves the linear equation $\mathbf{d} = \mathbf{G} \cdot \mathbf{m}$; we incorporate the effect of azimuthal anisotropy by modifying the shear wave speed anomaly to include a transversely isotropic and an azimuthally anisotropic part:

$$\delta\beta(\mathbf{r}) = \delta\beta_V(\mathbf{r}) + A_1(\mathbf{r}) \cos 2\Psi(\mathbf{r}) + A_2(\mathbf{r}) \sin 2\Psi(\mathbf{r}), \tag{2.7}$$

i.e., by augmenting the sensitivity matrix $\mathbf{G}$ to include the cosine and the sine of twice the azimuth Ψ of the ray in each surface cell. We try to explain the data with both heterogeneity and anisotropy, with an equal weighting. This leads to a threefold increase in the number of unknowns with respect to isotropic inversions. In Sec. 2.5.5 we evaluate the resolution of heterogeneity and anisotropic parameters as well as the importance of the trade-off of between them.

2.4.2 Model parameterization

We use inversions performed on different grids to test the robustness of the structures retrieved (with resolution-dependent grids) and to determine the extent to which the Australian continental structure at varying depths bears a relationship to the geology observed at the surface (with pure-path tectonic grids).

Due to the uneven sampling density the resolution is not uniform across the model domain, but the grid spacing can be adapted to the locally expected resolution in order to reduce the number of free parameters. Our approach to adaptive grids is akin to that of *Abers and Roecker* [1991] [see also *Spakman and Bijwaard*, 2001]. We used a fine (1°×1°) Cartesian grid upon which we define a coordinate transform that attempts to equalize the sampling density by clustering cells together. Starting from a maximum cell size of 8°×8° we allow a further subdivision, into cells sized 4°×4°, 2°×2° and 1°×1°, as long as this brings the resulting mean of the sampling density over these new cells closer to the global mean. As a measure of resolution we use the sum of path density (see Fig. 2-2a) and azimuthal variability (Fig. 2-2b) in the surface cells. Although this is not necessarily an optimal *a priori* estimate of the spatial resolution of the model [*Barth and Wunsch*, 1990; *Curtis and Snieder*, 1997] the similarity of the features of this measure with an exactly computed resolution matrix (see Sec. 2.5) is sufficient to warrant their use in the construction of an irregular grid. An example of an irregular grid that results from the clustering of the sampling density is plotted in Fig. 2-2c.

Pure-path methods are used to invert for the average velocity and anisotropy structure of predefined tectonic regions (Fig. 2-2d) [*Montagner*, 1985; *Nishimura and Forsyth*, 1985; *Souriau*, 1985]. Similarly to the construction of a resolution-dependent irregular grid we

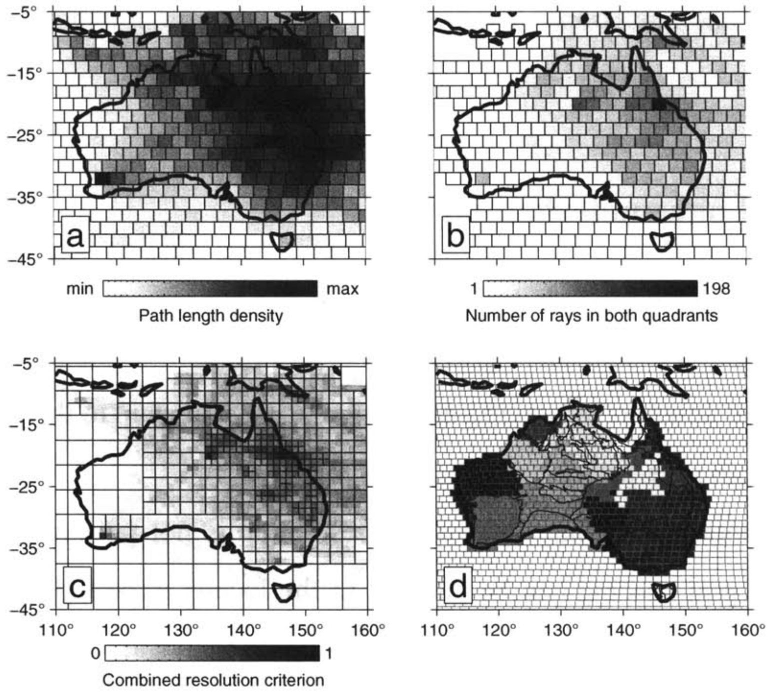


Figure 2-2: (a) Path density and (b) azimuthal coverage for the paths shown in Fig. 2-1, calculated on a 2°×2° equal-area grid. As a measure of azimuthal coverage, we took the number of pairs of rays in the cell coming into two different quadrants. (c) Design of a grid adapted to expected model resolution. Path density and azimuthal coverage calculated on a 1°×1° square grid were normalized and added to provide a measure of data quality. The cell size was adapted to allow a finer parameterization in areas where the resolution is expected to be better. (d) Tectonic subdivision. Based on an underlying equal-area grid, similarly shaded regions indicate approximate tectonic domains which are used as a single parameter in the pure-path inversion. Lines indicate the location of major sedimentary basins. For the definition of tectonic units, see *Simons et al.* [1999].

achieve this tectonic parameterization by a coordinate transform, such that all surface cells belonging to a tectonic domain are treated as a single parameter in the inversion.

2.4.3 Regularization or *a priori* information?

The discrete surface parameterization of our method leads to classic regularization approaches of systems of linear equations (Sec. 2.4.3). On the other hand, the continuous regionalization approach adopted by DK2000 draws solutions from a class of functions defined by their covariance structure (Sec. 2.4.3).

Discrete regularization

We construct a penalty function based on a weighted Euclidean norm of the solution error and a weighted norm in model parameter space as follows:

$$\begin{aligned}\Phi = \; & (\mathbf{d} - \mathbf{G} \cdot \mathbf{m}) \cdot \mathbf{A}^{-1} \cdot (\mathbf{d} - \mathbf{G} \cdot \mathbf{m}) \\ & + \mathbf{m} \cdot \mathbf{B}^{-1} \cdot \mathbf{m}. \end{aligned} \tag{2.8}$$

$\mathbf{A}^{-1}$ and $\mathbf{B}^{-1}$ are weighting matrices. Minimization of Eq. 2.8 in a least-squares sense is equivalent to solving the following modified set of matrix equations:

$$\begin{bmatrix} \mathbf{A}^{-1/2} \cdot \mathbf{G} \\ \mathbf{B}^{-1/2} \end{bmatrix} \cdot \mathbf{m} = \begin{bmatrix} \mathbf{A}^{-1/2} \cdot \mathbf{d} \\ 0 \end{bmatrix}, \tag{2.9}$$

The generalized inverse,

$$\mathbf{m} = (\mathbf{B}^{-1} + \mathbf{G}^{\mathrm{T}} \cdot \mathbf{A}^{-1} \cdot \mathbf{G})^{-1} \cdot \mathbf{G}^{\mathrm{T}} \cdot \mathbf{A}^{-1} \cdot \mathbf{d}, \tag{2.10}$$

can be approximated by the iterative LSQR algorithm [*Paige and Saunders*, 1982].

The optimal weighting matrix $\mathbf{A}^{-1}$ is given by the inverse of the data covariance matrix $\mathrm{C}_{\mathrm{d}}^{-1}$, so that the solution has minimum variance [*Draper and Smith*, 1981]. A covariance matrix, C_{d} is positive definite (hence invertible) and symmetric; in our implementation it is also diagonal because the individual parameters resulting from the waveform fitting procedure are independent by construction [*Nolet*, 1990; *Zielhuis and Nolet*, 1994a].

If we choose $\mathbf{B}^{-1} = \mathbf{I}$, the identity matrix, we are imposing a minimum norm constraint on the solution, by biasing the solution to zero in a least squares sense. This is appropriate if $\mathbf{m}$ represents perturbations to a reference model which is sufficiently close to the real solution so that linearity is preserved. Most inversion algorithms for the generalized inverse minimize the model norm implicitly [*Aki and Richards*, 1980; *van der Sluis and van der Vorst*, 1986]. In particular, the LSQR algorithm provides an intrinsic damping of the norm $\mathbf{m} \cdot \mathbf{m}$ when the iterations are truncated at some small finite value [*Nolet*, 1985; *van der Sluis and van der Vorst*, 1987; *Spakman and Nolet*, 1988].

Alternatively, we may propose $\mathbf{B}^{-1} = \mathbf{D}_1^2$, the square of the first finite difference matrix given by

$$\mathbf{D_1} = \begin{pmatrix} \dots & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & \dots \end{pmatrix} . \tag{2.11}$$

The effect of this "roughening" operator is to constrain the solution to be locally smooth [*Constable et al.*, 1987; *Nolet*, 1987]. Eq. 2.11 is the Toeplitz matrix of a high-pass filter in model space [*Strang and Nguyen*, 1997]. Hence, we are imposing that the high-passed model vector be minimal, which, after inversion, results in a smooth model.

In our inversion, we combine both approaches and minimize the penalty function

$$\begin{aligned} \Phi = \ & (\mathbf{d} - \mathbf{G} \cdot \mathbf{m}) \cdot \mathbf{C}_{\mathrm{d}}^{-1} \cdot (\mathbf{d} - \mathbf{G} \cdot \mathbf{m}) \\ & + \alpha \, \mathbf{m} \cdot \mathbf{I} \cdot \mathbf{m} + \beta \, \mathbf{m} \cdot \mathbf{D}_1^2 \cdot \mathbf{m}, \end{aligned} \tag{2.12}$$

where α and β are weights which can be adjusted arbitrarily. The solution

$$\mathbf{m} = (\alpha \mathbf{I} + \beta \mathbf{D}_1^2 + \mathbf{G}^{\mathrm{T}} \cdot \mathbf{C}_{\mathrm{d}}^{-1} \cdot \mathbf{G})^{-1} \cdot \mathbf{G}^{\mathrm{T}} \cdot \mathbf{C}_{\mathrm{d}}^{-1} \cdot \mathbf{d} \tag{2.13}$$

is calculated by LSQR.

Continuous regionalization

In the "Bayesian" approach [*Tarantola and Valette*, 1982a, b; *Montagner*, 1986] the model parameters $\mathbf{m}$ are distributed according to a Gaussian probability density function

$$\rho(\mathbf{m}) \propto \exp\left(-\frac{1}{2}\mathbf{m}\cdot\mathbf{C}_{\mathrm{m}}^{-1}\cdot\mathbf{m}\right). \tag{2.14}$$

The solution is obtained by maximizing the joint distribution of data and model parameters, and is given by

$$\mathbf{m} = \mathbf{C}_{\mathrm{m}}\cdot\mathbf{G}^{\mathrm{T}}\cdot(\mathbf{C}_{\mathrm{d}}+\mathbf{G}\cdot\mathbf{C}_{\mathrm{m}}\cdot\mathbf{G}^{\mathrm{T}})^{-1}\cdot\mathbf{d}. \tag{2.15}$$

The choice of the *a priori* model covariance matrix is important. DK2000 use

$$C_{\mathrm{m}}(\mathbf{r}_1,\mathbf{r}_2) = \sigma^2\exp\left(-\frac{|\mathbf{r}_1,\mathbf{r}_2|^2}{2L^2}\right), \tag{2.16}$$

where L denotes a correlation length, the determination of which is, like the level of regularization, at the discretion of the seismologist, even though some guidelines have been described [e.g. *Montagner and Jobert*, 1988; *Stutzmann and Montagner*, 1994]. It is possible, perhaps even sensible, to use different correlation lengths for velocity heterogeneity and anisotropy. In addition, the *a priori* standard deviations σ can be chosen independently.

Using the trivial matrix identity

$$\begin{aligned}&\mathbf{G}^{\mathrm{T}}\cdot\mathbf{C}_{\mathrm{d}}^{-1}\cdot(\mathbf{G}\cdot\mathbf{C}_{\mathrm{m}}\cdot\mathbf{G}^{\mathrm{T}}+\mathbf{C}_{\mathrm{d}}) =\\ &(\mathbf{G}^{\mathrm{T}}\cdot\mathbf{C}_{\mathrm{d}}^{-1}\cdot\mathbf{G}+\mathbf{C}_{\mathrm{m}}^{-1})\cdot\mathbf{C}_{\mathrm{m}}\cdot\mathbf{G}^{\mathrm{T}},\end{aligned} \tag{2.17}$$

we can rewrite Eq. 2.15 as

$$\mathbf{m} = (\mathbf{C}_{\mathrm{m}}^{-1} + \mathbf{G}^{\mathrm{T}} \cdot \mathbf{C}_{\mathrm{d}}^{-1} \cdot \mathbf{G})^{-1} \cdot \mathbf{G}^{\mathrm{T}} \cdot \mathbf{C}_{\mathrm{d}}^{-1} \cdot \mathbf{d}. \tag{2.18}$$

Equivalence of regularization and Bayesian approaches

Regularization of a discretely parameterized model and stochastic extensions using an *a priori* model covariance structure have been presented as distinct alternatives [see, e.g. *Scales and Snieder*, 1997] but the approaches are equivalent in many ways [see also *Ho-Liu et al.*, 1989; *Trampert*, 1990; *Montagner et al.*, 1990].

Comparison of Eq. 2.18 with Eq. 2.13 reveals that choosing norm and gradient regularization amounts to identifying

$$\mathbf{C}_{\mathrm{m}}^{-1} = \alpha \mathbf{I} + \beta \mathbf{D}_1^2 \tag{2.19}$$

as a subjective approximation to the inverse of the *a priori* model covariance matrix [*Spakman and Nolet*, 1988]. Damping criteria to stabilize inverse solutions induce a particular form of a model covariance matrix [*Marquardt*, 1970; *Jackson*, 1979; *VanDecar and Snieder*, 1994]. *Yanovskaya and Ditmar* [1990] show that for a Gaussian covariance function of the form of Eq. 2.16, the explicit equivalence holds:

$$\mathbf{m} \cdot \mathbf{C}_{\mathrm{m}}^{-1} \cdot \mathbf{m} = \frac{1}{2\pi} \frac{1}{(\sigma L)^2} \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{L^2}{2} \right)^n \boldsymbol{\nabla}^n \mathbf{m} \cdot \boldsymbol{\nabla}^n \mathbf{m} \tag{2.20}$$

With this choice of covariance function one obtains smooth, infinitely differentiable functions. Eq. 2.20 corresponds to a model norm weighted by an infinite number of finite difference matrices, starting with the zeroth difference, i.e. the traditional model norm. This is equivalent to Eq. 2.19 if the higher-order smoothing terms are neglected.

To illustrate this equivalence we conducted a synthetic experiment to compare the discrete regularization approach with the continuous regionalization for one particular depth.

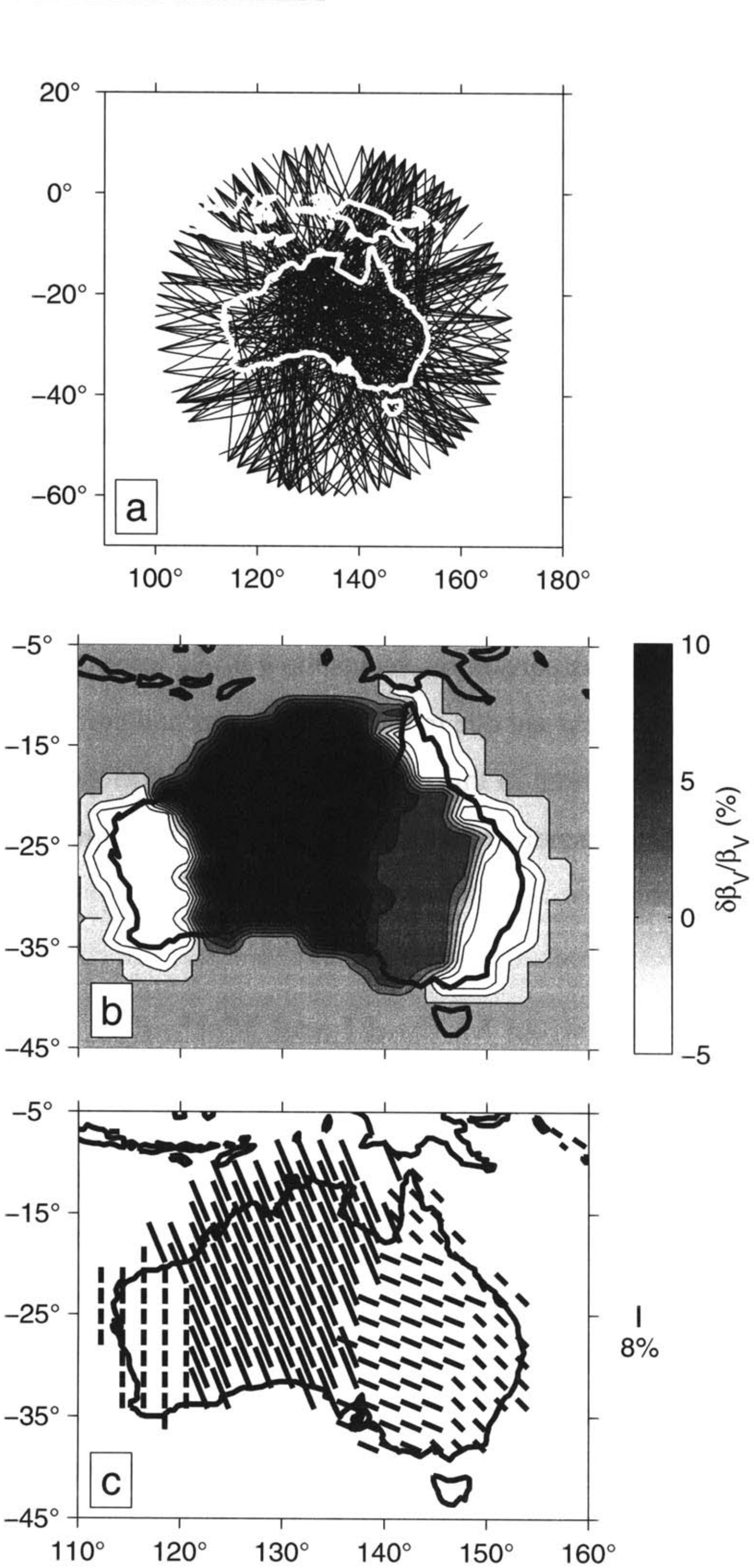


Figure 2-3: Synthetic experiment. (a) Distribution of 450 events and stations randomly distributed on a circle around Australia. (b) Input velocity heterogeneity, in %. (c) Input azimuthal anisotropy, direction of maximum velocity.

For a fabricated path coverage (Fig. 2-3a), (noiseless) data were synthesized from a model of shear wave speed heterogeneity (Fig. 2-3b), azimuthal anisotropy (Fig. 2-3c), or both. Fig. 2-4 shows the inversion results obtained with the discrete method (left column) and the continuous method (right column). Figs. 2-4a and 2-4b show the result for an input model with only heterogeneity (i.e. Fig. 2-3b). Figs. 2-4c and 2-4d show the solution for a model with only anisotropy (i.e. Fig. 2-3c), and Figs. 2-4e and 2-4f give the results for a model which has both heterogeneity and anisotropy. This test demonstrates that a set of inversion parameters (correlation lengths and *a priori* standard deviations) can be found with which both methods yield very similar results indeed. The exact values are unimportant given the synthetic nature of this example. Differences in amplitude recovery are due to the different implicit weighting of the data constraints versus the *a priori* information in both methods. The differences in smoothness are due to the chosen horizontal correlation length (chosen to maximize variance reduction; in this example we chose a correlation length of 175 km for both the velocity heterogeneity and the anisotropy) and the implicit correlation length provided by the damping.

2.4.4 Differences between Method I and Method II

While conceptually similar there are, however, theoretical and practical differences that set our method (Method I) apart from the method used by DK2000 (Method II).

In the previous section we have shown the theoretical equivalence of the Bayesian inversion approach with the regularization method, and that inversion parameters can be found so that both methods produce similar results (Figs. 2-3 and 2-4). One should realize, however, that we could find such parameters only because the input model was known. In inversions for unknown Earth structure this tuning is not possible. The challenge is to find *a priori* horizontal correlation lengths for both the velocity heterogeneity and the anisotropic parameters, and *a priori* standard deviations of the anomalies (σ_β for $\delta\beta_V$ and

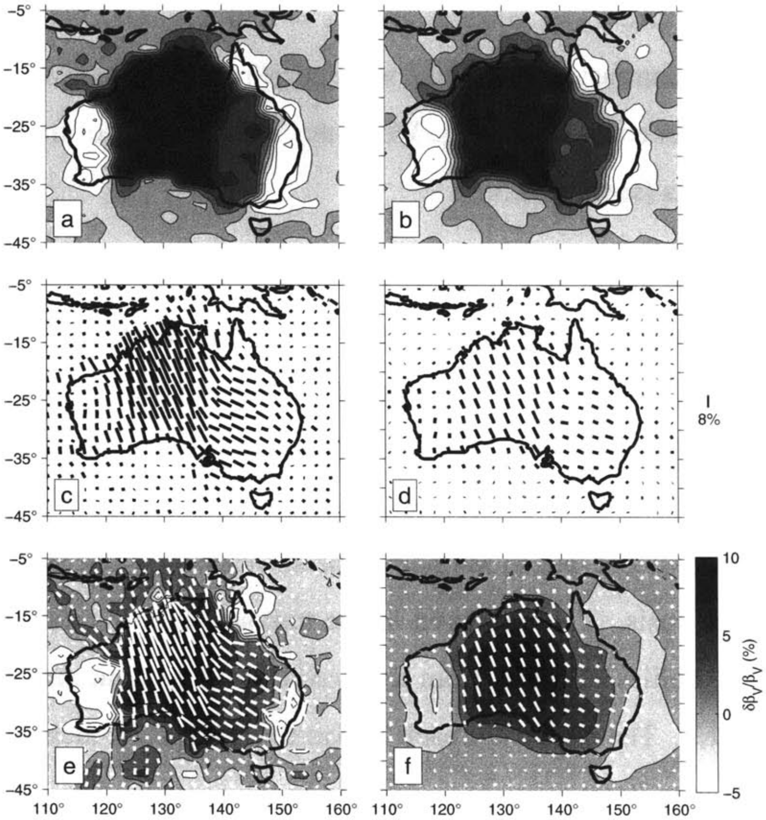


Figure 2-4: Comparison of the discrete iterative approach (left column) with the continuous regionalization inversion (right column). (a,b) Isotropic inversion result with the path distribution of Fig. 2-3a and input heterogeneity of Fig. 2-3b. (c,d) Anisotropy-only inversion with input anisotropy of Fig. 2-3c. (d,e) Joint inversion for velocity heterogeneity and anisotropy with input anomalies as in Figs. 2-3b and 2-3c.

σ_A for A_1 and A_2). The correlation length affects the smoothness of the resulting model and the *a priori* uncertainties dictate the amplitude. Moreover, the difference between σ_β and σ_A controls the trade-off between these quantities. The choice of σ does not affect the location but only the amplitude of the anomalies. Various authors using this technique have made wildly varying choices. *Lévêque et al.* [1998], for example, choose $\sigma_\beta = 100$ m/s and $\sigma_A = 5$ m/s. *Griot et al.* [1998] have $\sigma_\beta = 200$ m/s and $\sigma_A = 200$ m/s, and

Silveira et al. [1998] take $\sigma_\beta = 200$ m/s and $\sigma_A = 130$ m/s. DK2000 take $\sigma_\beta = 50$ m/s and $\sigma_A = 3$ m/s. In contrast, in Method I we incorporate the sensitivity to heterogeneity and anisotropy with an equal weighting. We evaluate the amount of trade-off *a posteriori* in various experiments (see Sec. 2.5.5).

Another difference is the treatment of the third dimension in the 3-D inversion. In Method I the third dimension is an integral part of the minimization problem, entering the equations through the depth-dependent linear constraints on the velocity and anisotropic structure [see *Zielhuis and Nolet*, 1994a]. With the continuous regularization technique of *Montagner* [1986] this could be achieved by retaining the full 3-D covariance matrices $C(\mathbf{r}, \mathbf{r}')$ in the development, but in Method II the shear-wave speed averages are inverted on a layer by layer basis. The data used in the tomographic inversion are given by the shear wave speed profiles obtained from the waveform inversion, and individual depths are treated separately. The results for each 25 km thick layer are then stacked into a 3-D model. While separate inversions are acceptable in the case the data are phase velocities (averages of Earth structure over depth), this "2.5-D" approach is inadequate with shear wave speeds at depth as data since now the depth variation of heterogeneity is determined solely by interpolation of the layer solutions, which ignores the actual depth variation of sensitivity. Furthermore, the only variation of the resolution matrix with depth is given by the errors on shear wave speed, which are, however, not independent from each other. As a consequence, with this method it is impossible to obtain direct estimates of the variation of lateral resolution with depth and the extent of radial smearing in the 3-D model.

Finally, in Method I, we use the inversion algorithm LSQR [*Paige and Saunders*, 1982], which speeds up the computations by several orders of magnitude, but no full resolution matrix can be computed directly [but see *Nolet et al.*, 1999]. In contrast, Method II performs the inversion directly, which is computationally much more expensive but enables the explicit computation of the resolution matrix, a capability we will use in the next section.

2.5 Resolution and robustness analysis

2.5.1 Data coverage

The best sampled regions are the easternmost two-thirds of the Australian continent (Fig. 2-1), with a maximum path density on the eastern rim of the continent (Fig. 2-2). Towards the west, the path density degrades. Fig. 2-2b depicts the number of pairs of rays that come into the cell in a different azimuthal quadrant. This score is more uniform across the continent, but it also diminishes in the western half. The northeastern portion of the continent has a maximum of about 200 such pairs of crossing rays in one cell. In Fig. 2-2c we combined both resolution criteria into a single value ranging from 0 to 1.

2.5.2 Exact calculation of the resolution matrix

The resolution kernel R is given by the reduction of the *a posteriori* covariance C_p with respect to the *a priori* covariance structure:

$$R(\mathbf{r},\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}') - \frac{C_\mathrm{p}(\mathbf{r},\mathbf{r}')}{C_\mathrm{m}(\mathbf{r},\mathbf{r}')}. \tag{2.21}$$

Using the direct Bayesian inversion we calculate the exact resolution matrix both for the actual (Fig. 2-1) and the synthesized uniform (Fig. 2-3a) data coverage and compare this to the qualitative resolution estimates from checkerboard tests. We input the path-average velocity anomaly at a single depth into the Bayesian inversion code, with a correlation length of 175 km for both velocity and anisotropy, and *a priori* model uncertainties of 120 m/s for $\delta\beta_V$, A_1 and A_2 (see Eq. 2.7). The value of these choices only affect the amplitude of the result; equating σ_{β_V} and σ_A best compares with the equal weighting used in Method I.

The results for the synthetic uniform path coverage of Fig. 2-3a are plotted in Fig. 2-5a (resolution of $\delta\beta_V$), Fig. 2-5b (resolution of A_1), and for the actual coverage of Fig. 2-1

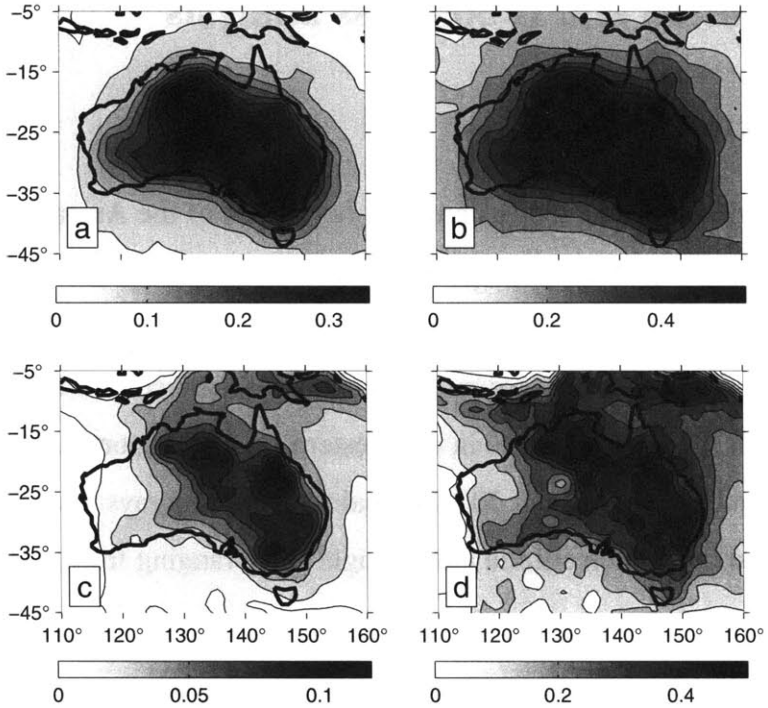


Figure 2-5: Resolution tests. Exact horizontal resolution computation for the synthetic path density experiment (a-b) and the real data coverage (c-d). Plotted is the diagonal value of $R(\mathbf{r}, \mathbf{r}')$ for the isotropic velocity solution $\delta\beta_V$ (a,c), and the cosine term A_1 (b,d). The resolution for the sine term A_2 is virtually indistinguishable from that of the cosine term.

in Fig. 2-5c–d. The resolution for A_2 is virtually identical to that of A_1 and is therefore omitted. The comparison of the synthetic resolution with the resolution from the actual data gives us a feel for the values we can expect where the model is well resolved. As anticipated, the solution in the eastern two-thirds of the continent is markedly better than in the west.

2.5.3 Checkerboard tests

A popular way of assessing the resolution in tomographic models is to calculate the recovery of a synthetic (e.g., checkerboard) input pattern. By themselves checkerboard tests are

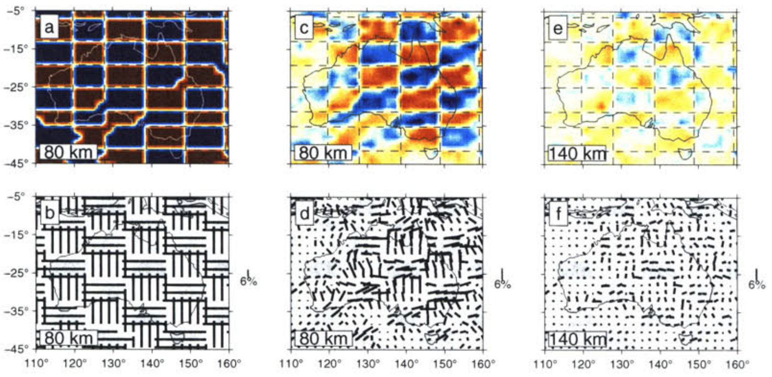


Figure 2-6: Checkerboard test for vertical smearing. An input pattern of velocity heterogeneity at 80 km (2-6a) superimposed on a pattern of azimuthal anisotropy (2-6b; shaded for clarity) offset by half a cell was used to generate synthetic data. The inversion for both wave speed anomalies and anisotropy yields a recovery at 80 km given in 2-6c and 2-6d. Due to the radial averaging properties of the surface wave kernels and the smearing in the inversion, some of the input structure at 80 km is also mapped into the next layer at 140 km (2-6e and 2-6f). Panels a, c, and e are plotted on the same color scale.

difficult to interpret. It is often hard to guess the exact pattern of the anomalies without knowledge of the input function, which strongly degrades their diagnostic value. Also, the damping parameters can be chosen to reproduce the input pattern optimally, a luxury which is not available when choosing the damping needed to model the actual data for unknown Earth structure. These and other, intrinsic, limitations of checkerboard tests have been pointed out by several authors [e.g. *Lévêque et al.*, 1993]. Hence, we will use the results of checkerboard tests in a qualitative way only, and compare them to the exact resolution computed in Section 2.5.

Since the sensitivity matrix maps the input anomalies onto the synthetic data according to the independently constrained depth-dependent kernel functions, the solution at one particular depth is related to the solution at adjacent depths in Method I. Therefore, we can use checkerboard tests to get a sense of the radial resolution and vertical smearing of the tomographic model (Fig. 2-6) as well as the horizontal resolution and its depth variation (Fig. 2-7).

On a $10 \times 10°$ grid, we have created two staggered checkerboard patterns, one for heterogeneity and one for azimuthal anisotropy (Fig. 2-6a,b). Synthetic data were generated for a model with this input pattern introduced at 80 km depth. The inversion result is shown at this same depth (80 km; Fig. 2-6c,d) as well as at the next level of the solution (140 km; Figs. 2-6e,f). Except for the diminished amplitudes the velocity input pattern is reasonably well recovered, including the sharp gradients between the cells. This and other tests suggest that large wave speed contrasts occurring over as little as 250 km can be resolved by our data. On the other hand this test does not tell us the amplitude of the minimum wave speed variation that can be detected at that same wavelength.

The N–S and E–W pattern of azimuthal anisotropy is fairly well recovered in the central and eastern part of the continent but the angular variability is at least 20° (Fig. 2-6d). Indeed, in large parts of our study region it is not obvious from Fig. 2-6d alone that the input pattern was a regular checkerboard pattern, which hampers the interpretation of the results of such tests significantly.

Radial smearing affects our results, but a comparison of the amplitudes of the solutions at both depths reveals that the level of contamination is small.

In Fig. 2-7 we illustrate the variation of horizontal resolution with depth. We have plotted the output to input ratio (in percent) of an isotropic checkerboard test with patches of about 550 km each ($5° \times 5°$, performed at 80 (Fig. 2-7a) and 200 km (Fig. 2-7b)). At 80 km, we observe that up to 80 % of the input anomaly is recovered in the best resolved regions; at 200 km this number has dropped to about 60 %, but in most of the study region the amplitude recovery is less than 50%. Fig. 2-7c and 2-7d plot the difference in angle between the synthetic input pattern of Fig. 2-6b in the presence of the velocity heterogeneity of Fig. 2-6a and the recovery of the anisotropy from a joint inversion for heterogeneity and

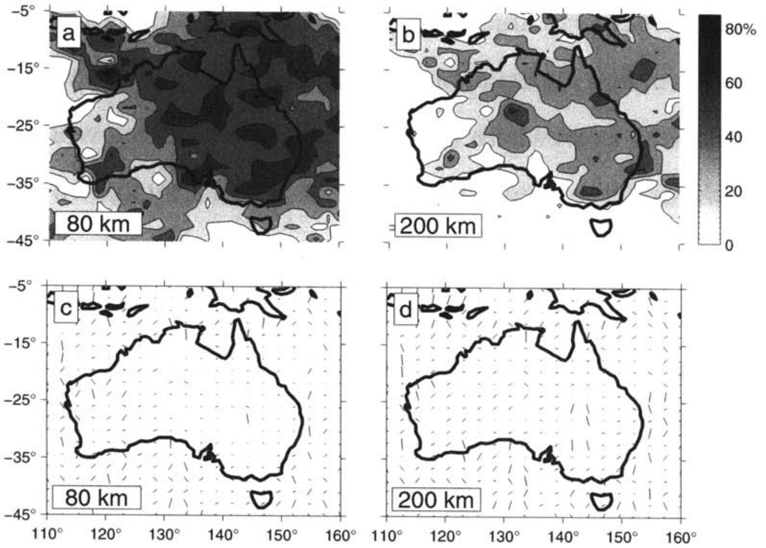


Figure 2-7: Checkerboard test for depth-varying horizontal resolution. (a,b) Percentage of input anomaly recovered of an isotropic checkerboard pattern at 80 and 200 km depth. (c,d) Difference in angle between input and output anisotropy in a joint inversion for a pattern of heterogeneity and anisotropy as in Fig. 2-6a,b at 80 and 200 km depth. Bars scale with the magnitude of the difference such that excellent recovery is depicted by a dot.

anisotropy, for an input pattern and model retrieval at 80 (also shown in Fig. 2-6d) and 200 km depth.

We conclude that, except in western Australia, our final model will give a fair representation of the velocity and anisotropy structure of the upper mantle under Australia.

2.5.4 Model robustness

To evaluate the robustness of the models we have performed an extensive series of tests with up to 40% of the data randomly removed. We also calculated the mean model obtained from 500 inversions with 5% of the data randomly removed, as well as the mean of 500 inversions performed on a data set to which noise was added (with signal-to-noise ratios up to 15 dB). The average models obtained in this way were visually indistinguishable from our final model. For this class of mixed-determined inversion problems it is not possible to use inversion experiments like these to interpret the variance of the model parameters quantitatively everywhere [*Nolet et al.*, 1999]. Areas with low resolution are always damped towards the reference model, and this perfectly known state of total uncertainty has therefore almost zero variance in such inversions. However, for the well resolved areas in the east, the standard deviation obtained is everywhere less than 0.8%, which is much smaller than the magnitude of the anomalies we obtain in the final model. This holds for all depths of our final model.

Furthermore, we used the different parameterizations shown in Fig. 2-2 to study the robustness of the patterns of heterogeneity and anisotropy.

2.5.5 Trade-off of heterogeneity and anisotropy

Several previous studies [e.g. *Vinnik et al.*, 1992; *Gaherty and Jordan*, 1995; *Clitheroe and van der Hilst*, 1998; *Girardin and Farra*, 1998; *Debayle and Kennett*, 2000a] have suggested that the upper mantle beneath Australia is anisotropic. However, we need to show that the inclusion of the azimuthal anisotropy into the inversion problem is required by our data.

The increase in variance reduction of isotropic-only (82.6 % in our model) versus azimuthally anisotropic models (87.5 % for our model) is an indication of the improvement of data fit when anisotropy is included, but its statistical significance is hard to assess. Methods to evaluate the improvement of data fit with an increasing number of parameters, such as Fischer tests, are only valid for overdetermined problems [*Menke*, 1989] and can thus not be used here. One indication of improvement of data fit is by comparing the increase in variance reduction from an isotropic model ($\delta\beta_V$) parameterized in 7059 surface cells of a certain area with an anisotropic inversion ($\delta\beta_V, A_1, A_2$) of 2353 cells of an area three times as large. The number of inversion parameters is the same in either case, but the variance

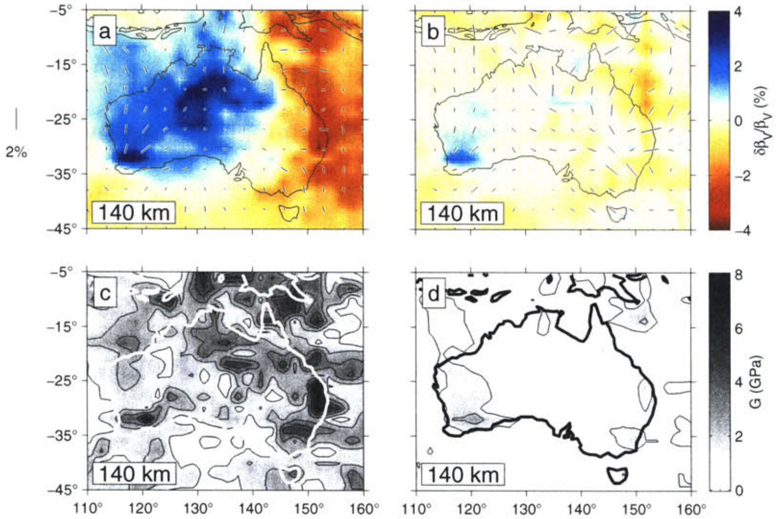


Figure 2-8: Trade-off of heterogeneity with anisotropy at 140 km. (a) Inversion for heterogeneity and anisotropy of synthetic data calculated from the isotropic part of our model. With perfect resolution and control of the trade-off, the anisotropy retrieved should be zero everywhere. (b) Inversion of synthetic data calculated from the anisotropic part of the model for heterogeneity and anisotropy. In this case perfect resolution should have resulted in zero heterogeneity. (c) Magnitude of anisotropy (in GPa) of the actual model (see Fig. 2-9). (d) Magnitude of the erroneous anisotropy of Fig. 2-8a. Contour lines plotted at 1 GPa interval for both panels.

reduction increases from 83.5 to 87.0%. The inclusion of anisotropy results in an increase of variance reduction even when the number of free parameters is left unchanged.

We explore the significance of the azimuthal anisotropy with another synthetic experiment designed to test to which extent uneven data coverage and heterogeneity can combine to produce an anisotropic bias. We represent our full solution by $(\delta\beta_V, A_1, A_2)$. In one test we inverted data synthesized from the heterogeneity part only, (i.e. $(\delta\beta_V, 0, 0)$) for both heterogeneity and anisotropy. From the result we can then infer how much anisotropy is produced by uneven data coverage and heterogeneity. In a second test we invert synthetic data from $(0, A_1, A_2)$ for heterogeneity and anisotropy to assess the projection of artificial anisotropy into heterogeneity.

Fig. 2-8 summarizes the test results for a depth of 140 km. Fig. 2-8a and 2-8b show the inversion of models which should have zero residual anisotropy and zero residual velocity heterogeneity, respectively. The magnitude of the model anisotropy, plotted in Fig. 2-8c, can be compared to the magnitude of the inferred anisotropy (Fig. 2-8d). For all depths considered, the average anisotropic bias is less than 20 % of the magnitude required to satisfy the data in the actual model.

In regions of uneven data coverage some artificial anisotropy is clearly produced. The direction of the artificially induced anisotropy is especially sensitive to uneven sampling, but, in general, the magnitude of the bias is small. In Western Australia, however, much of the anisotropy present in the actual model may well consist of bias due to the scarcity of crossing paths in that region. This makes it impossible to distinguish the fast velocity anomaly, which is clearly present in the data, from a possible alignment of the fast axis of anisotropy.

2.6 Results

In Figs. 2-9 and 2-10 we present models for heterogeneity and azimuthal anisotropy of Australia, using Method I. The inversions were performed on a regular (equal-area) grid.

The azimuthally isotropic component $\delta\beta_V$ shows many of the characteristics that were present in our earlier models [*Zielhuis and van der Hilst*, 1996; *Simons et al.*, 1999]. Low wave speed anomalies mark the eastern Phanerozoic rim of the continent. They are associated with thermal anomalies, recent volcanism and high attenuation [*Mitchell et al.*, 1998]. Wave speeds in the central and western region are high, but the correlation between the high velocities and the age of the overlying crust is scale-dependent [*Simons et al.*, 1999].

In most regions the anisotropy suggested by the data exceeds the anisotropic bias (see also Fig. 2-8). At shallow depths in the continental lithosphere the level of anisotropy exceeds 4% in many regions, and the lateral variability in directions is much larger than

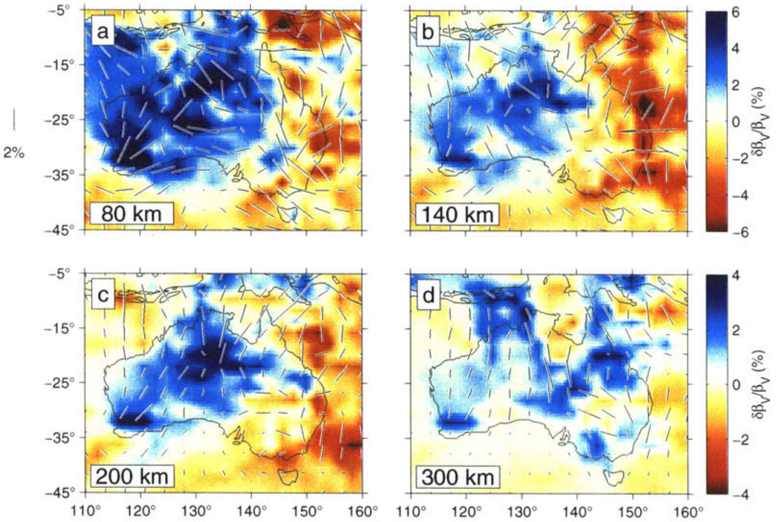


Figure 2-9: Inversion results at different depths. Percentages from reference model with background shear-wave velocities of 4500 m s^{-1} at 80 and 140 km depth; 4513.5 m s^{-1} at 210 km; 4635 m s^{-1} at 300 km.

the nominal uncertainty of 20°. The magnitude of the anisotropy (see, e.g., Fig. 2-8c) is consistent with laboratory estimates of upper mantle materials, which range from 2 to 7 GPa [*Peselnick and Nicolas*, 1978; *Estey and Douglas*, 1986]. There is no obvious relationship between wave speeds and anisotropy, neither in magnitude nor in direction, nor is there a clear correlation between the anisotropy and the large-scale geological structure (Fig. 2-11). In contrast, *Babuška et al.* [1998] have shown that there are systematic variations in radial anisotropy which are related to the age of continental provinces in North America and Eurasia.

The cross-sections shown in Fig. 2-10 suggest a depth to the base of the continental lithosphere of up to 250-300 km [see also *Simons et al.*, 1999; *Fischer and van der Hilst*, 1999], and Fig. 2-11 shows that the wave speed difference between the large-scale tectonic provinces is confined to the shallowest 300 km. The central-Australian region is underlain by the thickest high-velocity lid, which may be an effect of preferential preservation [*Simons et al.*, 1999; *Lee et al.*, 2001]. Similar lithospheric thicknesses are obtained for other

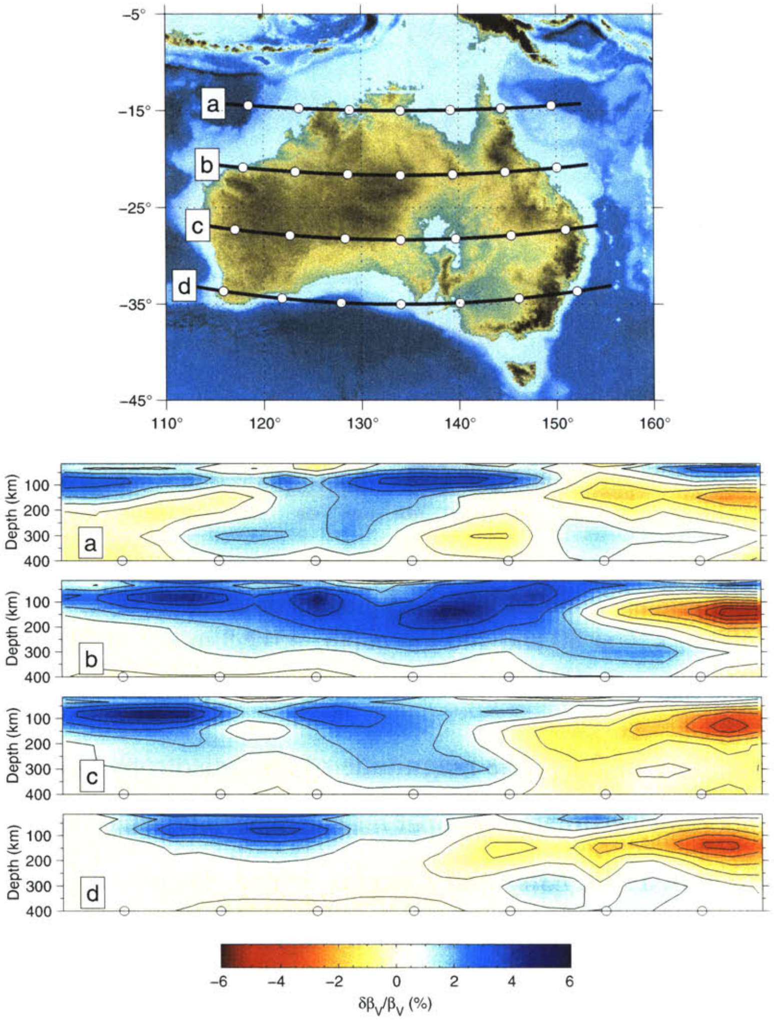


Figure 2-10: Cross-sections at various latitudes. Fifteen contour lines are plotted between -6 and 6%.

cratonic regions [for a review, see *Artemieva and Mooney*, 2001]. Even with the uncertainty estimates given in Sec. 2.5.5, the magnitude of the anisotropy does not fall off with depth as rapidly as the magnitude of the velocity heterogeneities. The average anisotropic magnitudes reach a maximum around 100 km (Fig. 2-11) and diminish at larger depth, in accordance with observations elsewhere [*Babuška et al.*, 1998].

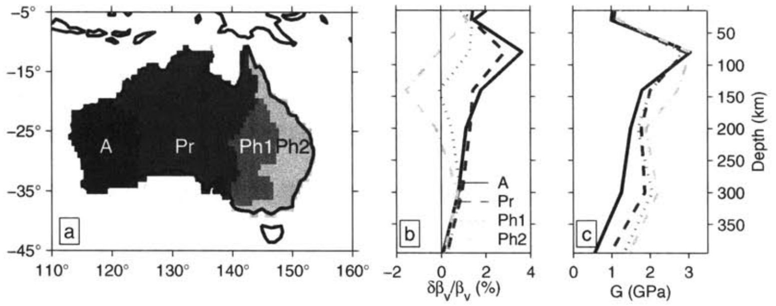


Figure 2-11: Average heterogeneity and anisotropy in function of crustal age. A coarse regionalization of the Australian continent consists of a predominantly Archean domain (A), the central Proterozoic region (Pr) and domains of Phanerozoic age (Ph1 and Ph2). While wave speed anomalies become progressively more positive with increasing crustal age [see also *Simons et al.*, 1999], no defining relationship is found in the magnitude of the anisotropy.

The pattern of anisotropy changes with increasing depth. The solutions at 80 and 140 km depth show a complex pattern of azimuthal anisotropy (short correlation wavelengths), whereas in the layers at 200 and 300 km a simpler, smoother pattern (long correlation wavelengths) is observed (Fig. 2-9). For instance, at shallow levels, the direction of fast axis is E–W to WNW–ESE in central Australia, which coincides with the structural trends in the Alice Springs orogeny, but this changes to N–S at larger depth. The character of the anisotropy and the length scale over which it varies seem to change rapidly across the 150 and 200 km depth interval (Fig. 2-11).

The velocity heterogeneity revealed in Fig. 2-9 was reproduced satisfactorily with the irregular parameterization (Fig. 2-2c), albeit with somewhat reduced amplitudes. In the western regions, where the path coverage is demonstrably worse, our fine equal-area grid over-parameterizes the structure but this does not introduce noticeable artifacts. Within the uncertainty of 20°, the different parameterizations yield similar directions of fast axes of anisotropy. The pure-path average solutions (on the grid shown in Fig. 2-2d) at 80, 140 and 200 km were very similar to the averages of the wave speeds imaged on the equal-area grid. Differences appear at larger depth, suggesting again that below 200 km most correspondence between wave speed structure and surface geology is lost. The central and northern region of Australia have stable N–S directions of maximum anisotropy, both under the equal-area and the tectonic parameterization, but the magnitudes are much reduced in the latter. In the eastern portion of the continent a tectonic parameterization reveals little anisotropy, at any depth. This could indicate that the anisotropy is small, i.e., at or below the level of the anisotropic bias, or that through much of the depth range the anisotropy varies on length scales smaller than the main tectonic units of the Australian continent.

2.7 Discussion

2.7.1 Comparison with other models

Global models of Rayleigh wave phase velocities [e.g., *Trampert and Woodhouse*, 1996; *Ekström et al.*, 1997; *van Heijst and Woodhouse*, 1999] predict faster than average wave speeds for the Archean and Proterozoic portions of the Australian continent (which correspond roughly to the western two-thirds of the continent). This pattern is corroborated by our regional model, but significant variations in heterogeneity and anisotropy exist at smaller scales than is apparent from global models. *Gaherty and Jordan* [1995] and *Debayle and Kennett* [2000b] found large polarization anisotropy in the Australian continent above 250 km. Our Rayleigh wave model does not contain information about polarization anisotropy.

Some global models suggest that the high wave speeds associated with ancient continents extend to depths exceeding 400 km [e.g. *Su et al.*, 1994; *Masters et al.*, 1996]. In our model, the fast ($>1\%$) velocity anomalies in the center of the continent form a continuous high wave speed lid down to at least 250 km, but in general the keels seem thinner than that. Eastern Australia, which is the most recently tectonized region of the continent, is characterized by a pronounced low velocity zone between 100-200 km depth underlain by fast anomalies at a depth of about 300 km (see Fig. 2-10b). The origin of the latter is uncertain and it may not be related to the overlying lithosphere. The wave speed heterogeneity of the Australian continent seems to correlate with crustal evolution and lithospheric age, at least to depths of about 200 km (Fig. 2-11) [see also *Simons et al.*, 1999; *Fischer and van der Hilst*, 1999]. At depths greater than 250 km there is no unambiguous relationship between crustal age and wave speed. The anisotropy does not seem to change on the length-scale of the major geological domains at any depth, although its magnitude remains higher at larger depth beneath the regions affected by the most recent tectonic activity (the easternmost parts of the continent). We conclude that the thickness of the Australian lithosphere does probably not exceed 275±50 km. This is in agreement with recent observations for the Canadian Shield [*Frederiksen et al.*, 2001] and South-Africa [*James et al.*, 2001].

Qualitatively, the wave speed and anisotropy maps presented in Fig. 2-9 argue for a change in regime between the first two depths (80 and 140) and the deeper parts of the solution (200 and 300 km). In terms of anisotropy, the first regime is characterized by a complex pattern of strong anisotropy, whereas the second presents more smooth variations of weaker anisotropy and a tendency towards a N–S fast axis in northern Australia. DK2000 observe a similar change at 150 km depth (but in our model the level of anisotropy beneath 200 km remains at a higher level), and this observation is generally consistent with global and regional studies [*Montagner and Tanimoto*, 1991; *Lévêque et al.*, 1998].

2.7.2 Comparison with results from shear-wave splitting

From the analysis of refracted *S* waves bottoming under North Australia, *Tong et al.* [1994] obtain shear wave splitting times exceeding 4 s in certain places. *Vinnik et al.* [1989] and *Barruol and Hoffman* [1999] report no perceptible splitting for *SKS* arrivals at station CAN in southeastern Australia, but *Girardin and Farra* [1998] interpret *P*-to-*S* converted phases to be indicative of a two-layer anisotropic system under that station with nearly perpendicular fast axes distributed over comparable depth intervals so the contributions to splitting cancel out. *Clitheroe and van der Hilst* [1998] present evidence for complex, frequency-dependent splitting from *SKS* and *SKKS* phases, with split times generally less than 1 s in the Precambrian central part of Australia. Their fast axis directions do not align with the direction of absolute plate motion, which is roughly SSW-NNE, but rather suggest alignment with crustal fabric or a complex lattice preferred orientation of olivine in the subcrustal lithosphere. However, *Özalaybey and Chen* [1999] find no convincing evidence for splitting at all. They argue that the analysis of a small amount of data over a limited azimuthal range and complex multipathing effects due to heterogeneities at depth may contribute to erroneous splitting measurements, especially at the frequencies higher than 0.3 Hz considered by *Clitheroe and van der Hilst* [1998].

SKS splitting times represent an integrated effect over a large depth range h due to the subvertical propagation of the phases. It is possible to relate our observations of surface-wave anisotropy to shear-wave splitting directly [*Montagner et al.*, 2000]. For simple media with a horizontal symmetry axis, the maximum splitting time is given by

$$\delta t = \sqrt{\left(\int_0^h \sqrt{\frac{\rho}{L}}\frac{G_c(z)}{L(z)}\,dz\right)^2 + \left(\int_0^h \sqrt{\frac{\rho}{L}}\frac{G_s(z)}{L(z)}\,dz\right)^2}\,, \tag{2.22}$$

and the splitting direction Ψ is calculated by

$$\tan 2\Psi = \int_0^h \sqrt{\frac{\rho}{L}}\frac{G_s(z)}{L(z)}\,dz \Big/ \int_0^h \sqrt{\frac{\rho}{L}}\frac{G_c(z)}{L(z)}\,dz. \tag{2.23}$$

Applying Eqs. 2.22 and 2.23 to our model (depicted in Fig. 2-9) yields considerable splitting times throughout the continent (Fig. 2-12a). The results of the surface wave inversion presented here suggest higher levels of azimuthal anisotropy than inferred from *SKS* by either *Clitheroe and van der Hilst* [1998] or *Özalaybey and Chen* [1999]. For compari-

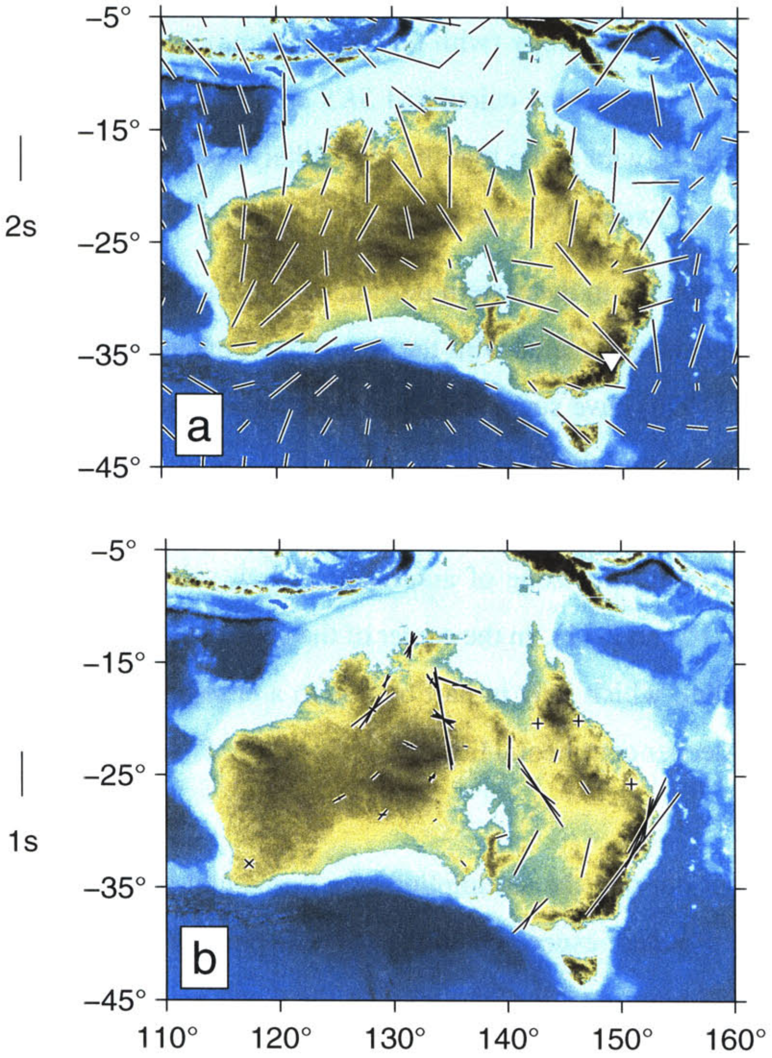


Figure 2-12: (a) *SKS* splitting directions predicted from the solution given in Fig. 2-9. Station CAN is represented by the white triangle. (b) Shear-wave splitting observations [from *Iidaka et al.*, 2001].

son, we have plotted a recent data set of splitting times and directions in Fig. 2-12b [from *Iidaka et al.*, 2001]. A mismatch between surface-wave and body wave anisotropy has been observed on other continents as well [*Saltzer et al.*, 2000; *Freybourger et al.*, 2001].

In addition to calculating *SKS* times due the entire depth range of anisotropy, we have investigated the incremental effect on the splitting times from successively deeper levels in our model. The transverse isotropy (with horizontal symmetry axis) between 140 and 400 km depth contributes most to the calculated *SKS* split times. This is in good agreement with *Tong et al.* [1994], who argue that the most likely source region of anisotropy is the asthenospheric zone between 210–410 km.

The fast axis directions in the eastern portions of the continent follow the rough outline of the high wave speeds associated with the Australian Precambrian (Fig. 2-9a). *Clitheroe and van der Hilst* [1998] have suggested this might be related to the crystal alignment by asthenospheric flow around the irregularly shaped eastern edge of the deep continent [see also *Fouch et al.*, 2000]. This is supported by the fact that the split times originate mainly at a depth near 140 km, in a zone of anomalously low velocities and, presumably, low viscosities (Figs. 2-10 and 2-11). In the center of the continent, the predicted splitting times are somewhat smaller, which may in part be due to a change from E–W in the shallower part to N–S in the deeper continental lithosphere [see, e.g. *Silver and Savage*, 1994; *Saltzer et al.*, 2000].

We should point out that the assumption of a horizontal orientation of the fast axis is not unlikely to be violated, especially in the shallow parts of the continental lithosphere. Further analysis will require taking into account the depth variation of fast axis directions, the possibility of dipping axes and frequency-dependent effects.

Our observations point to a complex geometry of the continental anisotropy at various length scales. We cannot rule out that strong anisotropy in the top 80–140 km exists if the direction of fast axes varies rapidly on a scale that is smaller than can be resolved by our data. With the magnitude of anisotropy increasing and the lateral correlation lengths decreasing towards the surface, it is conceivable that strong, rapidly changing azimuthal anisotropy in the top 140 km diminishes the effective birefringence of vertically propagating body waves without similarly influencing the long period surface wave data used in our study [*Saltzer et al.*, 2000].

2.8 Conclusions

Careful data selection and manual processing of stations operated by AGSO, IRIS and GEOSCOPE and data from the portable SKIPPY experiments have produced a data coverage that is superior to any other continental surface-wave study in terms of density and azimuthal distribution.

We have compared two approaches for the inversion of path-average Earth models for 3-D models of S-heterogeneity and azimuthal anisotropy. Method I [*Zielhuis and van der Hilst*, 1996; *Simons et al.*, 1999] performs a 3-D inversion for heterogeneity and anisotropy based on a discrete parameterization at the surface with depth basis functions representing independently constrained estimates of radial Earth structure. Method II, used by *Debayle and Kennett* [2000a], uses a continuous regionalization algorithm which performs a separate inversion for each depth layer. Discretely parameterized regularization methods and Bayesian continuous regionalization methods can lead to similar results. The differences between our results and the model by *Debayle and Kennett* [2000a] can be due to (1) the independence of the constraints from the waveform inversion used in Method I and the decorrelation of their errors; (2) the treatment of the third dimension (3-D inversion in Method I but a stack of layers in Method II); (3) somewhat different distribution of earthquake sources; (4) different data selection criteria (e.g., the frequency bands used); (5) balancing of fundamental-mode and higher-mode information through the use of one (Method II) or more (Method I) data windows; and (6) the crustal models used to apply shallow corrections.

The advantage of Method II is that a formal resolution matrix can be computed, but this comes at a substantial computational expense. We have shown that the information contained therein is qualitatively similar to *ad hoc* checkerboard response tests. Furthermore, the formal resolution is 2-D since the full 3-D covariance matrices of data and model are not used in Method II. As a result, variations in lateral resolution with depth and radial smearing can not be investigated efficiently with Method II. In contrast, Method I allows us to evaluate the uncertainty as well as the radial smearing in all parts of the 3-D model.

We have obtained a new shear wave speed model for the Australian upper mantle which includes azimuthal anisotropy. The inversion presented here was based on 30% more data than used by *Simons et al.* [1999], but we confirm their conclusions regarding the aspherical variation of isotropic wave speed in the Australian upper mantle. In particular, the conclusion that the high correlation between the westward increase in wave speed and lithospheric age hold for depths less than 200 km only is confirmed.

We have determined the spatial variation of azimuthal anisotropy (transverse anisotropy with a horizontal axis of symmetry) and investigated the trade off between heterogeneity and anisotropy. We conclude that with the exception of westernmost Australia, where data coverage is worst, the level of anisotropy that is produced by a combination of uneven data coverage and heterogeneity is at most 20% of the magnitude of anisotropy required by the data. In general the uncertainty in the angle of fast polarization is about 20°; locally the directions are more tightly constrained, but there are also regions where the uncertainty is much larger, in particular in western Australia. In the top 150 km the inversions yield fairly strong anisotropy (up to 4%) but the orientation of the fast axis changes in a complex manner. In contrast to the wave speed heterogeneity, there is no obvious relationship between anisotropy and the large scale geological domains that we considered, suggesting that the length scale for variation is much smaller than the crustal elements identified by *Shaw et al.* [1995]. The horizontal length scale over which anisotropy changes increases between 150 and 200 km depth.

In the northern central region we observe a predominance of N–S directions throughout much of the depth range. Within the substantial uncertainty on their direction, these N–S oriented domains are subparallel to the direction of absolute plate motion [*DeMets et al.*, 1990]. If shearing due to plate motion is the dominant process, then it is perhaps surprising that so much of the N–S anisotropic directions are confined to Central Australia and start at depths as shallow as 140 km or even less, i.e., well within what we interpret as continental lithosphere. If fossil anisotropy is the dominant process, then one needs to explain the change in length scale and behavior between 150 and 200 km depth for the regions outside of the Central Australian domain. Moreover, complex flow around stable continental interiors could yield anisotropic fabrics in relationships that are not straightforward to correlate with a single motion vector of the entire Indo-Australian plate. Our observations therefore call for much further study of the effect of plate motion on the shear fabric in the subcontinental asthenosphere.

We have used the approach suggested by *Montagner et al.* [2000] to predict *SKS* splitting times and directions from our 3-D anisotropic model. Much as the original inversion result for depths above 200 km, the predicted *SKS* splitting map does not reveal predominance of a direction parallel to plate motion. The first order difference between our result and observations of *SKS* and *SKKS* splitting is that the surface waves seem to prefer a higher level of horizontal transverse anisotropy than the body waves. This is true, in particular, for the Precambrian shields where shear wave birefringence yielded split times of less than 1 s whereas the surface wave inversion predicts times that are at least twice that large. The discrepancy needs to be investigated further, but we suggest frequency-dependent effects or the presence of dipping axes of symmetry might be responsible. We argue that the observed shear-wave splitting can be reconciled with the inferences from surface-wave propagation with a conceptual model of anisotropy in which the correlation wavenumbers and magnitude of anisotropy increase towards the surface.

2.9 Acknowledgments

Funding for the instruments and field campaigns involved in the SKIPPY and KIMBA experiments was provided by the Australian National University. We thank Eric Debayle, Don Forsyth, Hrafnkell Kárason, Brian Kennett, Sergei Lebedev and Albert Tarantola for stimulating discussions, Takashi Iidaka for providing his shear-wave splitting analyses and Mark Leonard and the IRIS Data Management Center for data access. Guust Nolet, Steve Ward, and an anonymous reviewer provided comments which improved the manuscript. RvdH and FJS thank the Institut de Physique du Globe in Paris, where parts of this paper were written. This research was supported by NSF (grant EAR-0001136), and it is IPGP contribution # 1793.

2.10 Appendix

In the following, we describe the construction of the linear tomographic inversion problem. With the exception of a few behind-the-scenes computational details, our method of *S* wave velocity tomography (Method I) differs little from the procedure described in *Zielhuis and Nolet* [1994a] (herein after ZN94). The present outline is added to make this paper self-contained.

Every seismogram yields a set of data, q_k^p, where p indexes the path $p = 1, ..., M$ and and $k = 1, ..., K(p)$ indexes the number of independent constraints that have been retained after applying some error criterion: the q_k^p are independent in that they have a diagonal covariance matrix; they have been scaled to unit variance. The statistical independence of the linear constraints is one of the key features in Nolet's method. The modal Fréchet kernels [for an illustration, see *Simons et al.*, 1999, Fig. 4] are not orthogonal, which is problematic when converting phase velocity measurements in a particular frequency band into depth-dependent shear velocity models.

Using the notation:

$$\langle f|g\rangle \equiv \int_0^a f(r)g(r)\,dr, \tag{2.24}$$

where the integration goes from the center ($r = 0$) to the surface ($r = a$) of the Earth, we can rewrite Eq. (16) of ZN94 as:

$$q_k^p = \int_{P_p} \langle G_k^p(r)|\delta\beta(r,\theta,\varphi)\rangle\; d\Delta, \tag{2.25}$$

where $d\Delta$ indicates the incremental epicentral distance along the great circle path P_p. The 3-D shear wave speed anomaly is denoted as $\delta\beta(\mathbf{r}) = \delta\beta(r,\theta,\varphi)$ and $G_k^p(r)$ is a kernel function which indicates how the data constraint q_k^p is related to a change in model parameter $\delta\beta$. The precise form of $G_k^p(r)$ is determined by the error structure of the parameters determined in the nonlinear waveform inversion; practically, as more (and better) information is available on the velocity structure, the total number k increases to some number $K(p)$ for a particular path. Insofar as this information is constrained by higher modes or lower-frequency fundamental modes of the seismogram, the sensitivity of the kernel $G_k^p(r)$ will shift to deeper structure as k increases. The value of q_k^p is influenced by the local path reference model that was used in obtaining the fit; the velocity anomaly is with respect to a regional 1-D background model common to all paths. Shallow structure was accounted for by finding the best crustal model for the source and receiver region and the average path, both on the basis of independent information from receiver functions and on the effect on the waveform fit (see text).

As in ZN94, we expand the velocity anomalies onto a set of equal area surface cells parameterized as $l_t(\theta, \varphi)$ for $t = 1, ..., T$ where $l_t(\theta, \varphi) = 1$ if the horizontal coordinates (θ, φ) are inside the cell, and 0 elsewhere. The radial basis functions are denoted as $h_j(r), j = 1, ..., J$; they can be triangular functions (as in ZN94 and this study), discrete layers, or splines. The surface and depth parameterization describe a set of basis functions

$$s_i(\mathbf{r}) = l_t(\mathbf{r}/r) h_j(r), \tag{2.26}$$

with $i = 1, ..., N$ and $N = TJ$ upon which to expand $\delta\beta(\mathbf{r})$ (Eq. (17) in ZN94):

$$\delta\beta(\mathbf{r}) = \sum_i^N b_i s_i(\mathbf{r}). \tag{2.27}$$

Inserting Eqs. 2.27 and 2.26 into Eq. 2.25 we obtain

$$\begin{aligned} q_k^p &= \int_{P_p} \langle G_k^p(r) | \sum_i^N b_i l_t(\theta, \varphi) h_j(r) \rangle \; d\Delta \\ &= \langle G_k^p(r) | \sum_i^N b_i \left\{ \int_{P_p} l_t(\theta, \varphi) \, d\Delta \right\} h_j(r) \rangle \end{aligned}$$

$$
\begin{aligned}
&= \langle G_k^p(r) | \sum_i^N b_i L_t^p h_j(r) \rangle \\
&= \sum_i^N b_i L_t^p \, \langle G_k^p(r) | h_j(r) \rangle \, . \qquad (2.28)
\end{aligned}
$$

We've introduced L_t^p, the path length of path p in surface cell t. To get as many equations as there are paths, we collapse the individual k constraints (this represents a departure from the implementation by ZN94) and write

$$
\sum_k^{K(p)} q_k^p = \sum_i b_i L_t^p \, \langle \sum_k^{K(p)} G_k^p(r) | h_j(r) \rangle \, . \qquad (2.29)
$$

All k constraints of one path sample the same cells on the surface. The $M \times N$ matrix $\mathbf{G}$ contains the sensitivity of the paths to the velocity parameters b_i; its elements are given by

$$
G_{pi} = L_t^p \, \langle \sum_k^{K(p)} G_k^p(r) | h_j(r) \rangle \, , \qquad (2.30)
$$

so $\mathbf{G}$ is dependent on the path lengths in the surface cells, and the sensitivity to depth structure which is influenced by the mode structure of the seismogram and the diagonalized error covariance matrix of the waveform inversion. Eq. 2.30 is the equivalent to Eq. (18) in ZN94. Obviously, $\sum_k^{K(p)} q_k^p$ defines an $M \times 1$ data matrix $\mathbf{d}$ and b_i form an $N \times 1$ matrix $\mathbf{m}$ containing the unknowns. We have now successfully cast our problem as a linear inverse problem:

$$
\mathbf{d} = \mathbf{G} \cdot \mathbf{m}. \qquad (2.31)
$$

To examine the behavior of the phase speeds of the modes for the group velocity-frequency windows under consideration, we calculate the modal eigenfunctions and the derivatives of their phase speeds with respect to elastic perturbations. Fig. 2-13 depicts an isotropic background model used in the waveform fitting [for examples, see *Zielhuis and* *van der Hilst*, 1996; *Simons et al.*, 1999] (Fig. 2-13a), along with the partial derivatives (Fréchet kernels) for the sum of all the surface modes in two group velocity-frequency ranges. The first window (Fig. 2-13b) captures most of the "surface wave" proper, whereas the latter (Fig. 2-13c) represents the "body-wave" equivalent modes such as *S*, *SS*, and *SSS* for increasing epicentral distance [for further information about these choices and their practical as well as theoretical limitations, see *Zielhuis and Nolet*, 1994a; *Zielhuis and van der Hilst*, 1996; *Simons et al.*, 1999]. Apart from the dependence, at shallow depth, on density and α, the Fréchet kernel for β_V (or equivalently, L) is many times larger than the other partials.

Furthermore, a parameterization as in Eq. 2.1 is only valid if the same dominance of the β_V partial is observed for realistic anisotropic *a priori* models for all the surface wave modes in both group velocity windows used in our analysis. In Fig. 2-13d we plot the model of *Gaherty and Jordan* [1995], which provides an independent constraint on the radially anisotropic structure of the Australian continent. The partials are plotted in Fig. 2-13e and Fig. 2-13f. The β_V-partial is always dominant, and we conclude that a parameterization in terms of β_V, G_c and G_s as in Eq. 2.1 or 2.6 is adequate when higher modes are included. When fitting an individual seismogram, it is reasonable to use the $\partial c/\partial\beta_V$ partial derivative with an isotropic reference model: the resulting 1-D path model of β_V is an adequate representation of the Earth when no further azimuthally anisotropic information is available. This is justified by the observation that the sensitivity kernels of Rayleigh wave phase velocities to perturbations in *S* velocity in isotropic and transversely isotropic Earth models are very similar [*Anderson and Dziewonski*, 1982] (see also Fig. 2-13), although a small error is probably made [*Lévêque and Cara*, 1985].

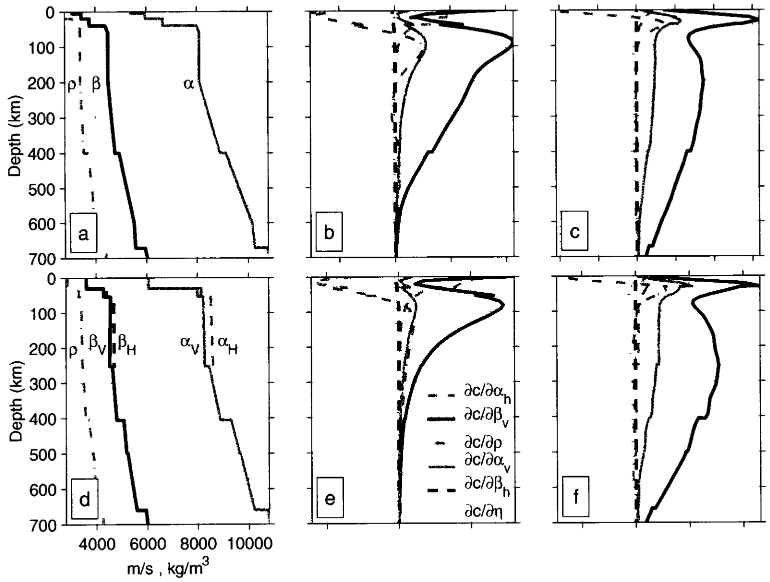


Figure 2-13: Earth models and sensitivity kernels. (a,b,c) One of the isotropic reference models used in the waveform inversion of an individual seismogram. The waveform inversion solves for β_V only, using vertical component seismograms. The different modes making up the two analysis windows of the seismogram have phase velocities that are sensitive to perturbations in the elastic parameters according to the Fréchet kernels plotted in Figs. 2-13b (representing predominantly the lower-frequency fundamental mode branch, in the group velocity range of 3.4–4.4 km/s and frequency range of 5–25 mHz) and 2-13c (higher modes at higher frequencies, in the range 4.2–4.8 km/s and 8–45 mHz). The dominant partial derivative is the β_V kernel. (d,e,f) Anisotropic model of continental Australia obtained independently by Gaherty and Jordan (1995), with the corresponding partials. The same mode selection criterion was used in 2-13b and 2-13e, and 2-13c and 2-13f. For both models and mode windows, the dominant partial is of β_V, which is taken as justification for our parameterization.

Chapter 3

Anisotropic Isostatic Response[1]

Abstract

Gravity and topography provide important insights regarding the degree and mechanisms of isostatic compensation. The azimuthally isotropic coherence function between the Bouguer gravity anomaly and topography evolves from high to low for increasing wavenumber, a diagnostic that can be predicted for a variety of lithospheric loading models and used in inversions for flexural rigidity thereof. In this study we investigate the isostatic response of continental Australia. We consider the effects of directionally anisotropic plate strength on the coherence. The anisotropic coherence function is calculated for regions of Australia that have distinctive geological and geophysical properties. The coherence estimation is performed by the Thomson multiple-Slepian-taper spectral analysis method extended to two-dimensional fields. Our analysis reveals the existence of flexural anisotropy in central Australia, indicative of a weaker N–S direction of lower T_e. This observation is consistent with the suggestion that the parallel faults in that area act to make the lithosphere weaker in the direction perpendicular to them. It can also be related to the N–S direction of maximum stress and possibly the presence of E–W running zones weakened due to differential sediment burial rates. We also demonstrate that the multitaper method has distinct advantages for computing the isotropic coherence function. The ability to make many independent estimates of the isostatic response that are minimally affected by spectral leakage results in a coherence that is more robust than with modified periodogram methods, particularly at low wavenumbers. Our analysis elucidates the reasons for discrepancies in previous estimates

---

[1] **Isostatic Response of the Australian Lithosphere: Estimation of Effective Elastic Thickness and Anisotropy Using Multitaper Spectral Analysis**, *J. of Geophysical Research* (**105**) B8, 19,163–19,184, 2000

of effective elastic thickness T_e of the Australian lithosphere. In isotropic inversions for T_e, we obtain values that are as much as a factor of 2 less than those obtained in standard inversions of the periodogram coherence using Bouguer gravity and topography but greater than those obtained by inversions that utilize free-air rather than Bouguer gravity and ignore the presence of subsurface loads. However, owing to the low spectral power of the Australian topography, the uncertainty on any estimate of T_e is substantial.

3.1 Introduction

3.1.1 Admittance and Coherence Calculations

Mountain belts are generally underlain by "roots" that are less dense than the surrounding mantle. In the most extreme, fully compensated case, they are in a state of near-Airy isostasy that corresponds to a lithosphere with no strength or zero thickness [*Turcotte and Schubert*, 1982]. In this case the free-air gravity anomaly is small and approaches zero for the longest wavelengths, and the Bouguer gravity anomaly is nonzero, reflecting the crustal root. Generally, the Bouguer anomaly is strongly correlated with the topography at long wavelengths. In other scenarios the lithosphere has more rigidity or strength and can support topographic loads without much of a compensating crustal root. For wholly uncompensated topography the free-air anomaly, not the Bouguer anomaly, will be correlated with topography [*Lambeck*, 1988; *Fowler*, 1990; *Blakely*, 1995]. In general, the correlation of the Bouguer anomaly to topography is wavelength-dependent. The wavelength range at which the transition from compensated to uncompensated topography occurs is diagnostic of the lithospheric rigidity. Thicker or more rigid lithospheres tend to undergo the transition from highly compensated to uncompensated topography at longer wavelengths than thinner or weaker lithospheres [e.g., *Karner*, 1982; *Forsyth*, 1985].

Admittance and coherence functions, spectral measures of the isostatic response of gravity to topography, can be used to invert for an effective elastic thickness T_e or flexural rigidity D [*Timoshenko and Woinowsky-Krieger*, 1959], assuming surface and/or subsurface loading of an elastic plate overlying a fluid substrate. This has been done for both oceanic and continental plates in a variety of settings [e.g., *McKenzie and Bowin*, 1976; *McNutt and Parker*, 1978; *Forsyth*, 1985; *Zuber et al.*, 1989; *McKenzie and Fairhead*, 1997].

3.1.2 Surface and Subsurface Loading

Loading of the lithosphere may occur at the surface $z = 0$ or at some subsurface depth $z = z_m$. For an elastic plate which is loaded at the top by the emplacement of topography, a density contrast at an interface located at $z = z_c$ provides compensation for the applied load. It is generally assumed that subsurface loads are to be applied in the form of relief at the Moho discontinuity (a seismic discontinuity indicative of large density contrasts) and that surface loads are compensated by deflecting the Moho. In other words: $z_m = z_c = z_{\mathrm{Moho}}$. An alternative approach is to find z_c independently from the slope of the shortest-wavelength piecewise linear segment of the log gravity power spectrum [*Karner and Watts*, 1983; *Bechtel et al.*, 1987; *Zuber et al.*, 1989]. However, although the inferred loads are sensitive to the assumed densities and thicknesses of the layers, the flexural rigidity from coherence analysis is fairly insensitive to those assumptions [*Forsyth*, 1985; *Lowry and Smith*, 1994]. A simple two-interface model is therefore preferred by most authors [e.g., *Ebinger and Hayward*, 1996; *Gwavava et al.*, 1996]. The amplitude of Moho relief is found from downward continuation of the Bouguer anomaly. If no subsurface loading is allowed, however, the effective depth of compensation z_c reenters the equation [*McKenzie and Fairhead*, 1997]. *McKenzie and Fairhead* [1997] determine z_c as a free parameter by minimizing a penalty function of the admittance of free-air gravity and topography which includes z_c and the effective elastic thickness T_e.

3.1.3 Statistical Independence of Top and Bottom Loading

One of the basic assumptions of the coherence method is that surface and subsurface loading are statistically independent. In many cases, however, surface and subsurface loading are likely to be tectonically related and therefore spatially correlated processes [*Forsyth*, 1985]. In general, as the degree of correlation increases, a downward bias in the estimated T_e is observed [*Macario et al.*, 1995]. The breakdown of the assumption of statistical independence was clearly observed by *Zuber et al.* [1989] in the case of central Australia. Correlated loads may reflect either a real (tectonic) relationship between surface and subsurface loading, or it may indicate that the uniform plate model was applied in an area of markedly nonuniform flexural rigidity [*Zuber et al.*, 1989]. In addition, the presence of anisotropy in the loading response may explain the breakdown of the uniform isotropic coherence method and the associated poor fits of observed to modeled values [*Stephenson and Lambeck*, 1985]. Finally, erosion processes that reduce the amplitude of the topography at all wavelengths without changing the shape may be responsible for the coherence of subsurface with surface topography [*McKenzie and Fairhead*, 1997].

3.1.4 Erosion

Erosion may play an important role in modifying the gravity to topography relationships, predominantly at short wavelengths [*Stephenson*, 1984; *Forsyth*, 1985]. The removal of the topographic expression of subsurface loads leaves only the Bouguer gravity anomaly and reduces the coherence with topography. The associated shift to longer transitional wavelengths of the coherence biases the T_e upward. To address this problem, *McKenzie and Fairhead* [1997] advocate using the free-air rather than the Bouguer anomaly in admittance calculations, arguing that, whether or not subsurface loads are present, surface topography must always be a load, and part of the free-air gravity anomaly must therefore always be coherent with topography. *McKenzie and Fairhead* [1997] find much lower values of elastic thickness than those obtained from inversions performed using the Bouguer anomaly in coherence calculations. In their analysis, they prefer values of $f = 0$ for the ratio of bottom to surface loading. As a result, their depth of compensation z_c shifts to mid-crustal values and the T_e estimates are reduced until none exceed 25 km. The assumption of top loading alone biases toward lower values of T_e.

3.1.5 Influence of the Spectral Estimation Technique

McKenzie and Fairhead [1997] advocate, with little explanation, the use of the multitaper spectral analysis method of *Thomson* [1982] rather than the traditional mirrored or windowed (in other words: modified) periodogram method of *Bechtel et al.* [1987] or the higher-resolution maximum entropy method developed by *Lowry and Smith* [1994]. In this chapter, we provide formal justification of the merits of the multitaper method and focus on two particular issues: the nature of bias at long wavelengths and the role of anisotropy in the loading response. To illustrate the technique, we analyze topography and Bouguer gravity data from continental Australia, where we identify evidence for anisotropy and interpret it in terms of geologic setting. Allowing for both surface and subsurface loading, we obtain estimates of the effective elastic thickness that are intermediate between the previous values of *Zuber et al.* [1989] and *McKenzie and Fairhead* [1997].

3.2 Coherence Studies of the Mechanical Lithosphere

3.2.1 Transfer Function Approach: Linear Filter Theory

The correlation between the Bouguer gravity field and surface topography requires a rigorous definition. *Dorman and Lewis* [1970] related topography to gravity anomalies as

$$\Delta g(\mathbf{r}_0) = \int_{\Sigma} h(\mathbf{r}) q(\mathbf{r} - \mathbf{r}_0)\, d\mathbf{r}^2 + n(\mathbf{r}_0). \tag{3.1}$$

Here $\Delta g(\mathbf{r}_0)$ represents the (Bouguer) gravity anomaly at position $\mathbf{r}_0$ due to topographic height $h(\mathbf{r})$, Σ is the surface of interest, and $q(\mathbf{r})$ is the unknown kernel function that relates the gravity anomaly to the load that causes it [*Phillips and Lambeck*, 1980; *Karner*, 1982; *Lambeck*, 1988]. All through the chapter, boldface letters will denote vectors of arbitrary dimension. In particular, $\mathbf{k} = (k_x, k_y)$ and $|\mathbf{k}| = k$. The system is characterized by "geologic" noise $n(\mathbf{r_0})$ that is not of isostatic origin and for which no further assumptions are made other than that it is uncorrelated with topography. The latter assumption, stated in the wavenumber domain as $S_{nh}(\mathbf{k}) = 0$, allows us to calculate the Fourier transform of the kernel as

$$\hat{q}(\mathbf{k}) = \frac{S_{h\Delta g}(\mathbf{k})}{S_{hh}(\mathbf{k})}, \tag{3.2}$$

where S represents the cross-spectral density of the two random variables identified by the subscripts (for an elaboration, see Appendix I). The function $\hat{q}(\mathbf{k})$ is the isostatic response or admittance function [*Bendat and Piersol*, 1986, 1993]. If the gravitational anomaly is in phase with the topography, $\hat{q}(\mathbf{k})$ is real.

We may renormalize the admittance by the power spectral density of the gravity anomaly and use the quantity

$$\gamma^2(\mathbf{k}) = \frac{|S_{h\Delta g}(\mathbf{k})|^2}{S_{\Delta g\Delta g}(\mathbf{k}) S_{hh}(\mathbf{k})} \quad ; \quad 0 \leq \gamma^2(\mathbf{k}) \leq 1, \tag{3.3}$$

instead of (3.2). This function is referred to as the coherence [*Bendat and Piersol*, 1986, 1993].

3.2.2 Spectral Estimation of Stochastic Signals: Averaging

The quantitative study of isostasy requires an estimation of the (cross-)spectral properties of two random variables. For any of N proposed spectral estimators $\hat{S}$ to be an ideal estimator of the true spectrum S, we would have, in function of the wave vector $\mathbf{k}$:

$$\forall\mathbf{k} : E\{\hat{S}(\mathbf{k})\} = S(\mathbf{k}), \tag{3.4a}$$

$$\lim_{N\to\infty} \mathrm{var}\{\hat{S}(\mathbf{k})\} = 0, \tag{3.4b}$$

$$\forall\mathbf{k}' \neq \mathbf{k} : \mathrm{cov}\{\hat{S}(\mathbf{k}'), \hat{S}(\mathbf{k})\} = 0. \tag{3.4c}$$

Here E denotes an expectation or averaging operator, and var and cov stand for variance and covariance, respectively. Thus we require that the estimate be unbiased (equation (3.4a)), have well-behaved estimation variance (equation (3.4b)), and map all power at the appropriate frequencies without leakage (equation (3.4c)).

The periodogram, the squared magnitude of the Fourier coefficients of the signal, is a first-order spectral estimator [*Tukey*, 1967; *Kay and Marple*, 1981; *Percival and Walden*, 1993]. With the periodogram the coherence function is approximated by [*Bendat and Piersol*, 1986, 1993; *Touzi et al.*, 1996, 1999]

$$\gamma^2(\mathbf{k}) = \frac{|E\{F(\mathbf{k})G^*(\mathbf{k})\}|^2}{E\{G(\mathbf{k})G^*(\mathbf{k})\}E\{F(\mathbf{k})F^*(\mathbf{k})\}}, \tag{3.5}$$

where F and G denote the Fourier transforms of the random variables f and g (e.g., gravity and topography), and the asterisk is used to denote complex conjugation.

Modifying the periodogram by mirroring or windowing the data set with a single window are both attempts to improve the estimation properties of the periodogram [*Tukey*, 1967; *Welch*, 1967; *Percival and Walden*, 1993]. The choice of the data window(s) primarily controls bias, while the smoothing or averaging operation reduces the estimation variance [see *Chave et al.*, 1987, and Appendix I].

It is important to reflect upon the physical nature of the coherence, and the role of averaging in its estimation. Indeed, without any averaging the numerator and denominator of (3.5) would be equal and the coherence exactly unity. By averaging over N different spectral estimates we are attempting to satisfy (3.4a), hoping that the estimate $\hat{S}$ (in the case of (3.5), this is the periodogram) is asymptotically unbiased in the sense

$$\lim_{N\to\infty} E\{\hat{S}(\mathbf{k})\} = S(\mathbf{k}). \tag{3.6}$$

The coherence function measures the consistency of the phase relationship between two fields, regardless of their amplitude. It is a measure of the variance of the phases associated with the individual terms whose expected values are obtained. For two fields that are uncorrelated at wave vector $\mathbf{k}_1$, the phases of the various cross-spectral estimates over which one averages are randomly distributed. For the component $\hat{S}(\mathbf{k}_1)$, averaging will then result in a cancellation, yielding $\gamma^2(\mathbf{k}_1) \to 0$. For signals that are correlated at $\mathbf{k}_2$, on the other hand, the interference of the phases will be constructive, and $\gamma^2(\mathbf{k}_2) \to 1$ upon averaging. Different representations or spectral estimates of the same process can be obtained by analyzing overlapping segments of the data, repeated measurements or, as we shall see, from data windowed with different orthogonal tapers.

Fig. 3-1 illustrates the importance of averaging. In Fig. 3-1a, two time series are given. Both have peaks in the power spectrum representing two distinct frequencies (Fig. 3-1b). At one frequency the signals are exactly in phase, while at the other they are off by $\pi/4$. A value of 15 % white noise was added to both signals. If no averaging is done and the periodogram estimate of the cross-spectral density is divided by the individual spectral densities (as prescribed by (3.3) and (3.5)), the coherence equals unity over the entire range of frequencies (Fig. 3-1c, dashed line). Providing multiple representations of the same stochastic process by splitting the signals in overlapping, windowed segments, and averaging the numerator and denominator of (3.5) before the division (a modified periodogram

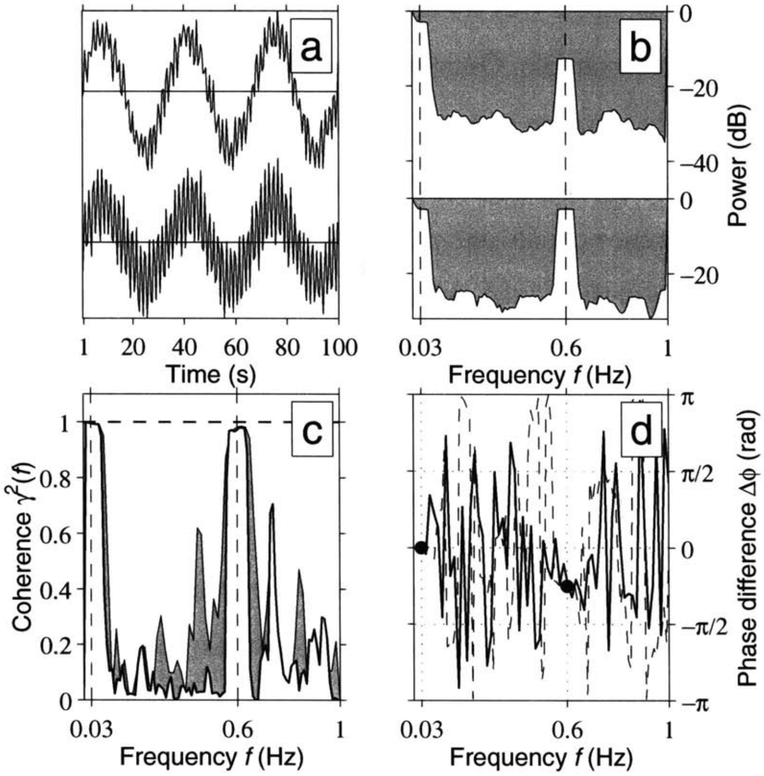


Figure 3-1: The importance of averaging in coherence estimation. (a) Top signal: $3\sin(3/50\pi t) + 1\cos(6/5\pi t)$. Bottom signal: $2\sin(3/100\pi t) + 2\cos(3/5\pi t + \pi/4)$; 15% white noise has been added to each. The phase relationship between the two time series is consistent over the length of the sequence. (b) Power spectral density. (c) Coherence estimates. Disregarding the noise level, the coherence should peak at 1 for both 0.03 and 0.6 Hz input frequencies. Thick dashed line, periodogram estimate without averaging; thin solid line, shaded area, with Welch overlapping segment algorithm, windowing with cosine tapers; thick solid line, open area, multitaper estimate. The resolution bandwidth is 80 mHz. (d) Phase of the cross-spectral density for multitaper (solid) and WOSA (dashed). Where the coherence is one, the correct phase difference between the components is retrieved (solid circles).

procedure called Welch Overlapping Segment Algorithm (WOSA) [*Welch*, 1967]), correctly yields coherence peaks at both input frequencies (Fig. 3-1c, shaded area). With the multitaper method averages are made over different windowed versions (described in detail in Appendix I) of the entire data set, and the coherence also peaks at both frequencies (Fig. 3-1c, open area). The broadness of the multitaper coherence peaks is due to the fixed bandwidth of the Slepian tapers (see *Riedel and Sidorenko* [1995] and Appendix I), which we set at 80 mHz for this example. Omitting further detail until later sections, we note that the multitaper coherence estimate is much less noisy than WOSA [see also *Bronez*, 1992]. The phase angle of the cross-spectrum correctly gives the phase difference between the two signals where the coherence equals unity. Both the Welch method (Fig. 3-1d, dashed line) and the multitaper method (Fig. 3-1d, solid line) pick out the input 0 and $-\pi/4$ phase shift at 30 and 600 mHz, respectively (Fig. 3-1d, solid circles).

3.2.3 Different Ways to Average Spectral Estimates

In coherence analyses applied to determining effective elastic lithosphere thickness, three main approaches of averaging to get the necessary reduction in estimation variance of the coherence have been followed.

Bin averaging

Assuming second-order stationarity in the data [*Welch*, 1967; *Kay and Marple*, 1981; *Bendat and Piersol*, 1986], the time series may be subdivided into overlapping segments, each of which is detrended and windowed and then Fourier transformed (such as the calculation that produced the thin solid gray filled line in Fig. 3-1c). The overlap of the data sequences assures that the down-weighted portions of sequence $N-1$ receive more weight in sequence N, so that a minimum of statistical information is discarded. However, section averaging is inappropriate for short signals. Also, the overlap between the sections needs to be sufficient to yield high efficiency, yet ensure approximate independence of the raw spectra [*Chave et al.*, 1987]. Finally, because the highest resolvable wavelength is the inverse of the fundamental (Rayleigh) frequency $1/(N\Delta t)$ [*Kay and Marple*, 1981] with N the data length and Δt the sampling interval, reducing N compromises long-wavelength resolution.

For applications of bin averaging to lithospheric studies, see, for example, *McKenzie and Bowin* [1976].

Ensemble averaging

If different sample functions are available (e.g., gravity and topography data acquired as multiple passes over the same geologic feature), each of them can provide an independent spectral estimate. The assumption of ergodicity guarantees that all the ensembles are realizations of the same random process [*Kay and Marple*, 1981; *Bendat and Piersol*, 1986]. Lithospheric studies utilizing this principle are, for example, those by *Watts* [1978] and *Detrick and Watts* [1979].

Smoothing in wave vector space

Averaging becomes more problematic in the two-dimensional case, as there can be only one data set for a given region, and the region itself is hard to subdivide. Usually, a single spectral estimate is calculated on a mirrored or windowed data set, but in the latter case the information contained in the downweighted edges of the data window is partially lost, while in the former spurious power is introduced at (long) wavelengths not present in the data. The spectrum can then be estimated as a moving average of the unsmoothed estimator [*Stephenson and Beaumont*, 1980], or the original estimate can be binned into annuli of wavenumber bands [*Bechtel et al.*, 1987]. Moving averaging can badly bias the spectrum if the phase changes over the averaging band [*Park et al.*, 1987b; *Kuo et al.*, 1990; *Mellors et al.*, 1998]. Moreover, with wavenumber binning the long-wavelength part of the spectrum will be based on fewer data points than are the higher frequencies. This leads to a positive long wavelength bias for the coherence [*Munk and Cartwright*, 1966]. And finally, if the variation of the coherence with direction is of interest to us we cannot average azimuthally.

3.3 Multitaper Method for 2-D Spectral Analysis

The multitaper technique, as originally proposed by *Thomson* [1982], reduces the estimation variance of the spectrum by calculating $\hat{S}$ as a weighted average over a number k of independent direct spectral estimates $\hat{S}_l$ with weights λ_l. The eigenspectra $\hat{S}_l$ that satisfy (3.4a) are based on the data windowed by a function $h_l(\mathbf{r})$ to be determined. If all $\hat{S}_l, l = 1, ..., k$ are pairwise uncorrelated with a common variance, then the variance associated with the average estimate will be smaller by $1/k$ [*Bronez*, 1992; *Riedel and Sidorenko*, 1995], meeting the objective of (3.4b). To minimize spectral leakage in the sense of (3.4c), we need to find data windows $h_l(\mathbf{r})$ whose spectral responses $\mathcal{H}_l(\mathbf{k})$ (see (3.13)) have the narrowest central lobe and the smallest possible sidelobe level (this is expressed as an energy concentration criterion). *Slepian* [1978] found that the ideal data windows $h_l(\mathbf{r})$ are given by discrete prolate spheroidal sequences (dpss). The width of the central lobe of $\mathcal{H}_l(\mathbf{k})$ is a measure of the resolution of the estimates (see (3.12) and (3.15)) and is defined at the discretion of the analyst. The half width W is usually an integer multiple j of the fundamental frequency: $W = j/(N\Delta t)$; commonly, j is quoted as NW (assuming $\Delta t = 1$). For every such choice of resolution bandwidth, there are $k = 2NW$ useful tapers; this number is known as the Shannon number in information theory. The dpss are the solutions to an eigenvalue problem (see Appendix I), and therefore the windows $h_l(\mathbf{r})$ are orthogonal and the resulting eigenspectra $\hat{S}_l$ (approximately) uncorrelated. The associated eigenvalues λ_l can be used as weights. The Slepian tapers minimize a particular concentration criterion (see (3.17) and (3.18)). Other criteria can be used and different sets of orthogonal tapers may be found that have similar performance as the Slepian tapers [*Riedel and Sidorenko*, 1995; *Walden et al.*, 1998; *Komm et al.*, 1999]. Throughout this chapter, however, we will call the multiple-taper method which uses Slepian windows "the" multitaper method.

The properties of a set of one-dimensional Slepian tapers are summarized in Fig 3-2. In Fig 3-2a the first five tapers of a sequence with $NW = 4$ are given. We assume $\Delta t = 1$ and

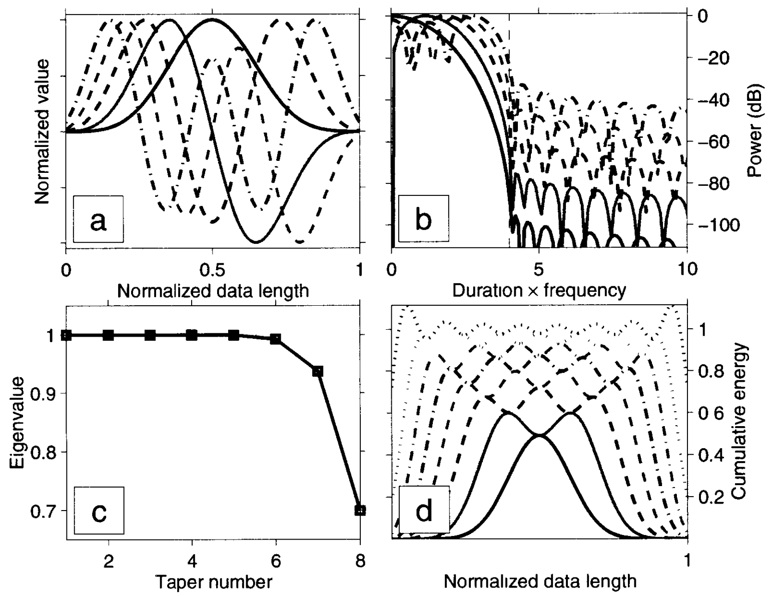
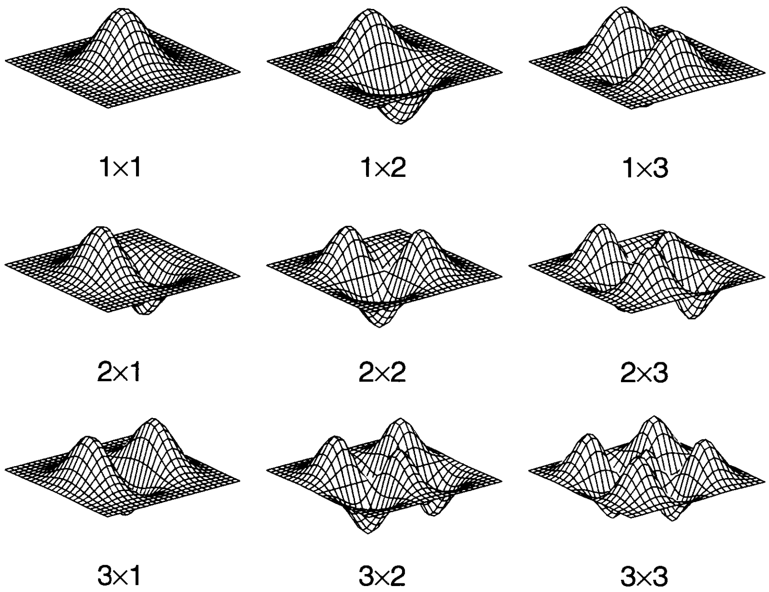


Figure 3-2: Properties of the dpss. (a) Data tapers $h_{t,k}$ for $NW = 4$ and $k = 1, ..., 5$ with normalized height and as a fraction of the data length N. (b) Spectral windows $\mathcal{H}_k(f)$ belonging to the tapers to the left. The sidelobe level, and hence the bias in the spectral estimation, of the higher order windows increases with order k. (c) Eigenvalues λ_k associated with the eigentapers. (d) Buildup of dpss taper energy defined as $\sum_{k=1}^{K} h_{t,k}^2$ for $K = 1, ..., 8$. Using only one taper, much information is lost. With an increasingly complete set of eigentapers, the information is extracted evenly from all samples.

non-dimensionalize the frequency axis by plotting Nf, i.e. duration times frequency, so the central lobe is confined within $[-NW, +NW]$ (Fig 3-2b, dashed vertical line at $Nf = 4$). The first spectral window assures over 100 dB of bias protection outside the inner domain of $NW = 4$ [see also *Chave et al.*, 1987]. As the order of the tapers increases they become more and more oscillatory. Their central resolution peak broadens and splits and the level of sidelobes increases. For $NW = 4$, the Shannon number $2NW$ indicates the order beyond which the eigenvalues λ_k drop steeply to zero (Fig 3-2c). Conventionally, only the first $2NW - 1$ eigentapers are used. Most popular data windows in non-multitaper applications are fairly similar to the zeroth-order multitaper window. The disadvantage of


1×1 1×2 1×3

2×1 2×2 2×3

3×1 3×2 3×3

Figure 3-3: A subset of the two-dimensional tapers used in the analysis, formed by taking the outer product of any two combinations $k \times k'$ (indicated in the legend) of the 1–D tapers $h_{t,k}$ of Fig 3-2.

having only one data window is that valuable statistical information is discarded, by virtue only of being on the edge of the data. Any information that is lost in this way is retrieved by the higher-order multitaper windows [*Walden*, 1990; *Komm et al.*, 1999]. An increasingly complete set of eigentapers extracts information more and more evenly from the entire signal, without this arbitrary down-weighting, as shown in Fig 3-2d.

Originally formulated for one-dimensional (1-D) time series, the extension to two-dimensional fields is readily made [*Liu and van Veen*, 1992; *Hanssen*, 1997]. The additivity of the Fourier transform allows us to calculate the 2-D transform by applying the algorithm to the rows of the data matrix first, tapered with one taper, and then performing the transform on the columns of the result, tapered with a different taper. This is equivalent to tapering the data with 2-D tapers given by the outer (dyadic) product of combinations of 1-D tapers. A subset of the 2-D tapers that we used is given in Fig. 3-3. The weights of the combination of two such tapers have to be given by the product of their individual eigenvalues [*Hanssen*, 1997]. In this way, using k different tapers in n dimensions, $K = k^n$ independent sets of windowed data are obtained. The spectral estimation variance can be calculated explicitly using (3.24).

The spectral windows $\mathcal{H}(\mathbf{k})$ associated with the tapers of Fig. 3-3 are plotted as contour lines in Fig. 3-4. As in Fig. 3-2b, the axis plots duration times frequency, which illustrates the effect of the resolution parameter NW. For the chosen set of tapers, $NW = 4$ and hence the central lobe is concentrated within the bounds of -4 to $+4$ (Fig. 3-4, dash-dotted lines). We are looking from above onto the (k_x, k_y) plane; hence a horizontal or vertical section through these plots would be identical to the spectral windows plotted in Fig. 3-2b. As in the 1-D case, with increasing order of the tapers the sidelobe level increases and hence the resolution decays. However, incorporating many independent tapers into the spectral estimation reduces the estimation variance. Resolution and variance trade off.

3.4 Tests With Synthetic Data

3.4.1 Motivation

The use of the multitaper method for determining admittance and coherence functions in T_e studies is not a novelty. *Scheirer et al.* [1995] reported its usefulness for the estimation of the elastic thickness of the Basin and Range province, and *McKenzie and Fairhead* [1997] present it as their method of choice. We believe that an application of this method should be accompanied by a rigorous theoretical discussion of its merits, as well as a thorough testing on realistic synthetic data sets. The remaining part of the discussion of T_e can then be devoted to the different assumptions involved in calculating the forward models, most notably the role of subsurface loading (see also Sec. 3.1.2). New to the topic of elastic thickness estimation, however, is the realization that only the multitaper method is able to produce reliable 2-D $\gamma^2(\mathbf{k})$ estimates, such that anisotropy in the loading response can be

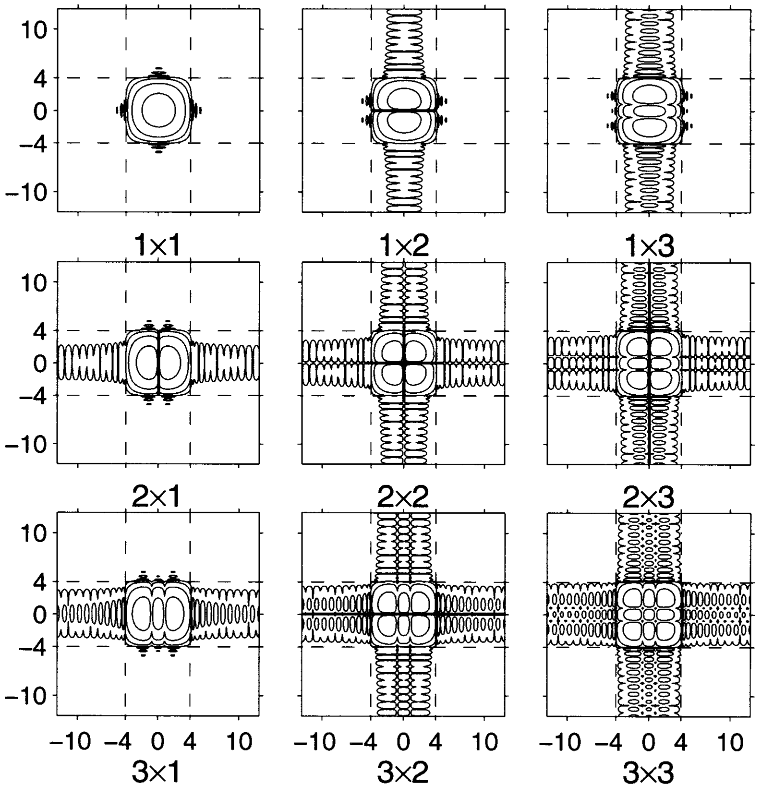


Figure 3-4: Spectral windows $\mathcal{H}_k(\mathbf{k})$ (see (3.13)) associated with the tapers of Fig. 3-3, plotted as functions of the duration × half bandwidth product. $NW = 4$. Contouring is at -105, -75, -45, and -15 dB.

investigated without resorting to smoothing or binning procedures in wave vector space, procedures which, as we have argued in Sec. 3.2.3, are inherently flawed.

3.4.2 Data Generation

As this is, to our knowledge, the first time the two-dimensional extension of the multitaper method is used in order to detect anisotropic power in the coherence, our approach needs to be tested with synthetic data sets. Besides, even in the isotropic or 1-D case some thought must be given to the selection of the resolution parameter NW.

As indicated in Sec. 3.2.2, the coherence function is a measure of the consistency of the phase relationship between individual measurements of a particular Fourier component, independent of its magnitude. This suggests a way of synthesizing two fields with a known coherence [*Lowry and Smith*, 1994]. Suppose we are able to get one direct cross-spectral density estimate $\hat{S}_{fg}$ between two fields f and g at a particular wave vector $\mathbf{k}_1$. As f and g are not necessarily in phase at $\mathbf{k}_1$, this is in general a complex number, represented as $r \exp i\varphi$. In order to obtain a stable estimate of the cross-spectral density of both stochastic random variables we need to perform such a measurement a number k times and average the results. If both fields are highly correlated, the $l = 1, ..., k$ individual terms will have the form $r \exp i(\varphi + \Delta\varphi_l)$, where for every l, $\Delta\varphi_l \ll \varphi$. Hence the average for this component adds up to a nonzero complex number, and, normalized, $\gamma^2 \rightarrow 1$. If, on the other hand, every measurement yields a widely varying phase vector so that $\Delta\varphi_l = \mathcal{O}(\varphi)$, the average will have a magnitude that is near zero, and $\gamma^2 \rightarrow 0$. Taking the Fourier representation of some field $H(\mathbf{k})$ and creating a second field $H(\mathbf{k}) + N(\mathbf{k})$ using some function $\gamma^2(\mathbf{k})$, as described by

$$|N(\mathbf{k})|^2 = |H(\mathbf{k})|^2 \frac{1 - \gamma^2(\mathbf{k})}{\gamma^2(\mathbf{k})}, \tag{3.7}$$

the coherence between $H(\mathbf{k})$ and $H(\mathbf{k})+N(\mathbf{k})$ is given by $\gamma^2(\mathbf{k})$. Thus to make $\gamma^2 \rightarrow 0$ at a certain wave vector, we add "noise" with a magnitude comparable to the Fourier component at the same wave vector but a random distribution of phases. To make $\gamma^2 \rightarrow 1$ we simply add random-phase noise with much less amplitude. Any field $H(\mathbf{k})$ can be taken as the start of this procedure, although it is desirable that its power spectrum have some resemblance to the actual data in the gravity-topography analysis (e.g., having a red spectrum).

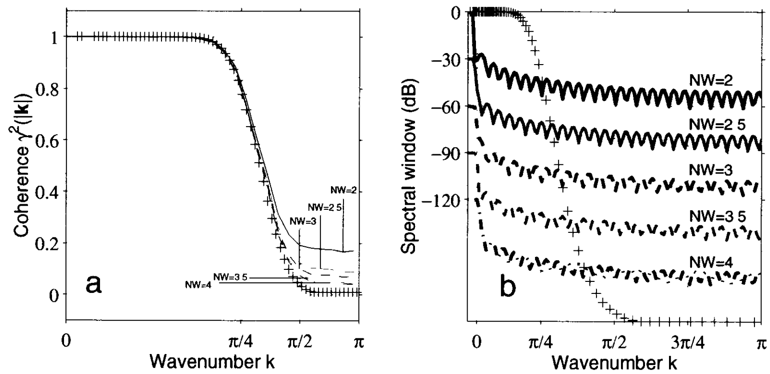


Figure 3-5: Influence of resolution bandwidth NW upon synthetic data with isotropic coherence. Frequency space is non-physical, hence contained within angular frequency between 0 and $+\pi$. (a) Isotropic input coherence (pluses). In shaded symbols, multitaper estimates for NW values of 2, 2.5, 3, 3.5, and 4. $NW = 4$ is the value we adopt in the subsequent analyses. The x scale is logarithmic. (b) Spectral windows $\mathcal{H}(f)$ (see (3.16)), offset for clarity. With increasing NW, the resolution degrades (the central peak broadens) but is still acceptable compared to the input coherence (pluses, scaled to fit the figure). The x scale is linear.

3.4.3 Tests With Synthetic (An)isotropic Coherence

Synthetic tests are presented in Fig. 3-5 and 3-6. In the tests the wavenumber ranges from $-\pi$ to $+\pi$, the Nyquist wavenumbers $\pm k_N$ for sampling intervals of unit step, $\Delta t = 1$: $k_N = 2\pi f_N$; $f_N = 1/(2\Delta t)$. Where indicated, the scale is in decibels (dB). Because of symmetry we will plot the results for $k = 0 \to \pi$.

Fig. 3-5 depicts the results of a first test. The input coherence function (pluses), given in Fig. 3-5a, is isotropic, so $\gamma^2(\mathbf{k}) = \gamma^2(|\mathbf{k}|)$. The multitaper estimates are compared to the input for values of the resolution parameter NW varying from 2 to 4 in 0.5 increments. As NW approaches 4, the final value, the estimate follows the synthetic curve more closely. The positive bias at high wavenumbers decreases monotonically with higher NW. This test enables us to identify an acceptable resolution parameter NW. Our final choice is $NW = 4$. All the following analyses presented in this chapter have used $NW = 4$ and

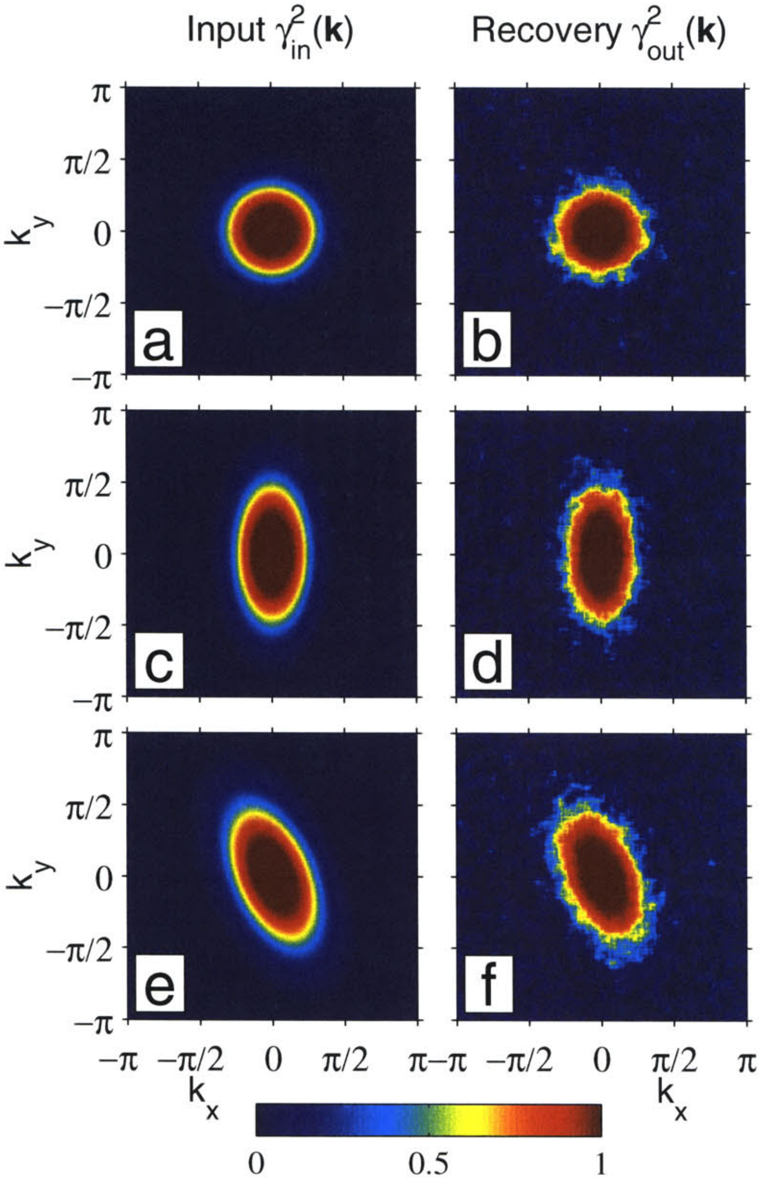


Figure 3-6: Synthetic (an)isotropic coherence with variable rotation: (a,c,e) input coherence, and (b,d,f) recovery with multitaper method, $NW = 4$.

$k = 7$ so at all times the resolution half bandwidth is given by $W = 4/(N\Delta t)$ with N the total data length and Δt the sampling interval. We have repeated our measurements with different values of NW. We will report this in Sec. 3.6.2.

For the set of tapers used, Fig. 3-5b gives the average spectral windows, $\bar{\mathcal{H}} = \sum_{l=1}^{k} \mathcal{H}_l(\mathbf{k})$ (see (3.15)) for the different values of NW and $k = 2NW - 1$. The half width of the central lobe $\bar{\mathcal{H}}$ gives the resolution of the coherence estimate. So, for $N = 100$ and $\Delta t = 1$, the wavenumber resolution $2k_W = 2\pi 2W = 2\pi j/N$, where $j = 2, 2.5, ..., 4$, of the different curves is $2/50\pi$, $2.5/50\pi$,$3/50\pi$, $3.5/50\pi$, and $4/50\pi$. For clarity, we have offset the different spectral windows in the y direction and plotted the synthetic coherence function used in Fig. 3-5a on top of them on the same linear x scale but scaled to fit the plot box in the y direction.

In Fig. 3-6 we test the ability of the 2-D multitaper method to correctly resolve isotropic and anisotropic input coherences, the latter with varying orientation. Based on Figs. 3-6a and 3-6b, we conclude that we will not unnecessarily reject the hypothesis of isotropy should we measure the coherence function between actual data. If the input is isotropic, then the measurements reveal nothing to the contrary. In Figs. 3-6c–3-6f, anisotropic input fields are analyzed. The input is found in Figs. 3-6c and 3-6e, whereas the multitaper measurements are found in Figs. 3-6d and 3-6f. From this test we deduce the ability of our implementation of the multitaper method to infer anisotropy in the coherence function and detect its major axis.

3.4.4 Other Methods to Compute 2-D Coherence

Stephenson and Beaumont [1980] were the first to consider the possibility of directional anisotropy in the loading response of continental lithosphere (the Canadian Shield). Later, *Stephenson and Lambeck* [1985] investigated the anisotropy of central Australia. More recently, *Lowry and Smith* [1995] give examples of anisotropic T_e for North America. The methods used in the above studies are similar in that one representation of the data is used, only mirrored, windowed or extended with a maximum-entropy criterion. The result is a 2-D coherence map that is nowhere near the true coherence spectrum: only after an averaging procedure is applied does this estimate approach the true coherence. We have verified that it is possible to obtain a certain form of stability by smoothing in wave vector space [as in *Stephenson and Beaumont*, 1980], averaging over an azimuthal wedge around a certain direction of interest [as in *Stephenson and Lambeck*, 1985], or performing binning over overlapping azimuthal wedges [the procedure used by *Lowry and Smith*, 1995] (A. R. Lowry, *personal communication*, 1999). All of these methods are ad hoc and the bias and instabilities present in the modified periodogram method are amplified by virtue of taking even fewer samples in the azimuthal averaging. It is in this light that we present the 2-D multitaper estimate as the only unbiased 2-D coherence estimate. At every wave vector, the coherence estimate is based upon the average of multiple orthogonal representations of the data, and its error is therefore not dependent on the position in wave vector space.

3.4.5 Comparison With Maximum Entropy Spectral Analysis

An alternative method of coherence estimation is Maximum Entropy Spectral Estimation (MESA) [*Burg*, 1975]. *Lowry and Smith* [1994] show the superior performance of MESA with respect to the periodogram method especially when the box size is small. Our tests compare favorably to the maximum entropy results shown by *Lowry and Smith* [1994]. Other authors have confirmed the multitaper performance with respect to MESA [e.g., *Lees and Park*, 1995]. The original maximum entropy approach of estimating spectra is a parametric method which assumes autoregressive signals [*Percival and Walden*, 1993]. Following *Burg* [1975], maximum entropy spectral analysis is based on choosing the spectrum which corresponds to the most random or the most unpredictable time series whose autocorrelation function agrees with the known values. However, many different spectral density functions can have the same autocovariance sequence up to a certain lag. Hence, in general, maximum entropy spectra tend to represent noise processes poorly [*Kay and Marple*, 1981; *Malik and Lim*, 1982; *Percival and Walden*, 1993].

The multitaper method is non-parametric. Tapering replaces the spectral windows of traditional (windowed) periodogram estimates with a unique and optimal set of windows that have better sidelobe properties. The amount of bias and leakage is exactly assessable [*Bronez*, 1992; *Riedel and Sidorenko*, 1995] in a non-subjective way (from the shape of the spectral windows in Figs. 3-2b, 3-4, and 3-5b), and given the multiplicity of the windows, the confidence intervals around the spectrum can be reduced dramatically compared to any other method.

3.5 Application to the Study of Australia

3.5.1 Makeup of the Australian Continent

Using various geophysical and geological criteria, Australia can be divided into three major structural units [*Veevers*, 1984; *Drummond*, 1991; *Braun et al.*, 1998; *Wellman*, 1998; *Simons et al.*, 1999]. At this scale the three domains represent the dominant age provinces as well. For a tectonic map, see Fig. 3-7.

Archean

The westernmost part, composed of Archean cratons (the Pilbara and Yilgarn cratons formed about 3500 and 3100 Ma ago), is geologically stable. It is characterized by high Moho depths (in places exceeding 50 km) [*Wellman*, 1979; *Shibutani et al.*, 1996; *Clitheroe et al.*, 2000a] and low heat flow values ($\sim$40–50 mW m^{-2}) [*Cull and Denham*, 1979; *Cull*, 1991]. The highest effective elastic thicknesses on the continent have previously been measured there ($T_e = 130$ km) [*Zuber et al.*, 1989].

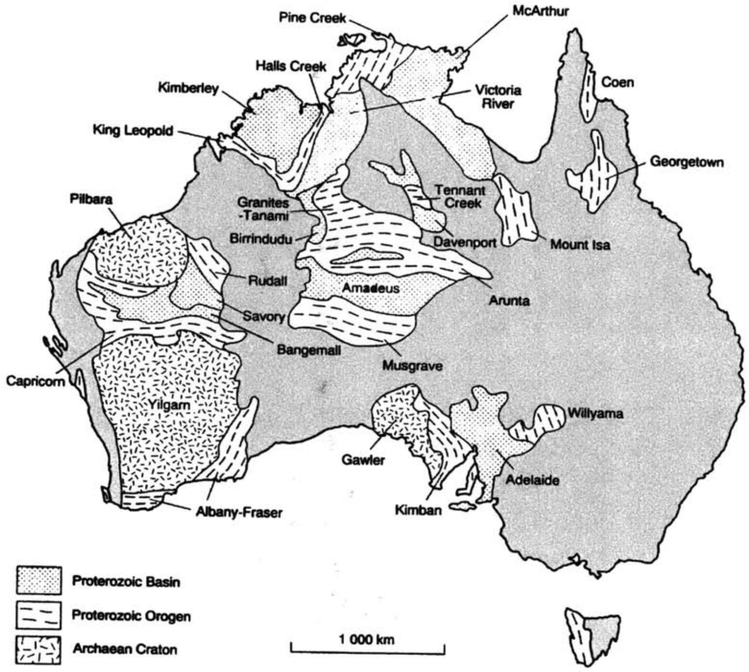


Figure 3-7: Tectonic map of Australia [from *Myers et al.*, 1996].

Proterozoic

Central Australia is made up of a structurally complex, heavily faulted Proterozoic block-and-basin structure. The extensive faulting has resulted in extreme Moho offsets with amplitude variations of more than 20 km [*Lambeck*, 1983; *Lambeck and Penney*, 1984; *Lambeck et al.*, 1988; *Goleby et al.*, 1989; *McQueen and Lambeck*, 1996]. Elastic thickness measurements have yielded values around $T_e = 90$ km [*Zuber et al.*, 1989]. Anisotropy in the isostatic response has been suggested [*Stephenson and Lambeck*, 1985].

Phanerozoic

Terranes to the east of the Tasman Line [*Murray et al.*, 1989] (roughly the easternmost third of the continent) are Phanerozoic in age. The easternmost rim of the continent is marked by a mountainous range which is in close to local isostatic equilibrium [*Murray et al.*,

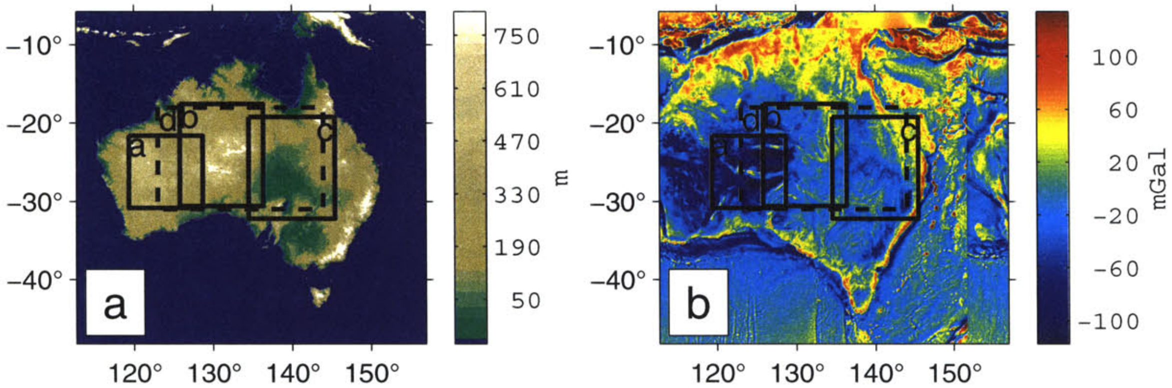


Figure 3-8: The Australian continent and the four regions we selected for the analysis: (a) GTOPO30 topography and (b) Bouguer (continent) and free air (ocean) gravitational anomaly from AGSO. The lettered boxes (lettering as in Figs. 3-9 and Figs. 3-10, 3-12 and 3-13) represent four sets of data used in the analysis.

1989], volcanism as recent as 10 Ma, high heat flow (∼70–100 mW m^{-2}) [*Cull and Denham*, 1979; *Cull*, 1991] and low mantle conductivity values [*Lilley et al.*, 1981; *Finlayson*, 1982]. Shear wave speeds from seismic tomography measurements are anomalously low [*Zielhuis and van der Hilst*, 1996; *Simons et al.*, 1999]: there is a pronounced low-velocity anomaly starting at about 80 km. The available T_e measurements were among the lowest of the continent: 30-80 km for the Phanerozoic Interior Lowlands and 15-35 km for the Eastern Highlands [*Zuber et al.*, 1989]. They have been interpreted in relation to the short time period since the last major thermal event and the deep extent of low mantle conductivity values [see also *Gough*, 1974; *Karner et al.*, 1983; *Bechtel*, 1989; *Mareschal et al.*, 1995]: the low rigidity and subsurface loading in southeastern Australia are consistent with crustal underplating, igneous intrusion and thermal perturbation of the upper mantle as mechanisms for the uplift and compensation of the highlands [*Zuber et al.*, 1989].

It is the aim of this chapter to include another complicating factor into the study of T_e: its anisotropy. We choose three regions of Australia (see the boxes labeled a through c in Fig. 3-8) that approximately correspond to the threefold characterization described above, and investigate the presence of isostatic anisotropy. A fourth region attempts to capture a broad average of the continent (box d in Fig. 3-8). To provide continuity with the earlier work by *Zuber et al.* [1989], we apply their mirrored periodogram method of spectral analysis alongside with the multitaper method to calculate an isotropic T_e. Any discrepancy of our results with those from *Zuber et al.* [1989] is due to the coherence measurement itself. In the discussion we will focus on the assumptions of the modeling, which account for the discrepancies between the present work and the study by *McKenzie and Fairhead* [1997].

3.5.2 Gravity and Topography Data

Gravity data were obtained from the Australian Geological Survey Organisation (AGSO; see Fig. 3-8b), which calculated Bouguer anomalies with a density of 2670 kg m^{-3}. No terrain correction was applied. According to AGSO the data precision is of the order of 0.1 mGal (1 mGal = 10^{-5} m s^{-2}), but no detailed statistics are available. It is to be expected that this uncertainty increases with the wavelength of the anomaly. At longer wavelengths the gravity field error from geoidal coefficients obtained from satellite tracking is ±0.5 mGal on the best-determined parts of the globe [*Lemoine et al.*, 1998]. Anisotropy in the data distribution was not a concern at the wavelengths that we consider.

Topography data (see Fig. 3-8a) were taken from a 30 arc sec (~1 km) digital elevation model, GTOPO30 [*Gesch et al.*, 1999]. The uncertainty for the Australian region is about ±60 m.

Both data sets were resampled at the same rate and projected onto a Cartesian grid. Prior to analyzing the data, a background gravity field obtained from a spherical harmonic expansion of satellite-derived geoidal coefficients [*Lemoine et al.*, 1998] was removed. We tapered the spherical harmonic spectrum of this background field with a cosine function which took a value of 1 at degree $l = 7$ ($\lambda \sim 5300$ km) to 0 at $l = 11$ ($\lambda \sim 3500$

km) (the degree-to-wavelength conversion is described by *Jeans* [1923] [see also *Brune*, 1964]). In this way, we attempted to filter out the gravitational contribution from large-scale mantle processes [*Dickey et al.*, 1998] (such as subducted slabs, whose gravity signature correlates well with the geoid up to about $l = 9$ [*Hager*, 1984]), while avoiding Gibbs phenomena [*Sandwell and Renkin*, 1988]. In the past, various authors have subtracted background gravity fields with different degree cutoffs: $l = 16$ [*McNutt and Parker*, 1978], $l = 22$ [*Stephenson and Beaumont*, 1980], or not at all [*Zuber et al.*, 1989; *McKenzie and Fairhead*, 1997]. We defer the discussion of the influence different choices of cutoff l have on the coherence to Sec. 3.6.2. Finally, before Fourier transformation, a best fitting plane trend was removed from the data.

3.6 Results

3.6.1 Anisotropy in the Isostatic Response

In Figs. 3-9a–3-9c, the results for each of the three geophysically distinct regions (labeled a through c in Fig. 3-8) are summarized. Fig. 3-9d analyzes the data box d marked by a dashed line in Fig. 3-8; it represents a broad average of the response of the entire continent. In the left column the two-dimensional, wave vector-dependent coherence function $\gamma^2(\mathbf{k})$ calculated with the two-dimensional multitaper method is given. On the right-hand side a comparison is made between the azimuthal average of the multitaper coherence, $\gamma^2(|\mathbf{k}|)$ (blue squares) and the isotropic mirrored periodogram method of *Zuber et al.* [1989] (red circles).

The periodogram method is standard and was applied to the data after mirroring. No window was applied. The quoted coherence values of the periodogram method have had a bias correction applied to them as follows [*Munk and Cartwright*, 1966]:

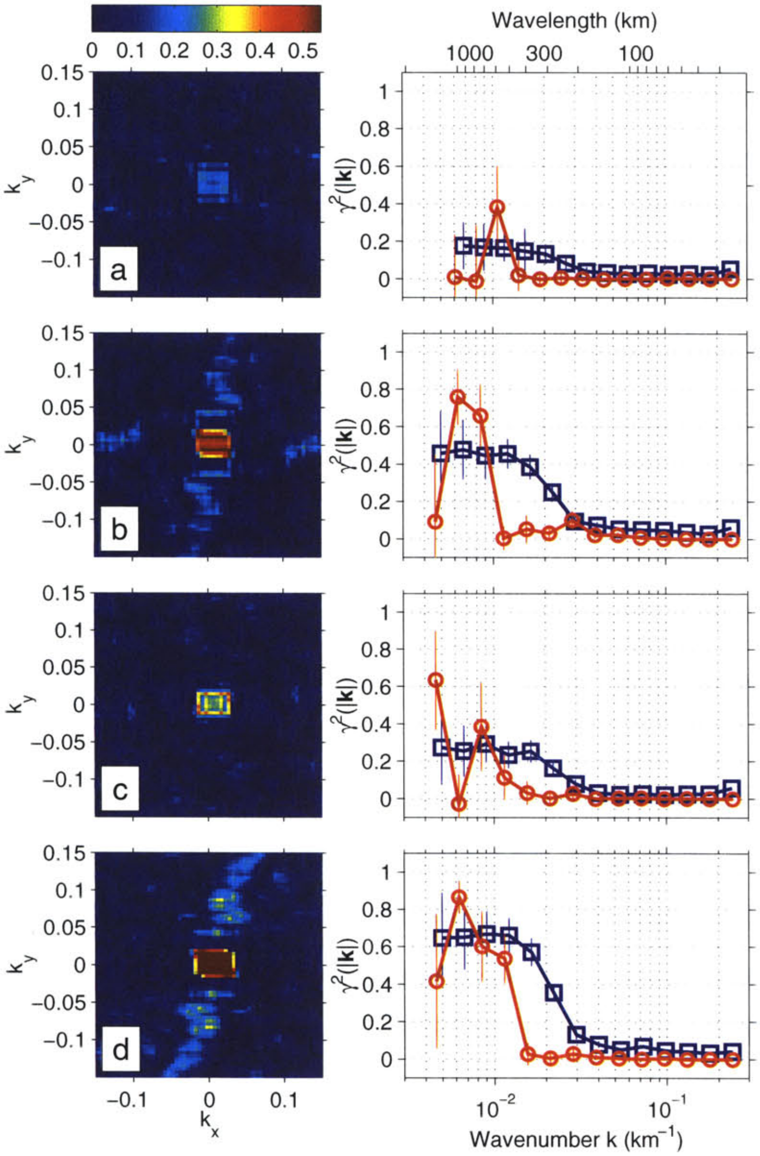


Figure 3-9: Coherence measurement of four regions of Australia. The letters a through d correspond to the labeled boxes in Fig. 3-8. (left) Multitaper 2–D coherence used as an indication of anisotropy in the isostatic response. We interpret the shape of the 2-D coherence of central Australia (Fig. 3-9b) as a direction of N–S weakness of the plate. (right) Azimuthally averaged multitaper (blue squares) and periodogram coherence (red circles). Confidence intervals are plotted at 2σ. See Secs. 3.6.2 and 3.6.2 for a discussion on the low coherence values at low wave numbers.

$$\gamma^{2'}(k) = \frac{n(k)\gamma^2(k) - 1}{n(k) - 1}. \tag{3.8}$$

The unprimed coherences refer to the those calculated before the bias correction, and $n(k)$ is the number of points in the wavenumber band. The standard error of the unbiased coherence $\gamma^{2'}$ is calculated from [*Bendat and Piersol*, 1986, 1993]

$$\Delta\gamma^{2'}(k) = (1 - \gamma^2(k))\sqrt{\frac{2\gamma^2(k)}{n(k)}}. \tag{3.9}$$

The error bars plotted correspond to $\pm 2\Delta\gamma^{2'}(k)$.

The variance of the coherence estimate $\gamma^2(\mathbf{k})$ obtained with the multitaper method is calculated using (3.24). The waveband-averaged $\gamma^2(|\mathbf{k}|)$ has this variance, normalized by the number of points in the annulus, $n(k)$. The error bars plotted are twice the standard deviation.

West and east Australia (Figs. 3-9a and 3-9c) show little or no indication of anisotropic isostatic response. In contrast, the central Australian domain (Fig. 3-9b) is characterized by a north–south direction where the coherence is higher than the azimuthal average. This feature extends over most of the wavenumber band. It is certainly more prominent than the slightly higher coherences in the E–W direction which become visible in the high-frequency range only. In Sec. 3.7.2 we will interpret this feature in terms of anisotropy in lithospheric strength. For region d, the N–S anisotropy is the most prominent feature, suggesting that it is an important factor in the isostatic response of the continent as a whole.

Consistency tests with randomly rotated data sets have shown no indication that any of the interpreted features is an artifact.

The following points can be made. (1) With the periodogram method and the bias correction of (3.8), the coherence values at long wavelengths can become negative. This occurs for low coherence values in the center of the spectral calculation, when the annuli-averaging method has at most tens of samples. A negative coherence is physically impossible and biases the solution as an inversion is performed. (2) Even after the bias correction of (3.8), the coherence is still biased toward higher values at the longer wavelengths. Again, a lack of terms in the averaging procedure is responsible for biasing the coherence toward 1. (3) The uncertainties of the estimate increase dramatically with wavelength as the number $n(k)$ decreases (see (3.9)). (4) We notice that the wavelengths at which the transition between high and low coherence occurs are longer for the periodogram method. As a result, the elastic thicknesses of the Australian continent given by *Zuber et al.* [1989] are likely to have been overestimated (see Sec. 3.6.2).

3.6.2 Isotropic Elastic Thickness Inversions

Power of uncompensated topography

McKenzie and Fairhead [1997] discuss the importance of comparing the spectral power present in the free-air gravitational anomaly with the power of the gravity anomaly due to the attraction of surface topography obtained by multiplying the elevation in meters by $2\pi\rho G$=0.11194 to give gravity anomalies in milligals (mGal). Fig. 3-10 shows that the power due to uncompensated topography is everywhere lower than that of the free-air anomaly. Erosional processes diminish the power of topography [*Stephenson*, 1984] and this limits the use of the Bouguer coherence method of *Forsyth* [1985]. In general, erosion can be treated as a noise process which randomizes the phase relationship between topography and gravity, and as such it annihilates the coherence. If this happens at the shorter wavelengths, the transitional wavelength is shifted to longer wavelengths, and the T_e estimates are biased upward. The presence of uncorrelated buried loads (demonstrably present in the Australian lithosphere [e.g., *Dooley*, 1977]), however, has the same effect of lowering the power of the uncompensated topography with respect to the free-air anomaly, and this effect is most appropriately modeled by the inclusion of buried loads in the analysis. (See

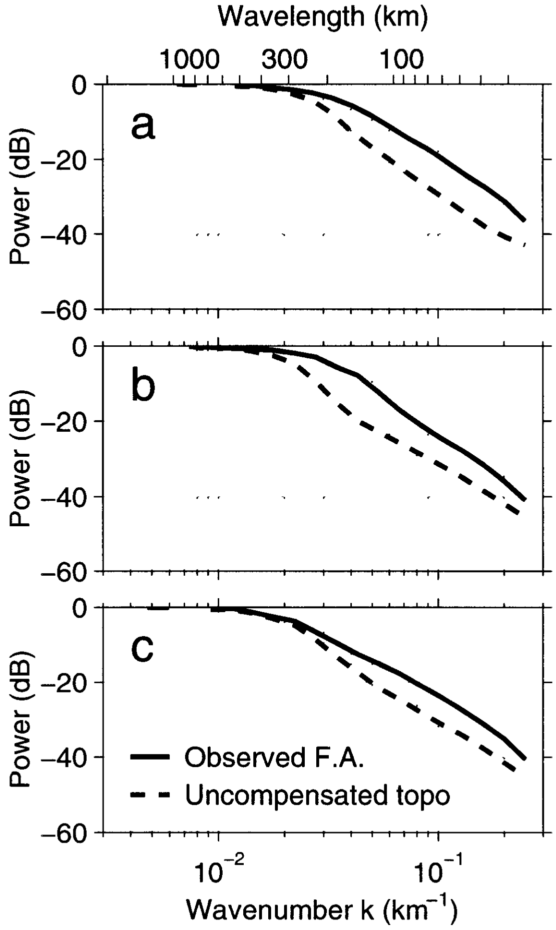


Figure 3-10: Power in the free-air gravity field compared to the power of the gravitational attraction due to uncompensated topography. The letters a, b, and c correspond to the labeled boxes in Fig. 3-8.

also the examples in the Canadian Shield [*Wang and Mareschal*, 1999] and Mars [*Zuber et al.*, 2000], and *Burov et al.* [1998]) Certainly, the fact that erosion subdues the topography severely limits the applicability of any correlation-based method for the estimation of the effective elastic thickness, but we do not believe that using the free-air anomaly admittance and neglecting the presence of subsurface loading as proposed by *McKenzie and Fairhead* [1997] is the answer. If erosion obfuscates the relationship between topography and gravity, then working with topography-related gravity anomalies only is not very likely to shed any more light on the true value of T_e. Moreover, the significant erosion has arguably largely occurred after the "true" value of the effective elastic thickness has left its signature on the Bouguer anomaly [but see also *Stephenson*, 1984]. We have therefore elected to retain the Bouguer coherence method but focus on the effect of the multitaper versus the modified periodogram method of coherence estimation for the isotropic T_e calculation.

Influence of the degree of the subtracted geoidal field

In Fig. 3-9 it can be observed that the coherence between Bouguer gravity and topography does not approach unity at the longest wavelengths. This may seem at odds with the commonly made assumption that those long wavelengths should be in complete Airy-isostatic equilibrium. Geodynamical modeling [e.g., *Zhong*, 1997] has suggested, however, that the elastic strength of the plate may not be the only mechanism to cause deviation from isostasy; this deviation may also have a dynamic origin. Long-wavelength gravity anomalies are then the surface expression of convective circulation in the mantle [*D'Agostino and McKenzie*, 1999]. We have experimented with different values of spectral cutoff ($l = 2$ through 8, 16 and 24) of the subtracted background gravity field (see Sec. 3.5.2) and recomputed the coherence functions. The resulting variations of the long-wavelength coherence ($\lambda \geq 300$ km) are small (within a few %) but real. However, given the substantial uncertainties on our final T_e values, the effects on the results and interpretations given in this chapter are minor.

Influence of resolution parameter *NW*

For a large part this chapter is concerned with the calculation of 2-D coherence functions in an attempt to detect anisotropy in the isostatic response. As a result, we have tried to obtain a maximum reduction of the estimation variance of $\gamma^2(\mathbf{k})$ before any azimuthal

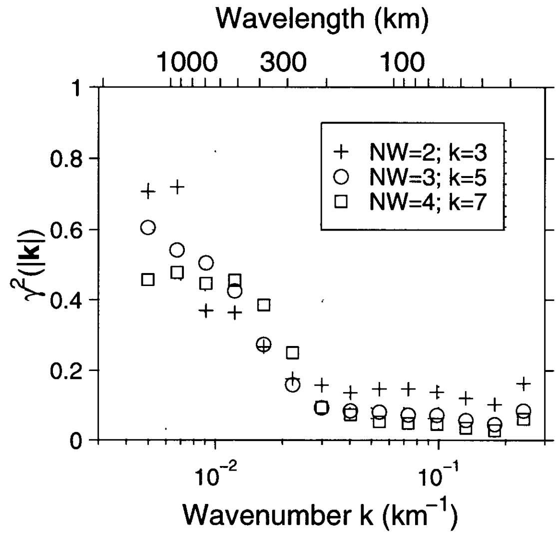


Figure 3-11: Influence of resolution parameter NW on the central Australian coherence data (box b in Fig. 3-8). The physical resolution $W = j/(N\Delta t)$ for $j = 2, 3, 4$ (quoted as $NW = 2, 3, 4$) depicted in this figure corresponds for this data set to a resolution in wavenumber space of 0.01, 0.015, and 0.02 rad/km.

averaging. The synthetic tests presented in Sec. 3.4.3 have indicated that an acceptable value for the resolution parameter is $NW = 4$. For the specific case of central Australia we have investigated how the value of $\gamma^2(\mathbf{k})$ is influenced by the choice of the resolution parameter NW and the associated multiplicity of the windows $k = 2NW - 1$. In Fig. 3-11 it can be seen that using smaller values of NW increases the coherence values at low wavenumbers. However, it can be verified that all these values fall within each other's uncertainties. In Fig. 3-5 we have inferred from the width of the central lobe of the spectral windows associated with the tapering that the wavenumber resolution does not suffer too greatly from increasing NW. In 2-D, the estimation variance is reduced as the square of the number of tapers, which is why we have adopted a value of $NW = 4$. As a consequence, we have to pay particular attention to the transitional wavelength, and downweight the influence of the low-wavenumber coherences when evaluating the model fits with predicted coherences.

Influence of box size

From the foregoing, it is apparent that the box sizes used in this study are adequate to capture the flexural wavelengths and the transition from high to low coherence. Increasing or decreasing the size of the data selection modestly, as we have verified, does not alter the coherence measurements in any appreciable way. *Lowry and Smith* [1994] have in fact pointed out that it is possible to make reliable measurements with box sizes much smaller than ours. Moreover, our selection of three dominant geophysical provinces has an appealing physical relevance.

Subsurface to surface loading ratios

The necessity of including buried loads can be substantiated by calculating the observed ratio of bottom to top loading f, as defined by *Forsyth* [1985]. This ratio is wavelength dependent. We have calculated this f ratio and found that it increases with wavelength but stays well within the range of 0.2 to 5.0 postulated by *Forsyth* [1985] for the transitional wavelengths. Similar behavior was found for the elastic thickness estimates of the Canadian Shield by *Wang and Mareschal* [1999]. The choice $f = 0$ made by *McKenzie and Fairhead* [1997] appears to be a philosophical one.

Calculation of errors

In the inversion for isotropic elastic thickness we follow a procedure identical to that used by *Zuber et al.* [1989]. Any discrepancy in our calculations is due to the intrinsic uncertainty of the coherence measurement. The influence of the gravity data itself and the revised Moho depth estimates (interpolated from *Wellman* [1979], *Shibutani et al.* [1996], and *Clitheroe et al.* [2000a]) are expected to be minor. In the inversion we use an L_2 norm of the model misfit divided by the standard error; this is identical to the work by *McKenzie and Fairhead* [1997]. The results for the three regions of Fig. 3-8 are given in Figs.

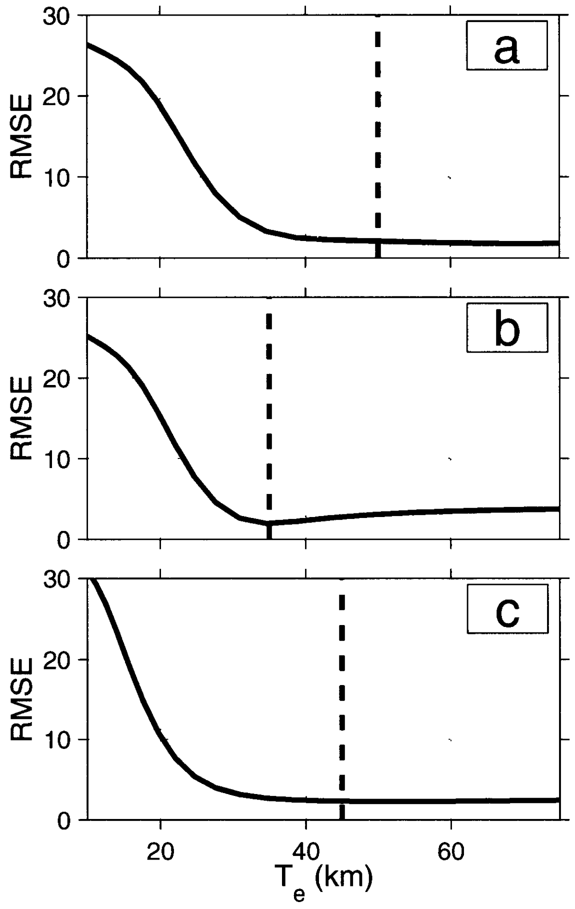


Figure 3-12: Model misfit in the inversion of coherence for effective elastic thickness for the three analysis regions corresponding to the lettered boxes in Fig. 3-8. RMSE, root-mean-square error.

3-12a–3-12c. We have excluded the first three wavenumbers from the error calculation. This is based on the fact that we are trying to capture the transitional wavelength, and the observation made from Fig. 3-11 that the first few wavenumbers are relatively sensitive to the choice of NW. Only the central Australian domain presents a clear minimum in the root-mean-square error (RMSE) (Fig. 3-12b). We have kept the value of that minimum error in choosing the optimal T_e for both western (Fig. 3-12a) and eastern Australia (Fig. 3-12c). From the plots, however, it should be clear that the range of allowable T_e values is fairly broad, and curves like this may be used to infer an absolute uncertainty of our T_e estimate of at least ±10 km. This uncertainty is purely based on the data and the quality of the fits; even more is to be expected when variations in Young's modulus or Poisson's ratio of the crust are to be incorporated [*Lowry and Smith*, 1995].

Elastic thickness values

Fig. 3-13, finally, represents the comparison of multitaper T_e estimates with the estimates derived from the mirrored periodogram method, which are essentially identical to the values given by *Zuber et al.* [1989], only we reprogrammed their codes and used more recent data sets. Figs. 3-13a–3-13c give the multitaper results (solid), for the three regions labeled a through c in Fig. 3-8. Figs. 3-13d–f show the corresponding periodogram estimates (shaded). In all analyses a Poisson ratio of 0.25, a Young's modulus of 10^{11} N m^{-2}, a crustal density of 2650 kg m^{-3}, and a mantle density of 3300 kg m^{-3} were maintained. Moho depths, taken from *Wellman* [1979], *Shibutani et al.* [1996] and *Clitheroe et al.* [2000a], were fixed at 50 km (Figs. 3-13a and 3-13d), 40 km (Figs. 3-13b and 3-13e) and 35 km (Figs. 3-13c and 3-13f).

As we have seen in Fig. 3-9, the transitional wavelengths obtained with our method are shorter than those from the periodogram method. Also, the new coherence estimates are better behaved. The only difference lies in the coherence estimation, but as can be seen, its effects on the effective elastic thickness are relatively profound. The values for T_e that we obtain here are as much as a factor of 2 lower than those found by *Zuber et al.* [1989].

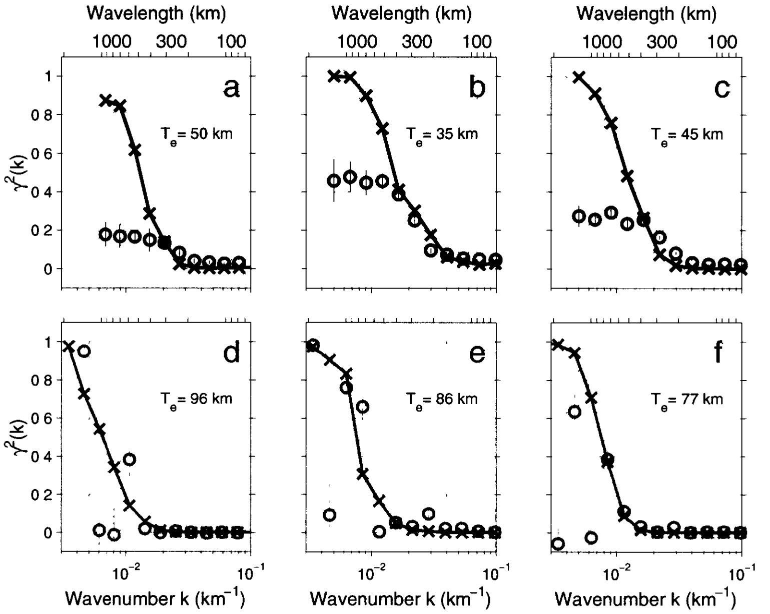


Figure 3-13: Isotropic T_e inversions of the three boxes lettered a through c in Fig. 3-8. (a–c) Inversion with multitaper coherence estimates. Circles, measurements with 2σ error bar; crosses, prediction. (d–f) Same for periodogram measurements. The inversion for both coherence measurement methods is identical: topographic loading and one interface of subsurface loading are considered [*Forsyth*, 1985; *Zuber et al.*, 1989], the interface being the local average Moho depth of 50 km for Figs. 3-13a and 3-13d, 45 km for Figs. 3-13b and 3-13e and 35 km for Figs. 3-13c and 3-13f.

3.7 Discussion

3.7.1 Anisotropic Mechanical Properties

A zeroth-order model for an elastically anisotropic lithosphere is one in which the average plate strength is made up of one direction which is "stronger" than the isotropic average, and another, "weaker" direction.

Stress state

The stress state of the continent is an important factor in the creation of anisotropic coherence functions [*Lowry and Smith*, 1995]. Flexural isostasy proposes that it is the stress distribution within the plate that supports the weight of the loads. In-plane tectonic (deviatoric) stresses (which are adequate representations of the lithospheric stress state [*Turcotte and Schubert*, 1982]) are predominantly associated with plate tectonics, through mechanisms of ridge push, slab pull, continental convergence, viscous drag at the base of the lithosphere or curvature changes with latitude [*Lambeck et al.*, 1984; *Coblentz et al.*, 1995, 1998]. Other processes are also at play, and moreover, the stress state may be inherent to the mechanical and thermal properties of the lithosphere [*Lambeck et al.*, 1984]. In viscoelastic models, stress relaxation rates may vary with direction. *Stephenson and Lambeck* [1985] proposed that regional compressive stresses help support near-surface loads in preferred directions. However, *Lowry and Smith* [1995] have shown that directions of both maximum compressive and extensional tectonic stress lower the T_e values in the same direction due to the shifting of the yield envelope which reduces the depth-integrated fiber stresses and hence the T_e. Unfortunately, very few extensive in situ stress measurements are available for Australia. Borehole break-out measurements, where available, are notorious for their scatter [*Hillis et al.*, 1998, 1999]. However, on the basis of focal mechanisms of the infrequent earthquakes in the area the direction of maximum compressive stress for central Australia scatters around N–S [*Lambeck*, 1983; *Lambeck et al.*, 1984; *Stephenson and Lambeck*, 1985; *Zoback*, 1992; *Spassov*, 1998]. For a N–S maximum compressive stress orientation we expect a T_e reduction in the same direction.

Moho depth

In an inversion for T_e, both surface and subsurface loading can and should be considered. However, subsurface interface depths need to be maintained throughout the region under study. Although the effect of uncertainties in the locations of the internal density stratification on the coherence is reported to be minor [*Forsyth*, 1985, Figure 9], azimuthal crustal thickness variations (such as those that characterize central Australia) will introduce a variation of mechanical properties with direction. The crustal thickness does influence T_e, if only because of the difference in strength of crust and mantle [*Burov and Diament*, 1995].

Temperature-induced strength variations

Differential burial rates can induce Moho temperature variations coupled to variations in lithospheric strength of 1–4 orders of magnitude [*Sandiford and Hand*, 1998]. For elongated basins such as the Amadeus basin in central Australia, the sediment infill has an E–W axis of symmetry, which has induced directional variations in strength.

Geologic factors

Furthermore, there are several geological reasons that act on widely differing scales, on the basis of which the mechanical properties of the lithosphere could vary azimuthally. Intrinsically, these include the anisotropy of mantle materials, and extrinsically, their preferential alignment or emplacement, or the geometry of weak zones in the crust exemplified by faulting or earthquake activity. See *Vauchez et al.* [1998] for a comprehensive discussion on mechanical anisotropy and deformation of continental lithospheres.

Faulting

In the case of a roughly parallel fault system of subvertical faults, loads varying along the strike of the faults would be supported by the strength of the plate, while loads perpendicular to the strike could be partially compensated by fault motion [*Bechtel*, 1989]. In such a case, the apparent elastic thickness of the lithosphere would be directionally anisotropic. Analyses for the western United States have indicated that the apparent rigidity is indeed azimuthally anisotropic with maximum rigidity parallel to the average orientation of faults in the region [*Bechtel*, 1989; *Lowry and Smith*, 1995].

Small-scale convection

Anisotropy in the response of the Canadian shield was proposed by *Stephenson and Beaumont* [1980] to indicate small-scale sublithospheric convective flow. This adds a dynamic component to the notion of the isostatic response [see also *Sandwell and Renkin*, 1988; *Zhong*, 1997; *D'Agostino and McKenzie*, 1999]. The presence of small-scale convection under central Australia appears unrealistic [*Simons et al.*, 2002a], as does our ability to detect it beneath thick cratons.

3.7.2 Anisotropy in the Central-Australian Isostatic Response

We propose that the observed reduction in T_e values in the N–S direction of the central Australian domain can be explained by a concurrence of three processes: (1) pervasive parallel faulting, (2) the regional stress field, and (3) intrinsic strength variations induced by temperature differences at the Moho.

For central Australia, fault directions are E–W, and the direction of maximum compressive stress is N–S. Regardless of the sign, the maximum stress direction lowers T_e in that direction [*Lowry and Smith*, 1995]. Our results are consistent with the idea that the E–W oriented parallel faults in the central-Australian regions act to weaken the plate in the N–S direction, and this is fully consistent with the observed N–S direction of maximum compressive stress.

Recently, *Sandiford and Hand* [1998] and *Hand and Sandiford* [1999] have postulated that thermal factors may have controlled the localization of deformation in central Australia during both the Petermann (600–520 Ma) and Alice Springs (400–300 Ma) orogenies. They invoke sediment thickness variations capable of producing Moho temperature disturbances which they equate with variations in effective lithospheric strain rate of 1–4 orders of magnitude. This thermally controlled localization of strain could help explain the extensive deformation during both orogenies, while at the same time the lithosphere is to this day strong enough to sustain substantial gravity anomalies. We point out that the sediment isopachs during the Alice Springs orogeny (after which the region has been tectonically stable [*Lambeck et al.*, 1984]) have an E–W axis of symmetry. We postulate that E–W-running zones of low lithospheric strength might have the same effect as the predominant E–W faulting direction in the region. This additional effect would further explain our observation of a weaker N–S direction of the central Australian region.

3.7.3 Elastic Thickness of the Australian Lithosphere

McKenzie and Fairhead [1997] call the elastic thicknesses obtained with Forsyth's method overestimated and ascribe this fact as solely due to the significant erosion of topography on the Australian continent. We agree, on statistical grounds, that the introduction of random noise by erosion can act to make the transitional wavelengths longer and thus the plate seem too thick. This happens when the magnitude of this noise becomes comparable with the magnitude of the topographical wavelengths it affects. Ideally, flexural response and erosion should be studied as coupled processes, but there is substantial uncertainty as to the precise wave vector dependence of the erosion time constant spectrum [*Stephenson*, 1984]. There is a priori no need for expecting high T_e-values to be unreasonable. A T_e of $\sim$100 km in regions of low (30–50 mW m^{-2}) surface heat flow implies a typical crustal rheology and a mantle activation energy of about 500 kJ mol^{-1}, perfectly consistent with experimental data for, e.g., olivine (A. R. Lowry, *personal communication*, 2000).

In the analysis by *McKenzie and Fairhead* [1997] the presence of buried loads is found unnecessary (they prefer $f = 0$), yet they are demonstrably present in Australia [e.g., *Dooley*, 1977; *Goncharov et al.*, 1998; *Zhang et al.*, 1998], and if not accounted for, they result in an underestimate of T_e [*Forsyth*, 1985]. Our results suggest that this shortcoming can be overcome by retaining the general method of *Forsyth* [1985] and thus the ability to include buried loads into the analysis but performing the analysis with the multitaper method. This diminishes the influence of the long-wavelength bias in the coherence and allows the identification of anisotropy, with which the elastic thickness value can be assessed.

When neglected, azimuthal anisotropy in the isostatic response may bias the result of inversions. It may provide a physical reason to be cautious about the interpretation of the various models and their mutual differences. Lithospheric strength is a complex geophysical quantity as it is always measured as a geographical and temporal (in other words, geological) average [*Burov and Diament*, 1995]. With the anisotropy we have determined in the response functions of the Australian continent, the concept of "strength" will need to be expanded to include the notion that it is a directional average as well. To quantify the precise effect of the azimuthal variability on the isostatic response will require the use of forward models.

Zuber et al. [1989, p. 9353] wrote that "Regions within the continent with different ages and tectonic histories exhibit different elastic thicknesses that increase with the time since lithospheric stabilization." [See also *Karner et al.*, 1983; *Bechtel et al.*, 1990; *Poudjom Djomani et al.*, 1999]. This picture is definitely a first-order one. In the past decade, much research has been done on the control on the effective elastic thickness of the thermal budget of the continent, sediment covering, the ambient stress field, strain rate, continental rheology, crust-mantle coupling, faulting, composition, plate curvature, crustal thickness and numerous other factors [see, e.g., *Burov and Diament*, 1995; *Lowry and Smith*, 1995; *Ebinger and Hayward*, 1996; *Lavier and Steckler*, 1997, 1998; *Burov et al.*, 1998, and references therein]. T_e values are integrations of all these processes through time and space and the correlation with surface observables such as crustal ages is not unambiguous.

On the whole, our isotropic T_e values are fairly similar for all three regions. Given the large number of competing factors influencing T_e and their intrinsic variability within those regions, this is not unexpected. Recent observations have indicated that even the direct surface observables are not as homogeneous throughout regions of equal crustal age as was once thought. To cite three examples, there are (1) large variations in heat flow values, even within similar tectonic and age domains due to significant crustal contributions [*Jaupart and Mareschal*, 1999]; (2) significant differences in upper mantle seismic structure within domains of seemingly similar geologic signature [*Simons et al.*, 1999] (in contrast to earlier inferences that the depth of high-velocity keels under stable continents was well correlated with formation age [*Polet and Anderson*, 1995]); and (3) large uncertainties in the state of stress of the continent [*Hillis et al.*, 1998] and analytical models thereof [*Coblentz et al.*, 1995].

We envisage that further analysis will include detailed comparisons with forward models and the extension of this method to models with higher spatial resolution (optimizing the method to account for smaller box sizes, as given by *Zuber et al.* [1989] and *Lowry and Smith* [1994]). Crust-mantle decoupling is expected to have a profound effect on the mechanisms of isostatic compensation [*ter Voorde et al.*, 1998], and subsequent work will investigate the potential interplay between the instantaneous elastic response from seismic shallow upper mantle studies [*Simons et al.*, 1999] and its azimuthal anisotropy [*Simons et al.*, 2002a] and the long-term mechanical isostatic response.

3.8 Conclusions

The focus of this chapter has been to reassess the nature of the isostatic response to loading in Australia. We extended the multitaper method of *Thomson* [1982] to two dimensions to derive an estimate of coherence that is less affected by bias than other methods of spectral analysis due to the ability to make more independent estimates of the response function.

This method also allows us to detect anisotropy when it is present. Our results show the presence of anisotropy in the response of Bouguer gravity to topography in central Australia. We interpret the anisotropic coherence as indicative of a weaker elastic lithosphere in the N–S direction, perpendicular to the E–W direction of faulting, consistent with the T_e-lowering effect of the N–S direction of maximum compressive stress [*Lowry and Smith*, 1995] and possibly also correlated with proposed E–W trending variations in Moho temperature associated with differential sediment loading [*Sandiford and Hand*, 1998].

Isotropic coherence estimates made by averaging out the anisotropy of our anisotropic models are more robust than those obtained by traditional isotropic methods. Using Bouguer gravity and topography, our multitaper coherences yield elastic thicknesses that are up to a factor of 2 lower than previous inversions that used periodogram coherences [*Zuber et al.*, 1989]. But our elastic thickness values are greater than suggested by admittance inversions based on free-air gravity that ignored the presence of subsurface loads [*McKenzie and Fairhead*, 1997]. The significant erosion of Australia is one likely cause for our T_e estimates to be on the high side; on the other hand, the bottom-to-top loading ratio of $f = 0$ preferred by *McKenzie and Fairhead* [1997] biases the T_e values toward lower values.

3.9 Appendix I: Multitaper Estimation of Coherence

Coherence is the wavenumber domain analogue of correlation. The squared coherence function $\gamma^2_{fg}(\mathbf{k})$ between two stochastic processes $\{f(\mathbf{r})\}$ and $\{g(\mathbf{r})\}$ is the magnitude squared of the cross-spectral density function S_{fg} of both fields, normalized by the power spectral density functions of the individual fields S_{ff} and S_{gg}:

$$\gamma^2_{fg}(\mathbf{k}) = \frac{|S_{fg}(\mathbf{k})|^2}{S_{ff}(\mathbf{k})S_{gg}(\mathbf{k})}. \tag{3.10}$$

Here $\mathbf{r} = (x, y)$ and $\mathbf{k} = (k_x, k_y)$ denote spatial coordinates and wave vectors, respectively.

The estimation of the power spectral density function (and by extension, of the cross-spectral density function) forms the subject of this section. For more detailed treatments of the multitaper method, see *Thomson* [1982], *Slepian* [1983], and *Percival and Walden* [1993]. A comprehensive discussion about quadratic spectrum estimators in general is given by *Mullis and Scharf* [1991]. Performance comparisons with other methods are given by *Bronez* [1992], *Lees and Park* [1995], *Riedel and Sidorenko* [1995], and *Komm et al.* [1999].

In the following, we will consider 1-D processes (time series). We drop the subscripts identifying the fields and give the dependence in terms of the temporal (or spatial) variable t and the frequency f, denoting the sampling interval with Δt. The Nyquist frequency $f_N = 1/(2\Delta t)$. The asterisk is used to denote complex conjugation. Let x_t be a portion of length N of a stationary process whose power spectral density function is given by $S(f)$. Suppose we were to window those data with a taper h_t, by forming the (point-by-point) product $y_t = h_t x_t$. The quantity $\hat{S}(f)$, given by

$$\hat{S}(f) = \Delta t \left| \sum_{t=1}^{N} h_t x_t e^{-i2\pi f t \Delta t} \right|^2 = \Delta t\, Y(f) Y^*(f), \tag{3.11}$$

is a direct spectral estimator for the stochastic process. This can be seen as follows. Applying the Cramér spectral representation theorem [*Priestley*, 1981], it follows that the first-moment properties of this spectral estimator are given by

$$E\{\hat{S}(f)\} = \int_{-1/(2\Delta t)}^{1/(2\Delta t)} \mathcal{H}(f-f')S(f')\,df', \tag{3.12}$$

so $\hat{S}(f)$ is indeed a spectral estimator; it represents the convolution of the true power spectral density function with the function $\mathcal{H}(f)$. $\mathcal{H}(f)$ is called the spectral window, obtained from the windowing coefficients by Fourier transforming

$$\mathcal{H}(f) = \frac{1}{\Delta t}|H(f)|^2 = \Delta t \left|\sum_{t=1}^{N} h_t e^{-i2\pi f t \Delta t}\right|^2. \tag{3.13}$$

It is the sidelobe level of $\mathcal{H}(f)$ that determines the extent to which $\hat{S}(f)$ is free of leakage. Ideally, $\mathcal{H}(f)$ is a delta function.

Next, suppose we define the multitaper method (MTM) estimate of the spectrum to be the (possibly weighted) average of k direct spectral estimators $\hat{S}_l, l = 1...k(f)$

$$\hat{S}^{\mathrm{MTM}}(f) = \frac{1}{k}\sum_{l=0}^{k-1} \hat{S}_l(f), \tag{3.14}$$

where each $\hat{S}_l$ is given by an equation of the kind of (3.11). From (3.12) it then follows that the first moment of this multitaper spectral estimate is given by

$$E\{\hat{S}^{\mathrm{MTM}}(f)\} = \int_{-f_N}^{f_N} \bar{\mathcal{H}}(f-f')S(f')\,df', \tag{3.15}$$

with

$$\bar{\mathcal{H}}(f) = \frac{1}{k}\sum_{l=0}^{k-1} \mathcal{H}_l(f). \tag{3.16}$$

Note that $1/(2\Delta t) = f_N$ corresponds to the Nyquist frequency. The degree to which the estimated spectrum is free of leakage is determined by the leakage characteristics of the preponderant data tapers. But more importantly, the estimation variance of the estimator $\hat{S}^{\mathrm{MTM}}(f)$ is much smaller than that of the individual spectra $\hat{S}_l(f)$. If the $\hat{S}_l(f)$ are mutually uncorrelated with common variance, then the variance of the resulting estimate $\hat{S}^{\mathrm{MTM}}(f)$ will be smaller than that of $\hat{S}_l(f)$ by a factor of $1/k$.

The sidelobe level of the $\mathcal{H}_l(f)$ (in other words, its deviation from a delta function) is crucial to the unbiasedness of the spectral estimate. We desire a set of tapers that, while limited in the time/spatial domain by the finite data size N, are optimally band limited in the sense that they have the majority of their energy concentrated in some frequency interval $[-W, W]$, where $-f_N < -W < f < W < f_N$. Clearly, $2W$ represents the resolution bandwidth of the spectral estimation using such a window. With the definition of $\mathcal{H}(f)$ as in (3.13), the frequency concentration $\beta^2(W)$ of the data tapers is measured as the fraction of energy in this frequency interval, by

$$\beta^2(W) = \int\limits_{-W}^{W} \mathcal{H}_l(f)\, df \Bigg/ \int\limits_{-f_N}^{f_N} \mathcal{H}_l(f)\, df \,, \tag{3.17}$$

and, desirably, $\beta^2(W) > 1/2$. For concentration in the time domain we define

$$\alpha^2(N) = \sum_{t=0}^{N-1} |h_{t,l}|^2 \Bigg/ \sum_{t=-\infty}^{\infty} |h_{t,l}|^2 \,, \tag{3.18}$$

and again, if $\alpha^2(N) > 1/2$, then the majority of the energy is concentrated in the index range 0 to $N-1$.

A nontrivial band limited signal cannot be time limited as well. We may ask how large $\alpha^2(N)$ can be for $h_{t,l}$ band limited to $|f| \leq W < f_N$. The concentration problem amounts to an optimization. It is straightforward to rewrite the concentration measure of (3.18) in the frequency domain as

$$\begin{aligned}\alpha^2(N) = &\left(\int_{-W}^{W} |H_l(f)|^2\, df\right)^{-1} \\ &\int_{-W}^{W}\int_{-W}^{W} H_l(f) N\mathcal{D}(f, f'; N) e^{i\pi(N-1)(f-f')} H_l^*(f')\, df\, df',\end{aligned} \tag{3.19}$$

where $\mathcal{D}(f, f'; N)$ denotes Dirichlet's kernel given by

$$\mathcal{D}(f, f'; N) = \frac{1}{N}\frac{\sin[N\pi(f' - f)]}{\sin[\pi(f' - f)]}. \tag{3.20}$$

We can recognize (3.19) as a formulation of Rayleigh's principle of variational analysis [*Dahlen and Tromp*, 1998], from which we deduce that the value of $\alpha^2(N, W)$ in (3.19) will be extremized if the following equation is satisfied:

$$\int_{-W}^{W} N\mathcal{D}(f, f'; N) e^{i\pi(N-1)(f-f')} H_l(f)\, df = \alpha_l^2 H_l(f'), \tag{3.21}$$

for $|f'| \leq W$. It is convenient to reformat (3.21) to the time domain. After some manipulation (as given by *Park et al.* [1987a] and *Lindberg and Park* [1987]), we can rewrite it as

$$\mathbf{C} \cdot \mathbf{h} = \alpha^2(N)\mathbf{h}, \tag{3.22}$$

where $\mathbf{C}$ is the $N \times N$ Toeplitz matrix whose elements are given by

$$\mathbf{C}_{tt'} = \frac{\sin[2\pi W(t - t')]}{\pi(t - t')} \quad ; \quad t, t' = 0, 1, ..., N - 1, \tag{3.23}$$

and $\mathbf{h}$ is the matrix containing the various data tapers. The matrix $\mathbf{C}$ has a finite set of eigenvalues and eigenfunctions, and the spectra $\hat{S}_l(f)$ are appropriately known as eigenspectra. Any of numerous numerical calculation schemes can be used to perform the required diagonalization. The data windows are known as discrete prolate spheroidal sequences (dpss) or Slepian sequences [*Slepian*, 1978, 1983]. The first $2WN\Delta t$ (i.e., the Shannon number) eigenvalues are close to unity, and then the eigenvalues fall off rapidly to zero (see Fig. 3-2).

The data tapers are orthogonal. This causes the individual eigenspectra of (3.12) to be approximately uncorrelated, from which the desired variance reduction in the multitaper estimate follows. It is left to the analyst to select the resolution bandwidth $2W$. Typically, W is taken to be a small multiple $j > 1$ of the fundamental frequency, such that $W = j/(N\Delta t)$ for $j = 4$. The trade-off between resolution and variance is nicely illustrated by the following: as W increases, so does the number of tapers with good leakage properties, and the variance of the resulting estimate decreases. On the other hand, with the width of $2W$, the resolution also diminishes.

Hence the multitaper method consists of (1) finding the eigentapers for a given resolution bandwidth by diagonalizing matrix $\mathbf{C}$ (equation (3.23)), (2) windowing the data, (3) taking the Fourier transform of the windowed data and calculating eigenspectra as in (3.11), and (4) performing the average as in (3.14), weighted with the eigenvalues. The calculation of the cross-spectral density function in order to arrive at the coherence function of (3.10) is as in (3.11), except the product involves both data sets. The estimated squared coherence function $\hat{\gamma}^2_{fg}$ will have almost the same distribution as is usually assigned to squared coherence estimates for k independent realizations of a time series. *Thomson* [1982] gives the estimation variance σ^2 as

$$\sigma^2 = \frac{k-1}{k(k+1)} \left[\frac{1}{k} + 2\frac{k-2}{k+2}\sqrt{\hat{\gamma}^2_{fg}} \right] . \tag{3.24}$$

This measure gives the variance of the 2-D coherence estimation. In case the coherence is averaged azimuthally, the average variance is further divided by the number of points in the wave band. The standard deviation $\sigma = \sqrt{\sigma^2}$ is used as our error estimate.

The extension of the one-dimensional multitaper technique is easily made [*Liu and van Veen*, 1992; *Hanssen*, 1997]. The multidimensional eigentapers can be constructed as outer products of all the possible combinations of their 1-D counterparts (see Fig. 3-3). So for n dimensions, the number of tapers is $K = k^n$. The associated spectral windows have been given in Fig. 3-4. The corresponding eigenvalues are obtained as the products of the original ones.

We have indicated how the expected value of the direct spectral estimates are equal to the true spectrum convolved with the power spectrum of some window function (equation 3.12). Subsequently, we have shown how data windows of length N can be found in a manner that concentrates the majority of their energy in a (narrow) frequency band $2W$. The resulting eigenspectra are independent, providing “new” realizations (or optimal projections) of the same data set. Variance reduction of the estimate is obtained by averaging the eigenspectra.

3.10 Appendix II: Predicting admittance and coherence

3.10.1 Flexure of the elastic lithosphere

The equation describing the deflection $n(x)$ of a thin elastic plate due to a load $l(x)$ is given by the one-dimensional flexure equation [*Turcotte and Schubert*, 1982]:

$$\left(\frac{d^4}{dx^4} + \frac{\Delta_{\mathrm{R}}}{D} g\right) n(x) = -\frac{\Delta_{\mathrm{F}} g}{D} l(x). \tag{3.25}$$

Here, Δ_{R} represents the density contrast causing the flexed interfaces to exert an upward restoring force, and Δ_{F} is the driving density contrast given by the density contrast across the interface being loaded. The flexural rigidity is denoted by D, and g is the gravitational acceleration. The effective elastic thickness, T_e, relates to D by

$$D = \frac{ET_e^3}{12(1-\nu^2)}, \tag{3.26}$$

where E is Young's modulus and ν is Poisson's ratio.

For wavenumber $k = 2\pi/\lambda$, the Fourier transform of the load $L(k)$ is related to the Fourier transform of the deflection $N(k)$ by

$$N(k) = -L(k)\frac{\Delta_{\mathrm{F}}}{\frac{Dk^4}{g} + \Delta_{\mathrm{R}}}. \tag{3.27}$$

Let's examine the driving and restoring forces in the four-layer (three interface) density case composed of a water layer (density ρ_w), an upper (ρ_u) and a lower crust (ρ_l), and a half-space mantle (ρ_m). We represent the topography loading interface i by W_i and the topographic surface expression caused by loading at interface i by ${}_iH$. The density jumps at the interfaces are given by $\rho_u - \rho_w = \Delta_1$, $\rho_l - \rho_u = \Delta_2$ and $\rho_m - \rho_l = \Delta_3$.

For loading at the surface (the fluid-solid interface, $i = 1$), the load is identified with the topography. The deflection due to loading on interface $i > 1$ is assumed to be expressed

i	Δ_F	Δ_R	Surface	Load	Deflection	Initial
1	Δ_1	$\Delta_2 + \Delta_3$	${}_1H$	$W_1 = {}_1H$	$W_2 = W_3$	$[{}_1H - W_2]g\Delta_1$
2	Δ_2	$\Delta_1 + \Delta_3$	${}_2H$	W_2	${}_2H$	$[W_2 - {}_2H]g\Delta_2$
3	Δ_3	$\Delta_1 + \Delta_2$	${}_3H$	W_3	${}_3H$	$[W_3 - {}_3H]g\Delta_3$

Table 3.1: Forcing and restoring density differences for loading on interface i.

equally at all interfaces $j \neq i$. All these loading scenarios, therefore, represent "coherent" loading.

We note that all values represent equilibrium values; the "applied loads" are then given by the sum of the load (from Table 3.1) and the resulting deflection.

In the assumption that any topography can be represented by a surface density of equivalent mass, the gravity anomaly associated with a warped interface W_i located at depth z_i is given by $\Delta_i 2\pi G W_i(k) e^{-kz_i}$. If the anomaly is free-air, the surface contribution is taken into account by $\Delta_1 2\pi G_i H(k)$, but not for the Bouguer anomaly.

3.10.2 Admittance: The response of gravity to topography

From Eq. 3.27 and Table 3.1, loading at the surface (interface $i = 1$) produces warped interfaces, $W_{i>1}$ at z_i, related to the loading topography, ${}_1H$, by the factor

$$W_i = -{}_1H \frac{\Delta_1}{\left[\frac{Dk^4}{g} + \sum\limits_{j>1} \Delta_j\right]} \quad , \quad i > 1. \tag{3.28}$$

The Fourier transform of the Bouguer gravity anomaly, ${}_1G_B$, is the sum of the individual gravity anomalies $\Delta_i 2\pi G W_i e^{-kz_i}$ resulting from the surface loading and hence the **Bouguer admittance**, Z_1 (the ratio of gravity anomaly to surface expression) is expressed by:

$$\boxed{Z_1 = -2\pi G \frac{\Delta_1}{\left[\frac{Dk^4}{g} + \sum\limits_{i>1} \Delta_i\right]} \sum_{i>1} \Delta_i e^{-kz_i}.} \tag{3.29}$$

Similarly, the **free-air admittance** is given by $Z_1 + 2\pi G \Delta_1$. From Eq. 3.27 and Table 3.1, loading $W_{i>1}$ on an interface different from the surface and the subsequent deflection on other interfaces $W_{j \neq i}$ (expressed identically at the surface as ${}_iH$) are related as:

$${}_iH = -W_i \frac{\Delta_i}{\left[\frac{Dk^4}{g} + \sum\limits_{j \neq i} \Delta_j\right]} \quad , \quad i > 1. \tag{3.30}$$

For such subsurface loading scenarios, we get a contribution to the gravity anomaly from the loading term $\Delta_i 2\pi G W_i e^{-kz_i}$ as well as contributions from the deflections, which, by assumption, are all identical to ${}_iH$, $\Delta_j 2\pi G {}_iH e^{-kz_j}$. The Bouguer admittance thus becomes:

$$\boxed{Z_{i>1} = 2\pi G \left(\sum_{j \neq i, j>1} \Delta_j e^{-kz_j} - e^{-kz_i} \left[\frac{Dk^4}{g} + \sum_{j \neq i} \Delta_j \right] \right).} \tag{3.31}$$

The free-air admittance is given by $Z_{i>1} + 2\pi G \Delta_1$. Fig. 3-14 shows the admittance behavior for surface and subsurface loading individually.

3.10.3 Combining surface and internal loading

The initial applied loads are given by adding the static values of the loading/deflection term to the topography, as per Table 3.1. Let the total topography from different loading processes be

$$H = \sum_i H_i, \tag{3.32}$$

and express the initial applied loads as fractions of the total initial applied load:

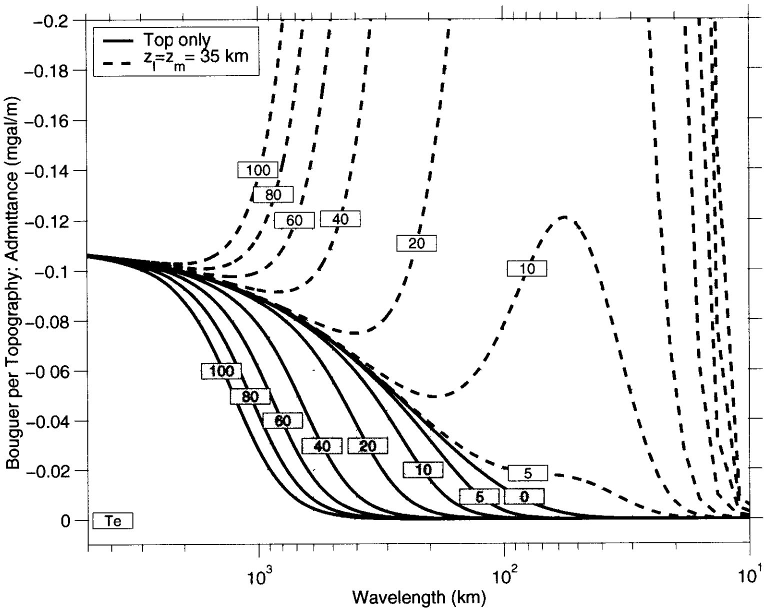


Figure 3-14: Admittance curves in function of the effective elastic thickness T_e, for a two-layer model ($\rho_1 = 2650$, $\rho_2 = 3300$ kgm^{-}3, $E = 1 \times 10^{11}$ Pa and $\nu = 0.25$) loaded at the surface (solid lines) and internally at 35 km (dashed lines). The depth of bottom loading z_l is the same as the compensation depth z_m.

$$F_i I = I_i \quad , \quad \sum_i F_i = 1. \tag{3.33}$$

From Eqs. 3.28 and 3.30, the fractional initial applied loads are expressed as:

$$I_1 \;=\; {}_1H \left(\frac{\frac{Dk^4}{g} + \sum\limits_i \Delta_i}{\frac{Dk^4}{g} + \sum\limits_{i>1} \Delta_i} \right) g\Delta_1, \tag{3.34}$$

$$I_i \;=\; -{}_iH \left(\frac{Dk^4}{g} + \sum_j \Delta_j \right) g\Delta_i \quad , \quad i > 1. \tag{3.35}$$

Combining Eqs. 3.33, 3.34 and 3.35, we can express:

$${}_1H \;=\; Y_1 F_1 I, \tag{3.36}$$

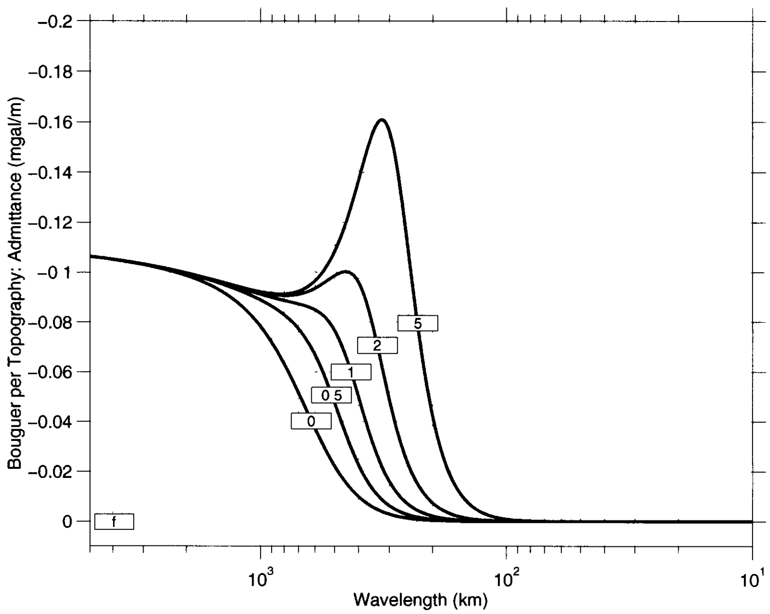


Figure 3-15: Admittance curves in function of the ratio of subsurface to surface loading f, for T_e= 40 km, for the same model as in Fig. 3-14.

$$ {}_iH = Y_i F_i I \quad , \quad i > 1, \tag{3.37}$$

with

$$ Y_1 = \left(\frac{\frac{Dk^4}{g} + \sum\limits_{i>1} \Delta_i}{\frac{Dk^4}{g} + \sum\limits_{i} \Delta_i} \cdot \frac{1}{g\Delta_1} \right), \tag{3.38}$$

$$ Y_i = -\left(\frac{1}{\frac{Dk^4}{g} + \sum\limits_{j} \Delta_j} \cdot \frac{1}{g} \right) \quad , \quad i > 1. \tag{3.39}$$

The total topography and gravity anomaly are given by, respectively, $H = I \sum_i F_i Y_i$ and $G_B = I \sum_i Z_i F_i Y_i$. If the loading processes and the noise are uncorrelated, then we can estimate the admittance from the cross-spectral densities of the gravity anomaly and the total observed topography:

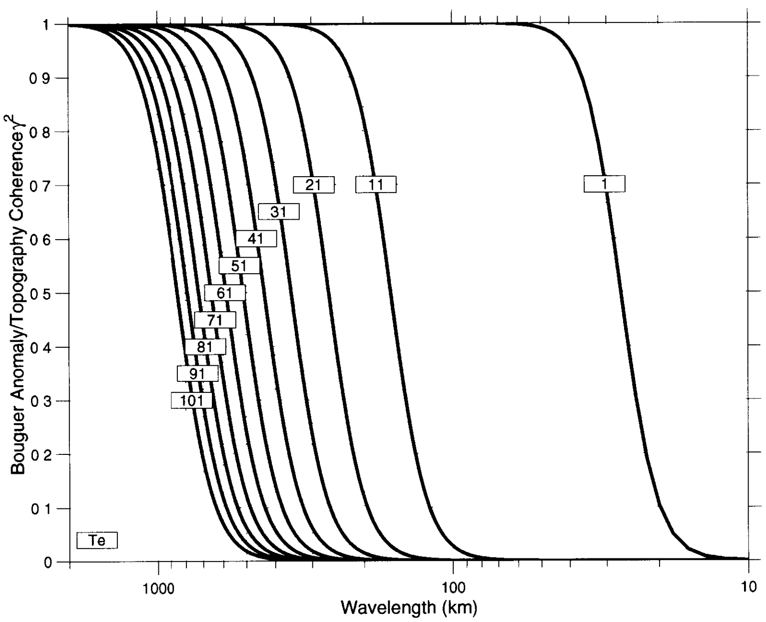


Figure 3-16: Coherence curves in function of the effective elastic thickness T_e, for the same model as in Fig. 3-14, with equal amounts of surface and subsurface loading.

$$Z = \frac{\langle G_B | H^* \rangle}{\langle H | H^* \rangle} = \frac{\sum_i F_i^2 Y_i^2 Z_i}{\sum_i F_i^2 Y_i^2}. \tag{3.40}$$

The coherence square is then given by:

$$\gamma^2 = \frac{| \langle G_B | H^* \rangle |^2}{\langle H | H^* \rangle \langle G_B | G_B^* \rangle} = \frac{\left(\sum_i F_i^2 Y_i^2 Z_i \right)^2}{\left(\sum_i F_i^2 Y_i^2 \right) \left(\sum_i F_i^2 Y_i^2 Z_i^2 \right)}. \tag{3.41}$$

Fig. 3-15 plots admittance curves for varying ratios of subsurface-to-surface loading f as defined by *Forsyth* [1985] (Note that this is an alternative formulation to the load fractions F_i, suitable for models in which only one density interface exists). Fig. 3-16 shows the coherence in function of the effective elastic thickness, for a model in which the surface and subsurface are loaded equally $f = 1$. Fig. 3-17 plots how the coherence

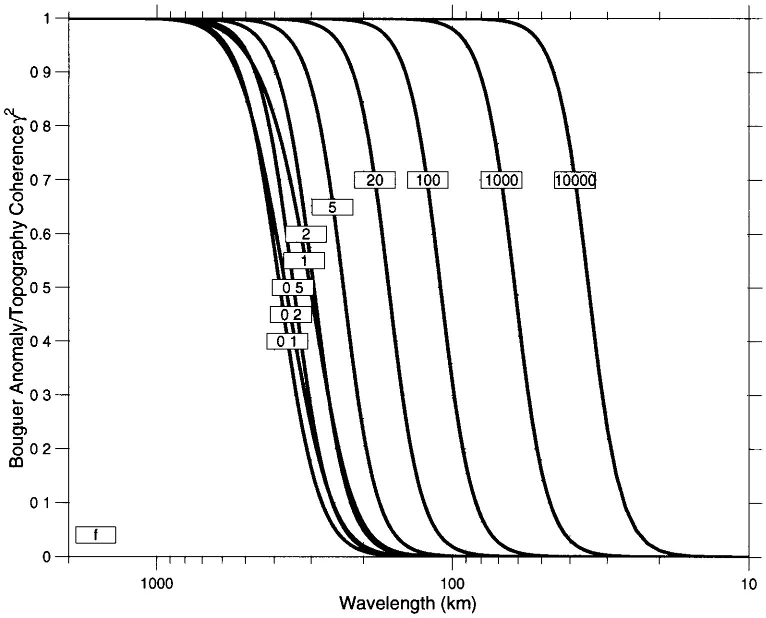


Figure 3-17: Coherence curves in function of the ratio of subsurface to surface loading f, for T_e= 40 km, for the same model as in Fig. 3-14.

for a given effective elastic thickness varies as the loading fraction f changes. Fig. 3-18 gives the behavior of free-air and Bouguer admittance and coherence, for varying loading fractions F_2.

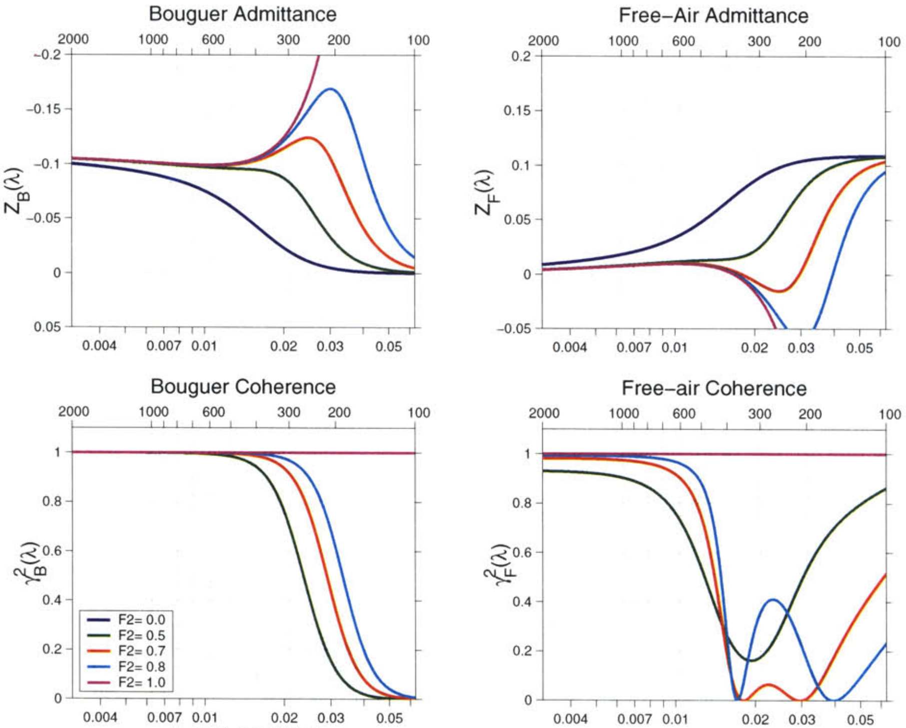
For a model with one density contrast, and a subsurface-to-surface loading ratio f, the predicted coherence is given by [*Forsyth*, 1985]

$$\gamma^2 = \frac{(\xi\Delta\rho^2 + f^2\rho^2\phi)^2}{(\xi^2\Delta\rho^2 + f^2\rho^2)(\Delta\rho^2 + f^2\rho^2\phi^2)}, \tag{3.42}$$

where

$$\xi = 1 + \frac{Dk^4}{\Delta\rho g}, \tag{3.43}$$

$$\phi = 1 + \frac{Dk^4}{\rho g}. \tag{3.44}$$



© Frederik J Simons, MIT, 06 Feb 2002

Figure 3-18: Admittance and coherence functions for Bouguer and free-air anomaly with surface topography. The four-density layer model has a top layer with $\rho_w = 0$, $\rho_u = 2600$, $\rho_l = 2900$, and $r_m = 3300$ kg/m^3, respectively. The interfaces are located at depths $z_1 = 0$, $z_2 = 15$, and $z_3 = 35$ km. The effective elastic thickness is $T_e = 17$ km. Loading at three interfaces happens with fractions $F_3 = 0$, $F_2 = 0, 0.5, 0.7, 0.8, 1$ and $F_1 = 1 - F_2$. Young's modulus $E = 1.78\times10^1 1$ Pa and Poisson's ratio $\nu = 1/4$. The scale is logarithmic in the wavenumber k.

The wavelength at which the Bouguer coherence $\gamma^2(k) = 0.5$ can be calculated explicitly. An analytic expression for the transition wavenumber is:

$$k^{1/2} = \frac{2^{-1/4}}{Df}\left(\left(-f\Delta\rho - f\rho + f^2\rho + \Delta\rho + \beta^{1/2}\right) gD^3 f^3\right)^{1/4}, \qquad (3.45)$$

where

$$\begin{aligned}\beta = {} & f^2\Delta\rho^2 + 4f^2\Delta\rho\rho - 2f^3\Delta\rho\rho + 2f\Delta\rho^2 \\ & + f^2\rho^2 + 2f^3\rho^2 - 2f\rho\Delta\rho + f^4\rho^2 + \Delta\rho^2. \end{aligned} \tag{3.46}$$

The transition wavelength is commonly used as a diagnostic of the effective elastic thickness and is plotted in Fig. 3-19. We note that the ratio of subsurface-to-surface loading f and the effective elastic thickness T_e cannot be determined uniquely.

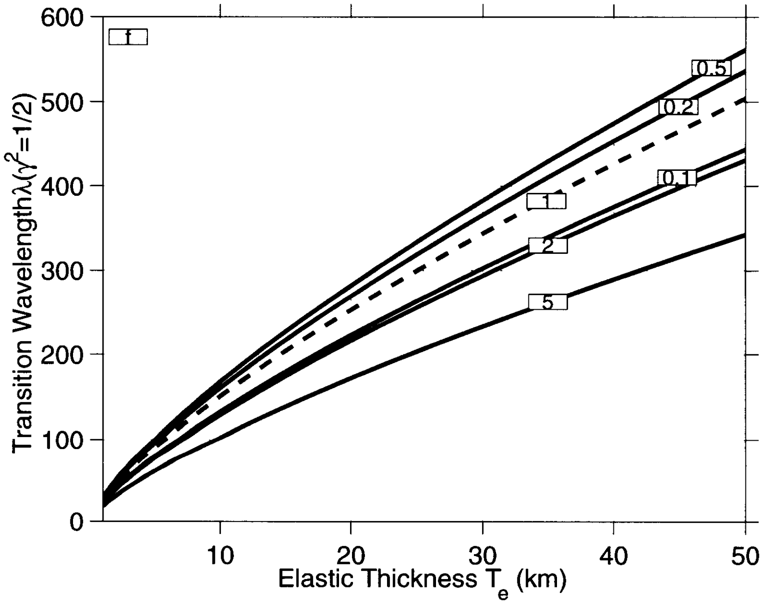


Figure 3-19: Wavelength at which the Bouguer coherence reaches half of its long-wavelength value, for different loading ratios f, defined by *Forsyth* [1985] as the ratio of subsurface-to-surface loading.

Chapter 4

Mechanical and Seismic Anisotropy[1]

Abstract

The coherence of gravity anomalies and topography is indicative of the strength of the lithosphere. Any dependence of the coherence on azimuth represents a departure from the commonly assumed isotropy of the isostatic response. This mechanical anisotropy is likely to be related to the fossil strain field and is, therefore, of relevance to geologic and seismological studies. The robust and unbiased estimation of spatially, directionally (and wavelength) dependent coherence functions between two-dimensional fields is not trivial. If the fields are assumed to be spatially stationary, the Thomson multi-taper method using Slepian windows is particularly successful at estimating the coherence and its anisotropy. In order to study spatial variations (non-stationary analysis) of spectral properties, a localization both in space and in frequency is required. Space-scale concentration can be achieved with a class of wavelets known as Slepian wavelets. Equivalently, orthonormal Hermite spectral windowing functions can provide concentration in the space-frequency domain. We discuss the properties and advantages of the Hermite multi-spectrogram method with respect to the wavelet-based method, and apply the former to the coherence estimation between Bouguer anomalies and topography of Australia. We identify regions on the continent marked by preferential directions of isostatic compensation in two wavelength ranges. With few exceptions, the short-wavelength coherence anisotropy is nearly perpendicular to the major trends of the suture zones between stable continental domains, supporting the geological observation that such zones are mechanically weak. The correlation of isostatic anisotropy

---

[1] **Isostatic and Seismic Anisotropy of Australia from Non-Stationary Gravity-Topography Coherence Analysis and Surface Wave Tomography**, *Journal of Geophysical Research*, submitted, 2002.

with directions of fast wave propagation gleaned from the azimuthal anisotropy of surface waves as well as shear-wave splitting results belies a first-order relationship between mechanical and seismic anisotropy. Mechanical anisotropy reflects lithospheric strain accumulation, and its presence must be related to the deformational processes integrated over time. The absence of agreement between the directions of mechanical and seismic anisotropy argues against the hypothesis that internally coherent deformation of the entire lithosphere is responsible for the anisotropy inferred from surface wave analysis and shear wave splitting observed at the surface.

4.1 Introduction

On opposite ends of the range of mechanical properties of the Earth's lithosphere are their instantaneous local elastic response to a stress field associated with seismic waves and the long-term elastic response of the lithosphere due to loading caused by topography at various interfaces. The former is characterized by the elasticity tensor, which represents seismic wave speeds and their dependence on propagation and polarization direction. The long-term mechanical properties are reflected in the level of coherence between gravity anomalies and topography [*Forsyth*, 1985].

Many authors have studied the seismic anisotropy of continental lithospheres [e.g. *Silver and Chan*, 1988; *Savage*, 1999; *Debayle and Kennett*, 2000a; *Savage and Sheehan*, 2000; *Simons et al.*, 2002a], and a connection between seismic anisotropy and the long-term mechanical properties of the lithosphere has been suggested [*Silver and Chan*, 1991; *Griot and Montagner*, 1998; *Chen and Özalaybey*, 1998; *Vauchez et al.*, 1998]. Perhaps the most far-reaching hypothesis argues that seismic anisotropy (manifest in, e.g., shear-wave splitting) is predominantly caused by "fossil" strain in the subcontinental mantle due to the last major episode of tectonic activity [*Silver and Chan*, 1988]. In this hypothesis, lithospheric deformation is thought to be internally coherent and continental surface structural geology is a reasonable predictor for the seismic anisotropy. *Simons et al.* [2002a] have shown, however, that shear wave splitting gives an oversimplified view of continental anisotropy. The seismic anisotropy of Australia is more complex than simple end-member models of vertically coherent deformation [*Silver and Chan*, 1988] or a predominant influence of present-day mantle deformation [*Vinnik et al.*, 1995] would predict.

The isotropic mechanical strength of continents is well studied [e.g. *Zuber et al.*, 1989; *McKenzie and Fairhead*, 1997]. Even though mechanical anisotropy is well-documented in the geological literature [*Vauchez et al.*, 1998; *Tommasi and Vauchez*, 2001], only a few studies have investigated the mechanical anisotropy of the lithosphere with spectral methods [*Stephenson and Beaumont*, 1980; *Stephenson and Lambeck*, 1985; *Lowry and Smith*, 1995; *Escartin and Lin*, 1998; *Simons et al.*, 2000]. The challenge is to extract directional information from the gravity and topography data. We show that this can be achieved by measuring the azimuthal variation of the coherence (or isostatic response) of gravity with topographic and interface loading,

Lateral and directional resolution come naturally to seismic surface waves, since they sample the Earth in a spatially well-defined manner. Spectral properties, on the other hand, are by definition not localized in space (pure frequencies have infinite extent) and the phase-averaging schemes used to obtain stable coherence estimates (e.g. wavenumber binning [*Bechtel et al.*, 1987]) obliterate the directional signal. The optimal estimation of azimuthally varying isostatic response functions by multi-taper spectral analysis was the subject of a previous chapter [*Simons et al.*, 2000], but with their method spatial resolution was achieved in the *ad hoc* manner of boxing the data and assuming local stationarity. Also in the context of isostatic response function analysis, *Simons et al.* [1997] formulated a wavelet-based spatial localization of the admittance (rather than the coherence) function between gravity and topography, but their method is not suited for measuring mechanical anisotropy.

Time-frequency [*Flandrin*, 1998] and time-scale (wavelet) analysis [*Mallat*, 1998] find the compromise between resolution in the spatial and spectral domains. Wavelet methods are used pervasively for the univariate time-scale characterization and representation of geophysical fields [*Kumar and Foufoula-Georgiou*, 1997; *Torrence and Compo*, 1998; *Bergeron et al.*, 2000], including their anisotropy [*Kumar*, 1995], but the calculation of multivariate coherence functions by wavelet transforms has eluded many attempts [*Liu*, 1994; *Santoso et al.*, 1997; *Meredith*, 1999]. Primarily responsible is the lack of reasonable smoothing operations of the cross-spectrum prior to normalization.

Only a few authors have developed methods, based on multi-wavelets, that yield adequate time and frequency resolution for the calculation of coherence functions between one-dimensional signals [*Daubechies and Paul*, 1988; *Lilly and Park*, 1995]. We will show here that the most natural extension to two-dimensional problems (and hence the retrieval of anisotropic qualities) is not by space-scale (scalogram) methods but by space-frequency (spectrogram) analysis. This conclusion follows from a comparison of the concentration domains in the space-frequency plane. The Hermite-multi-spectrogram method we propose to use in this context builds on the work by *Daubechies* [1988] and *Flandrin* [1988] and was applied in one dimension for univariate analysis by *Bayram and Baraniuk* [1996]. The multivariate method that we describe here is capable of detecting both the anisotropy and the spatial variations of the spectral coherence function.

Because of the novelty of our method we discuss in some detail its development and characteristics in Part I of this chapter. Part II is concerned with the practical application to the study of the spatial and azimuthal variability of the isostatic response of the Australian continent. We study this mechanical anisotropy by calculating coherence functions between topography and Bouguer gravity anomalies. We discuss the mechanical properties of the Australian lithosphere in comparison with the fast directions obtained from high-resolution seismic surface-wave tomography with azimuthal anisotropy [*Simons et al.*, 2002a] and shear wave splitting [*Iidaka et al.*, 2001]. In this vein, we test the simple hypothesis that seismic anisotropy is reflected in the anisotropy of fossil strain (recorded by the gravitational anomalies and their topographic expression and represented by the coherence function).

4.2 Part I: Estimation of non-stationary coherence

4.2.1 General ideas and basic theory

The calculation of coherence functions is of particular importance in many areas of science and engineering [*Bendat and Piersol*, 1993]. For the calculation of isostatic response functions the coherence method due to *Forsyth* [1985] is by far the most widely embraced. This approach assumes stationary processes. For stationary coherence analysis multiple options exist to perform the necessary phase averaging to avoid that the coherence equals unity [*Simons et al.*, 2000]. Unfortunately, the calculation of time-frequency (i.e. non-stationary) coherence functions has remained problematic [*Torrence and Compo*, 1998], and for wavelet analysis with just one mother wavelet, no satisfactory averaging procedures exist. Averaging in time or in frequency [as by *Santoso et al.*, 1997] defeats the purpose of time-frequency analysis. The lack of averaging capabilities has precluded the development of successful wavelet-based coherence methods [*Liu*, 1994; *Kumar and Foufoula-Georgiou*, 1997], with the exception of the Slepian wavelets by *Lilly and Park* [1995]. This recognition is of central importance to understanding the utility of multi-window methods. The Slepian multi-scalogram and Hermite multi-spectrogram approaches described in this section have in common that independent spectral estimates are made with each different wavelet or window.

For two non-stationary random processes $\{X\}$ (in our application the Bouguer gravity anomaly) and $\{Y\}$ (the topography), defined on $\mathbf{r}$ in the spatial domain and on $\mathbf{k}$ in the Fourier domain, the coherence-square function relating both fields, γ^2_{XY}, is defined as the ratio of their cross-spectral density, S_{XY}, normalized by the individual power spectral densities, S_{XX} and S_{YY} [*Bendat and Piersol*, 2000]:

$$\gamma^2_{XY}(\mathbf{r},\mathbf{k}) = \frac{|S_{XY}(\mathbf{r},\mathbf{k})|^2}{S_{XX}(\mathbf{r},\mathbf{k})S_{YY}(\mathbf{r},\mathbf{k})} = \frac{\left|E\{\tilde{X}(\mathbf{r},\mathbf{k})\tilde{Y}^*(\mathbf{r},\mathbf{k})\}\right|^2}{E\{\tilde{X}(\mathbf{r},\mathbf{k})\tilde{X}^*(\mathbf{r},\mathbf{k})\}E\{\tilde{Y}(\mathbf{r},\mathbf{k})\tilde{Y}^*(\mathbf{r},\mathbf{k})\}}. \tag{4.1}$$

Here, E denotes an expectation operator, the tildes ($\sim$) refer to the Fourier transformed signal, and the asterisk ($*$) is used to represent the complex conjugate. The coherence is a measure of the consistency of the phase relationship between both processes. Therefore, it cannot simply be estimated from the ratios of the periodograms (Eq. 4.1 without the averaging operators), as they do not record phase information. Averaging operations are required to avoid that $\gamma^2_{XY} = 1$ everywhere [*Bendat and Piersol*, 1993]. A variety of averaging methods used for isostatic response analysis was discussed by *Simons et al.* [2000].

The qualities desired in estimators, $\hat{S}$, of spectral densities, S, are unbiasedness (Eq. 4.2a), a low estimation variance (Eq. 4.2b), and the minimization of leakage from adjacent frequencies (Eq. 4.2c):

$$\begin{aligned} \forall\mathbf{k} : E\{\hat{S}(\mathbf{k})\} &= S(\mathbf{k}), && \text{(4.2a)} \\ \lim_{N\to\infty} \mathrm{var}\{\hat{S}(\mathbf{k})\} &= 0, && \text{(4.2b)} \\ \forall\mathbf{k}' \neq \mathbf{k} : \mathrm{cov}\{\hat{S}(\mathbf{k}'),\hat{S}(\mathbf{k})\} &= 0. && \text{(4.2c)} \end{aligned}$$

In the following, for the clarity of the presentation we will write the relations for one-dimensional signals, $x(t)$, defined in the time-frequency plane, (t,f). For *stationary* signals, the criteria represented by Eqs. (4.2a)-(4.2c) are met by a class of quadratic spectrum estimators [*Mullis and Scharf*, 1991] consisting of an average of K modified periodograms acting on data windowed by different tapers, h_k. The multi-taper spectral estimate, $\hat{S}^{\mathrm{MT}}(f)$, is defined as [*Thomson*, 1982]:

$$\hat{S}^{\mathrm{MT}}(f) = \frac{1}{K}\sum_{k=0}^{K-1}\left|\int_{-\infty}^{\infty} h_k(\tau)x(\tau)e^{-i2\pi f\tau}\,d\tau\right|^2 . \tag{4.3}$$

Satisfying Eq. (4.2a), the first moment of $\hat{S}^{\mathrm{MT}}$ is the true spectrum, S, smoothed in frequency space by a kernel, $\mathcal{H}$, which is the average periodogram of the windowing functions $h_k(\tau)$ [*Percival and Walden*, 1993]:

$$E\{\hat{S}^{\mathrm{MT}}(f)\} = S(f)\otimes\mathcal{H}(f). \tag{4.4}$$

The convolution is in the frequency-domain, and the explicit form of the kernel function is:

$$\mathcal{H}(f) = \frac{1}{K}\sum_{k=0}^{K-1}\left|\int_{-\infty}^{\infty} h_k(\tau)e^{-i2\pi f\tau}\,d\tau\right|^2 . \tag{4.5}$$

The desired (as per Eq. 4.2b) reduction of estimation variance is obtained by the averaging of more than one direct spectral estimate, while the requirements contained in Eq. (4.2c) can be met by an appropriate choice of the windowing functions $h_k(\tau)$ (see Sec. 4.2.2).

For the estimation of *non-stationary* spectra, we must first introduce three time-frequency distributions [*Mallat*, 1998]. They are the Wigner-Ville transform (WV) (Eq. 4.6a), the windowed Fourier transform (WF) (Eq. 4.6b), and the wavelet transform (WT) (Eq. 4.6c):

$$P_{\mathrm{WV}}[x](t,f) = \int_{-\infty}^{\infty} x\left(t+\frac{\tau}{2}\right)x\left(t-\frac{\tau}{2}\right)e^{-i2\pi f\tau}\,d\tau, \tag{4.6a}$$

$$P_{\mathrm{WF}}[x](t,f) = \langle x|h_{t,f}\rangle = \int_{-\infty}^{\infty} h(\tau-t)x(\tau)e^{-i2\pi f\tau}\,d\tau, \tag{4.6b}$$

$$P_{\mathrm{WT}}[x](t,s) = \langle x|\psi_{t,s}\rangle = \int_{-\infty}^{\infty} x(\tau)\frac{1}{\sqrt{s}}\psi\left(\frac{\tau-t}{s}\right)d\tau. \tag{4.6c}$$

Here, s is the scaling parameter which can be related to the frequency [*Kumar and Foufoula-Georgiou*, 1997]. The WV is a quadratic energy distribution. The energy density function associated with the windowed Fourier transform is called the *spectrogram* and given by

$$|P_{WF}(t, f)|^2. \tag{4.7}$$

The wavelet energy density function is termed the *scalogram* and is given by

$$|P_{WT}(t, s)|^2. \tag{4.8}$$

In analogy with Eq. (4.3), *Bayram and Baraniuk* [1996] have proposed a multi-taper spectrogram, $\hat{S}^{\mathrm{WF}}(t, f)$:

$$\hat{S}^{\mathrm{WF}}(t, f) = \frac{1}{K} \sum_{k=0}^{K-1} \left| \int_{-\infty}^{\infty} x(\tau) h_k(\tau - t) e^{-i2\pi f \tau} \, d\tau \right|^2 . \tag{4.9}$$

On the other hand, *Lilly and Park* [1995] proposed a multi-wavelet scalogram, $\hat{S}^{\mathrm{WT}}(t, s)$, to estimate $S(t, f)$:

$$\hat{S}^{\mathrm{WT}}(t, s) = \frac{2}{K} \sum_{k=0}^{K-1} |x \otimes \psi_{k,s}(-\tau)|^2 . \tag{4.10}$$

Note that we have rewritten the continuous wavelet transform in terms of the convolution as which it is most conveniently calculated. The parallelism between Eq. (4.10) and Eq. (4.4) can be extended as follows: the time-varying spectrum, $S(t, f)$, is the expectation of the WV [*Mallat*, 1998]:

$$S(t, f) = E\{P_{\mathrm{WV}}(t, f)\}. \tag{4.11}$$

Both the scalogram and the spectrogram are members of a Cohen's class [*Cohen*, 1989]: they can be obtained by time-frequency or time-scale (affine) smoothing of the WV [*Rioul and Flandrin*, 1992; *Frazer and Boashash*, 1994]. For any such estimator $\hat{S}$,

$$\hat{S} = P_{\mathrm{WV}}[x] \otimes\oplus \frac{1}{K} \sum_{k=0}^{K-1} P_{\mathrm{WV}}[h_k]. \tag{4.12}$$

Consequently, the expectation of either Eq. (4.9) or (4.10) is the true spectrum, S, smoothed by a kernel which is the average WV of the windowing functions $h_k(\tau)$ or wavelets $\psi_k(\tau)$. The convolution is both in the time and in the frequency or scale domain. Comparing Eqs. (4.11) and (4.12) with Eq. (4.4), we see that the WV takes the role of the periodogram in stationary spectrum analysis. In order to study the performance of non-stationary spectrum estimators, we need to study the properties of the WV of the wavelets (see Sec. 4.2.3) or windows (see Sec. 4.2.4) used.

4.2.2 Optimal windows for stationary spectrum estimation

Slepian [1978] discovered that windowing functions $h_k(\tau)$ could be found that, notwithstanding the finite data length, are optimally concentrated in the frequency domain. Their narrow central lobe and low sidelobe level assure low bias and leakage, while their multiplicity reduces the variance of the estimate when used as in Eq. (4.3).

It is straightforward to see how a finite data length, T, implies the action of a time-projection operator $\mathbf{T}$:

$$\mathbf{T}\{h(t)\} = h(t) R_{-\mathrm{T}/2}^{\mathrm{T}/2}(t), \tag{4.13}$$

where $R_{-\mathrm{T}/2}^{\mathrm{T}/2}$ denotes a boxcar function of length T. A low-pass projection within a bandwidth W is achieved by the operator $\mathbf{L}$ [*Flandrin*, 1988].

$$\mathbf{L}\{h(t)\} = \int_{-W/2}^{W/2} \hat{h}(f) e^{i2\pi ft}\, df. \tag{4.14}$$

As is easily shown, the combined effect of both operators is:

$$\mathbf{TL}\{h(t)\} = \int_{-T/2}^{T/2} \frac{\sin \pi W(t-\tau)}{\pi(t-\tau)} h(\tau)\, d\tau. \tag{4.15}$$

The Shannon number, TW, is the number of optimally concentrated windows $h_k(\tau)$ that can be found as the eigenfunctions of Eq. (4.15). These $h_k(\tau)$ are known as *prolate spheroidal wave functions* (pswf) or Slepian functions [*Slepian*, 1983]. In our previous chapter, *Simons et al.* [2000], we have outlined their properties and for more information we refer to *Percival and Walden* [1993].

4.2.3 Optimal wavelets

By construction, the spectral windows $\mathcal{H}(f)$ associated with the pswf $h_k(\tau)$ are low-pass filters concentrated in a domain $|f| \leq W/2$. As shown by *Lilly and Park* [1995], wavelets are obtained when Eq. (4.15) is rewritten to find band-pass filters centered around a frequency f_c as $|f - f_c| \leq W/2$. From symmetry considerations (compare with Eq. 4.15), they are the eigenfunctions of:

$$\int_{-T/2}^{T/2} \left\{ \frac{\sin 2\pi(f_c + W/2)(t-\tau)}{\pi(t-\tau)} - \frac{\sin 2\pi(f_c - W/2)(t-\tau)}{\pi(t-\tau)} \right\} \psi(\tau)\, d\tau \tag{4.16}$$

.

The Shannon number is now $2TW$. (The factor of 2 arises from using both positive and negative frequency intervals in the optimization condition; it is also present in Eq. 4.10). The $2TW$ solutions $\psi_k(T, f_c, W; \tau)$ are indeed wavelets: they are asymptotically self-

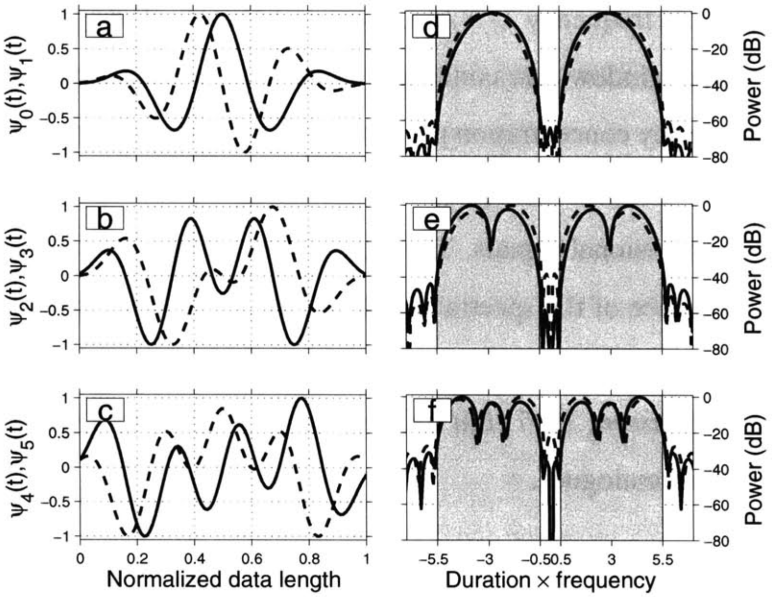


Figure 4-1: Six prolate spheroidal (Slepian) wavelets [*Lilly and Park*, 1995] in the time (a-c) and frequency (d-f) domain, grouped in three pairs (solid and dashed lines) with similar frequency response. The $\Psi_k(t)$, $k = 0 \to 5$, have a center frequency of three times the fundamental frequency: $f_c = 3/N/\Delta t$, and a half-bandwidth $W/2 = 2.5/N/\Delta t$. Their concentration properties are comparable to those of the prolate spheroidal wave functions of *Slepian* [1978] [*Simons et al.*, 2000], but the band-pass magnitude response enables time-frequency analysis.

similar, and as their length increases their sensitivity shifts to lower frequencies and smaller bandwidths. *Lilly and Park* [1995] have called them *Slepian wavelets*. Examples of their usage are shown by *Bear and Pavlis* [1997, 1999].

To compare the properties of the pswf, or Slepian *functions* [see *Simons et al.*, 2000, their Fig. 2] with the Slepian *wavelets*, we have calculated six Slepian wavelets with a central frequency $f_c = 3/N/\Delta t$ and a half-bandwidth $W/2 = 2.5/N/\Delta t$. In Fig. 4-1 we have plotted the wavelets and their magnitude response in the time (Fig. 4-1a-c) and frequency domain (Fig. 4-1d-f). Six wavelets are plotted, grouped in pairs (solid and dashed lines) with similar magnitude response. The frequency axis of duration×frequency is in terms of the smallest resolvable frequency $1/N/\Delta t$ [*Kay and Marple*, 1981], so for every wavelet of the set the spectral windows are concentrated in two symmetric bands of width 5, centered at 3. The frequency concentration properties of the Slepian wavelets are fairly similar to those of the pswf, and their multiplicity makes them suitable for coherence analysis of time-varying one-dimensional signals. The eigenvalues, a measure of the concentration of energy in the central lobe of the spectral windows, are extremely close to unity [*Lilly and Park*, 1995]. However, unlike the 2-D pswf, which can be formed by a simple outer product of two 1-D tapers [*Hanssen*, 1997; *Liu and van Veen*, 1992], the Slepian wavelets do not have two-dimensional analogues.

4.2.4 Optimal windows for non-stationary spectrum estimation

The pswf and Slepian wavelets are optimally localized in a rectangular domain $[-T/2, T/2]\times[-W/2, W/2]$ of the (t, f)-plane [*Lilly and Park*, 1995; *Kumar and Foufoula-Georgiou*, 1997]. This tiling of the time-frequency plane with rectangular concentration regions is, however, not the most natural choice. It is preferable to treat the time-frequency space as a whole, by looking for operators that maximally concentrate simultaneously in time and frequency. If the concentration domain can be made circularly symmetric in the (t, f)-phase plane, time and frequency resolution will not be at the expense of each other. In the formalism of Eq. (4.12) the estimated spectrum will have isotropic distortion in time and frequency. This quality is responsible for the superior performance of methods based on disk-shaped concentration domains compared to traditional (i.e. single-window) spectrogram methods or methods which employ the pwsf in a sliding-window fashion [*Frazer and Boashash*, 1994; *Bayram and Baraniuk*, 1996; *Mellors et al.*, 1998]. Orthonormalized Hermite functions [*Daubechies*, 1988; *Flandrin*, 1988] possess such concentration regions that are functions of $t^2 + f^2 \leq R^2$. Their analytic form is given by

$$h_k(t) = \pi^{-1/2}(2^k k!)^{-1/2}\left(t - \frac{d}{dt}\right)^k e^{-t^2/2}, \tag{4.17}$$

i.e., they are Hermite polynomials modulated by a Gaussian function. The polynomials can be calculated by recursion. The eigenvalues are dependent on the width, R, of the concentration region and are given by

$$\lambda_k(R) = \frac{1}{k!}\gamma(R^2/2, k+1), \tag{4.18}$$

where γ denotes the incomplete gamma function.

The remarkable property of the Hermite functions and their eigenvalues is that the eigenfunctions of the concentration operator do not depend on the width R of the domain; the R-dependence is contained in the eigenvalues. This is not the case for the pswf or Slepian wavelets, where an eigenvalue problem needs to be solved for every choice of bandwidth and signal length or central frequency. This makes the Hermite method computationally very fast.

In Fig. 4-2a, five orthonormal Hermite functions are plotted [compare with the pswf shown by *Simons et al.*, 2000, their Fig. 2]. We let $t : -5 \to +5$ in Eq. (4.17). The concentration region R, expressed in the same units, then corresponds to a time-concentration within $R/10$ multiplied by the actual physical length of the window. The spectral windows $\mathcal{H}(f; R)$ are plotted in Fig. 4-2b for varying values of R. They are calculated as

$$\mathcal{H}(f; R) = \frac{1}{K}\sum_{k=0}^{K-1} \lambda_k(R)\left|\int_{-\infty}^{\infty} h_k(\tau)e^{-i2\pi f\tau}\, d\tau\right|^2. \tag{4.19}$$

On the dimensionless duration× frequency scale, the corresponding frequency concentration (around 10 dB of bias reduction) lies within the gray shaded regions. The associated eigenvalues λ_k for the three choices of R are plotted in Fig. 4-2c. The decreasing number

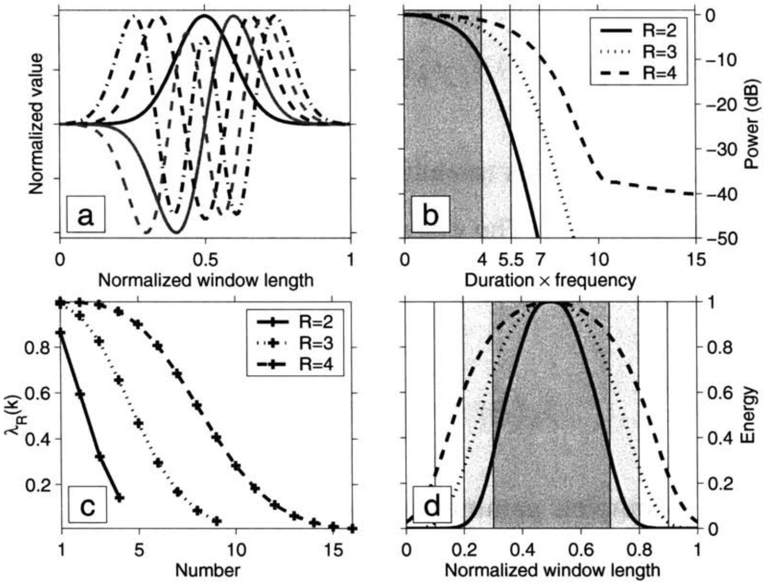


Figure 4-2: Hermite functions and their eigenvalues. (a) Orthonormal $h_k(t)$, $k = 0 \rightarrow 4$. (b) Frequency concentration: spectral windows $\mathcal{H}(f; R)$ (see Eq. (4.19)). (c) Eigenvalues $\lambda_R(k)$ for three different radii of concentration. (d) Time concentration: energy $U(t; R)$ (see Eq. (4.20))

of significantly contributing window functions with increased resolution (smaller R) is a manifestation of the duality of resolution and variance. In practice, we take $K = R^2$ tapers.

As a measure of the concentration of the windows in the time domain we define their energy as:

$$U(t; R) = \frac{1}{K} \sum_{k=0}^{K-1} \lambda_k(R) h_k^2(t), \tag{4.20}$$

which we plotted in Fig. 4-2d. We have normalized the time axis by the entire data length, so the energy of the windows is contained within $0.5 \pm R/10$ (gray shaded regions). The even extraction of energy from the signal is achieved by the calculation of the spectrogram at specific time intervals corresponding to an effective overlap of the average window.

Besides the ease with with they can be calculated and the elegance of their construction, perhaps the most attractive characteristic of the Hermite functions is the fact that the eigenfunctions of n-dimensional circularly symmetric concentration operators are outer products of the same Hermite functions [*Daubechies*, 1988]. The extension from time-frequency to multi-dimensional space-frequency analysis therefore becomes practical. Multi-dimensional extensions to space-scale (wavelet) approaches are difficult [*Foufoula-Georgiou and Kumar*, 1994]. Concentration operator formalisms have been derived for the space-scale phase plane, but rather than the circularly symmetric concentration regions of the time-frequency case, the associated wavelets have a complex, non-symmetrical concentration domains [*Daubechies and Paul*, 1988], and their usability remains a field of active research (A.T. Walden, *pers. comm.*).

4.2.5 Comparison of approaches

Concentration regions

As indicated in Sec. 4.2.1, the best way of characterizing windowing or wavelet functions when used for space-frequency or space-scale analysis is not by their average periodograms or energy functions (as in Figs. 4-1 and 4-2) but by their average Wigner-Ville transform. In Fig. 4-3, we have plotted the alias-free Wigner-Ville distribution [*Jeong and Williams*, 1992] of both classes of functions. The box-tiling of the phase plane of the Slepian wavelets and the disk-shaped concentration domains of the Hermite functions are well captured by their WV. For the Slepian multi-wavelet scalogram, the center frequency and bandwidth are determined subjectively (two cases are shown), whereas for the Hermite function the radius of the concentration disk is given by R (two cases are shown in physical units). The symmetry and smoothness of the Hermite function kernels make them attractive for their use in time-frequency coherence analysis.

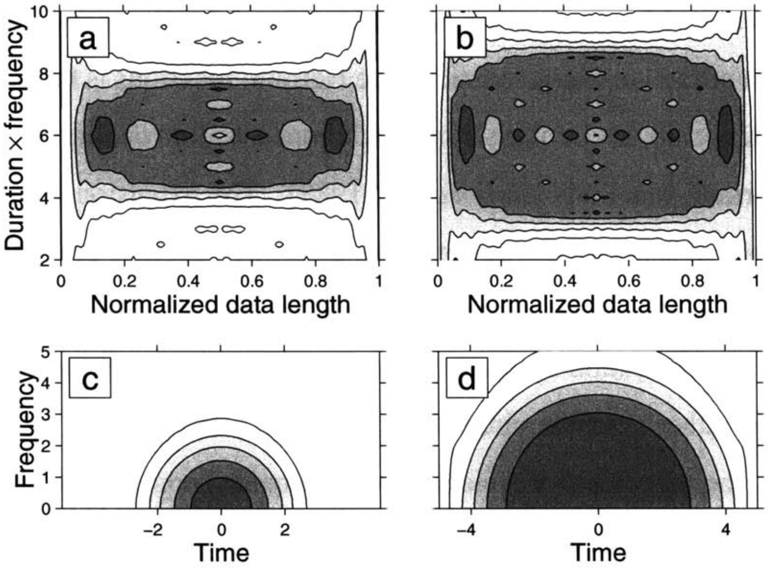


Figure 4-3: Concentration of Slepian wavelets and Hermite functions in the time-frequency plane by their normalized average Wigner-Ville energy distribution. Contours are drawn at [0.1:0.2:1]. (a) Slepian wavelets with $f_c = 6$ and $W/2 = 2$, and (b) with $f_c = 6$ and $W/2 = 3$ in dimensionless units normalized by the fundamental frequency. (c) Concentration of Hermite functions in the time-frequency plane for radii $R = 2$ and (d) $R = 4$.

Time-scale versus time-frequency analysis: Univariate case

We compare the wavelet and the Hermite function approaches for the characterization of the synthetic signal (see Fig. 4-4a) used by *Lilly and Park* [1995]. The signal consists of a sine-modulated linear chirp and a separate notch. In Fig. 4-4b, the Slepian multiwavelet scalogram is plotted (where scale was converted to frequency), calculated with wavelets of center frequency $f_c = 7/N/\Delta t$ and half-width $W/2 = 3/N/\Delta t$, i.e. the parameters used by *Lilly and Park* [1995] (their Fig. 5). The number of wavelets used per scale was smaller or equal to the Shannon number of $2W$. This determines the wavelet length and the maximum resolvable frequency. In Fig. 4-4c, we have plotted the Hermite multi-spectrogram, with $R = 1$. The window length was 0.13 in normalized time units,

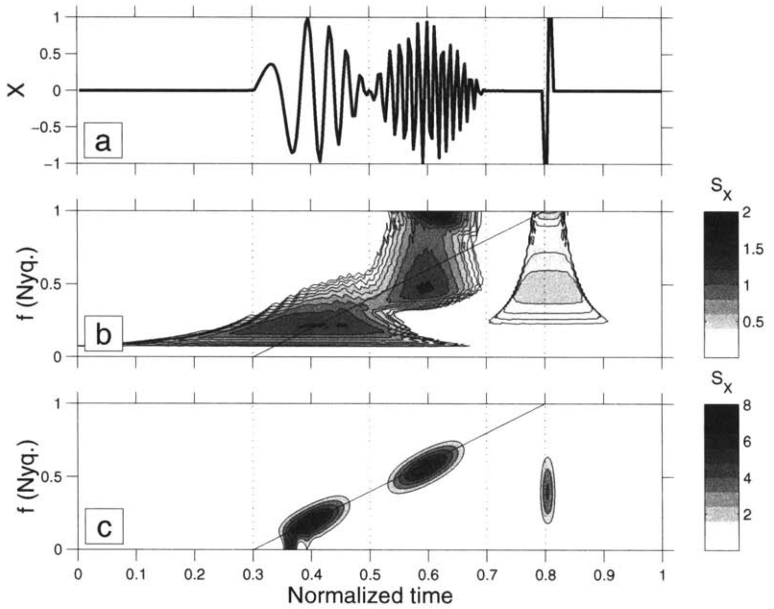


Figure 4-4: Comparison of the Hermite multiple-spectrogram method with the Slepian multi-wavelet method. (a) Synthetic signal, also used by *Lilly and Park* [1995]. (b) Power spectrum based on the Slepian multi-wavelet transform. Slepian wavelets with $f_c = 7/N/\Delta t$ and $W/2 = 3/N/\Delta t$ were used. (c) Power spectrum from the multi-spectrogram method, the concentration region $R = 1$, window length $L = 32\Delta t = 0.13N\Delta t$, overlap 99%, nfft = 256. The thin solid lines in (b) and (c) correspond to the instantaneous frequency of the input chirp shown in (a).

and the values were calculated at time steps corresponding to an overlap of the windows of 99 %.

An important difference is that in the case of wavelet analysis, the time-resolution is inversely proportional to the frequencies at which the signal is analyzed: the high-frequency components are analyzed with much shorter wavelets than the low-frequency portions of the signal. This explains the typical pyramidal form generated by discontinuities in the signal. It also explains the particular efficiency of the wavelet transform for the representation of signals with a small number of coefficients [*Strang and Nguyen*, 1997]. For the analysis, rather than representation of signals, on the other hand, we prefer the capabilities of the

Hermite method. The temporal and spectral resolution of the estimates can be controlled by choosing the length, the concentration radius R and the effective overlap of the windows, and, likewise, the estimation variance with which it trades off. The interpretation of the spectral estimate as a filtered or blurred version of the input properties is particularly intuitive: the true spectrum is seen through a lens which has isotropic distortion in the time and frequency space.

Bivariate 1-D time-frequency analysis

We take the synthetic example proposed by *Santoso et al.* [1997] to show that the Hermite function approach is capable of retrieving both the magnitude and the phase of the coherence between both signals; we can distinguish variations in time as well as frequency. *Santoso et al.* [1997] were unable to provide frequency resolution. The synthetic signals, plotted in Fig. 4-5a and 4-5b, consist of two time segments with (a) a sinusoid at 1/4 and 1/2 of the Nyquist frequency, and (b) another at 1/4 and 1/1 of the Nyquist, $-\pi/2$ radians out of phase with the first. Uniform random noise (with a signal-to-noise ratio of 5 dB) was added. The localization properties of the Hermite functions are excellent, both in time and in frequency. This is evident from the individual power spectral estimations of both signals, plotted in Fig. 4-5c and 4-5d, and in the strong peak of unit coherence in Fig. 4-5e. Where the coherence is unity the measurement of its phase correctly yields the input phase difference between both segments.

Bivariate 2-D space-frequency analysis

In Fig. 4-6 we present the results of a two-dimensional test. As in *Simons et al.* [2000] (their Plate 1), we follow the method of *Lowry and Smith* [1994] and create pairs of real synthetic fields with a known (an)isotropic coherence between them. Four pairs of coherent fields were joined together, forming a spatial input pattern of coherence shown in Fig. 4-6.

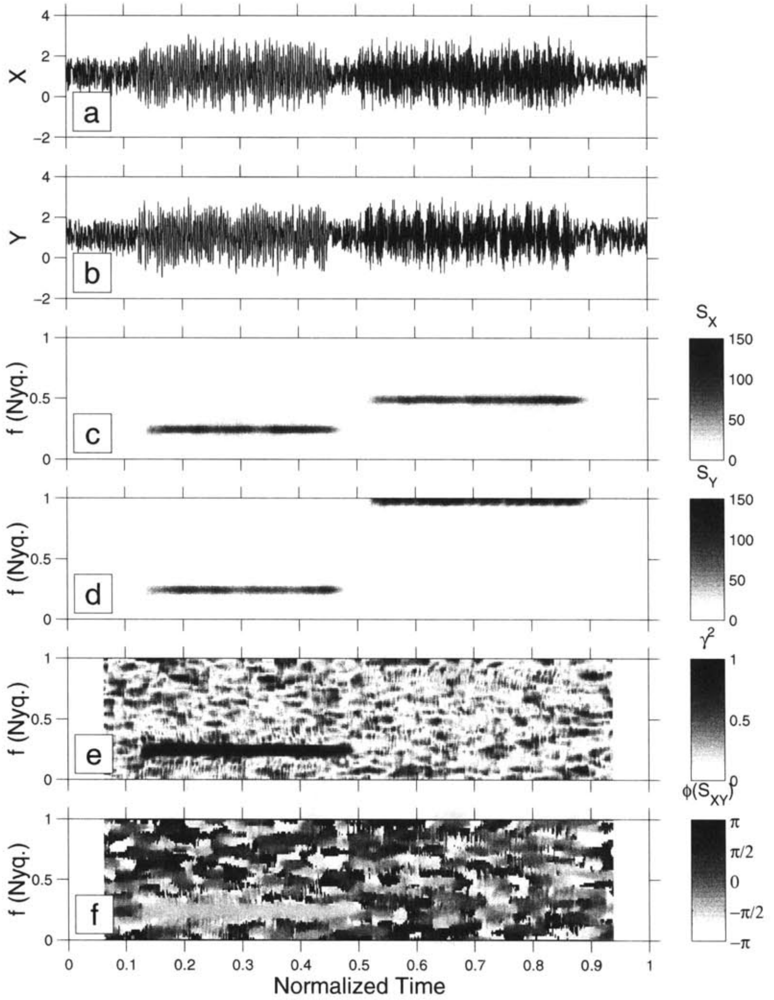


Figure 4-5: Coherence estimation with the Hermite method. (a-b) Synthetic signals as in *Santoso et al.* [1997] (see text). (c-d) Time-varying power spectra of both signals, obtained with the Hermite multi-spectrogram method. The coherence function is plotted in (e), and its phase in (f). Where the input signals are coherent, between 0.125 and 0.45, the measured coherence equals unity and the input phase difference of $-\pi/2$ is correctly retrieved. (Concentration region $R = 3$, window length $L = 256\Delta t = 0.12N\Delta t$, overlap 99%, nfft=128.)

The Hermite multi-spectrogram method was applied on spatially overlapping tiles of the data. The concentration region $R = 3$. Fig. 4-6 illustrates how spatially varying anisotropic coherence functions between two 2-D fields may be retrieved with confidence.

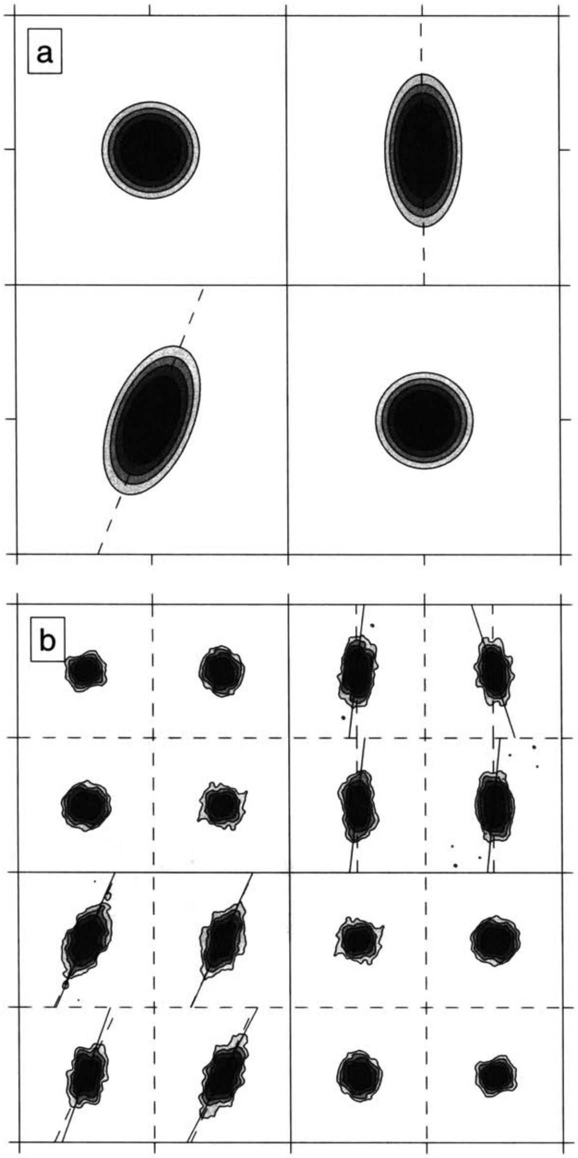


Figure 4-6: Retrieval of spatially varying anisotropic coherence functions between two synthetic fields. The two input fields are made of 4 adjacent tiles with a particular coherence function relating the two. (a) Spatial variation of the input coherence. The top left and bottom right quarters of the 2-D field show isotropic coherence, whereas the top right and bottom left are anisotropically coherent (axis of anisotropy plotted as dotted lines). For each tile, the coherence function is plotted in wave vector space, with the long wavelengths at the center of the tile. (b) Retrieval with Hermite multi-spectrogram method. Concentration radius $R = 3$, tiles overlap 25% spatially. The measured coherence functions are centered on the location of the measurement, but in reality, the measurement regions overlap spatially. Four measurements are made in each quarter, and the inferred anisotropic directions (solid lines) are compared to the input directions (dotted lines).

In summary, the Hermite multi-spectrogram method allows for unbiased, high-resolution, low-variance coherence calculations on $(\mathbf{r}, \mathbf{k})$, and this is the case for any number of dimensions. It compares favorably to traditional and Slepian multi-taper spectrogram and multi-wavelet scalogram methods of time-frequency analysis. The $(\mathbf{r}, \mathbf{k})$-space is treated as a whole without sacrificing temporal or spatial resolution to frequency resolution. Resolution and variance are at the discretion of the analyst, who can choose the radius of the concentration region. The Hermite windows and their eigenvalues are easy to calculate recursively, and no matrix diagonalization is required.

4.2.6 Variance of the estimate: Error analysis

In order to calculate the standard error on the coherence-square estimate one may resort to non-parametric methods such as jack-knifing [*Thomson and Chave*, 1991]. These methods, however, are time-consuming. Besides, the statistics of the coherence square estimator have been studied extensively. *Munk and Cartwright* [1966] and *Bendat* [1978] derive *ad hoc* expressions for the bias and variance of the estimator $\hat{\gamma}^2$. For an estimate, $\hat{\gamma}^2$, based on the average of K independent direct estimates (in our case from the multi-taper spectrograms from Eq. 4.9), the bias and variance are, in terms of the true coherence square function γ^2:

$$E\{\hat{\gamma}^2\} = \gamma^2 + \frac{1-\gamma^2}{K}, \tag{4.21}$$

and

$$\sigma^2\{\hat{\gamma}^2\} = 2\gamma^2\frac{(1-\gamma^2)^2}{K}. \tag{4.22}$$

Eq. 4.21 [misprinted in *Bendat and Piersol*, 2000, p. 333] implies a positive bias which is accentuated for low coherences based on few independent estimates. The effect of the overestimation of the coherence, which with traditional windowed periodogram methods occurs mostly in the long-wavelength range of the spectrum, provided one of the reasons for the re-evaluation with the multiple-taper technique by *Simons et al.* [2000] of the single-taper continental coherence measurements of Australia made by *Zuber et al.* [1989]. The more tapers go into the calculation of the spectrogram, the lower the variance of the estimate. However, as we have seen, this $1/K$ decrease in the estimation variance is achieved at the expense of widening the concentration region $R = \sqrt{K}$, which degrades the spectral resolution.

The approximate result stated in Eq. 4.22 is actually the expression for the well-known Cramer-Rao *lower* bound on the variance associated with the maximum-likelihood estimation of any unbiased estimate [*Seymour and Cumming*, 1994; *Touzi et al.*, 1999]. More detailed expressions for bias and variance are based on exact distribution theory [*Carter et al.*, 1973; *Touzi et al.*, 1996]. The distribution functions for the coherence square estimator of two Gaussian processes are complicated expressions involving gamma functions and generalized hypergeometric functions. Their validity has been verified experimentally [*Guarnieri and Prati*, 1997; *Touzi et al.*, 1999].

In order for us to place error bounds on the coherence calculations, we compare the exact and the approximate (Cramer-Rao) expressions for the bias (plotted in Figs. 4-7a and 4-7c) and standard deviation (Figs. 4-7b and 4-7d) of the coherence-square estimate, as a function of the number of independent estimates K. Note that, in two dimensions, a concentration radius of $R = 3$ gives rise to $K = R^2$ tapers per dimension, which amounts to $K = R^4$ independent spectral estimates from which the coherence function is calculated.

From Fig. 4-7 we conclude the following. First, our estimates based on several independent tapered estimates are nearly unbiased, and this holds for all coherence values and for all regions in the wave vector plane. Second, when basing our estimates on 9 different windowing functions in 2 dimensions, we may safely quote the easily computed approximate error of Eq. 4.22. However, as we shall see in a later section, the formal error of the

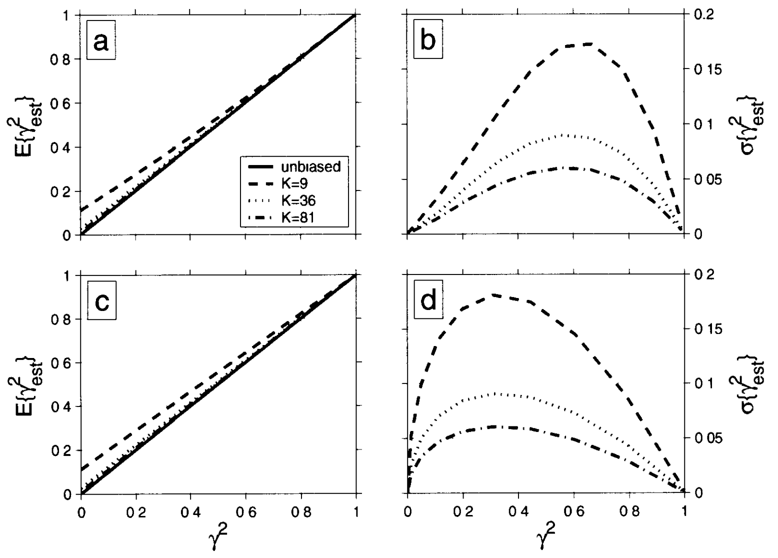


Figure 4-7: Statistics of the coherence-square estimator. (a) Expectation value and (b) standard deviation of the estimator $\hat{\gamma}^2$ in function of the true coherence square γ^2, based on exact expressions for the distribution of the estimator. (c) Expected value and (d) standard deviation of $\hat{\gamma}^2$ based the approximations discussed in the text.

coherence square estimator is of secondary importance in the assessment of the significance of our results. For this, we will turn to experiments with random and rotated data sets.

4.3 Part II: Application to isostasy in Australia

4.3.1 Gravity and topography data

The continental gravity and topography data sets used in this study are identical to the ones used by *Simons et al.* [2000]. Topography and bathymetry data (see Fig. 4-8) are from the compilation by *Smith and Sandwell* [1997], which is identical to the GTOPO30 data set [*Gesch et al.*, 1999] over the continental areas. The oceanic bathymetry was added to provide continuity with the continental topography. However, the intricacies involved in the generation of the bathymetry data from the integration of ship-sounding data with satellite

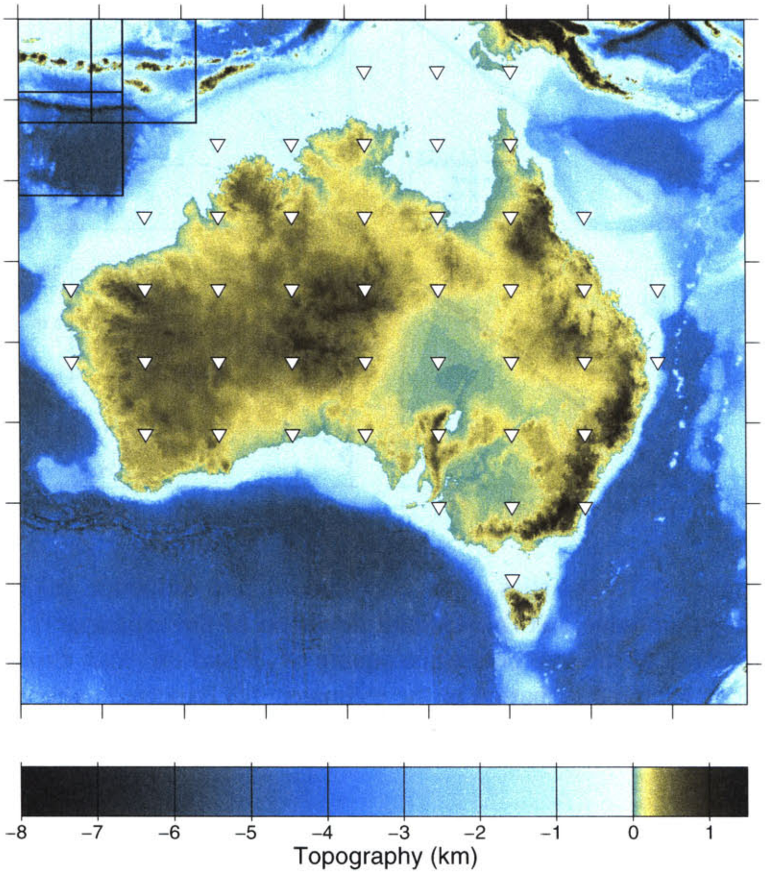


Figure 4-8: Topography and bathymetry of Australia and its surrounding areas, in km. Data from *Smith and Sandwell* [1997]. Inverted triangles indicate center of analysis boxes. Box size and overlap are plotted top left.

gravity field measurements (in which an isotropic coherence function was implicitly assumed) prevent our coherence analyses from being meaningful over the oceanic areas. The topography and gravity data were projected onto a Cartesian grid with a sampling interval of $\sim$ 5.5 km in both the x and y direction. On the continent, the AGSO (Australian Geological Survey Organisation) data set contains Bouguer anomalies calculated with a density for the crust of 2670 kg m^{-3}. Using the method of *Parker* [1972] we computed Bouguer anomalies for the oceanic areas with a crustal thickness of 6 km and adjusted the data sets to have the same baseline. The resulting oceanic and continental Bouguer anomalies are

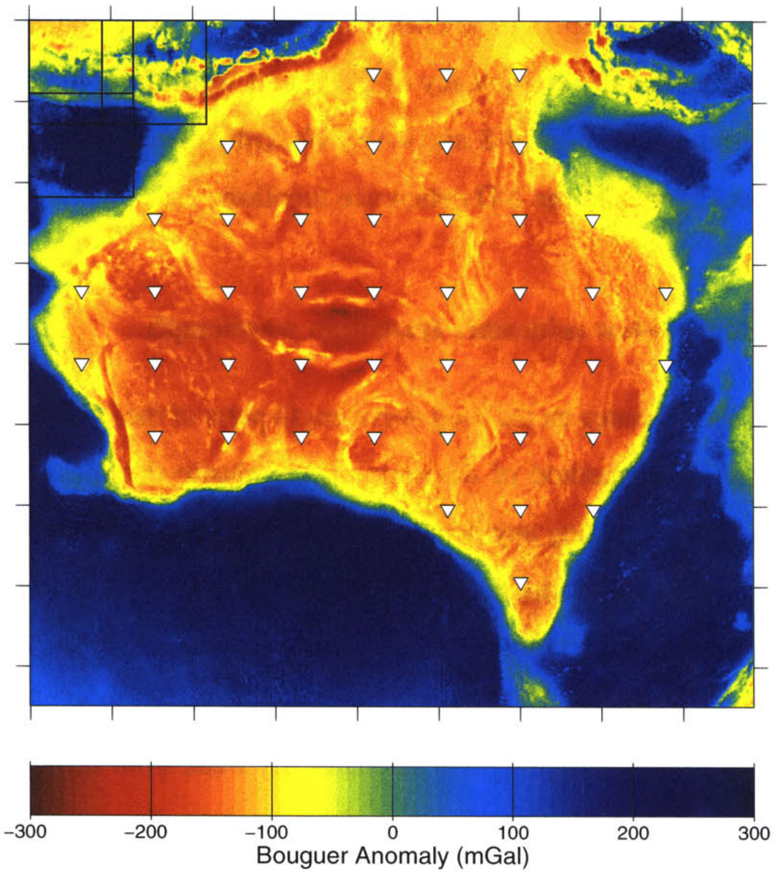


Figure 4-9: Oceanic and continental Bouguer gravity anomalies, in mGal (1 mGal = 10^{-5} m s^{-2}). Continental Bouguer anomaly data from the Australian Geological Survey Organisation. Over the oceans, Parker's (1972) method was used to compute the Bouguer correction with an average crustal thickness of 6 km.

plotted in Fig. 4-9. Figs. 4-8 and 4-9 also show the centers, size, and overlap of the analyses boxes used to generate the results.

4.3.2 Estimating the anisotropic directions

We applied the Hermite multi-spectrogram method to characterize the coherence between the data sets shown in Figs. 4-8 and 4-9. The full extent of the data was subdivided into square boxes with a size of 720 km, and an overlap of 30%. The data were analyzed using 81 different Hermite tapers (9 in each dimension) having a concentration region of $R = 3$. We have checked the consistency of the results by shifting the centers of the data boxes (inverted triangles in Figs. 4-8 and 4-9). We follow two approaches for extracting the directional dependence of the 2-D coherence function.

Estimating the high-to-low coherence transition wavelength

The first method concentrates on the long-wavelength parts of the coherence and is directly related to the traditional method of determining the elastic thickness of the continental plate, without, however, making any detailed assumptions about the governing parameters (Young's modulus, Poisson's ratio, ratio of top to bottom loading and compensation depths) which complicate the interpretation of T_e [*Burov and Diament*, 1995; *Lowry and Smith*, 1995; *McKenzie and Fairhead*, 1997; *Simons et al.*, 2000; *Banks et al.*, 2001].

The primary diagnostic to distinguish mechanically weak from strong plates is the transition wavelength at which the change occurs from isostatically compensated loads (and, thus, high coherence of the Bouguer anomaly with the topography) to uncompensated loads supported by the elastic strength of the plate (low or zero coherence) [*Forsyth*, 1985]. For every coherence estimate $\gamma^2(\mathbf{k})$ we determine, as a function of the azimuth θ of the profile defined in (k_x, k_y)-space, the wavelength $\lambda^{1/2}(\theta)$ at which the coherence drops to half of its maximum (long-wavelength) value. This transition wavelength $\lambda^{1/2}(\theta)$ is well-defined if the coherence falls off as predicted by simple compensation modeling of elastic plates, and the function $\lambda^{1/2}(\theta)$ should show a clear minimum if there is a dominant direction of weakness in the plate. The interpretation with respect to standard curves for varying T_e can then be verified by comparing $\gamma^2(k, \theta_s)$ (in the strongest direction θ_s) and $\gamma^2(k, \theta_w)$ (in the weakest direction θ_w).

In Fig. 4-10, we give three examples of results obtained using this approach. From left to right, rows a through c in Fig. 4-10 consist of a 720×720 km^2 tile of topography,

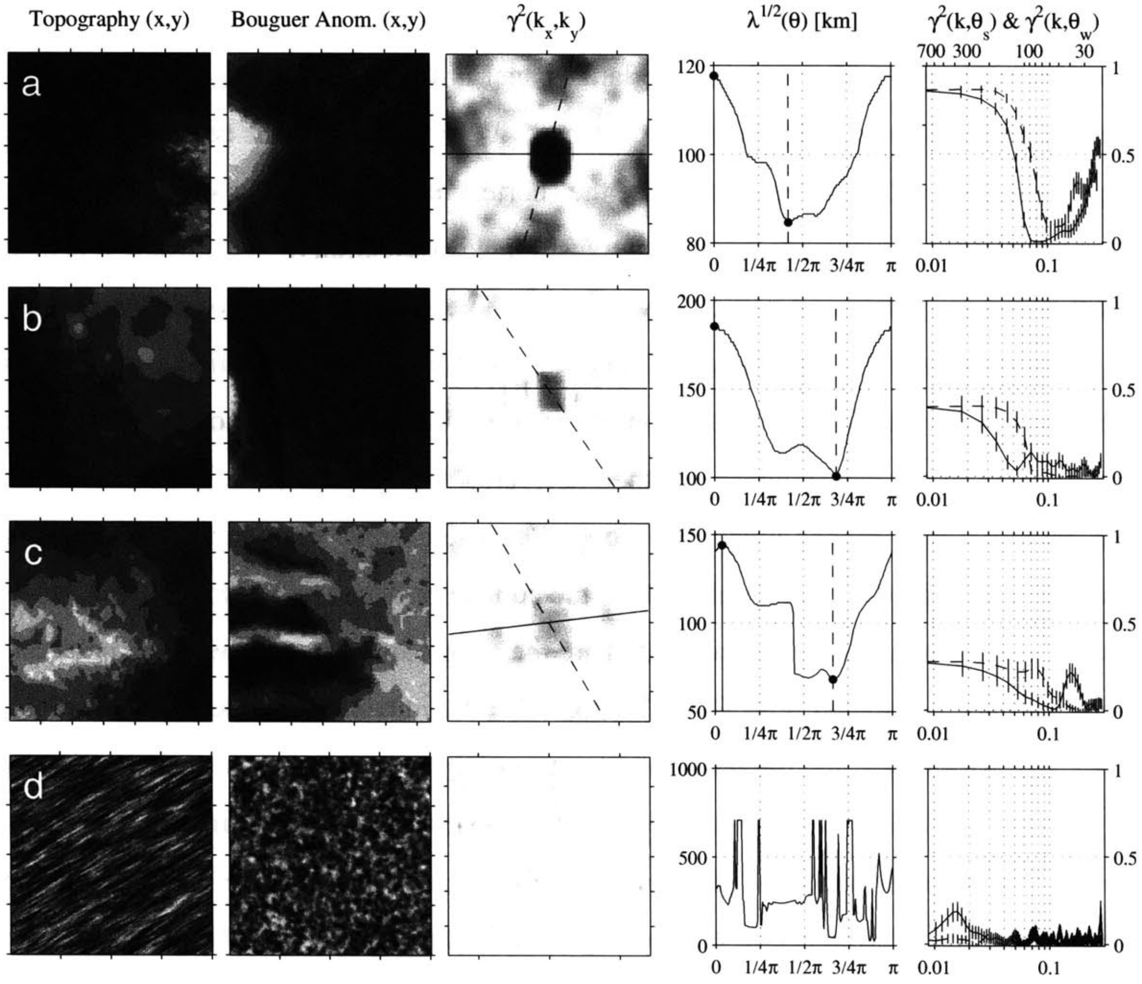


Figure 4-10: Coherence anisotropy. Rows a-c. Real data examples. Row d: synthetic example. The topography, Bouguer anomaly, and coherence-square function are plotted in the first three panels. The fourth panel shows the transition wavelength, in km, in function of the azimuth θ. The minimum wavelength corresponds to θ_w, the weakest direction of the plate (dashed lines), whereas the maximum indicates θ_s, the strongest direction (solid lines). The fifth panel shows the coherence along with its error in function of the wavenumber. The noticeable difference in transitional wavelength can be interpreted as a directional dependence of the effective elastic thickness, but due to the complexity of the assumptions involved, we have preferred to show only the coherence data rather than calculate actual values for T_e in various directions. Row d indicates that if the relation of topography to gravity was purely random, even if they had anisotropic power spectra, we would not be able to see directional differences or pick out a transitional wavelength.

$h(x,y)$, the corresponding Bouguer gravity anomaly, $B(x,y)$, and the coherence square function, $\gamma^2(k_x,k_y)$. The scale of the latter panel is linear in (k_x,k_y), with the longest resolvable (also known as Rayleigh) wavelength, $\lambda_R = 720$ km, in the center, and a wavelength of $\pm 2\lambda_N = 20$ km, equal to twice the shortest resolvable (Nyquist) wavelength, on the sides. Next is the azimuth-dependent transitional wavelength, $\lambda^{1/2}(\theta)$, where λ is in km and θ in radians, measured anti-clockwise from the equatorial line through the center of the previous panel. The maximum and minimum $\lambda^{1/2}(\theta)$ are used to define a strong (solid line) and a weak (dashed line) direction of the plate. The last panel shows the coherence as a function of the wavenumber, for the interpreted strongest ($\gamma^2(k,\theta_s)$, solid lines) and weakest ($\gamma^2(k,\theta_w)$, dashed lines) directions. Because the 2-D plots in (k_x,k_y) space are symmetric through the center, only half of the coherence needs to be plotted (here we use positive wavenumbers only). Also plotted are the error bars associated with the coherence measurements, calculated with Eq. 4.22, for $K = 81$. From the shape of the minimum of $\lambda^{1/2}(\theta)$, we can see that the direction corresponding to the minimum transitional wavelength is not always unambiguous; a conservative estimate of its uncertainty can be as large as 45°. On the other hand, the plots of $\gamma^2(k)$ for the strongest and the weakest direction show a significant difference in the coherence, and such coherence behavior corresponds to an interpretable difference in T_e of several km [for sample curves see *Forsyth*, 1985, his Fig. 7b]. We also note that the error bars were obtained for a unique azimuth without averaging in (k_x,k_y)-space. They could be further reduced by considering the coherence in a small azimuth range, as opposed to a single azimuth, but we have not found it necessary to do so.

We may assess the significance of our results by comparing them with random-phase synthetic topography and gravity data with a particular realistic power spectrum [we've used the stochastic model by *Goff and Jordan*, 1988]. Two such realizations are shown in row d of Fig. 4-10. The coherence shows no particular pattern. The transitional wavelengths are erratic and no unique minimum can be found. The wavenumber coherence is near-zero for all wavelengths.

The analysis described above was carried out on all topography/gravity tiles, and "good" measurements (such as the ones plotted in Fig. 4-10a-c) were separated from "fair" measurements (with less well defined minima). "Bad" measurements were rejected. The results will be presented in a later section.

The repeatability of our measurements derived in this way was checked by rotating a circularly tapered data set over a set of angles θ, and verifying that the coherence pattern as well as the interpreted directions rotated over approximately the same angle.

Azimuthal coherence variations at short wavelengths

The transitional wavelengths do not tell the entire story. We can further extract directional variability from the shorter-wavelength parts of the coherence function. Certain azimuths in the (k_x, k_y)-plane are characterized by higher-than average coherence values over the wavelength range shorter than about 150 km. This anisotropy is not easily quantified in terms of a single number such as a transitional wavelength, but the isostatic compensation has a preferred direction for shorter wavelengths that show a markedly higher coherence than the isotropic average. These short wavelength features are more likely caused by shallow, crustal processes such as faulting. In contrast, the anisotropy measured by the transitional wavelengths, as described in the previous section, also involve the deeper parts of the elastic lithosphere.

The approach we took was to disregard the central parts of the coherence spectrum, in the wavelength range of 150-720 km, and find the direction in which the average coherence exceeds the value of the other direction.

Three examples are given in Fig. 4-11. Again, rows a through c represent three real data examples (note that they represent different regions from the examples in Fig. 4-10),

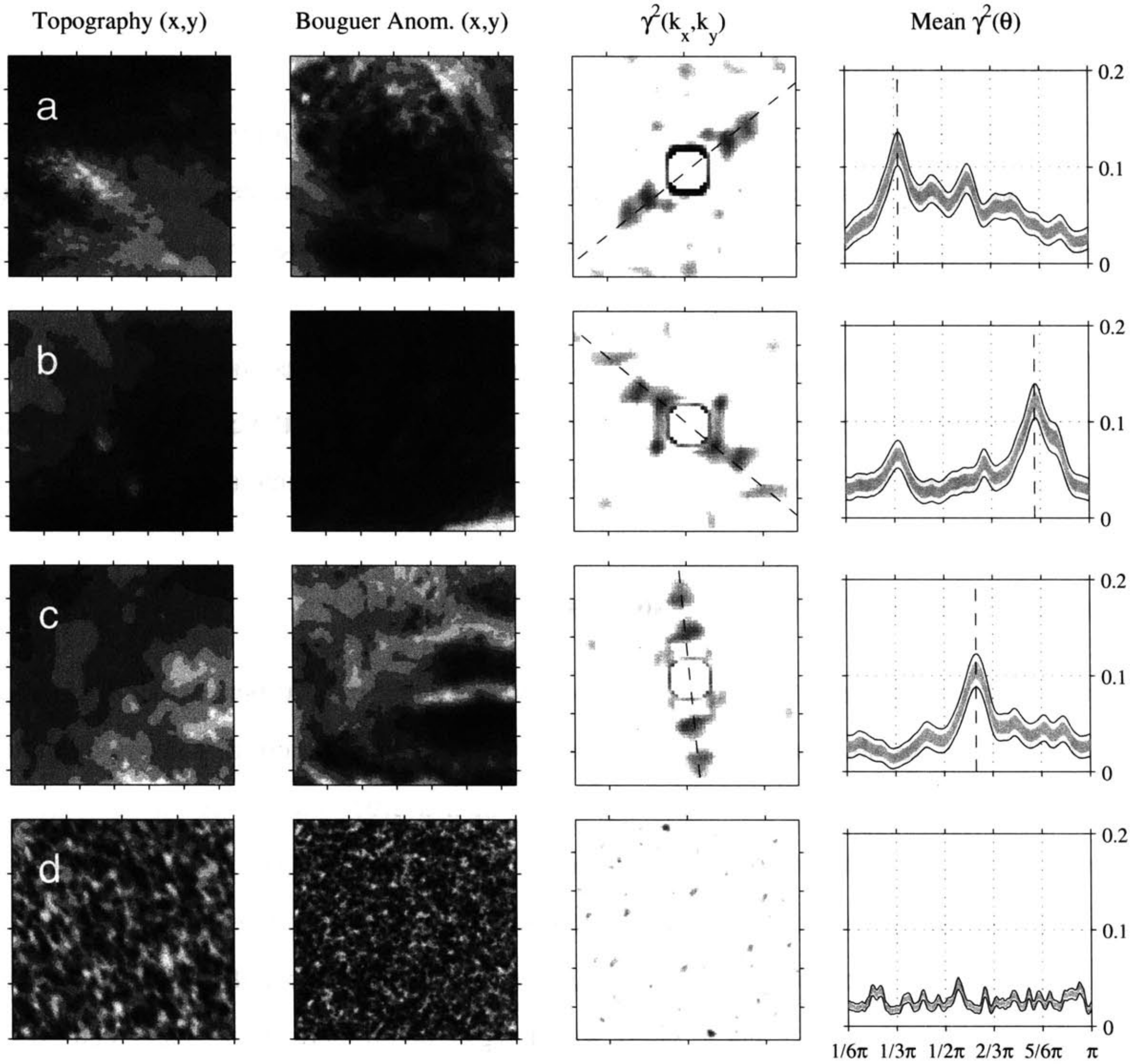


Figure 4-11: Coherence anisotropy in the wavelength ranges from 20-150 km. Rows a-c show real data (different regions than Fig. 4-10a-c), row d has a synthetic example with random-phase data. Topography and gravity fields are plotted together with the short-wavelength part of their coherence function. The average coherence per azimuth is plotted with their 1σ and 2σ error bars (dark and light shadings), and a "weak" direction is found by looking for a well-identified peak and direction (dashed lines)

which are compared with a random-phase model in row d. The first and second panels in each row show the topography and gravity fields, whereas the third shows the coherence-square function in wave vector space, without the long-wavelength part. The last panel in every row gives the average coherence as a function of azimuth. The 68% confidence interval on the average coherence is shown in dark gray. The 95% interval is bounded by solid black lines. The direction of maximum coherence is marked by a dashed line. As this represents again the weaker or more easily isostatically compensated direction of the continental plate, the same direction is also plotted on top of the coherence function in the third panel. As with the long-wavelength part of the coherence, we are looking for directions corresponding to easily identified peaks in the coherence averaged over an azimuth ("good" measurements), along with more ambiguous peaks ("fair"), and we reject "bad" measurements. The results are summarized in the next section.

4.3.3 Results: Space-varying anisotropic directions in the coherence

The weak directions based on the (long-wavelength) transitional wavelength method are plotted in Fig. 4-12a, and the directions where the coherence is higher than average in the short wavelength range are shown in Fig. 4-12b. High-quality measurements are shown with thick solid lines, and fair ones with thin solid lines.

The western two-thirds of the Australian continent are of Precambrian origin [*Rutland*, 1976], but rather than being a single stable shield they are the result of the amalgamation of numerous continental fragments [*Myers et al.*, 1996] that are joined by weak zones characterized by extensively reworked crust, which are often manifest as strong gradients in the gravity and magnetic anomaly maps [*Wellman*, 1998]. Precambrian Australia constitutes an Archean Western Australian craton (with the Yilgarn and the Pilbara craton as two distinct elements), a stable North Australian craton, which, based on its wave speeds, extends into the offshore areas of the continental platform [*Zielhuis and van der Hilst*, 1996; *Simons et al.*, 1999], and a South Australian craton comprising the Archean to Proterozoic Arunta and Gawler units. The eastern third of the Australian continent was accreted during the Paleozoic, and has been the scene of intense orogenic activity with predomi-

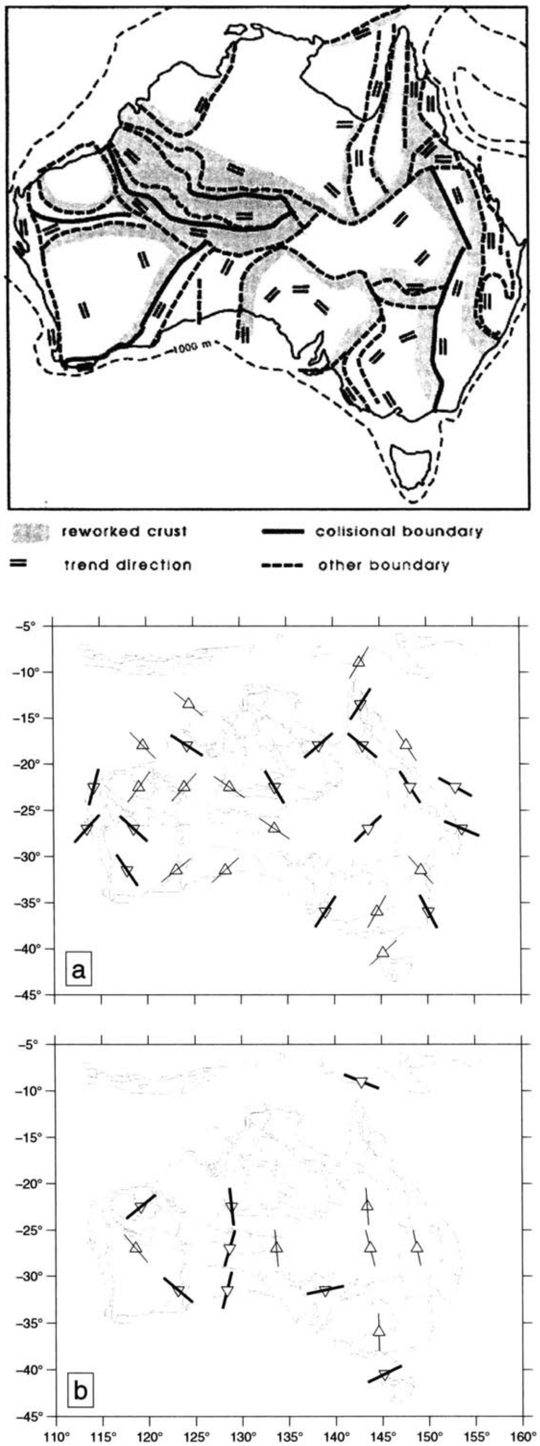


Figure 4-12: (Top) Major trend directions on the Australian continent, from an analysis of gravity and magnetic anomalies. From [*Wellman*, 1998]. (a-b) Spatial variation of coherence anisotropy between continental Bouguer anomalies and topography. (a) Long-wavelength weak directions based on the direction of minimum transitional wavelength as described in the text. (b) Short-wavelength weak directions based on a consistently higher than average coherence in the range from 20 to 150 km. "Good" measurements in thick solid lines; "fair" measurements in thin solid lines.

nantly N–S oriented strike. The Central Australian region, which extends to the NW and separates the Western and Southern from the Northern Australian mega-elements, is an extensively reworked zone with major collisional and other boundaries, which are oriented predominantly E–W in the central parts of Australia [*Wellman*, 1998; *Shaw et al.*, 1995]. Fig. 4-12 (after *Wellman* [1998]) shows the location of the major geological and geophysical boundaries on the Australian continent. Such zones represent substantial rheological heterogeneities, and as pre-existing weak zones they have been invoked to explain the relative stability of cratonic regions [*Vauchez et al.*, 1998; *Lenardic et al.*, 2000; *Tommasi and Vauchez*, 2001]. In our short-wavelength anisotropy map of Fig. 4-12b, we notice a good correlation between the zones in which we observe strongly anisotropic coherence and the suture zones between the major geological and geophysical domains in Australia. With few exceptions, the weak direction is indeed oriented nearly perpendicularly to the trend directions mapped in Fig. 4-12. Furthermore, with few exceptions, the long-wavelength weak directions evident from Fig. 4-12a are oriented at a high angle to the ocean-continent boundary, which may be a manifestation of the weakness associated with the junction of oceanic and continental crust.

The clear N–S anisotropy of the central Australian lithosphere evident in the short-wavelength coherence anisotropy from Fig. 4-12b corroborates the findings of *Simons et al.* [2000], who tentatively related the direction of weakness to the presence of pervasive E–W running zones of faulting [*Lambeck et al.*, 1988] and rheological weakness due to differential sediment burial rates [*Sandiford and Hand*, 1998]. It is also consistent with the predicted effect of the N–S regional pattern of maximum compressive stress [*Lambeck et al.*, 1984] on the isostatic compensation mechanism [*Lowry and Smith*, 1995].

4.3.4 Discussion

The entire Australian plate is in a state of compression [*Zoback*, 1992; *Coblentz et al.*, 1998]. This compression is oriented in the N–S to SW–NE direction in northern and central Australia, and E–W in southwest and southeast Australia, which is consistent with its probable origin in plate-boundary forces [*Coblentz et al.*, 1995]. The correspondence of isostatic anisotropy and contemporaneous lithospheric stress is, however, far from unequivocal [as it is in other parts of the world, see *Lowry and Smith*, 1995]. This may have several reasons. Foremost, several regional stress indicators are only partially consistent with each other, as shown by comprehensive measurements of bore-hole break-outs [*Hillis et al.*, 1998, 1999] and focal mechanisms [*Lambeck et al.*, 1984; *Spassov*, 1998; *Spassov and Kennett*, 2000]. Moreover, the seismicity of Australia is very low. As a consequence, it is virtually impossible to find a unique stress direction responsible for the anisotropic isostatic compensation. Second, loading of an elastic plate and the creation of gravity anomalies might have reflected the stress state of the plate at the moment the loading and deformation took place, but from the preserved gravity structure it is not possible to derive the present-day stress field [unless, as shown by *Lowry and Smith*, 1995, for certain areas in the continental United States, present-day stress and fossil strain are still related]. Attempts to find a direct and unambiguous correspondence between our measurements and the principal directions of stress are thus hampered by the uncertainty in the measurement and modeling of the intraplate stress field, and the absence of an *a priori* genetic relationship between both.

We turn to a discussion of the directions of mechanical anisotropy along with the seismic anisotropy inferred from surface-wave tomography. The internally coherent deformation hypothesis due to *Silver* [1996] postulates that the strain field observed in the structural geology of the Earth's surface adequately represents the mantle strain, which, in turn, produces a lattice preferred orientation (LPO) of anisotropic upper mantle minerals, such as olivine, that affects the propagation of seismic waves.

Interpretations of lithospheric structure and tectonic history based solely on surface geology can be seriously flawed [*Goleby et al.*, 1989; *Lucas et al.*, 1993]. Large-scale lithospheric deformation processes [e.g. *Stephenson and Cloetingh*, 1991], on the other hand, are well documented by the gravity anomalies they create [e.g. *Chamberlain et al.*, 1993; *Cloetingh et al.*, 1999]. We argue that the anisotropy in the isostatic compensation mechanism provides a more robust handle on the principal strain directions than surface mapping of geologic trends alone.

Seismic anisotropy can result from the finite strain-induced preferred orientation of mantle minerals. Based on a variety of experimental and theoretical studies, the [100] axis of olivine is predicted to be parallel to the extensional direction for nearly all kinds of finite strain, or perpendicular to the shortening direction [*Nicolas and Poirier*, 1976]. Hence, for example, if our coherence analysis shows that a plate is weaker in a certain direction, and deformation causes it to deform, fold or buckle more easily in that direction, the fast polarization direction from *SKS* splitting or the direction of maximum propagation speed from the azimuthal anisotropy of surface waves (as predicted by a model of internally coherent deformation) is perpendicular to the inferred collisional episodes, or parallel to the structural trends [*Silver and Chan*, 1991].

Our analysis allows us to confirm or reject if this first-order hypothesis about mantle deformation and anisotropy is applicable to Australia. We compare the weak plate direction, which has accommodated the most strain, with the direction of maximum wave propagation speed from seismic studies, and investigate to which extent they are parallel or at right angles to each other.

There are two differences between our approach and those followed elsewhere. First, rather than basing our predictions on surface geology [strike of faults, province boundaries, stress directions, etc. as was done by *Silver and Chan*, 1991] we take the mechanical anisotropy from coherence analysis as a predictor for the fast polarization angle. Second, rather than utilizing seismic anisotropy measurements from *SKS* splitting, which provide no depth resolution [the most likely depth regions for their origins being still very much debated; see *Savage*, 1999] we use the direction of fast wave propagation from the analysis of azimuthal anisotropy of surface waves. Surface-wave waveform tomography enables us to detect horizontal as well as vertical variations in anisotropy and wave speed [*Simons et al.*, 2002a]. The advantages of surface-wave studies are two-fold. First, they have good resolution at depth, in contrast to the vertically propagating body waves, so we can present the results as a function of depth. Second, the connection with *SKS*-type splitting analysis can still be made because the formalism exists to relate body-wave and surface-wave anisotropy [*Rümpker and Silver*, 1998; *Saltzer et al.*, 2000; *Montagner et al.*, 2000]. This is important as the splitting observations made in Australia by *Clitheroe and van der Hilst* [1998] and *Özalaybey and Chen* [1999] are at odds with each other.

Simons et al. [2002a] have investigated the three-dimensional S-wave heterogeneity and azimuthal anisotropy from Rayleigh waves propagating in the Australian upper mantle. They presented the results in terms of a (presumed horizontal) fast axis of anisotropy. In their model, there is little relation between the anisotropy in the upper 200 km and surface geological units. The pattern at larger depths is smoother, but no significant correspondence with the direction of absolute plate motion could be detected. Therefore, the seismic anisotropy of Australia appears to be more complex than simple end-member models of vertically coherent deformation [*Silver and Chan*, 1988] or a predominant influence of present-day mantle deformation [*Vinnik et al.*, 1995] would predict.

We have averaged the anisotropy model of *Simons et al.* [2002a] over equally sized, similarly overlapping regions of the continent as we used in the coherence analysis. The fast axes for various depths are shown in Fig. 4-13 (black lines scaled to the magnitude of the anisotropy). Overlaying these results, we plot the long-wavelength directions of mechanical weakness (gray lines) in panels a through d (corresponding to the thick solid lines in Fig. 4-12a), and the short-wavelength results in panels e through h (the thick solid lines of Fig. 4-12b). We show this comparison only for the locations of the best measurements of the coherence anisotropy.

In most cases the mechanical and seismic anisotropy are at large angles. We quantify this by calculating which percentage of the weak directions (from both estimation regimes) fall within ±22.5° of the perpendicular direction to the seismic anisotropy (and which ones fall within ±22.5° of the parallel direction). We plot the answer to this question in Fig. 4-14, for different depths. From this figure, we may conclude that there is indeed some agreement between the predictions based on the long-wavelength mechanical anisotropy and the seismically anisotropic directions. At 30 km depth, the seismic directions are perpendicular to the directions of mechanical weakness for about 60% of the measurements, while only 10% of the weak directions which are parallel to the fast axis of anisotropy. With increasing depth, these figures change to about 30% (perpendicular) and 20% (parallel) at 200 km depth. For the mechanical anisotropy derived from the high-frequency component of the coherence, the relation with seismic anisotropy is more ambiguous. As we have argued before, its origin probably lies in the shallow crust.

In addition to comparing the directions of mechanical weakness with the depth-dependent directions of surface-wave anisotropy, we have compared them with splitting directions from *SKS*, *SKKS* and *PKS* phases [*Clitheroe and van der Hilst*, 1998; *Iidaka et al.*, 2001]. Shear-wave splitting times in Australia are, in general, small. Of the 89 measurements reported in the above two studies, only 13 display splitting times equal than or larger than 1 s. At only 4 stations a corresponding high-quality measurement of coherence anisotropy could be made, and for 3 out of those 4 cases, the mechanical anisotropy was very nearly perpendicular to the splitting direction. As shear-wave splitting measurements are predom-

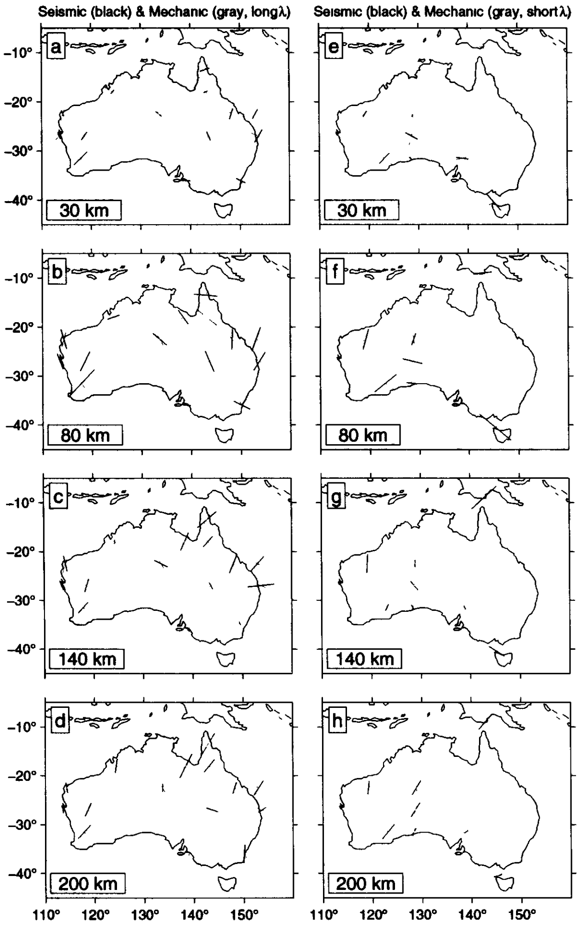


Figure 4-13: The relation between seismic and mechanical anisotropy. (a-d) Long-wavelength mechanical anisotropy (weak directions from Fig. 4-12a, in gray) compared to fast-axis directions at different depths from the azimuthally anisotropic surface-wave model by *Simons et al.* [2002a]. The orientation of the bars represents the fast direction of anisotropy; their length is proportional to the magnitude of the anisotropy. (e-h) Short-wavelength mechanical anisotropy (weak directions from Fig. 4-12b, in gray) compared to seismic fast-axis directions.

inantly influenced by shallow structures [*Saltzer et al.*, 2000], this perpendicular relationship is in agreement with the inference made from Fig. 4-14 that such a relationship holds predominantly for the uppermost upper mantle.

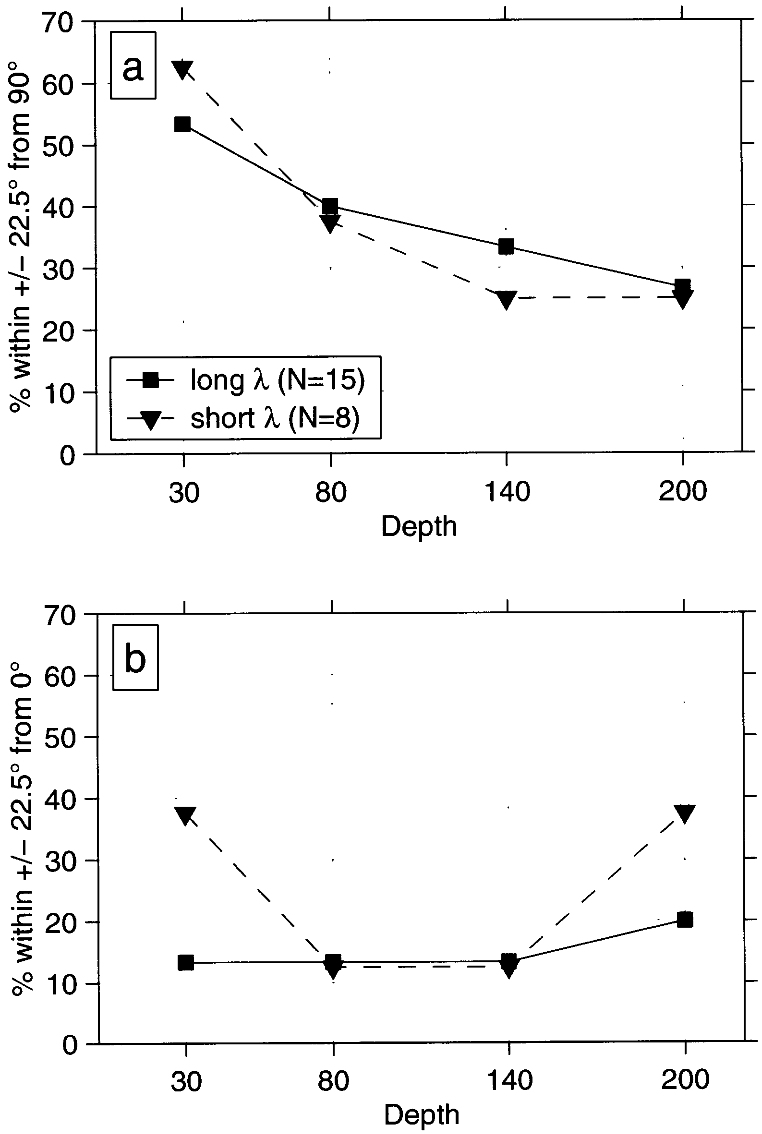


Figure 4-14: Relationship between mechanically weak directions and fast axis of seismic anisotropy. (a) Percentage of long (squares, solid lines) and short-wavelength (triangles, dashed lines) weak directions plotting within 45° perpendicular to the seismic fast axis. (b) Percentage of long and short-wavelength mechanical weak directions that are within 45° from being parallel to the fast seismic directions.

4.4 Conclusions

Whether the origin of continental seismic anisotropy lies in fossil strain from wholesale deformation processes [*Silver and Chan*, 1988] or in contemporaneous mantle deformation [*Vinnik et al.*, 1995] due to plate motion is a subject of much active debate [see, e.g. *Mooney*, 1995]. Either interpretation is inherently limited by two effects. First, the fast polarization directions obtained from shear-wave splitting analysis are commonly taken to be representative of a narrow depth range of the upper mantle (deeper than 40 but shallower than 400 km) even though this origin depth is still very poorly constrained and widely debated. Second, features that are observable exclusively at the surface are taken as proxy for the deformation that affects the entire lithosphere, but this deformation arguably is not uniform with depth.

Seismic surface waves are better candidates to investigate the depth-dependent anisotropy of the lithosphere. In addition, it is important to realize that besides being expressed geologically at the surface, continental deformation produces topography and gravity anomalies. The relation between gravity and topography contains valuable information about fossil lithospheric strain. Thus, the study of gravity-topography coherence (the isostatic response), and, in particular, its variation with azimuth, may supplant the necessarily limited and ambiguous mapping of surface trend directions in order to infer the strain direction dominant over time and integrated over depth.

For determining the anisotropic coherence of two-dimensional fields (to reveal directional differences in mechanical strength or directions of preferential isostatic compensation and strain accumulation) multi-window methods of spectral analysis are needed to provide the necessary phase averaging. To study variations of coherence with space, multiwavelet or multi-spectrogram methods can be used. The concentration properties of windows or wavelets depend on their average periodogram (in the stationary case) or on their average Wigner-Ville transform (for non-stationary fields). Slepian sequences and Slepian wavelets tile the $(\mathbf{r}, \mathbf{k})$-space with rectangular concentration regions. The optimal way of analyzing space-varying spectral properties is using Hermite windowing functions. Their concentration regions are spherical and hence treat the phase-space as a whole, without trading off spatial for spectral resolution. We have described a multi-spectrogram method of coherence analysis based on Hermite windows and shown that it should be the method of choice for retrieving spatial, azimuthal and wavelength-dependent variations of coherence.

We have applied this method to the study of isostasy on the Australian continent. The azimuth with the shortest transitional wavelength at which the coherence between Bouguer gravity and topography occurs is identified as the direction of mechanical weakness (and hence strain concentration). In the wavelength range below 150 km, we further identify weak directions of preferential isostatic compensation by their anomalously high coherence with respect to the isotropic average.

Several possible causes for mechanical anisotropy have been explored. Some of these measurements are consistent with the notion that the boundaries between stable continental cratons are mechanically weaker. However, the most significant conclusions may be deduced from a systematic analysis of the angle between the mechanically weak and seismically fast directions. If the hypothesis of internally coherent deformation is correct, this relationship is predicted to be perpendicular. The weak directions inferred from the transitional wavelengths seem to corroborate this hypothesis to a certain extent; about 60 % of the angle differences cluster around 90°, but only for shallow depths. However, the degree of this correspondence drops off quickly with increasing depth. At a depth of 200 km, not significantly more weak directions are perpendicular to the fast polarization direction than are parallel to it.

Not only have we proposed an adequate means of testing models that attempt to relate continental deformation with the development of seismic anisotropy, but we have also shown that simple end-member models are not universally substantiated: it is unlikely that either basal drag due to plate motion or fossil deformation alone are responsible for the development of seismic anisotropy.

Acknowledgments

We thank Oded Aharonson, Mark D. Behn, Felix Herrmann, Hrafnkell Kárason, Gilbert Strang, Paul Tapponnier and Andrew T. Walden for help and advice. We thank Takashi Iidaka for supplying his shear-wave splitting measurements. This research was supported by NSF grant EAR-0001136.

Chapter 5

Age-Dependent Lithospheric Thickness[1]

Abstract

We present seismic and mechanical constraints on the regional variations of the thickness of the Australian lithosphere. We infer the seismic thickness from a waveform tomographic model of *S* wave speed, and as a proxy for the elastic thickness we use the wavelength at which the coherence of surface topography and Bouguer gravity drops below half of its long-wavelength maximum. Our results show that on scales larger than ~1000 km the relationship between the age of the crust and the thickness of the lithosphere is more complicated than longer wavelength global averages suggest. Recent geochemical and geodynamical evidence for small-scale secular variations of the composition and stability of continental cratons further illustrates the complexity of the age dependence of seismo-mechanical lithospheric properties on regional scales smaller than ~1000 km.

---

[1] **Age-Dependent Seismic Thickness and Mechanical Strength of the Australian Lithosphere**, *Geophysical Research Letters*, submitted, 2002.

5.1 Introduction

In global compilations the age of the crust correlates with its bulk rock composition [*Griffin et al.*, 1999] and elastic thickness [*Zuber et al.*, 1989; *Bechtel et al.*, 1990], surface heat flow [*Artemieva and Mooney*, 2001], and seismic wave speed anomalies in the subcontinental lithospheric mantle down to about 200 km [*Polet and Anderson*, 1995]. As low temperatures and refractory mantle compositions contribute to high S wave speeds, these observations suggest that progressive depletion of basaltic constituents by tectonic processes stabilizes the lithosphere over time into a seismologically well-defined thick and strong thermo-chemical boundary layer or tectosphere [*Jordan*, 1988].

Recent studies have shown regional departures from the global correlation of geological, geochemical and geophysical lithospheric properties with age. Detailed heat flow studies manifest the difficulty in representing an entire tectonic province by a single average crustal model [*Jaupart and Mareschal*, 1999]. The predicted remobilization of cratonic lithosphere [*Conrad and Molnar*, 1997] is geochemically documented in South Africa [*Jacob et al.*, 2000], the Proterozoic Sierra Nevada [*Lee et al.*, 2000], and the Archean Wyoming province [*Eggler et al.*, 1988]. A thick lithosphere might be an effect of preferential preservation inside orogens [*Lenardic et al.*, 2000] or of its initial composition and geochemical fertility [*Lee et al.*, 2001]. Hence, lithospheric thickness may be poorly correlated with the age of the overlying crust. Crust and mantle need not have the same age.

Elastic properties depend on the time scale. Wave speeds reflect the elastic response on short time scales (seconds to hours). In contrast, the effective elastic thickness (T_e) estimated from the coherence between gravity anomalies and topography is a measure of its long-term mechanical strength (millions of years).

Our knowledge of the velocity structure of the Australian continent has greatly improved since the Skippy project [*van der Hilst et al.*, 1994], which enabled the analysis of

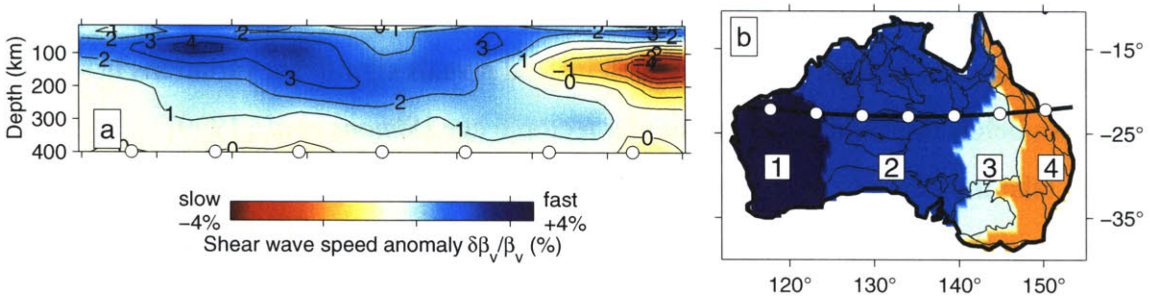


Figure 5-1: Wave speed anomalies beneath Australia. (a) Profile (see (b) for location) and (b) map, at 80 km, of average anomalies over crustal age domains. The Phanerozoic ((3-4), <542 Ma) is characterized by a thin lithosphere, whereas in the Proterozoic ((2), 542 to 2500 Ma) and Archean ((1), >2500 Ma) parts of the continent, lithospheric keels extend to 200–250 km depth.

large broadband datasets [*Zielhuis and van der Hilst*, 1996; *Debayle and Kennett*, 2000a; *Simons et al.*, 2002a]. In addition, recent advances in coherence analysis now allow robust high-resolution measurements of the mechanical properties of the lithosphere [*Simons et al.*, 2002b].

Building on the improvements in both fields we investigate the relationship between the seismic thickness and the long-term elastic strength of the lithosphere and the age of the crust. These correlations are more complex than was previously assumed on the basis of global analyses and reveal a difference in lithospheric properties at length scales above and below ∼1000 km.

5.2 Seismic thickness

On a length scale of ∼1000 km or more the propagation speed of seismic waves in the Australian lithosphere correlates very well with the age of the crust (Fig. 5-1). The eastward decrease in crustal age coincides with a reduction in seismic wave speed, at least to a depth of 200 km. However, below 200 km, the realm of deep-penetrating stable continental keels in global models [see, e.g., *Masters et al.*, 1996], this correlation is not borne out by high resolution seismic modeling which instead reveals significant variations in seismic velocities below crustal domains of similar tectonic age [*Simons et al.*, 1999].

We take the depth to a particular velocity anomaly contour as an estimate of seismic lithospheric thickness [see, e.g., *Frederiksen et al.*, 2001]. Fig. 5-2 shows the depth at which the *S* wave speed deviation from the average continental reference model [*Simons et al.*, 2002a] drops below +1.0%. The choice of contour is arbitrary but we will show that the pattern of seismic thickness variations obtained from the 1.5% and 2.0% contours is consistent with that inferred from the 1.0% contour. The Archean and Proterozoic parts of Australia are marked by thick lithosphere, whereas the Phanerozoic crust overlays a much thinner high velocity lid or none at all. The transition occurs near the eastern extent of Precambrian outcrop delineated by the Tasman Line [*Veevers*, 1984]. Some Proterozoic regions of Australia exhibit high velocity keels to greater depth than the Archean portions of the continent. The structure of the mantle beneath easternmost Australia resembles the structure of the adjacent oceanic mantle. In the offshore regions the model shows a large difference in thickness between the old oceanic lithosphere of the Indo-Australian plate northwest and the much younger lithosphere east of Australia.

5.3 Long-term elastic strength

The relative mechanical strength or T_e of the lithosphere can be obtained from the coherence (γ^2) between Bouguer gravity anomalies and surface topography [*Forsyth*, 1985]. Theoretically, γ^2 evolves from 1 to 0 with decreasing wavelength, and the wavelength at which $\gamma^2 = 0.5$ varies positively, but nonlinearly, with T_e. Numerous factors, such as the unknown ratio of buried to surface loads [*Banks et al.*, 2001], the method used in the spectral calculations [*Simons et al.*, 2000], erosion [*Ito and Taira*, 2000], and the assumed value of several rheological variables [*Lowry and Smith*, 1995] influence the estimation of T_e.

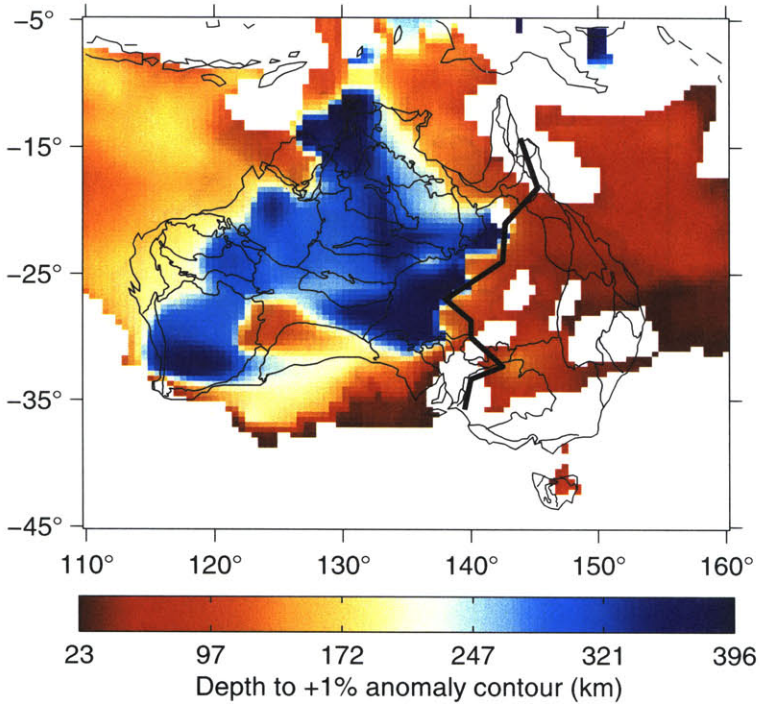


Figure 5-2: Depth to the +1.0% velocity anomaly contour, a proxy for the seismic thickness of the lithosphere. Thick solid: Tasman Line [after *Veevers*, 1984].

Differences in T_e of up to 100% can result from varying methods and modeling assumptions. There is, however, a direct physical significance to wavelengths of loads that are isostatically compensated (high γ^2) and those of loads that are supported by the strength of the plate (lowγ^2). Because of the uniqueness concerns mentioned above we calculate the wavelength ($\lambda^{1/2}$) that marks the transition between these regimes without converting it into a single T_e value.

The mechanical strength of continents correlates with the time since the last tectonothermal event [*Zuber et al.*, 1989; *Bechtel et al.*, 1990; *Poudjom Djomani et al.*, 1999]. We use the coherence method of *Simons et al.* [2002b] to investigate this correlation in more detail. We depict the values of the $\lambda^{1/2}$ at the midpoints of the analysis windows, which have a constant dimension of $\sim$ 700 km.

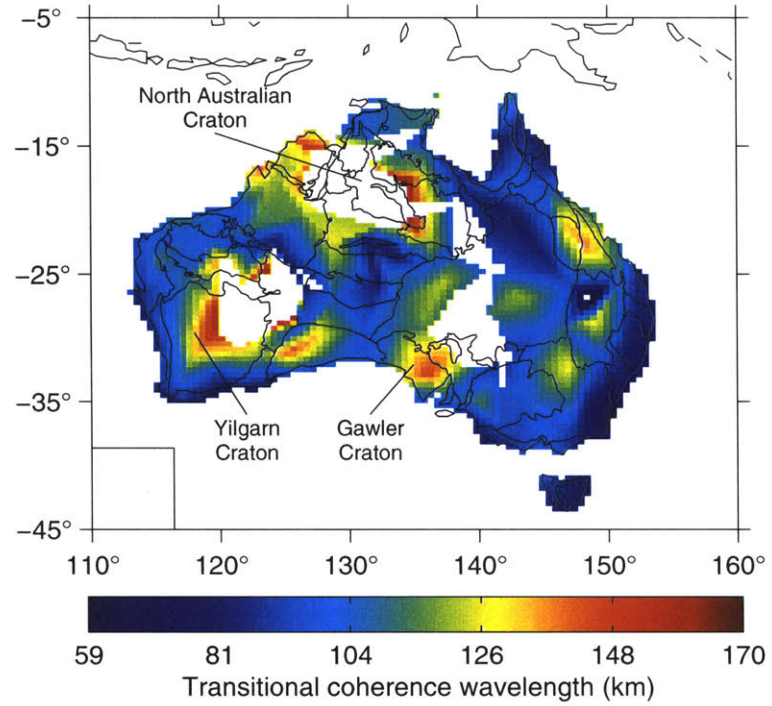


Figure 5-3: Map of transitional coherence wavelengths $\lambda^{1/2}$, an indicator of the mechanical thickness. Bottom left: window size used.

Fig. 5-3 shows the lateral variation of the relative mechanical strength of Australia. Longer $\lambda^{1/2}$ mostly occur in the western (Archean) and northern (Proterozoic) parts of Australia. In addition, there is a high in the South centered over the Archean Gawler craton, whereas East Australia encompasses regions with low and high $\lambda^{1/2}$. Fig. 5-3 suggests a large-scale trend of $\lambda^{1/2}$ decreasing eastward with the age of the crust. A low value of γ^2 at the wavelength corresponding to the window size indicates a failure of the linear isostatic compensation model, and no $\lambda^{1/2}$ is plotted in the regions where $\gamma^2 < 0.2$. We assessed the robustness of the trends observed in Fig. 5-3 by varying window sizes and using more conservative cut-off values for γ^2, i.e. higher than 0.2.

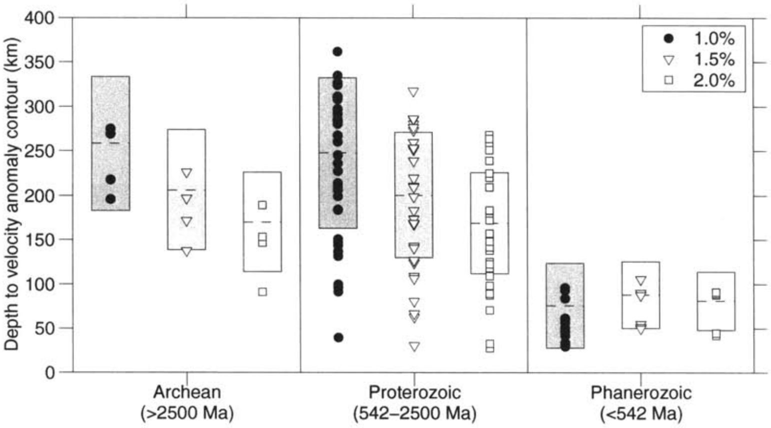


Figure 5-4: Average seismic thickness of tectonic subdomains (symbols), and area-normalized mean (dashed line) and standard deviation (box) over the eon to which they belong. Based on $+1.0\%$, $+1.5\%$, and $+2.0\%$ anomaly contours.

5.4 Age dependence

Fig. 5-4 depicts seismic lithospheric thickness as a function of eon, using the depth to the $+1.0\%$, $+1.5\%$, and $+2.0\%$ wave speed anomaly contours. We show average values over the tectonic domains outlined in Figs. 5-2 and 5-3 and grouped in function of the eon to which these domains belong (as in Fig. 5-1b). The overall decrease in lithospheric thickness with decreasing age, previously seen in longer wavelength global models, is corroborated by our age-averaged values, but the improved resolution of our regional surface-wave modeling reveals substantial variation about this long-wavelength trend. Using different wave speed contours leads to qualitatively similar conclusions, but the thickness variation from one contour to another relates to the velocity gradient, which is much steeper in the Phanerozoic than in the Precambrian.

An averaging of $\lambda^{1/2}$ per tectonic domain and eon is shown in Fig. 5-5. The mean $\lambda^{1/2}$ increases slightly with age. Although this indicates a thicker or stronger elastic lithosphere, even in the most favorable of loading scenarios the difference in T_e calculated from coherence curves with the same range of $\lambda^{1/2}$ will not exceed 10 km. The variation of mechanical

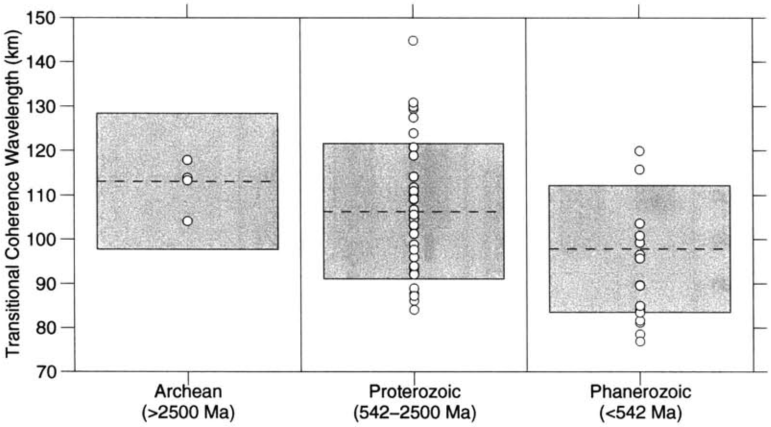


Figure 5-5: Average transitional coherence wavelength averaged over tectonic subdomains and grouped per eon.

strength within domains formed in the same eon is far greater than the slight thickening of the lithosphere with age. We caution that the absence of large strength variations overall may be as much related to intrinsic properties of the lithosphere as to our inability to detect them [see e.g., *McKenzie and Fairhead*, 1997].

5.5 Discussion

Both the seismic thickness and the mechanical strength of the lithosphere generally correlate with the age of the overlying crust over length scales of thousands of kilometers, but there is significant scatter on scale lengths smaller than $\sim$1000 km. Do the seismic and the mechanical thickness covary on these length scales, as assumed in many studies [see *Nyblade*, 2001]? Does the strength of the crust control the stability of the lithosphere? Fig. 5-6 shows the correlation of seismic with mechanical thickness for the Australian Archean, Proterozoic, and Phanerozoic. Fig. 5-6 also shows the standard deviation of the thickness within the domain under consideration. The (few) Archean regions of Australia cluster within a narrow range of transition wavelengths and show little variation in their seismic thickness. The seismic thickness and mechanical strength of Proterozoic Australia

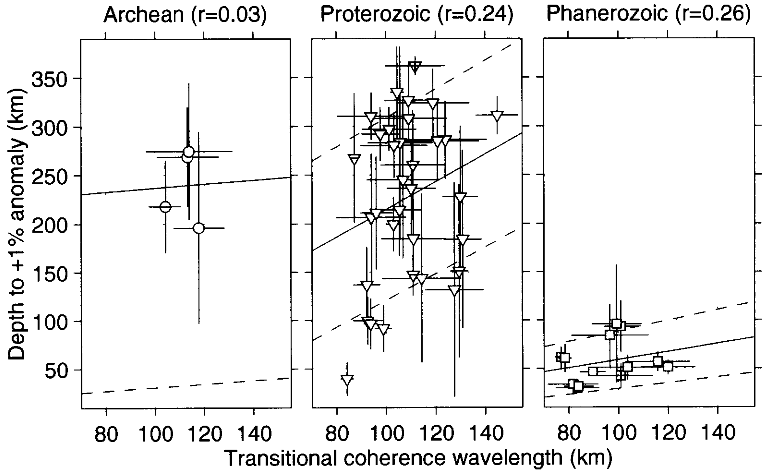


Figure 5-6: Transitional wavelengths, seismic thickness, and their standard deviations, as a function of age. Regression lines (solid) and their 50% confidence intervals (dashed) are shown, but the correlations are not statistically significant.

encompass the values for the Archean and scatter around a poorly defined positive linear correlation. The range of thicknesses seen in the Proterozoic is larger than for the Archean of Phanerozoic parts of Australia. Phanerozoic Australia displays small if any high velocity lids that, again, show essentially no correlation with their elastic strength. East Australia is not underlain by a thick tectospheric root, which cautions against making comparisons of the seismic thickness with the elastic strength. We stress that there is no statistical significance to any of the three regression lines shown in Fig. 5-6, as deduced from the low values of the linear correlation coefficient.

There are plausible geodynamical and geochemical mechanisms to preclude a straightforward age relationship of seismic or mechanical thicknesses on regional length scales. The geodynamically sheltered central location of the Australian Proterozoic [*Simons et al.*, 1999] or the geometry of orogens [*Lenardic et al.*, 2000] may be responsible in maintaining a thick tectosphere under Proterozoic Australia, which is sometimes thicker than under the Archean. Convection processes causing the break-up of Antarctica/Australia from India by 100 Ma [*Storey*, 1995] may have removed parts or all of the western Archean tectosphere. Furthermore, the breakdown of the correlation between thickness and age occurs at scale lengths compatible with lithospheric instabilities[*Conrad and Molnar*, 1997]. Alternatively, not all Archean crust may have developed a strong and thick boundary layer due to intrinsic compositional differences such as have been identified in the southwestern United States [*Lee et al.*, 2001].

5.6 Conclusion

Using the seismic wave speed model of *Simons et al.* [2002a] we have estimated the depth to the base of the high velocity lid underlying the Australian continent. We compared this seismic thickness to an estimate of the elastic thickness from isotropic coherence analysis using the method of *Simons et al.* [2002b]. On the scales resolved by our model, the variations in seismic thickness with crustal age within broad subdivisions of the Australian continent are larger than the differences between them. To first order, the mechanical strength of the lithosphere increases slightly with age, and the (Archean) Yilgarn and Gawler blocks as well as the (Proterozoic) North Australian craton stand out as mechanically strong. There are substantial strength differences within domains of equal crustal age, however. The seismic thickness of the lithosphere, which represents the short-term elastic response, follows a different pattern and does not seem to be a good proxy for its long-term mechanical properties. Our detailed study of Australia suggests that the notion that old cratons are both mechanically strong and have thick seismic lids is an inaccurate oversimplification.

Acknowledgments

We thank Alet Zielhuis for many years of collaboration on the surface wave inversions, and Maria Zuber for stimulating our interest in problems of coherence and elastic thickness. This work was supported by NSF grant EAR-0001136.

Chapter 6

Structure and Deformation of Australia[1]

Abstract

A high-resolution tomographic model of the azimuthally anisotropic wave speed structure of the Australian upper mantle is jointly interpreted with information on lithospheric deformation from the orientation of gravity anomalies in relation to topography. The use of surface waves tightly constrains the depth variation of seismic anisotropy which provides an advantage over the much used *SKS* waves, which accumulate splitting delay times anywhere on their near-vertical propagation path. Since it reflects the degree and mechanism of isostatic compensation, the spectral coherence between gravity anomalies and topography provides better constraints on the fossilized strain field averaged over depth and integrated over time than geological trend indicators observed at the surface. The relationship between the long-term elastic and the instantaneous seismic anisotropy above 200 km is consistent with a model in which seismic anisotropy has developed perpendicular to the direction in which the lithosphere deforms more easily under compression, whereas the fast axes of seismic anisotropy at depths greater than 200 km are predominantly aligned in the direction of current absolute plate motion. That the fossil strain field preserved in the gravity field bears a relation to the seismic anisotropy suggests that past deformation episodes have coherently affected the upper 200 km of the lithosphere, whereas the cause for seismic anisotropy below 200 km may lie in present-day asthenospheric flow.

---

[1] **Structure and Deformation of the Australian Lithosphere**, *manuscript in preparation.*

Whether the origin of continental seismic anisotropy lies in contemporaneous mantle deformation resulting from current plate motion or is due to fossil strain from wholesale deformation processes is actively debated [*Mooney*, 1995]. In the first end-member scenario [*Vinnik et al.*, 1995] the fast axis of anisotropy recorded at the surface is interpreted to be aligned with the texture that forms from shearing of the asthenosphere [*Mainprice and Silver*, 1993; *Tommasi*, 1998; *Fouch et al.*, 2000] below a lithospheric plate that moves with respect to a deeper mantle which is presumed to be stationary. In the second view [*Silver and Chan*, 1988] intraplate deformation affects the crust and the lithospheric mantle coherently. The strain-induced lattice-preferred orientation of anisotropic mantle minerals causes the fast axis of seismic anisotropy to align with the axis of greatest horizontal elongation or perpendicular to the axis of greatest horizontal shortening [*Ribe*, 1992; *Mainprice and Silver*, 1993]. Neglecting the effects of recrystallization [*Zhang and Karato*, 1995], the fast axis will be parallel to the tectonic trends, for both transcurrent and transpressional features [*Silver*, 1996; *Tommasi et al.*, 1999]. Weak zones, large-scale faults or pre-existing fabric may cause the lithosphere to display mechanically anisotropic behavior [*Vauchez et al.*, 1998; *Tommasi and Vauchez*, 2001]. If there exists a direction in which strain is preferentially accommodated the seismic fast axis will be perpendicular to the azimuth in which the lithosphere appears the weakest. Interpretations advocating a combination of asthenospheric and frozen lithospheric anisotropy [*Kubo and Hiramatsu*, 1998] can be explained with the presence of a rheological boundary which is seismically identified as the Lehmann discontinuity [*Karato*, 1992; *Gaherty and Jordan*, 1995; *Gu et al.*, 2001].

Much of the ongoing debate is based on the pattern of shear-wave splitting observed at the surface. The commonly studied *SKS* phases leave the fluid outer core radially polarized and are split by an anisotropic layer, whereby the component polarized parallel to the fast direction accumulates a delay time over the orthogonally polarized component [*Savage*, 1999]. Traditionally, maps of shear wave splitting polarization directions are compared

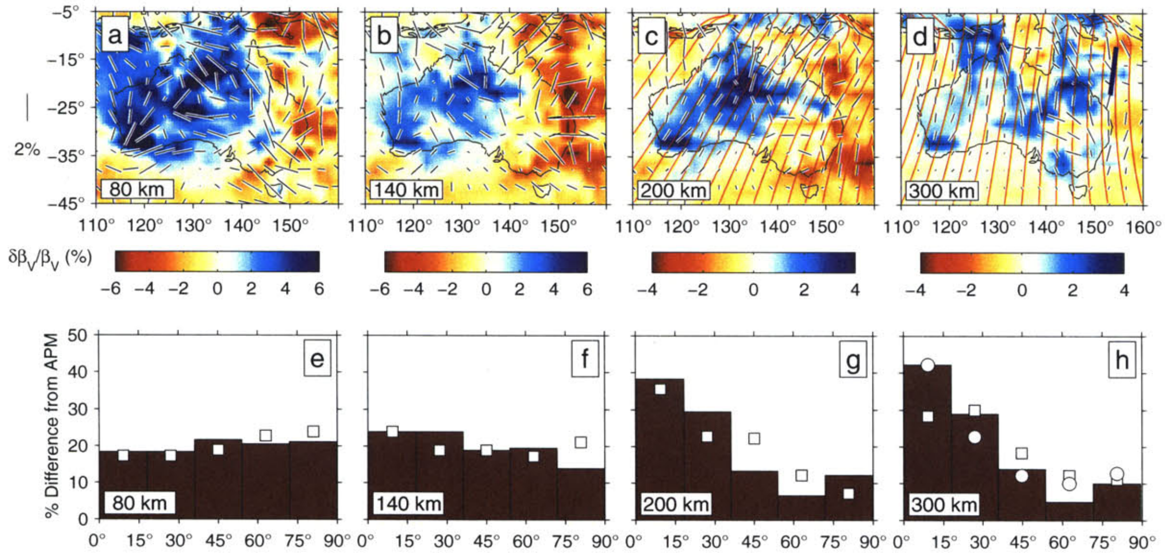


Figure 6-1: Wave speed structure and azimuthally anisotropic fabric of the Australian lithosphere (a–d) and histograms of the difference of the seismic fast axes (e–h) with respect to the direction of absolute plate motion in a hot-spot reference frame [*Gripp and Gordon*, 1990] (solid line; solid bars); a no-net-rotation frame [*Argus and Gordon*, 1991] (dashed line; filled squares) and our best estimate at 300 km (blue solid line; filled circles).

to the direction of absolute plate motion [*Vinnik et al.*, 1992; *Barruol et al.*, 1997], to the tectonic trends observed at the surface [*Silver and Chan*, 1991; *Silver et al.*, 2001] or to directions inferred from forward models of finite deformation [*Griot and Montagner*, 1998; *Davis et al.*, 1997; *Fouch et al.*, 2000] to argue the validity of either end-member case. These comparisons have inherent limitations. First, one has to assume that the fast polarization directions obtained from shear-wave splitting are representative of the narrow depth range of the upper mantle of interest. Significant delay times likely originate in the interval between 40 and 400 km [*Mainprice and Silver*, 1993; *Silver*, 1996] but with vertically propagating waves it is difficult to constrain this depth range more precisely [*Girardin and Farra*, 1998; *Silver and Savage*, 1994]. Second, surface features such as fold axes, foliation planes, and lineations are a poor proxy for the deformation history that could affect the entire lithosphere, as deformation need not be uniform with depth [*Goleby et al.*, 1989; *Lucas et al.*, 1993].

Our work addresses both concerns. First, we investigate the depth dependence of anisotropy of the lithosphere with horizontally propagating seismic surface waves. This allows us to identify the depth at which the alignment of seismic anisotropy with the direction of absolute plate motion becomes the dominant feature in the model. Second, we propose to identify the time- and depth-integrated dominant direction of strain accumulation from the resulting gravity anomalies, rather than from surface strain indicators. Instead of testing the parallelism of seismic anisotropy with tectonic trends we test to what extent the seismic fast axis is indeed perpendicular to the azimuth in which the lithosphere has accumulated the most gravity anomalies for a given amount of topography. This direction in which the lithosphere appears weaker than the isotropic average is revealed from the two-dimensional isostatic response characterized by the coherence between the Bouguer gravity anomalies and topography [*Forsyth*, 1985]. We developed a new method of spectral analysis [*Simons et al.*, 2002b] in order to extract the directional dependence of the isostatic compensation.

Fig. 6-1(a–d) shows a tomographic model of isotropic shear-wave speed variations and azimuthal anisotropy of the Australian lithosphere obtained by us [*Simons et al.*, 2002a] from the inversion of a large set of surface-wave waveforms using partitioned waveform inversion [*Nolet*, 1990] in combination with a tomographic inversion for azimuthal anisotropy [*Smith and Dahlen*, 1973]. We have performed numerous tests on the robustness, resolution and variance of our model [*Simons et al.*, 2002a] and in particular, on the necessity of including anisotropy to better explain the data. The large-scale structure of our model is in good agreement with other studies [*Debayle and Kennett*, 2000a; *Montagner and Tanimoto*, 1991]. In the top 200 km of the lithosphere the anisotropy is strong and varies

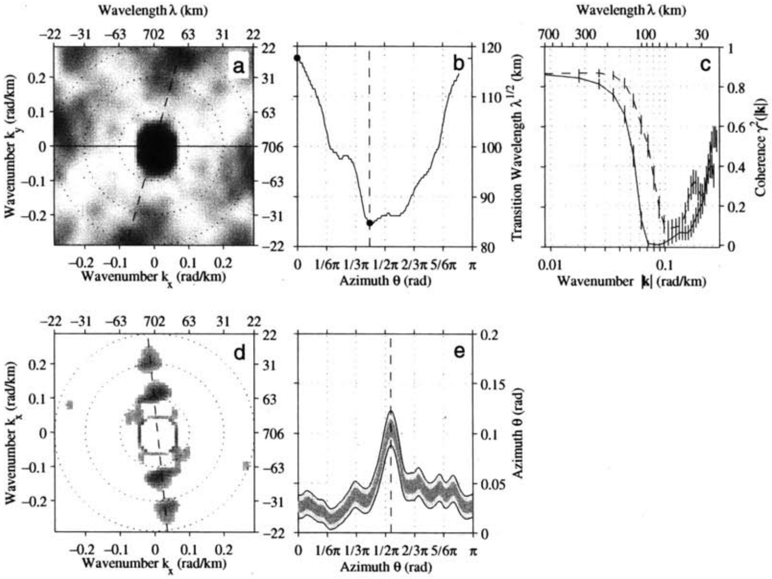


Figure 6-2: Detection of long- and short-wavelength coherence anisotropy. (a–c) Two-dimensional coherence plot (a) and the evolution of the transition wavelength with azimuth (b). Coherence of the mechanically weakest (dashed) and strongest (solid) direction, with 1σ error bars (c). (d—e) Two-dimensional coherence of a different region, with the long-wavelength portion not displayed (d) and the weak direction (dashed) identified from the peak in the average coherence in function of the azimuth (e). Shaded regions indicate 99% and 95% confidence intervals on the average coherence.

on short length scales, whereas in the bottom 200 km of the model a smoother pattern of slightly weaker anisotropy is observed. Absolute plate motions were calculated in a hotspot reference frame [*Gripp and Gordon*, 1990] and in a no-net-rotation frame [*Argus and Gordon*, 1991]. With increasing depth, a progressively larger proportion of anisotropic fast directions trend with the direction of absolute plate motion (Fig. 6-1(e–f)). Fig. 6-1 also shows the direction of absolute plate motion most compatible with our seismic model at 300 km depth. In profile [*Fischer and van der Hilst*, 1999; *Simons et al.*, 1999, 2002a], the Australian subcontinental lithospheric keel is characterized by high-velocity anomalies that extend to a depth of 250-300 km. The pattern of North-South directions of anisotropy

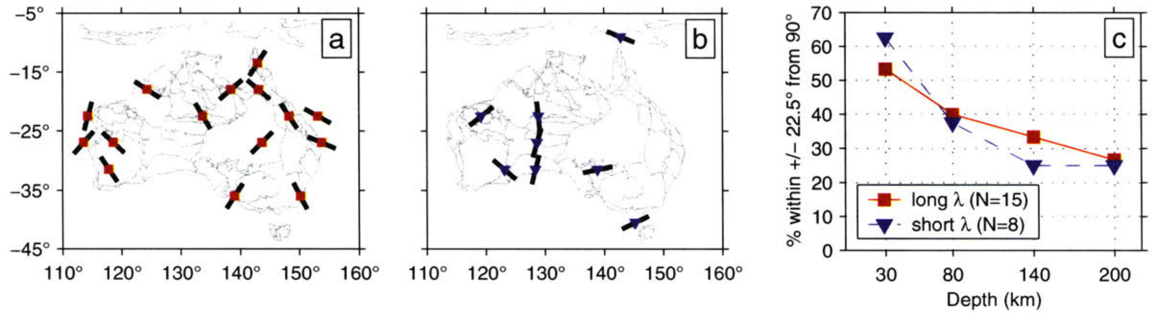


Figure 6-3: Mechanical anisotropy of Australia and the distribution of weak directions perpendicular to the fast axis of anisotropy. (a) Weak directions based on the long-wavelength (as in Fig. 6-2a–c) and (b) short-wavelength portions of the coherence (as in Fig. 6-2d–e). (c) Proportion of mechanically weak directions perpendicular to the fast seismic direction at the same location as a function of depth.

observed at 300 km is consistent with the view that the bottom of the lithosphere is indeed being deformed by the actively flowing mantle. In how far this might affect the stability of the lithosphere is at present unexplored.

Theoretically, the isotropic coherence or transfer function between Bouguer gravity anomalies and topography drops off with decreasing wavelength. The wavelength at which the coherence reaches half of its long-wavelength maximum varies positively, but nonlinearly, with the apparent strength of the plate [*Forsyth*, 1985]. The two-dimensional coherence between gravity and topography resolves directions in which isostatic compensation preferentially accommodates continental deformation, expressed in the gravity anomalies. We identify the direction of mechanical weakness from the azimuth with the shortest coherence transition wavelength (Fig. 6-2a–c). In the wavelength range below 150 km, we further identify weak directions of preferential isostatic compensation by their anomalously high coherence with respect to the isotropic average (Fig. 6-2d–e). The accuracy of our spectral estimates is guaranteed by our use of a multitaper spectrogram technique with windowing functions [*Daubechies*, 1988] optimized to capture the spatial variation of coherence anisotropy [*Simons et al.*, 2002b], and the stability of the coherence was tested under random rotation of the data set.

A comparison of the angle between the mechanically weak and seismically fast directions shows that about 60% of the angle differences cluster around 90° at shallow depths, and that the degree of this correspondence drops off with increasing depth (Fig. 6-3). At a depth of 200 km, not significantly more weak directions are perpendicular to the fast polarization direction than are parallel to it.

We have proposed a means to relate continental deformation with the development of seismic anisotropy and distinguish between the influence of past and present lithospheric deformation. Our measurements suggests that the top 200 km of the subcontinental lithospheric mantle under Australia is affected by vertically coherent deformation, whereas in the bottom 200 km active processes related to current plate motion can explain the observed anisotropy.

Acknowledgments

We thank Oded Aharonson, Thorsten Becker, Alet Zielhuis and Maria Zuber for valuable discussions.

Bibliography

Abers, G. A., and S. W. Roecker, Deep structure of an arc-continent collision: Earthquake relocation and inversion for upper mantle *P* and *S* wave velocities beneath Papua New Guinea, *J. Geophys. Res.*, *96*, 6379–6401, 1991.

Aki, K., and P. G. Richards, *Quantitative Seismology*, Freeman, San Francisco, Calif., 1980.

Anderson, D. L., Elastic wave propagation in layered anisotropic media, *J. Geophys. Res.*, *66*, 2953–2963, 1961.

Anderson, D. L., and J. D. Bass, Mineralogy and composition of the uppper mantle, *Geophys. Res. Lett.*, *11*, 637–640, 1984.

Anderson, D. L., and A. M. Dziewonski, Upper mantle anisotropy: Evidence from free oscillations, *Geophys. J. R. Astron. Soc.*, *69*, 383–404, 1982.

Argus, D. F., and R. G. Gordon, No-net-rotation model of current plate velocities incorporating plate motion model Nuvel-1, *Geophys. Res. Lett.*, *18*, 2039–2042, 1991.

Artemieva, I. M., and W. D. Mooney, Thermal thickness and evolution of Precambrian lithosphere: A global study, *J. Geophys. Res.*, *106*, 163,387–16,414, 2001.

Artyushkov, E. V., On the origin of the seismic anisotropy of the lithosphere, *Geophys. J. R. Astron. Soc.*, *76*, 173–178, 1984.

Babuška, V., J.-P. Montagner, J. Plomerová, and N. Girardin, Age-dependent large-scale fabric of the mantle lithosphere as derived from surface-wave velocity anisotropy, *Pure Appl. Geophys.*, *151*, 257–280, 1998.

Banks, R. J., S. C. Francis, and R. G. Hipkin, Effects of loads in the upper crust on estimates of the elastic thickness of the lithosphere, *Geophys. J. Int.*, *145*, 291–299, 2001.

Bannister, S. C., B. Ruud, and E. S. Husebye, Tomographic estimates of sub-Moho seismic velocities in Fennoscandia and structural implications, *Tectonophysics*, *189*, 37–53, 1991.

Barruol, G., and R. Hoffman, Upper mantle anisotropy beneath the GEOSCOPE stations, *J. Geophys. Res.*, *104*, 10,757–10,773, 1999.

Barruol, G., G. Helffrich, and A. Vauchez, Shear wave splitting around the northern Atlantic: Frozen Pangaean lithospheric anisotropy?, *Tectonophysics*, *279*, 135–148, 1997.

Barth, N., and C. Wunsch, Oceanographic experiment design by simulated annealing, *J. Phys. Oceanography*, *20*, 1249–1263, 1990.

Bayram, M., and R. G. Baraniuk, Multiple window time-frequency analysis, in *Proc. IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, pp. 173–176, IEEE, 1996.

Bear, L., and G. Pavlis, Multi-channel estimation of time residuals from broadband seismic data using multi-wavelets, *Bull. Seism. Soc. Am.*, *89*, 681–692, 1999.

Bear, L. K., and G. Pavlis, Estimation of slowness vectors and their uncertainties using multi-wavelet seismic array processing, *Bull. Seism. Soc. Am.*, *87*, 755–769, 1997.

Bechtel, T. D., Mechanisms of isostatic compensation in East Africa and North America, Ph.D. thesis, Brown Univ., Providence, R. I., 1989.

Bechtel, T. D., D. W. Forsyth, and C. J. Swain, Mechanisms of isostatic compensation in the vicinity of the East African Rift, Kenya, *Geophys. J. R. Astron. Soc.*, *90*, 445–465, 1987.

Bechtel, T. D., D. W. Forsyth, V. L. Sharpton, and R. A. F. Grieve, Variations in effective elastic thickness of the North-American lithosphere, *Nature*, *343*, 636–638, 1990.

Bendat, J. S., Statistical errors in measurement of coherence functions and input/output quantities, *J. Sound Vibration*, *59*, 405–421, 1978.

Bendat, J. S., and A. G. Piersol, *Measurement and Analysis of Random Data*, 2nd ed., John Wiley, New York, 1986.

Bendat, J. S., and A. G. Piersol, *Engineering Applications of Correlation and Spectral Analysis*, 2nd ed., John Wiley, New York, 1993.

Bendat, J. S., and A. G. Piersol, *Random data: Analysis and Measurement Procedures*, 3rd ed., John Wiley, New York, 2000.

Bergeron, S. Y., D. A. Yuen, and A. P. Vincent, Capabilities of 3-D wavelet transforms to detect plume-like structures from seismic tomography, *Geophys. Res. Lett.*, *27*, 3433–3436, 2000.

Blakely, R. J., *Potential Theory in Gravity and Magnetic Applications*, Cambridge Univ. Press, New York, 1995.

Bostock, M., Seismic imaging of lithospheric discontinuities and continental evolution, *Lithos*, *48*, 1–16, 1999.

Boyd, F. R., Compositional distinction between oceanic and cratonic lithosphere, *Earth Planet. Sci. Lett.*, *96*, 15–26, 1989.

Boyd, F. R., and J. J. Gurney, Diamonds and the African lithosphere, *Science*, *232*, 472–477, 1986.

Braun, J., J. C. Dooley, B. Goleby, R. D. van der Hilst, and C. Klootwijk (Eds.), *Structure and Evolution of the Australian Continent*, vol. 26 of *Geodyn. Ser.*, AGU, Washington, D. C., 1998.

Bronez, T. P., On the performance advantage of multitaper spectral analysis, *IEEE Trans. Signal Process.*, *40*, 2941–2946, 1992.

Brune, J. N., Travel times, body waves, and normal modes of the Earth, *Bull. Seism. Soc. Am.*, *54*, 2099–2128, 1964.

Burg, J., Maximum Entropy Spectral Analysis, Ph.D. thesis, Stanford Univ., Stanford, Calif., 1975.

Burke, K., and W. Kind, Were Archean continental geothermal gradients much steeper than those of today?, *Nature*, *272*, 240–241, 1978.

Burov, E. B., and M. Diament, The effective elastic thickness (T_e) of continental lithosphere: What does it really mean?, *J. Geophys. Res.*, *100*, 3905–3927, 1995.

Burov, E. B., C. Jaupart, and J.-C. Mareschal, Large-scale crustal heterogeneities and lithospheric strength in cratons, *Earth Planet. Sci. Lett.*, *164*, 205–219, 1998.

Cara, M., Lateral variations of S-velocity in the upper mantle from higher Rayleigh modes, *Geophys. J. R. Astron. Soc.*, *57*, 649–670, 1979.

Cara, M., and J.-J. Lévêque, Waveform inversion using secondary observables, *Geophys. Res. Lett.*, *14*, 1046–1049, 1987.

Carter, G. C., C. H. Knapp, and A. H. Nuttal, Statistics of the estimate of the magnitude-coherence function, *IEEE Trans. Audio Electroacoust.*, *AU21*, 388–389, 1973.

Chamberlain, K. R., S. C. Patel, B. R. Frost, and G. L. Snyder, Thick-skinned deformation of the Archean Wyoming province during Proterozoic arc-continent collision, *Geology*, *21*, 995–998, 1993.

Chave, A. D., D. J. Thomson, and M. E. Ander, On the robust estimation of power spectra, coherences, and transfer functions, *J. Geophys. Res.*, *92*, 633–648, 1987.

Chen, W.-P., and S. Özalaybey, Correlation between seismic anisotropy and Bouguer anomalies in Tibet and its implications for lithospheric structures, *Geophys. J. Int.*, *135*, 93–101, 1998.

Chevrot, S., J.-P. Montagner, and R. K. Snieder, The spectrum of tomographic Earth models: Correction, *Geophys. J. Int.*, *135*, 311, 1998.

Cichowicz, A., and R. Green, Tomographic study of upper-mantle structure of the South-African continent, using wave-form inversion, *Phys. Earth Planet. Inter.*, *72*, 276–285, 1992.

Clitheroe, G., and R. D. van der Hilst, Complex anisotropy in the Australian lithosphere from shear-wave splitting in broad-band *SKS*-records, in *Structure and Evolution of the Australian Continent*, edited by J. Braun, J. C. Dooley, B. Goleby, R. D. van der Hilst, and C. Klootwijk, vol. 26 of *Geodyn. Ser.*, pp. 73–78, AGU, Washington, D. C., 1998.

Clitheroe, G., Ó. Guðmundsson, and B. L. N. Kennett, The crustal thickness of Australia, *J. Geophys. Res.*, *105*, 13,697–13,713, 2000a.

Clitheroe, G., Ó. Guðmundsson, and B. L. N. Kennett, Sedimentary and upper crustal structure of Australia from receiver functions, *Aust. J. Earth Sci.*, *47*, 209–216, 2000b.

Cloetingh, S., E. Burov, and A. Poliakov, Lithosphere folding: Primary response to compression? (from central Asia to Paris basin, *Tectonics*, *18*, 1064–1083, 1999.

Coblentz, D. D., M. Sandiford, R. M. Richardson, S. Zhou, and R. Hillis, The origins of the intraplate stress field in continental Australia, *Earth Planet. Sci. Lett.*, *133*, 299–309, 1995.

Coblentz, D. D., S. Zhou, R. R. Hillis, R. M. Richardson, and M. Sandiford, Topography, boundary forces, and the Indo-Australian intraplate stress field, *J. Geophys. Res.*, *103*, 919–931, 1998.

Cohen, L., Time-frequency distributions — A review, *Proc. IEEE*, *77*, 941–981, 1989.

Conrad, C. P., and P. Molnar, The growth of Rayleigh-Taylor-type instabilities in the lithosphere for various rheological and density structures, *Geophys. J. Int.*, *129*, 95–112, 1997.

Constable, S. C., R. L. Parker, and C. G. Constable, Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data, *Geophysics*, *52*, 289–300, 1987.

Cull, J., Heat flow and regional geophysics in Australia, in *Terrestrial Heat Flow and the Lithosphere Structure*, edited by V. Cermak and L. Rybach, pp. 486–500, Springer-Verlag, New York, 1991.

Cull, J. P., and D. Denham, Regional variations in Australian heat flow, *BMR J. Aust. Geol. Geophys.*, *4*, 1–13, 1979.

Curtis, A., and R. Snieder, Reconditioning inverse problems using the genetic algorithm and revised parameterization, *Geophysics*, *62*, 1524–1532, 1997.

D'Agostino, N., and D. McKenzie, Convective support of long-wavelength topography in the Apennines (Italy), *Terra Nova*, *11*, 234–238, 1999.

Dahlen, F. A., Models of the lateral heterogeneity of the Earth consistent with eigenfrequency splitting data, *Geophys. J. R. Astron. Soc.*, *44*, 77–105, 1976.

Dahlen, F. A., and J. Tromp, *Theoretical Global Seismology*, Princeton Univ. Press, Princeton, N. J., 1998.

Daubechies, I., Time-frequency localization operators: A geometric phase space approach, *IEEE Trans. Inform. Theory*, *34*, 605–612, 1988.

Daubechies, I., and T. Paul, Time-frequency localization operators: A geometric phase space approach: II. The use of dilations, *Inv. Probl.*, *4*, 661–680, 1988.

Davis, P., P. England, and G. Houseman, Comparison of shear wave splitting and finite strain from the India-Asia collision zone, *J. Geophys. Res.*, *102*, 27,511–27,522, 1997.

De Smet, J. H., A. P. van den Berg, and N. J. Vlaar, The evolution of continental roots in numerical thermo-chemical mantle convection models including differentiation by partial melting, *Lithos*, *48*, 153–170, 1999.

Debayle, E., *SV*-wave azimuthal anisotropy in the Australian upper mantle: Preliminary results from automated Rayleigh waveform inversion, *Geophys. J. Int.*, *137*, 747–754, 1999.

Debayle, E., and B. L. N. Kennett, The Australian continental upper mantle: Structure and deformation inferred from surface waves, *J. Geophys. Res.*, *105*, 25,423–25,450, 2000a.

Debayle, E., and B. L. N. Kennett, Anisotropy in the Australasian upper mantle from Love and Rayleigh waveform inversion, *Earth Planet. Sci. Lett.*, *184*, 339–351, 2000b.

DeMets, C., R. G. Gordon, D. F. Argus, and S. Stein, Current plate motions, *Geophys. J. Int.*, *101*, 425–478, 1990.

Detrick, R. S., and A. B. Watts, An analysis of isostasy in the world's oceans, 3, Aseismic ridges, *J. Geophys. Res.*, *84*, 3637–3653, 1979.

Dickey, J. O., et al., Satellite gravity: Insights into the solid Earth and its fluid envelope, *Eos Trans. AGU*, *79*, 237, 242–243, 1998.

Dooley, J. C., Implications of Australian seismic and gravity measurements for the structure and composition of the upper mantle, *BMR J. Aust. Geol. Geophys.*, *2*, 1–5, 1977.

Dorman, L. M., and B. T. Lewis, Experimental isostasy, 1, Theory of the determination of the Earth's isostatic response to a concentrated load, *J. Geophys. Res.*, *75*, 3357–3365, 1970.

Draper, N., and H. Smith, *Applied Regression Analysis*, 2nd ed., John Wiley, New York, 1981.

Drummond, B. (Ed.), *The Australian Lithosphere*, Geol. Soc. Aust., Sydney, 1991.

Dziewonski, A. M., and D. L. Anderson, Preliminary Reference Earth Model, *Phys. Earth Planet. Inter.*, *25*, 297–356, 1981.

Dziewonski, A. M., T.-A. Chou, and J. H. Woodhouse, Determination of earthquake source parameter from waveform data for studies of global and regional seismicity, *J. Geophys. Res.*, *86*, 2825–2852, 1981.

Ebinger, C., and N. Hayward, Soft plates and hot spots: Views from Afar, *J. Geophys. Res.*, *101*, 21,859–21,876, 1996.

Eggler, D. H., J. K. Keen, F. Welt, F. O. Dudas, K. P. Furlong, M. E. McCallum, and R. W. Carlson, Tectonomagmatism of the Wyoming Province, *Colorado School of Mines Quarterly*, *83*, 25–40, 1988.

Ekström, G., J. Tromp, and E. W. F. Larson, Measurements and global models of surface wave propagation, *J. Geophys. Res.*, *102*, 8137–8157, 1997.

Engdahl, E., R. D. van der Hilst, and R. Buland, Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, *Bull. Seism. Soc. Am.*, *88*, 722–743, 1998.

Escartin, J., and J. Lin, Tectonic modification of axial crustal structure: Evidence from spectral analyses of residual gravity and bathymetry of the Mid-Atlantic Ridge flanks, *Earth Planet. Sci. Lett.*, *154*, 279–293, 1998.

Estey, L. H., and B. J. Douglas, Upper mantle anisotropy: A preliminary model, *J. Geophys. Res.*, *91*, 11,393–11,406, 1986.

Finlayson, D. M., Geophysical differences in the lithosphere between Phanerozoic and Precambrian Australia, *Tectonophysics*, *84*, 287–312, 1982.

Fischer, K. M., and R. D. van der Hilst, A seismic look under the continents, *Science*, *285*, 1365–136, 1999.

Flandrin, P., Maximum signal energy concentration in a time-frequency domain, in *Proc. IEEE Int. Conf. Acoust. Speech Signal Proc.*, pp. 2176–2179, IEEE, 1988.

Flandrin, P., *Temps-Fréquence*, 2 ed., Hermès, Paris, 1998.

Forsyth, D. W., The early structural evolution and anisotropy of the oceanic upper mantle, *Geophys. J. R. Astron. Soc.*, *43*, 103–162, 1975.

Forsyth, D. W., Subsurface loading and estimates of the flexural rigidity of continental lithosphere, *J. Geophys. Res.*, *90*, 12,623–12,632, 1985.

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

Foufoula-Georgiou, E., and P. Kumar (Eds.), *Wavelets in Geophysics*, Academic Press, San Diego, Calif., 1994.

Fowler, C., *The Solid Earth*, Cambridge Univ. Press, New York, 1990.

Frazer, G., and B. Boashash, Multiple window spectrogram and time-frequency distributions, in *Proc. IEEE Int. Conf. Acoust. Speech Signal Proc.*, vol. IV, pp. 293–296, IEEE, 1994.

Frederiksen, A. W., M. G. Bostock, and J. F. Cassidy, *S*-wave velocity structure of the Canadian upper mantle, *Phys. Earth Planet. Inter.*, *124*, 175–191, 2001.

Freybourger, M., J. B. Gaherty, and T. H. Jordan, Structure of the Kaapvaal craton from surface waves, *Geophys. Res. Lett.*, *28*, 2489–2492, 2001.

Gaherty, J. B., and T. H. Jordan, Lehmann discontinuity as the base of an anisotropic layer beneath continents, *Science*, *268*, 1468–1471, 1995.

Gaherty, J. B., T. H. Jordan, and L. S. Gee, Seismic structure of the upper mantle in a central Pacific corridor, *J. Geophys. Res.*, *101*, 22,291–22,309, 1996.

Gesch, D., K. L. Verdin, and S. K. Greenlee, New land surface digital elevation model covers the Earth, *Eos Trans. AGU*, *80*, 69–70, 1999.

Girardin, N., and V. Farra, Azimuthal anisotropy in the upper mantle from observations of *P*-to-*S* converted phases: Application to southeast Australia, *Geophys. J. Int.*, *133*, 615–629, 1998.

Goff, J. A., and T. H. Jordan, Stochastic modeling of seafloor morphology: Inversion of sea beam data for second-order statistics, *J. Geophys. Res.*, *93*, 13,589–13,608, 1988.

Goleby, B. R., R. D. Shaw, C. Wright, and B. L. N. Kennett, Geophysical evidence for "thick-skinned" crustal deformation in central Australia, *Nature*, *337*, 325–330, 1989.

Goncharov, A. G., M. D. Lizinsky, C. D. N. Collins, K. A. Kalnin, T. N. Fomin, B. J. Drummond, B. R. Goleby, and L. N. Platonenkova, Intra-crustal "seismic isostasy" in the Baltic Shield and Australian Precambrian cratons from deep seismic profiles and the Kola Superdeep Bore Hole data, in *Structure and Evolution of the Australian Continent*, edited by J. Braun, J. C. Dooley, B. Goleby, R. D. van der Hilst, and C. Klootwijk, vol. 26 of *Geodyn. Ser.*, pp. 119–138, AGU, Washington, D. C., 1998.

Goncz, J., and J. R. Cleary, Variations in the structure of the upper mantle beneath Australia, from Rayleigh wave observations, *Geophys. J. R. Astron. Soc.*, *44*, 507–516, 1976.

Goncz, J. H., A. L. Hales, and K. J. Muirhead, Analysis to extended periods of Rayleigh and Love wave dispersion across Australia, *Geophys. J. R. Astron. Soc.*, *41*, 81–108, 1975.

Gough, D. I., Electrical conductivity under western North America in relation to heat flow, seismology and structure, *J. Geomagn. Geoelectr.*, *26*, 105–123, 1974.

Griffin, W. L., S. Y. O'Reilly, C. G. Ryan, O. Gaul, and D. A. Ionov, Secular variation in the composition of subcontinental lithospheric mantle: Geophysical and geodynamic implications, in *Structure and Evolution of the Australian Continent*, edited by J. Braun, J. C. Dooley, B. Goleby, R. D. van der Hilst, and C. Klootwijk, no. 26 in Geodyn. Ser., pp. 1–26, AGU, Washington, D. C., 1998.

Griffin, W. L., S. Y. O'Reilly, and C. G. Ryan, The composition and origin of subcontinental lithospheric mantle, in *Mantle Petrology: Field Observations and High-Pressure Experimentation*, edited by Y. Fei, C. M. Bertka, and B. O. Mysen, 6, pp. 13–43, The Geochemical Society Special Publication,, 1999.

Griot, D.-A., and J.-P. Montagner, Confrontation of mantle seismic anisotropy with two extreme models of strain, in Central Asia, *Geophys. Res. Lett.*, *25*, 1447–1450, 1998.

Griot, D.-A., J.-P. Montagner, and P. Tapponnier, Phase velocity structure from Rayleigh and Love waves in Tibet and its neighboring regions, *J. Geophys. Res.*, *103*, 21,215–21,232, 1998.

Gripp, A. E., and R. G. Gordon, Current plate velocities relative to the hotspots incorporating the Nuvel-1 global plate motion model, *Geophys. Res. Lett.*, *17*, 1109–1112, 1990.

Gu, Y. J., A. M. Dziewonski, and G. Ekström, Preferential detection of the Lehmann discontinuity beneath continents, *Geophys. Res. Lett.*, *28*, 4655–4658, 2001.

Guarnieri, A. M., and C. Prati, SAR interferometry: A "quick and dirty" coherence estimator for data browsing, *IEEE Trans. Geosci. Remote Sens.*, *35*, 660–669, 1997.

Guðmundsson, Ó., and M. Sambridge, A regionalized upper mantle (RUM) seismic model, *J. Geophys. Res.*, *103*, 7121–7136, 1998.

Gurnis, M., Large-scale mantle convection and the aggregation and dispersal of supercontinents, *Nature*, *332*, 695–699, 1988.

Gurnis, M., R. D. Müller, and L. Moresi, Cretaceous vertical motion of Australia and the Australian-Antarctic discordance, *Science*, *279*, 1499–1504, 1998.

Gwavava, O., C. J. Swain, and F. Podmore, Mechanisms of isostatic compensation of the Zimbabwe and Kaapvaal cratons, the Limpopo belt and the Mozambique basin, *Geophys. J. Int.*, *127*, 635–650, 1996.

Hager, B. H., Subducted slabs and the geoid: Constraints on mantle rheology and flow, *J. Geophys. Res.*, *89*, 6003–6015, 1984.

Hand, M., and M. Sandiford, Intraplate deformation in central Australia, the link between subsidence and fault reactivation, *Tectonophysics*, *305*, 121–140, 1999.

Handler, M. R., V. Bennett, and T. Esat, The persistence of off-cratonic lithospheric mantle: Os isotopic systematics of variably metasomatised southeast Australian xenoliths, *Earth Planet. Sci. Lett.*, *151*, 61–75, 1997.

Hanssen, A., Multidimensional multitaper spectral estimation, *Signal Process.*, *58*, 327–332, 1997.

Herman, G. T. (Ed.), *Image Reconstruction from Projections*, vol. 32 of *Topics in applied physics*, Springer-Verlag, New York, 1979.

Hess, H. H., Seismic anisotropy of the uppermost mantle under oceans, *Nature*, *203*, 629–630, 1964.

Hillis, R. R., J. J. Meyer, and S. D. Reynolds, The Australian stress map, *Explor. Geophys.*, *29*, 420–427, 1998.

Hillis, R. R., J. R. Enever, and S. D. Reynolds, In situ stress field of eastern Australia, *Aust. J. Earth Sci.*, *46*, 813–825, 1999.

Ho-Liu, P., J.-P. Montagner, and H. Kanamori, Comparison of iterative back-projection inversion and generalized inversion without blocks: Case studies in attenuation tomography, *Geophys. J. Int.*, *97*, 19–29, 1989.

Humphreys, E. D., and R. Clayton, Adaptation of back projection tomography to seismic travel time problems, *J. Geophys. Res.*, *93*, 1073–1085, 1988.

Iidaka, T., R. D. van der Hilst, and G. Clitheroe, Regional variation of the shear-wave splitting in the central part of Australia, *Phys. Earth Planet. Inter.*, 2001, manuscript in preparation.

Ito, G., and A. Taira, Compensation of the Ontong Java Plateau by surface and subsurface loading, *J. Geophys. Res.*, *105*, 11,171–11,183, 2000.

Jackson, D. D., The use of a priori data to resolve non-uniqueness in linear inversion, *Geophys. J. R. Astron. Soc.*, *57*, 137–157, 1979.

Jacob, D. E., K. S. Viljoen, N. Grassineau, and E. Jagoutz, Remobilization in the cratonic lithosphere recorded in polycrystalline diamond, *Science*, *289*, 1182–1185, 2000.

James, D. E., M. J. Fouch, J. C. VanDecar, S. van der Lee, and Kaapvaal Seismic Group, Tectospheric structure beneath southern Africa, *Geophys. Res. Lett.*, *28*, 2485–2488, 2001.

Jaupart, C., and J.-C. Mareschal, Thermal structure and thickness of continental roots, *Lithos*, *48*, 93–114, 1999.

Jeans, J., The propagation of earthquake waves, *Philos. Trans. R. Soc. London, Ser. A*, *102*, 554–574, 1923.

Jeong, J., and W. J. Williams, Alias-free generalized discrete-time time-frequency distributions, *IEEE Trans. Signal Process.*, *40*, 2757–2765, 1992.

Johnson, R. W. (Ed.), *Intraplate Volcanism in Eastern Australia and New Zealand*, Cambridge Univ. Press, New York, 1989.

Jordan, T. H., Lateral heterogeneity and mantle dynamics, *Nature*, *257*, 745–750, 1975a.

Jordan, T. H., The continental tectosphere, *Rev. Geophys.*, *13*, 1–12, 1975b.

Jordan, T. H., Composition and development of the continental tectosphere, *Nature*, *274*, 544–548, 1978.

Jordan, T. H., Continents as a chemical boundary layer, *Philos. Trans. R. Soc. London, Ser. A*, *301*, 359–373, 1981a.

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

Jordan, T. H., Structure and formation of the continental tectosphere, *J. Petrol.*, *Special Lithosphere Issue*, 11–37, 1988.

Karato, S., On the Lehmann discontinuity, *Geophys. Res. Lett.*, *22*, 2255–2258, 1992.

Karner, G. D., Spectral representation of isostatic models, *BMR J. Aust. Geol. Geophys.*, *7*, 55–62, 1982.

Karner, G. D., and A. B. Watts, Gravity-anomalies and flexure of the lithosphere at mountain ranges, *J. Geophys. Res.*, *88*, 449–477, 1983.

Karner, G. D., M. S. Steckler, and J. A. Thorne, Long-term thermo-mechanical properties of the continental lithosphere, *Nature*, *304*, 250–253, 1983.

Kaula, W. M., Geophysical applications of satellite determinations of the Earth's gravitational field, *Space Sci. Rev.*, *7*, 769–794, 1967.

Kay, S. M., and S. L. Marple, Spectrum analysis: A modern perspective, *Proc. IEEE*, *69*, 1380–1419, 1981.

Kennett, B. L. N., Guided wave propagation in laterally varying media, I, Theoretical development, *Geophys. J. R. Astron. Soc.*, *79*, 235–255, 1984.

Kennett, B. L. N., Approximations for surface-wave propagation in laterally varying media, *Geophys. J. Int.*, *122*, 470–478, 1995.

Kennett, B. L. N., and G. Nolet, The interaction of the S-wavefield with upper mantle heterogeneity, *Geophys. J. Int.*, *101*, 751–762, 1990.

Kennett, B. L. N., E. Engdahl, and R. Buland, Constraints on seismic velocities in the Earth from travel-times, *Geophys. J. Int.*, *122*, 108–124, 1995.

Komm, R. W., Y. Gu, F. Hill, P. B. Stark, and I. K. Fodor, Multitaper spectral analysis and wavelet denoising applied to helioseismic data, *Astroph. J.*, *519*, 401–421, 1999.

Kubo, A., and Y. Hiramatsu, On presence of seismic anisotropy in the asthenosphere beneath continents and its dependence on plate velocity: Significance of reference frame selection, *Pure Appl. Geophys.*, *151*, 281–303, 1998.

Kumar, P., A wavelet-based methodology for scale-space anisotropic analysis, *Geophys. Res. Lett.*, *22*, 2777–2780, 1995.

Kumar, P., and E. Foufoula-Georgiou, Wavelet analysis for geophysical applications, *Rev. Geophys.*, *35*, 385–412, 1997.

Kuo, C., C. Lindberg, and D. J. Thomson, Coherence established between atmospheric carbon dioxide and global temperature, *Nature*, *343*, 709–714, 1990.

Lambeck, K., Structure and evolution of the intracratonic basins of central Australia, *Geophys. J. R. Astron. Soc.*, *74*, 843–886, 1983.

Lambeck, K., *Geophysical Geodesy*, Oxford Univ. Press, New York, 1988.

Lambeck, K., and C. Penney, Teleseismic travel time anomalies and crustal structure in central Australia, *Phys. Earth Planet. Inter.*, *34*, 46–56, 1984.

Lambeck, K., H. W. S. McQueen, R. A. Stephenson, and D. Denham, The state of stress within the Australian continent, *Ann. Geophys.*, *2*, 723–742, 1984.

Lambeck, K., G. Burgess, and R. D. Shaw, Teleseismic travel-time anomalies and deep crustal structure in central Australia, *Geophys. J. Int.*, *94*, 105–124, 1988.

Laske, G., and G. Masters, Constraints on global phase velocity maps from long-period polarization data, *J. Geophys. Res.*, *101*, 16,059–16,075, 1996.

Lavier, L. L., and M. S. Steckler, The effect of sedimentary cover on the flexural strength of continental lithosphere, *Nature*, *389*, 476–479, 1997.

Lavier, L. L., and M. S. Steckler, The effect of sedimentary cover on the flexural strength of continental lithosphere, Correction, *Nature*, *392*, 843, 1998.

Lay, T., and T. C. Wallace, *Modern Global Seismology*, Academic Press, San Diego, Calif., 1995.

Lee, C.-T., Q. Yin, R. L. Rudnick, J. T. Chesley, and S. B. Jacobson, Osmium isotopic evidence for Mesozoic removal of lithospheric mantle beneath the Sierra Nevada, California, *Science*, *289*, 1912–1916, 2000.

Lee, C.-T., Q. Yin, R. L. Rudnick, and S. B. Jacobson, Preservation of ancient and fertile lithospheric mantle beneath the southwestern United States, *Nature*, *411*, 69–73, 2001.

Lees, J. M., and J. Park, Multiple-taper spectral analysis: A stand-alone C-subroutine, *Comput. Geosci.*, *21*, 199–236, 1995.

Lemoine, F., et al., *The Development of the Joint NASA GSFC and the National Imagery and Mapping Agency (NIMA) Geopotential Model EGM96*, NASA Goddard Space Flight Cent., Greenbelt, Md., 1998.

Lenardic, A., On the partitioning of mantle heat loss below oceans and continents over time and its relationship to the Archaean paradox, *Geophys. J. Int.*, *134*, 706–720, 1998.

Lenardic, A., L. Moresi, and H. Mühlhaus, The role of mobile belts for the longevity of deep cratonic lithosphere: The crumple zone model, *Geophys. Res. Lett.*, *27*, 1235–1238, 2000.

Lévêque, J.-J., and M. Cara, Inversion of multimode surface wave data: Evidence for sublithospheric anisotropy, *Geophys. J. R. Astron. Soc.*, *83*, 753–773, 1985.

Lévêque, J.-J., L. Rivera, and G. Wittlinger, On the use of the checker-board test to assess the resolution of tomographic inversions, *Geophys. J. Int.*, *115*, 313–318, 1993.

Lévêque, J.-J., E. Debayle, and V. Maupin, Anisotropy in the Indian Ocean upper mantle from Rayleigh- and Love-waveform inversion, *Geophys. J. Int.*, *133*, 529–540, 1998.

Li, A., K. M. Fischer, M. E. Wysession, and T. J. Clarke, Mantle discontinuities and temperature under the North American continental keel, *Nature*, *395*, 160–163, 1998.

Li, X.-d., and T. Tanimoto, Waveforms of long-period body waves in a slightly aspherical Earth model, *Geophys. J. Int.*, *112*, 92–102, 1993.

Lilley, F. E. M., D. V. Woods, and M. N. Sloane, Electrical conductivity from Australian magnetometer arrays using spatial gradient data, *Phys. Earth Planet. Inter.*, *25*, 202–209, 1981.

Lilly, J. M., and J. Park, Multiwavelet spectral and polarization analyses of seismic records, *Geophys. J. Int.*, *122*, 1001–1021, 1995.

Lindberg, C. R., and J. Park, Multiple-taper spectral analysis of terrestrial free oscillations, II, *Geophys. J. R. Astron. Soc.*, *91*, 795–836, 1987.

Liu, P. C., Wavelet spectrum analysis and ocean wind waves, in *Wavelets in Geophysics*, edited by E. Foufoula-Georgiou and P. Kumar, pp. 151–166, Academic Press, San Diego, Calif., 1994.

Liu, T.-C., and B. D. van Veen, Multiple window based minimum variance spectrum estimation for multidimensional random fields, *IEEE Trans. Signal Process.*, *40*, 578–589, 1992.

Lowry, A. R., and R. B. Smith, Flexural rigidity of the Basin and Range-Colorado Plateau-Rocky Mountain transition from coherence analysis of gravity and topography, *J. Geophys. Res.*, *99*, 20,123–20,140, 1994.

Lowry, A. R., and R. B. Smith, Strength and rheology of the western U. S. Cordillera, *J. Geophys. Res.*, *100*, 17,947–17,963, 1995.

Lucas, S. B., A. green, Z. Hajnal, D. White, J. Lewry, K. Ashton, W. Weber, and R. Clowes, Deep seismic profile across a Proterozoic collision zone: Surprises at depth, *Nature*, *363*, 339–342, 1993.

Macario, A., A. Malinverno, and W. F. Haxby, On the robustness of elastic thickness estimates obtained using the coherence method, *J. Geophys. Res.*, *100*, 15,163–15,172, 1995.

Mainprice, D., and P. G. Silver, Interpretation of *SKS*-waves using samples from the subcontinental lithosphere, *Phys. Earth Planet. Inter.*, *78*, 257–280, 1993.

Malik, N., and J. S. Lim, Properties of two-dimensional maximum entropy power spectrum estimates, *IEEE Trans. Acoust. Speech Signal Process.*, *30*, 788–798, 1982.

Mallat, S., *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, Calif., 1998.

Mareschal, M., R. L. Kellett, R. D. Kurtz, J. N. Ludden, S. Ji, and R. C. Bailey, Archaean cratonic roots, mantle shear zones and deep electrical anisotropy, *Nature*, *375*, 134–137, 1995.

Marquardt, D. W., Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation, *Technom.*, *12*, 1970.

Marquering, H., and R. Snieder, Surface-wave mode coupling for efficient forward modelling and inversion of body-wave phases, *Geophys. J. Int.*, *120*, 186–208, 1995.

Marquering, H., and R. Snieder, Shear-wave velocity structure beneath Europe, the Northeastern Atlantic and Western Asia from waveform inversions including surface-wave mode coupling, *Geophys. J. Int.*, *127*, 283–304, 1996.

Marquering, H., R. Snieder, and G. Nolet, Waveform inversions and the significance of surface-wave mode coupling, *Geophys. J. Int.*, *124*, 258–278, 1996.

Masters, G., S. Johnson, G. Laske, and H. Bolton, A shear-velocity model of the mantle, *Philos. Trans. R. Soc. London, Ser. A*, *354*, 1385–1411, 1996.

McKenzie, D., and C. Bowin, The relationship between bathymetry and gravity in the Atlantic Ocean, *J. Geophys. Res.*, *81*, 1903–1915, 1976.

McKenzie, D., and J. D. Fairhead, Estimates of the effective elastic thickness of the continental lithosphere from Bouguer and free air gravity anomalies, *J. Geophys. Res.*, *102*, 27,523–27,552, 1997.

McNutt, M. K., and R. L. Parker, Isostasy in Australia and the evolution of the compensation mechanism, *Science*, *199*, 773–775, 1978.

McQueen, H. W. S., and K. Lambeck, Determination of crustal structure in central Australia by inversion of traveltime residuals, *Geophys. J. Int.*, *126*, 645–662, 1996.

Mellors, R., F. Vernon, and D. J. Thomson, Detection of dispersive signals using multitaper dual-frequency coherence, *Geophys. J. Int.*, *135*, 146–154, 1998.

Menke, W., *Geophysical Data Analysis: Discrete Inverse Theory*, vol. 45 of *International Geophysics Series*, 2nd ed., Academic Press, San Diego, Calif., 1989.

Meredith, R. W., Coherence estimation for high-frequency narrowband cw pulsed signals in shallow water, *J. Acoust. Soc. Am.*, *106*, 828–836, 1999.

Mitchell, B. J., S. Baqer, A. Akinci, and L. Cong, L_g coda Q in Australia and its relation to crustal structure and evolution, *Pure Appl. Geophys.*, *153*, 639–653, 1998.

Montagner, J.-P., Seismic anisotropy of the Pacific Ocean inferred from long-period surface-waves dispersion, *Phys. Earth Planet. Inter.*, *38*, 28–50, 1985.

Montagner, J.-P., Regional three-dimensional structures using long-period surface waves, *Ann. Geophys.*, *4*, 283–294, 1986.

Montagner, J.-P., Surface waves on a global scale, Influence of anisotropy and anelasticity, in *Seismic Modelling of Earth Structure*, edited by E. Boschi, G. Ekström, and A. Morelli, pp. 81–148, Editrice Compositori,, Bologna, Italy, 1996.

Montagner, J.-P., and D. L. Anderson, Petrological constraints on seismic anisotropy, *Phys. Earth Planet. Inter.*, *54*, 82–105, 1989.

Montagner, J.-P., and N. Jobert, Vectorial tomography – II. Application to the Indian Ocean, *Geophys. J. Int.*, *94*, 309–344, 1988.

Montagner, J.-P., and H.-C. Nataf, A simple method for inverting the azimuthal anisotropy of surface waves, *J. Geophys. Res.*, *91*, 511–520, 1986.

Montagner, J.-P., and H.-C. Nataf, Vectorial tomography – I. Theory, *Geophys. J. Int.*, *94*, 295–307, 1988.

Montagner, J.-P., and T. Tanimoto, Global upper mantle tomography of seismic velocities and anisotropies, *J. Geophys. Res.*, *96*, 20,337–20,351, 1991.

Montagner, J.-P., P. Ho-Liu, and H. Kanamori, Reply to comment by J. Trampert, *Geophys. J. Int.*, *103*, 757–758, 1990.

Montagner, J.-P., D.-A. Griot-Pommera, and J. Lavé, How to relate body wave and surface wave anisotropy?, *J. Geophys. Res.*, *105*, 19,015–19,027, 2000.

Mooney, W. D., Continental roots go with the flow, *Nature*, *375*, 15, 1995.

Mullis, C. T., and L. L. Scharf, Quadratic estimators of the power spectrum, in *Advances in Spectrum Analysis and Array Processing*, edited by S. Haykin, vol. 1, chap. 1, pp. 1–57, Prentice Hall, Englewood Cliffs, N. J., 1991.

Munk, W. H., and D. E. Cartwright, Tidal spectroscopy and prediction, *Philos. Trans. R. Soc. London, Ser. A*, *259*, 533–581, 1966.

Murray, C. G., E. Scheibner, and R. N. Walker, Regional geological interpretation of a digital coloured residual Bouguer gravity image of eastern Australia with a wavelength cut-off of 250 km, *Aust. J. Earth Sci.*, *36*, 423–449, 1989.

Myers, J. S., R. D. Shaw, and I. M. Tyler, Tectonic evolution of Proterozoic Australia, *Tectonics*, *15*, 1431–1446, 1996.

Nicolas, A., and J.-P. Poirier, *Crystalline plasticity and solid state flow in metamorphic rocks*, Wiley, New York, 1976.

Nishimura, C. E., and D. W. Forsyth, Anomalous Love-wave phase velocities in the Pacific: Sequential pure-path and spherical harmonic inversion, *Geophys. J. R. Astron. Soc.*, *81*, 389–407, 1985.

Nishimura, C. E., and D. W. Forsyth, Rayleigh wave phase velocities in the Pacific with implications for azimuthal anisotropy and lateral heterogeneities, *Geophys. J. Int.*, *94*, 479–501, 1988.

Nishimura, C. E., and D. W. Forsyth, The anisotropic structure of the upper mantle in the Pacific, *Geophys. J. Int.*, *96*, 203–229, 1989.

Nolet, G., Solving or resolving inadequate and noisy tomographic systems, *J. Comput. Phys.*, *61*, 463–482, 1985.

Nolet, G. (Ed.), *Seismic Tomography*, Reidel, Hingham, MA, 1987.

Nolet, G., Partitioned waveform inversion and two-dimensional structure under the Network of Autonomously Recording Seismographs, *J. Geophys. Res.*, *95*, 8499–8512, 1990.

Nolet, G., and A. Zielhuis, Low *S*-velocities under the Tornquist-Teisseyre zone, Evidence for water injection into the transition zone by subduction, *J. Geophys. Res.*, *99*, 15,813–15,820, 1994.

Nolet, G., J. van Trier, and R. Huisman, A formalism for nonlinear inversion of seismic surface waves, *Geophys. Res. Lett.*, *13*, 26–29, 1986.

Nolet, G., S. P. Grand, and B. L. N. Kennett, Seismic heterogeneity in the upper mantle, *J. Geophys. Res.*, *99*, 23,753–23,766, 1994.

Nolet, G., R. Montelli, and J. Virieux, Explicit, approximate expressions for the resolution and a posteriori covariance of massive tomographic systems, *Geophys. J. Int.*, *138*, 36–44, 1999.

Nyblade, A. A., Heat flow and the structure of Precambrian lithosphere, *Lithos*, *48*, 81–91, 1999.

Nyblade, A. A., Earth science — Hard-cored continents, *Nature*, *2001*, 6833, 2001.

Okal, E. A., The effect of intrinsic oceanic upper-mantle heterogeneity on regionalization of long-period Rayleigh-wave phase velocities, *Geophys. J. R. Astron. Soc.*, *49*, 357–370, 1977.

Özalaybey, S., and W. P. Chen, Frequency-dependent analysis of *SKS-SKKS* waveforms observed in Australia: Evidence for null birefringence, *Phys. Earth Planet. Inter.*, *114*, 197–210, 1999.

Paige, C. C., and M. A. Saunders, LSQR: An algorithm for sparse linear equations and sparse least sqaures, *ACM Trans. Math. Softw.*, *8*, 43–71, 1982.

Pari, G., and W. R. Peltier, The free-air gravity constraint on subcontinental mantle dynamics, *J. Geophys. Res.*, *101*, 28,105–28,132, 1996.

Park, J., C. R. Lindberg, and D. J. Thomson, Multiple-taper spectral analysis of terrestrial free oscillations, I, *Geophys. J. R. Astron. Soc.*, *91*, 755–794, 1987a.

Park, J., C. R. Lindberg, and F. L. Vernon, Multitaper spectral analysis of high-frequency seismograms, *J. Geophys. Res.*, *92*, 12,675–12,684, 1987b.

Parker, R. L., The rapid calculation of potential anomalies, *Geophys. J. R. Astron. Soc.*, *31*, 447–455, 1972.

Passier, M. L., R. D. van der Hilst, and R. K. Snieder, Surface wave waveform inversions for local shear-wave velocities under eastern Australia, *GRL*, *24*, 1291–1294, 1997.

Pearson, D. G., The age of continental roots, *Lithos*, *48*, 171–194, 1999.

Pearson, D. G., G. Snyder, S. Shirey, L. Taylor, R. W. Carlson, and N. Sobolev, Archean Re-Os age for Siberian eclogites and constraints on Archean tectonics, *Nature*, *374*, 711–713, 1995.

Percival, D. B., and A. T. Walden, *Spectral Analysis for Physical Applications, Multitaper and Conventional Univariate Techniques*, Cambridge Univ. Press, New York, 1993.

Peselnick, L., and A. Nicolas, Seismic anisotropy in an ophiolite peridotite. Application to oceanic upper mantle, *J. Geophys. Res.*, *83*, 1227–1235, 1978.

Phillips, R. J., and K. Lambeck, Gravity fields of the terrestrial planets: Long-wavelength anomalies and tectonics, *Rev. Geophys.*, *18*, 27–76, 1980.

Plomerová, J., R. C. Liebermann, and V. Babuška (Eds.), *Geodynamics of Lithosphere and Earth's Mantle*, vol. 151 of *Pure Appl. Geophys.*, Birkhäuser, Boston, Mass., 1998.

Plumb, K. A., The tectonic evolution of Australia, *Earth Sci. Rev.*, *14*, 205–249, 1979.

Polet, J., and D. L. Anderson, Depth extent of cratons as inferred from tomographic studies, *Geology*, *23*, 205–208, 1995.

Poudjom Djomani, Y. H., J. D. Fairhead, and W. L. Griffin, The flexural rigidity of Fennoscandia: Reflection of the tectonothermal age of the lithospheric mantle, *Earth Planet. Sci. Lett.*, *174*, 139–154, 1999.

Powell, C., Break-up of Rodinia in Australia; Implications for the western margin of Laurentia, in *Abstracts with Programs, Geological Society of America, 1998 annual meeting*, vol. 30, p. 47, Geological Society of America, 1998.

Priestley, K., Velocity structure of the continental upper mantle: Evidence from South Africa, *Lithos*, *48*, 45–56, 1999.

Priestley, M., *Spectral Analysis and Time Series*, Academic Press, San Diego, Calif., 1981.

Ribe, N. M., On the relation between seismic anisotropy and finite strain, *J. Geophys. Res.*, *97*, 8737–8747, 1992.

Richardson, S. H., J. J. Gurney, A. J. Erlank, and J. W. Harris, Origin of diamonds in old enriched mantle, *Nature*, *310*, 198–202, 1984.

Richter, F. M., Regionalized models for the thermal evolution of the Earth, *Earth Planet. Sci. Lett.*, *68*, 471–484, 1984.

Richter, F. M., Models for the Archean thermal regime, *Earth Planet. Sci. Lett.*, *73*, 350–360, 1985.

Richter, F. M., A major change in the thermal state of the Earth at the Archean-Proterozoic boundary: Consequences for the nature and preservation of continental lithosphere, *J. Petrol.*, *Special Lithosphere Issue*, 39–52, 1988.

Riedel, K. S., and A. Sidorenko, Minimum bias multiple taper spectral estimation, *IEEE Trans. Signal Process.*, *43*, 188–195, 1995.

Rioul, O., and P. Flandrin, Time-scale energy distributions: A general class extending wavelet transforms, *IEEE Trans. Signal Process.*, *40*, 1746–1757, 1992.

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

Rutland, R. W. R., Orogenic evolution of Australia, *Earth Sc. Rev.*, *12*, 161–196, 1976.

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

Sandiford, M., and M. Hand, Controls on the locus of intraplate deformation in central Australia, *Earth Planet. Sci. Lett.*, *162*, 97–110, 1998.

Sandwell, D. T., and M. L. Renkin, Compensation of swells and plateaus in the North Pacific: No direct evidence for mantle convection, *J. Geophys. Res.*, *93*, 2775–2783, 1988.

Santoso, S., E. J. Powers, R. D. Bengston, and A. Ouroua, Time-series analysis of nonstationary plasma fluctuations using wavelet transforms, *Rev. Sci. Instrum.*, *68*, 898–901, 1997.

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

Savage, M. K., and A. F. Sheehan, Seismic anisotropy and mantle flow from the Great Basin to the Great Plains, western United States, *J. Geophys. Res.*, *105*, 13,715–13,734, 2000.

Scales, J. A., and R. Snieder, To Bayes or not to Bayes?, *Geoph.*, *62*, 1045–1046, 1997.

Scheirer, D., D. W. Forsyth, and A. Hosford, Multitaper estimates of the lithospheric strength of the Basin and Range province (abstract), *Eos Trans. AGU*, *76(17)*, Spring Meet. Suppl., S282, 1995.

Sclater, J. G., B. Parsons, and C. Jaupart, Oceans and continents: Similarities and differences in the mechanisms of heat loss, *J. Geophys. Res.*, *86*, 11,535–11,552, 1981.

Seymour, M. S., and I. G. Cumming, Maximum likelihood estimation for SAR interferometry, *Proc. 1994 Int. Geosci. Remote Sensing Symp., IGARSS '94*, pp. 2272–2275, 1994.

Shapiro, S. S., B. H. Hager, and T. H. Jordan, Stability and dynamics of the continental tectosphere, *Lithos*, *48*, 115–133, 1999a.

Shapiro, S. S., B. H. Hager, and T. H. Jordan, The continental tectosphere and Earth's long-wavelength gravity field, *Lithos*, *48*, 135–152, 1999b.

Shaw, R. D., P. Wellman, P. Gunn, A. J. Whitaker, C. Tarlowski, and M. P. Morse, Australian crustal elements map, *AGSO Res. Newslett.*, *23*, 1–3, 1995.

Shibutani, T., M. Sambridge, and B. L. N. Kennett, Genetic algorithm inversion for receiver functions with application to crust and uppermost mantle structure beneath eastern Australia, *Geophys. Res. Lett.*, *23*, 1829–1832, 1996.

Silveira, G., E. Stutzmann, D.-A. Griot, J.-P. Montagner, and L. M. Victor, Anisotropic tomography of the Atlantic Ocean from Rayleigh surface waves, *Phys. Earth Planet. Inter.*, *106*, 257–273, 1998.

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

Silver, P. G., and W. Chan, Shear-wave splitting and subcontinental mantle deformation, *J. Geophys. Res.*, *96*, 16,429–16,454, 1991.

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

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

Silver, P. G., S. S. Gao, K. H. Liu, and Kaapvaal Seismic Group, Mantle deformation beneath southern Africa, *Geophys. Res. Lett.*, *28*, 2493–2496, 2001.

Simons, F. J., A. Zielhuis, and R. D. van der Hilst, The deep structure of the Australian continent from surface-wave tomography, *Lithos*, *48*, 17–43, 1999.

Simons, F. J., M. T. Zuber, and J. Korenaga, Isostatic response of the Australian lithosphere: Estimation of effective elastic thickness and anisotropy using multitaper spectral analysis, *J. Geophys. Res.*, *105*, 19,163–19,184, 2000.

Simons, F. J., R. D. van der Hilst, J.-P. Montagner, and A. Zielhuis, Multimode Rayleigh wave inversion for heterogeneity and azimuthal anisotropy of the Australian upper mantle, *Geophys. J. Int.*, 2002a, in press.

Simons, F. J., R. D. van der Hilst, and M. T. Zuber, Isostatic and seismic anisotropy of Australia from non-stationary gravity-topography coherence analysis and surface wave tomography, *J. Geophys. Res.*, 2002b, in press.

Simons, M., S. C. Solomon, and B. H. Hager, Localization of gravity and topography: Constraints on the tectonics and mantle dynamics of Venus, *Geophys. J. Int.*, *131*, 24–44, 1997.

Sipkin, S. A., Rapid determination of global moment-tensor solutions, *Geophys. Res. Lett.*, *21*, 1667–1670, 1994.

Sipkin, S. A., and T. H. Jordan, Lateral heterogeneity of the upper mantle determined from the travel times of *ScS*, *J. Geophys. Res.*, *80*, 1474–1484, 1975.

Sipkin, S. A., and T. H. Jordan, Lateral heterogeneity of the upper mantle determined from the travel times of multiple *ScS*, *J. Geophys. Res.*, *81*, 6307–6320, 1976.

Slepian, D., Prolate spheroidal wave functions, Fourier analysis and uncertainty, V, The discrete case, *Bell Syst. Tech. J.*, *57*, 1371–1429, 1978.

Slepian, D., Some comments on Fourier-analysis, uncertainty and modeling, *SIAM Rev.*, *25*, 379–393, 1983.

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

Smith, W. H. F., and D. T. Sandwell, Global sea floor topography from satellite altimetry and ship depth soundings, *Science*, *277*, 1956–1962, 1997.

Souriau, A., On the retrieval of pure path velocities from great circle data, *Geophys. J. R. Astron. Soc.*, *80*, 783–790, 1985.

Spakman, W., and H. Bijwaard, Optimization of cell parameterization for tomographic inverse problems, *Pure Appl. Geophys.*, *158*, 1401–1423, 2001.

Spakman, W., and G. Nolet, Imaging algorithms, accuracy and resolution in delay time tomography, in *Mathematical Geophysics: A Survey of Recent Developments in Seismology and Geodynamics*, edited by N. J. Vlaar, G. Nolet, M. J. R. Wortel, and S. A. P. L. Cloetingh, pp. 155–187, Reidel, Dordrecht, The Netherlands, 1988.

Spassov, E., The stress field in Australia from composite fault plane solutions of the strongest earthquakes in the continent, *J. Seismol.*, *2*, 173–178, 1998.

Spassov, E., and B. L. N. Kennett, Stress and faulting in southeast Australia as derived from the strongest earthquakes in the region, *J. Asian Earth Sc.*, *18*, 17–23, 2000.

Stephenson, R., Flexural models of continental lithosphere based on the long-term erosional decay of topography, *Geophys. J. R. Astron. Soc.*, *77*, 385–413, 1984.

Stephenson, R., and C. Beaumont, Small-scale convection in the upper mantle and the isostatic response of the Canadian shield, in *Mechanisms of Continental Drift and Plate Tectonics*, edited by P. A. Davies and S. K. Runcorn, pp. 111–122, Academic Press, San Diego, Calif., 1980.

Stephenson, R., and K. Lambeck, Isostatic response of the lithosphere with in-plane stress: Application to central Australia, *J. Geophys. Res.*, *90*, 8581–8588, 1985.

Stephenson, R. A., and S. A. P. L. Cloetingh, Some examples and mechanical aspects of continental lithospheric folding, *Tectonophysics*, *188*, 27–37, 1991.

Storey, B. C., The role of mantle plumes in continental breakup: case histories from gondwanaland, *Nature*, *377*, 301–308, 1995.

Strang, G., and T. Nguyen, *Wavelets and Filter Banks*, 2nd ed., Wellesley-Cambridge Press, Wellesley, Mass., 1997.

Strong, D., and R. K. Stevens, Possible thermal explanation of contrasting Archean and Proterozoic geological regimes, *Nature*, *249*, 545–546, 1974.

Stutzmann, E., and J.-P. Montagner, An inverse technique for retrieving higher mode phase velocity and mantle structure, *Geophys. J. Int.*, *113*, 669–683, 1993.

Stutzmann, E., and J.-P. Montagner, Tomography of the transition zone from the inversion of higher-mode surface-waves, *Phys. Earth Planet. Inter.*, *86*, 99–115, 1994.

Su, W.-j., R. Woodward, and A. M. Diewonski, Degree-12 model of shear velocity heterogeneity in the mantle, *J. Geophys. Res.*, *99*, 6945–6980, 1994.

Takeuchi, H., and M. Saito, Seismic surface waves, in *Seismology: Surface Waves and Earth Oscillations*, edited by B. A. Bolt, vol. 11 of *Methods In Computational Physics*, pp. 217–295, Academic Press, San Diego, Calif., 1972.

Tarantola, A., and B. Valette, Generalized nonlinear inverse problems solved using the least squares criterion, *Rev. Geophys.*, *20*, 219–232, 1982a.

Tarantola, A., and B. Valette, Inverse problems — Quest for information, *J. Geoph.*, *50*, 159–170, 1982b.

ter Voorde, M., R. T. van Balen, G. Bertotti, and S. A. P. L. Cloetingh, The influence of a stratified rheology on the flexural response of the lithosphere to (un)loading by extensional faulting, *Geophys. J. Int.*, *134*, 721–735, 1998.

Thomson, D. J., Spectrum estimation and harmonic analysis, *Proc. IEEE*, *70*, 1055–1096, 1982.

Thomson, D. J., and A. D. Chave, Jackknifed error estimates for spectra, coherences, and transfer functions, in *Advances in Spectrum Analysis and Array Processing*, edited by S. Haykin, vol. 1, chap. 2, pp. 58–113, Prentice Hall, Englewood Cliffs, N. J., 1991.

Timoshenko, S., and S. Woinowsky-Krieger, *Theory of Plates and Shells*, 2nd ed., McGraw-Hill, New York, 1959.

Toksöz, M. N., and D. L. Anderson, Phase velocities of long-period surface waves and structure of the upper mantle, I. great-circle Love and Rayleigh wave data, *J. Geophys. Res.*, *71*, 1649–1658, 1966.

Tommasi, A., Forward modeling of the development of seismic anisotropy in the uppper mantle, *Earth Planet. Sci. Lett.*, *160*, 1–13, 1998.

Tommasi, A., and A. Vauchez, Continental rifting parallel to ancient collisional belts; An effect of the mechanical anisotropy of the lithospheric mantle, *Earth Planet. Sci. Lett.*, *185*, 199–210, 2001.

Tommasi, A., B. Tikoff, and A. Vauchez, Upper mantle tectonics: Three-dimensional deformation, olivine crystallographic fabrics and seismic properties, *Earth Planet. Sci. Lett.*, *168*, 1733–186, 1999.

Tong, C., Ó. Guðmundsson, and B. L. N. Kennett, Shear wave splitting in refracted waves returned from the upper mantle transition zone beneath northern Australia, *J. Geophys. Res.*, *99*, 15,783–15,797, 1994.

Torrence, C., and G. P. Compo, A practical guide to wavelet analysis, *Bull. Am. Meter. Soc.*, *79*, 61–78, 1998.

Touzi, R., A. Lopes, and P. Vachon, Estimation of the coherence function for interferometric SAR applications, in *European Conference on Synthetic Aperture Radar 1996*, pp. 241–244, EUSAR, Königswinter, Germany, 1996.

Touzi, R., A. Lopes, J. Bruniquel, and P. Vachon, Coherence estimation for SAR imagery, *IEEE Trans. Geosci. Remote Sens.*, *37*, 135–149, 1999.

Trampert, J., Comparison of iterative back-projection inversion and generalized inversion without blocks —- Case-studies in attenuation tomography — Comment, *Geophys. J. Int.*, *103*, 755–756, 1990.

Trampert, J., and J. H. Woodhouse, Global phase-velocity maps of Love and Rayleigh-waves between 40 and 150 seconds, *Geophys. J. Int.*, *122*, 675–690, 1995.

Trampert, J., and J. H. Woodhouse, High resolution global phase velocity distributions, *Geophys. Res. Lett.*, *23*, 21–24, 1996.

Tukey, J. W., An introduction to the calculations of numerical spectrum analysis, in *Advanced Seminar on Spectral Analysis of Time Series*, edited by B. Harris, pp. 25–46, John Wiley, New York, 1967.

Turcotte, D. L., and G. Schubert, *Geodynamics, Application of Continuum Physics to Geological Problems*, John Wiley, New York, 1982.

van der Hilst, R. D., B. L. N. Kennett, D. Christie, and J. Grant, Project SKIPPY explores the lithosphere and mantle beneath Australia, *Eos Trans. AGU*, *75*, 177–181, 1994.

van der Hilst, R. D., B. L. N. Kennett, and T. Shibutani, Upper mantle structure beneath Australia from portable array deployments, in *Structure and Evolution of the Australian Continent*, edited by J. Braun, J. C. Dooley, B. Goleby, R. D. van der Hilst, and C. Klootwijk, vol. 26 of *Geodyn. Ser.*, pp. 39–57, AGU, Washington, D. C., 1998.

van der Lee, S., and G. Nolet, Seismic image of the subducted trailing fragments of the Farallon plate, *Nature*, *386*, 266–269, 1997a.

van der Lee, S., and G. Nolet, Upper mantle *S*-velocity structure of North America, *J. Geophys. Res.*, *102*, 22,815–22,838, 1997b.

van der Sluis, A., and H. A. van der Vorst, The rate of convergence of conjugate gradients, *Numer. Math.*, *48*, 543–560, 1986.

van der Sluis, A., and H. A. van der Vorst, Numerical solution of large sparse linear algebraic systems arising from tomographic problems, in *Seismic Tomography*, edited by G. Nolet, chap. 3, pp. 49–83, Reidel, Hingham, Mass., 1987.

van Heijst, H. J., and J. H. Woodhouse, Measuring surface-wave overtone phase velocities using a mode-branch stripping technique, *Geophys. J. Int.*, *131*, 209–230, 1997.

van Heijst, H. J., and J. H. Woodhouse, Global high-resolution phase velocity distributions of overtone and fundamental-mode surface waves determined by mode branch stripping, *Geophys. J. Int.*, *137*, 601–620, 1999.

VanDecar, J. C., and R. Snieder, Obtaining smooth solutions to large, linear inverse problems, *Geophysics*, *59*, 818–829, 1994.

Vauchez, A., A. Tommasi, and G. Barruol, Rheological heterogeneity, mechanical anisotropy and deformation of the continental lithosphere, *Tectonophysics*, *296*, 61–86, 1998.

Veevers, J. J. (Ed.), *Phanerozoic Earth History of Australia*, Oxford Univ. Press, New York, 1984.

Vinnik, L. P., V. Farra, and B. Romanowicz, Azimuthal anisotropy in the Earth from observations of *SKS* at GEOSCOPE and NARS broad-band stations, *Geophys. J. Int.*, *79*, 1542–1558, 1989.

Vinnik, L. P., L. I. Makeyeva, A. Milev, and A. Y. Usenko, Global patterns of azimuthal anisotropy and deformations in the continental mantle, *Geophys. J. Int.*, *111*, 433–447, 1992.

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

Walden, A., Improved low-frequency decay estimation using the multitaper spectral-analysis method, *Geophys. Prospect.*, *38*, 61–86, 1990.

Walden, A. T., D. B. Percival, and E. J. McCoy, Spectrum estimation by wavelet thresholding of multitaper estimators, *IEEE Trans. Signal Process.*, *46*, 3153–3165, 1998.

Wang, Y., and J.-C. Mareschal, Elastic thickness of the lithosphere in the central Canadian shield, *Geophys. Res. Lett.*, *26*, 3033–3036, 1999.

Wang, Z., and F. A. Dahlen, Validity of surface-wave ray theory on a laterally heterogeneous Earth, *Geophys. J. Int.*, *123*, 757–773, 1995.

Watts, A. B., An analysis of isostasy in the world's oceans, 1, Hawaiian-Emperor seamount chain, *J. Geophys. Res.*, *83*, 5989–6004, 1978.

Welch, P. D., The use of Fast Fourier Transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms, *IEEE Trans. Audio Electroacoust.*, *15*, 70–73, 1967.

Wellman, P., On the isostatic compensation of Australian topography, *BMR J. Aust. Geol. Geophys.*, *4*, 373–382, 1979.

Wellman, P., Mapping of geophysical domains in the Australian continental crust using gravity and magnetic anomalies, in *Structure and Evolution of the Australian Continent*, edited by J. Braun, J. C. Dooley, B. Goleby, R. D. van der Hilst, and C. Klootwijk, vol. 26 of *Geodyn. Ser.*, pp. 59–71, AGU, Washington, D. C., 1998.

Wenk, H. R., and C. N. Tomé, Modeling dynamic recrystallization of olivine aggregates deformed in simple shear, *J. Geophys. Res.*, *104*, 25,513–25,527, 1999.

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

Woodhouse, J. H., and A. M. Dziewonski, Mapping the upper mantle: Three-dimensional modeling of Earth structure by inversion of seismic waveforms, *J. Geophys. Res.*, *89*, 5,953–5,986, 1984.

Yanovskaya, T. B., and P. G. Ditmar, Smoothness criteria in surface wave tomography, *Geophys. J. Int.*, *102*, 63–72, 1990.

Zhang, S. Q., and S. Karato, Lattice preferred orientation of olivine aggregates deformed in simple shear, *Nature*, *375*, 774–777, 1995.

Zhang, Y., E. Scheibner, B. E. Hobbs, A. Ord, B. J. Drummond, and S. J. D. Cox, Lithospheric structure in Southeast Australia: A model based on gravity, geoid and mechanical analyses, in *Structure and Evolution of the Australian Continent*, edited by J. Braun, J. C. Dooley, B. Goleby, R. D. van der Hilst, and C. Klootwijk, vol. 26 of *Geodyn. Ser.*, pp. 89–108, AGU, Washington, D. C., 1998.

Zhong, S., Dynamics of crustal compensation and its influences on crustal isostasy, *J. Geophys. Res.*, *102*, 15,287–15,299, 1997.

Zielhuis, A., *S*-wave velocity below Europe from delay-time and waveform inversions, Ph.D. thesis, Utrecht Univ., Utrecht, The Netherlands, 1992.

Zielhuis, A., and G. Nolet, Shear-wave velocity variations in the upper mantle beneath central Europe, *Geophys. J. Int.*, *117*, 695–715, 1994a.

Zielhuis, A., and G. Nolet, Deep seismic expression of an ancient plate boundary in Europe, *Science*, *265*, 79–81, 1994b.

Zielhuis, A., and R. D. van der Hilst, Upper-mantle shear velocity beneath eastern Australia from inversion of waveforms from SKIPPY portable arrays, *Geophys. J. Int.*, *127*, 1–16, 1996.

Zoback, M. L., First- and second-order patterns of stress in the lithosphere: The World Stress Map project, *J. Geophys. Res.*, *97*, 11,703–11,728, 1992.

Zuber, M. T., T. D. Bechtel, and D. W. Forsyth, Effective elastic thicknesses of the lithosphere and mechanisms of isostatic compensation in Australia, *J. Geophys. Res.*, *94*, 9353–9367, 1989.

Zuber, M. T., et al., Internal structure and early thermal evolution of Mars from Mars Global Surveyor topography and gravity, *Science*, *287*, 1788–1793, 2000.