JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 105, NO. B8, PAGES 19,163–19,184, AUGUST 10, 2000

Isostatic response of the Australian lithosphere: Estimation of effective elastic thickness and anisotropy using multitaper spectral analysis

Frederik J. Simons, Maria T. Zuber, and Jun Korenaga
Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge

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 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.

1. Introduction

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.

Paper number 2000JB900157.
0148-0227/00/2000JB900157$09.00

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].

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: $z_m = z_c = z_{\text{Moho}}$. 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.

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].

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 midcrustal 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.

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 paper, 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].

2. Coherence Studies of the Mechanical Lithosphere

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

$$\Delta g(\mathbf{r}_0) = \int_{\Sigma} h(\mathbf{r}) q(\mathbf{r} - \mathbf{r}_0)\, d\mathbf{r}^2 + n(\mathbf{r}_0). \qquad (1)$$

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 paper, 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

$$\hat{q}(\mathbf{k}) = \frac{S_{h\Delta g}(\mathbf{k})}{S_{hh}(\mathbf{k})}, \tag{2}$$

where S represents the cross-spectral density of the two random variables identified by the subscripts (for an elaboration, see the Appendix). 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

$$\gamma^2(\mathbf{k}) = \frac{|S_{h\Delta g}(\mathbf{k})|^2}{S_{\Delta g\Delta g}(\mathbf{k})S_{hh}(\mathbf{k})} \quad ; \quad 0 \leq \gamma^2(\mathbf{k}) \leq 1, \tag{3}$$

instead of (2). This function is referred to as the coherence [*Bendat and Piersol*, 1986, 1993].

2.2. Spectral Estimation of Stochastic Signals: Importance of 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 $\mathbf{k}$:

$$\forall \mathbf{k} : E\{\hat{S}(\mathbf{k})\} = S(\mathbf{k}), \tag{4a}$$

$$\lim_{N\to\infty} \mathrm{var}\{\hat{S}(\mathbf{k})\} = 0, \tag{4b}$$

$$\forall \mathbf{k}' \neq \mathbf{k} : \mathrm{cov}\{\hat{S}(\mathbf{k}'), \hat{S}(\mathbf{k})\} = 0. \tag{4c}$$

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 (4a)), have well-behaved estimation variance (equation (4b)), and map all power at the appropriate frequencies without leakage (equation (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]

$$\gamma^2(\mathbf{k}) = \frac{|E\{F(\mathbf{k})G^*(\mathbf{k})\}|^2}{E\{G(\mathbf{k})G^*(\mathbf{k})\}E\{F(\mathbf{k})F^*(\mathbf{k})\}}, \tag{5}$$

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 the Appendix].

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 (5) would be equal and the coherence exactly unity. By averaging over N different spectral estimates we are attempting to satisfy (4a), hoping that the estimate $\hat{S}$ (in the case of (5), this is the periodogram) is asymptotically unbiased in the sense

$$\lim_{N\to\infty} E\{\hat{S}(\mathbf{k})\} = S(\mathbf{k}). \tag{6}$$

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.

Figure 1 illustrates the importance of averaging. In Figure 1a, two time series are given. Both have peaks in the power spectrum representing two distinct frequencies (Figure 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) and (5)), the coherence equals unity over the entire range of frequencies (Figure 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 (5) before the division (a modified periodogram procedure called Welch Overlapping Segment Algorithm (WOSA) [*Welch*, 1967]), correctly yields coherence peaks at both input frequencies (Figure 1c, shaded area). With the multitaper method averages are made over different windowed versions (described in detail in the Appendix) of the entire data set, and the coherence also peaks at both frequencies (Figure 1c, open area).

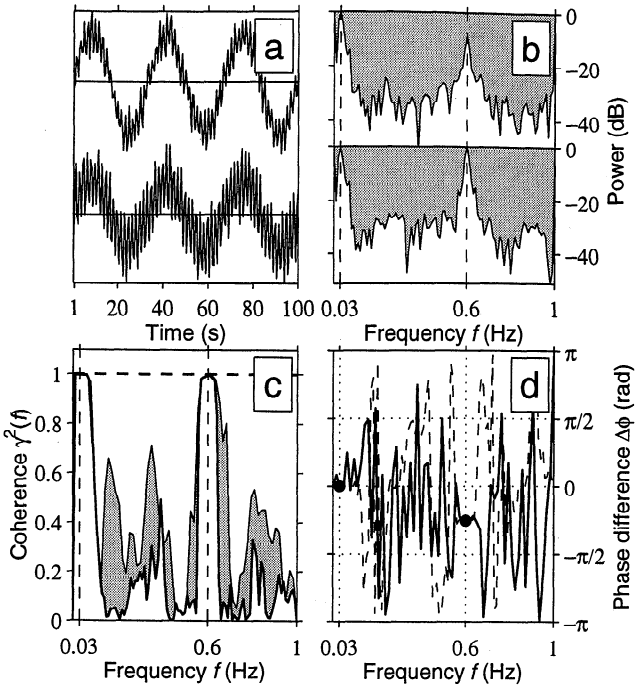

Figure 1. The importance of averaging in coherence estimation. (a) Top signal: $3\sin(3/50\pi t) + 1\cos(6/5\pi t)$. Bottom signal: $2\sin(3/100\pi t) + 2\cos(3/5\pi t + \pi/4)$; 15% white noise has been added to each. The phase relationship between the two time series is consistent over the length of 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).

The broadness of the multitaper coherence peaks is due to the fixed bandwidth of the Slepian tapers (see *Riedel and Sidorenko* [1995] and the Appendix), 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 (Figure 1d, dashed line) and the multitaper method (Figure 1d, solid line) pick out the input 0 and $-\pi/4$ phase shift at 30 and 600 mHz, respectively (Figure 1d, solid circles).

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.

2.3.1. 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 Figure 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].

2.3.2. 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].

2.3.3. 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 down-weighted 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. Multitaper Method for Two-Dimensional 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

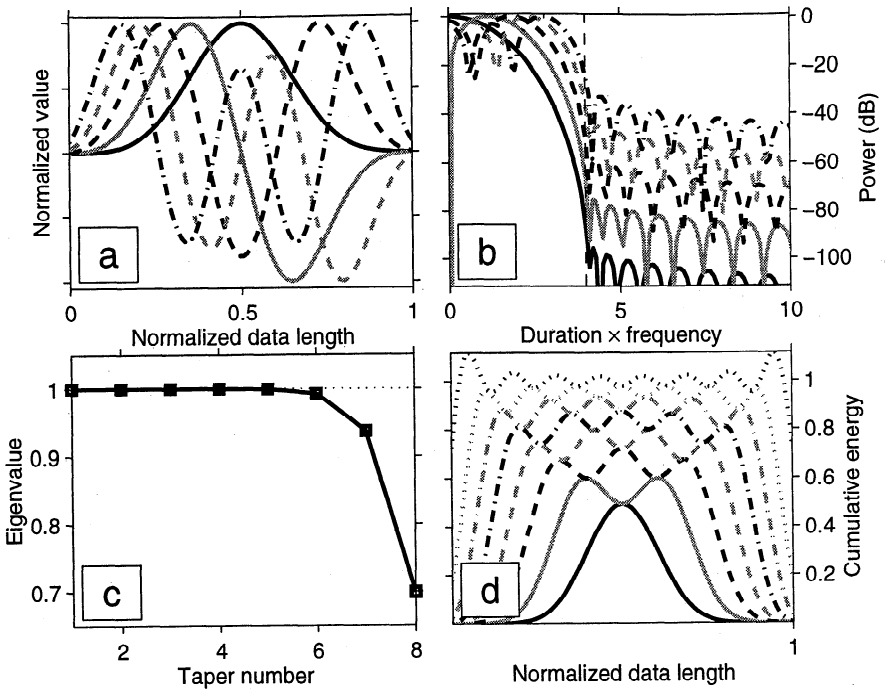

Figure 2. Properties of the dpss. (a) Data tapers $h_{t,k}$ for $NW = 4$ and $k = 1, ..., 5$ with normalized height and as a fraction of the data length N. (b) Spectral windows $\mathcal{H}_k(f)$ belonging to the tapers to the left. The sidelobe level, and hence the bias in the spectral estimation, of 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.

k of independent direct spectral estimates $\hat{S}_l$ with weights λ_l. The eigenspectra $\hat{S}_l$ that satisfy (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 (4b). To minimize spectral leakage in the sense of (4c), we need to find data windows $h_l(\mathbf{r})$ whose spectral responses $\mathcal{H}_l(\mathbf{k})$ (see (A4)) 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 (A3) and (A6)) 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 the Appendix), 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 (A8) and (A9)). 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 paper, 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 Figure 2. In Figure 2a the first five tapers of a sequence with $NW = 4$ are given. We assume $\Delta t = 1$ and nondimensionalize the frequency axis by plotting Nf, i.e. duration times frequency, so the central lobe is confined within $[-NW, +NW]$ (Figure 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 (Figure 2c). Conventionally, only the first $2NW - 1$ eigentapers are used. Most popular data windows in nonmultitaper applications are fairly similar to the zeroth-order multitaper window. The disadvantage of 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

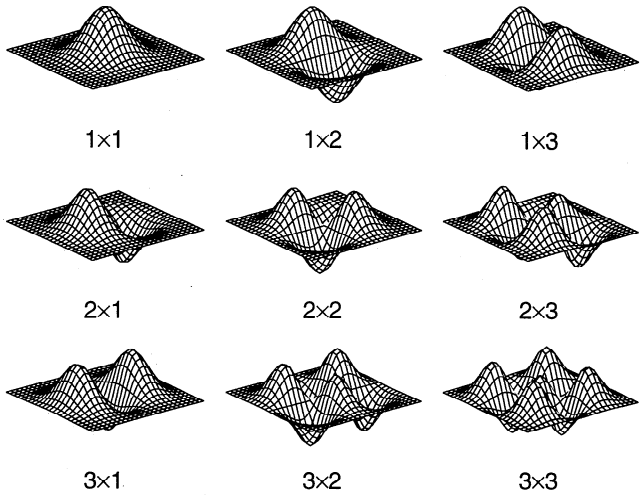

Figure 3. A subset of the two-dimensional tapers used in the analysis, formed by taking the outer product of any two combinations $k \times k'$ (indicated in the legend) of the 1–D tapers $h_{t,k}$ of Figure 2.

extracts information more and more evenly from the entire signal, without this arbitrary down-weighting, as shown in Figure 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 Figure 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 (A15).

The spectral windows $\mathcal{H}(\mathbf{k})$ associated with the tapers of Figure 3 are plotted as contour lines in Figure 4. As in Figure 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 $+4$ (Figure 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 Figure 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.

4. Tests With Synthetic Data

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 section 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 investigated without resorting to smoothing or binning procedures in wave vector space, procedures which, as we have argued in section 2.3.3, are inherently flawed.

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

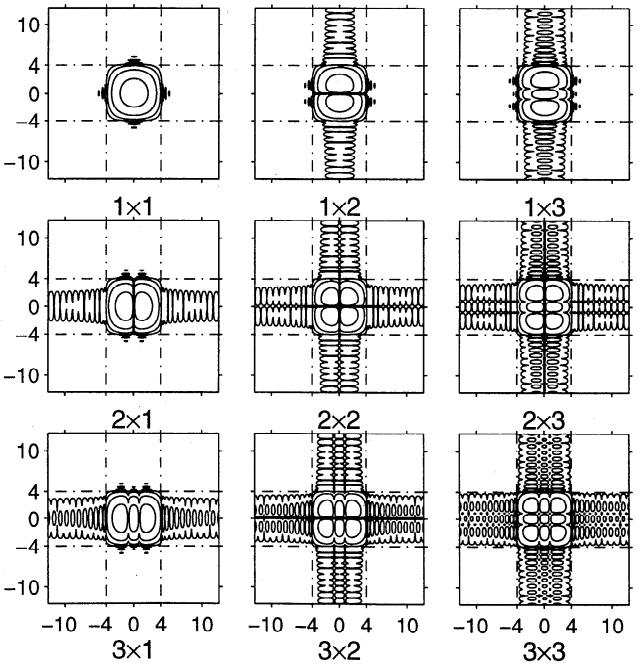

Figure 4. Spectral windows $\mathcal{H}_k(\mathbf{k})$ (see (A4)) associated with the tapers of Figure 3, plotted as functions of the duration × half bandwidth product. $NW = 4$. Contouring is at -105, -75, -45, and -15 dB.

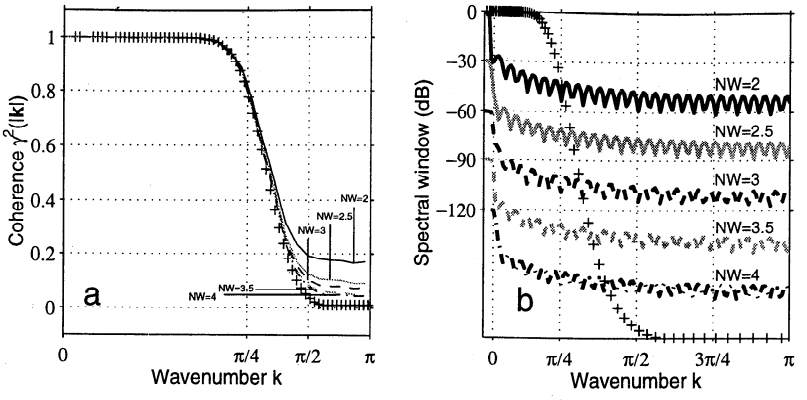

Figure 5. Influence of resolution bandwidth NW upon synthetic data with isotropic coherence. Frequency space is nonphysical, hence contained within angular frequency between 0 and $+\pi$. (a) Isotropic input coherence (pluses). In shaded symbols, multitaper estimates for NW values of 2, 2.5, 3, 3.5, and 4. $NW = 4$ is the value we adopt in the subsequent analyses. The x scale is logarithmic. (b) Spectral windows $\bar{\mathcal{H}}(f)$ (see (A7)), 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.

the isotropic or 1-D case some thought must be given to the selection of the resolution parameter NW.

As indicated in section 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, $\gamma^2 \to 1$. 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 $\gamma^2 \to 0$. 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

$$|N(\mathbf{k})|^2 = |H(\mathbf{k})|^2 \frac{1 - \gamma^2(\mathbf{k})}{\gamma^2(\mathbf{k})}, \tag{7}$$

the coherence between $H(\mathbf{k})$ and $H(\mathbf{k}) + N(\mathbf{k})$ is given by $\gamma^2(\mathbf{k})$. Thus to make $\gamma^2 \to 0$ 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 $\gamma^2 \to 1$ 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).

4.3. Tests With Synthetic (An)isotropic Coherence

Synthetic tests are presented in Figure 5 and Plate 1. 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 $k = 0 \to \pi$.

Figure 5 depicts the results of a first test. The input coherence function (pluses), given in Figure 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 paper have used $NW = 4$ and $k = 7$ 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 section 6.2.3.

For the set of tapers used, Figure 5b gives the average spectral windows, $\bar{\mathcal{H}} = \sum_{l=1}^{k} \mathcal{H}_l(\mathbf{k})$ (see (A6)) 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 Figure 5a on top of them on the same linear x scale but scaled to fit the plot box in the y direction.

In Plate 1 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 Plates 1a and 1b, 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 Plates 1c–1f, anisotropic input fields are analyzed. The input is found in Plates 1c and 1e, whereas the multitaper measurements are found in Plates 1d and 1f. 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.

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.

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 nonparametric. 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 nonsubjective way (from the shape of the spectral windows in Figures 2b, 4, and 5b), and given the multiplicity of the windows, the confidence intervals around the spectrum can be reduced dramatically compared to any other method.

5. Application to the Study of Australia

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.*, 1999b]. At this scale the three domains represent the dominant age provinces as well. For a tectonic map, see Figure 6.

5.1.1. 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.*, 2000] 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].

5.1.2. 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].

5.1.3. 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

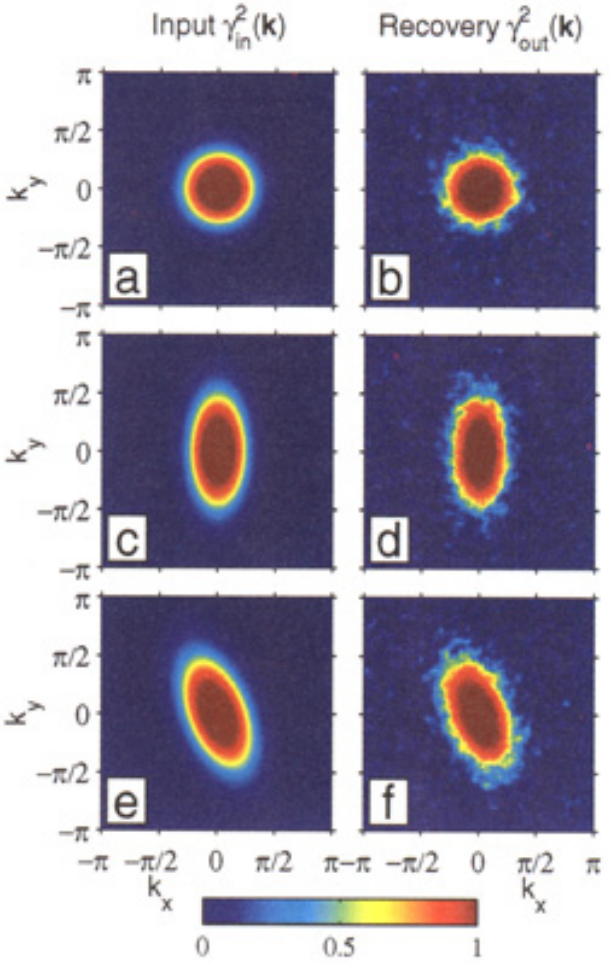

Plate 1. Synthetic (an)isotropic coherence with variable rotation: (a,c,e) input coherence, and (b,d,f) recovery with multitaper method, $NW = 4$.

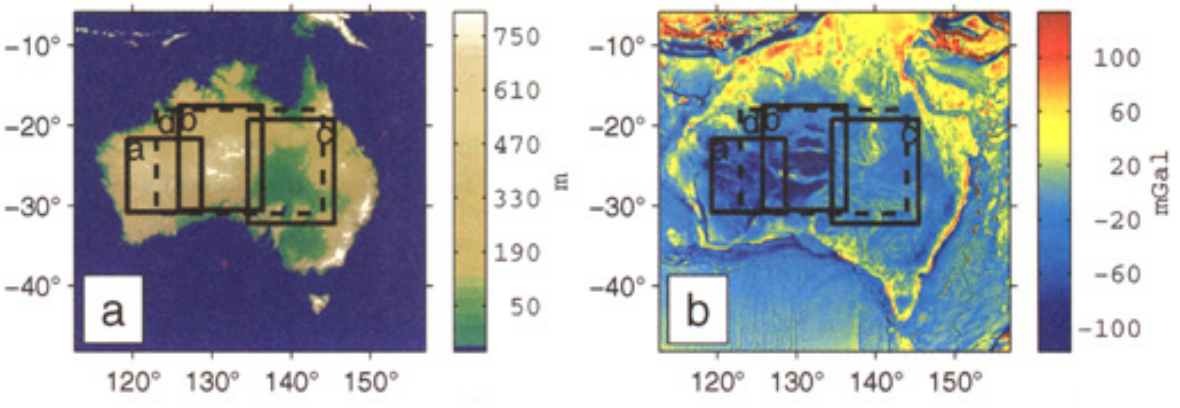

Plate 2. The Australian continent and the four regions we selected for the analysis: (a) GTOPO30 topography and (b) Bouguer (continent) and free air (ocean) gravitational anomaly from AGSO. The lettered boxes (lettering as in Plate 3 and Figures 7, 9 and 10) represent four sets of data used in the analysis.

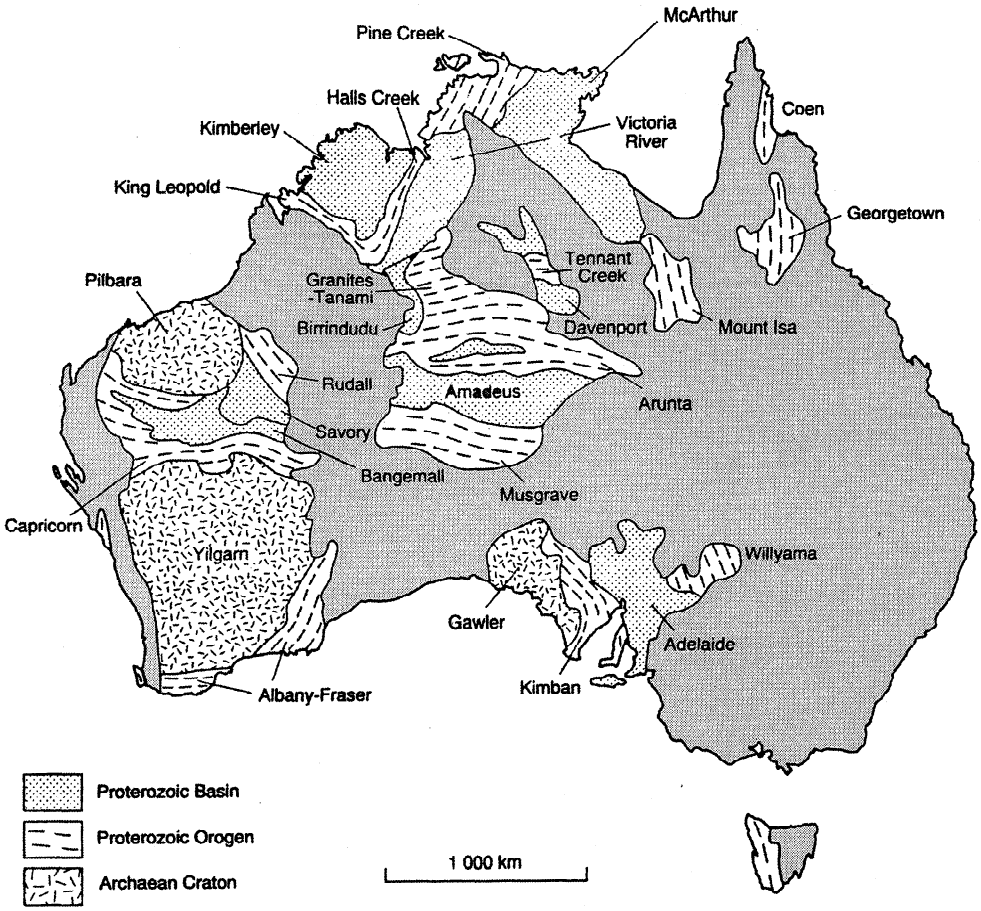

Figure 6. Tectonic map of Australia [from *Myers et al.*, 1996].

in close to local isostatic equilibrium [*Murray et al.*, 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.*, 1999b]: 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 paper 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 Plate 2) 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 Plate 2). 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].

5.2. Gravity and Topography Data

Gravity data were obtained from the Australian Geological Survey Organisation (AGSO; see Plate 2b), 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.

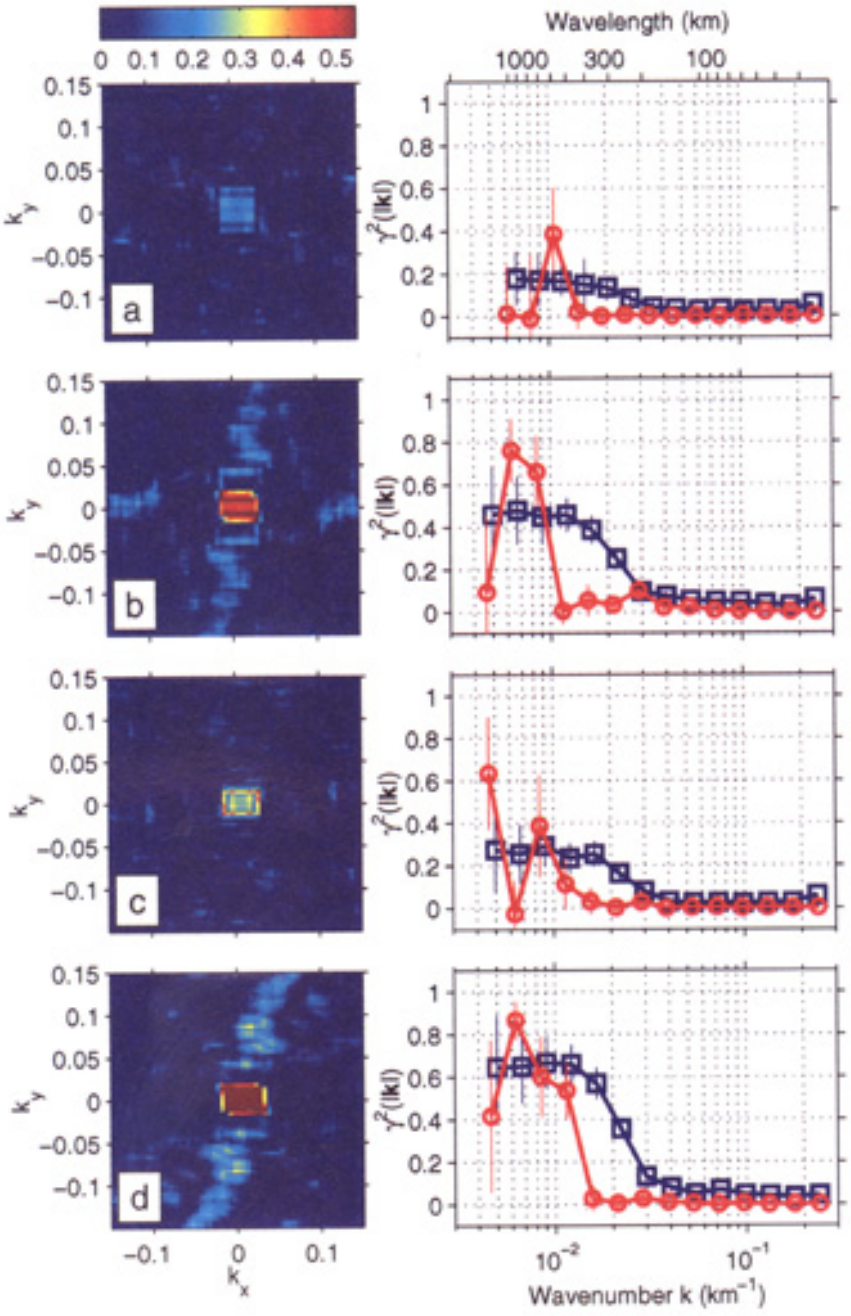

Plate 3. Coherence measurement of four regions of Australia. The letters a through d correspond to the labeled boxes in Plate 2. (left) Multitaper 2–D coherence used as an indication of anisotropy in the isostatic response. We interpret the shape of the 2-D coherence of central Australia (Plate 3b) as 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 sections 6.2.2 and 6.2.3 for a discussion on the low coherence values at low wave numbers.

Topography data (see Plate 2a) 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 section 6.2.2. Finally, before Fourier transformation, a best fitting plane trend was removed from the data.

6. Results

6.1. Anisotropy in the Isostatic Response

In Plates 3a–3c, the results for each of the three geophysically distinct regions (labeled a through c in Plate 2) are summarized. Plate 3d analyzes the data box d marked by a dashed line in Plate 2; 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]:

$$\gamma^{2'}(k) = \frac{n(k)\gamma^2(k) - 1}{n(k) - 1}. \tag{8}$$

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]

$$\Delta\gamma^{2'}(k) = (1 - \gamma^2(k))\sqrt{\frac{2\gamma^2(k)}{n(k)}}. \tag{9}$$

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 (A15). 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 (Plates 3a and 3c) show little or no indication of anisotropic isostatic response. In contrast, the central Australian domain (Plate 3b) 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 section 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 (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 (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 (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 section 6.2.7).

6.2. Isotropic Elastic Thickness Inversions

6.2.1. 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). Figure 7 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 uncorre-

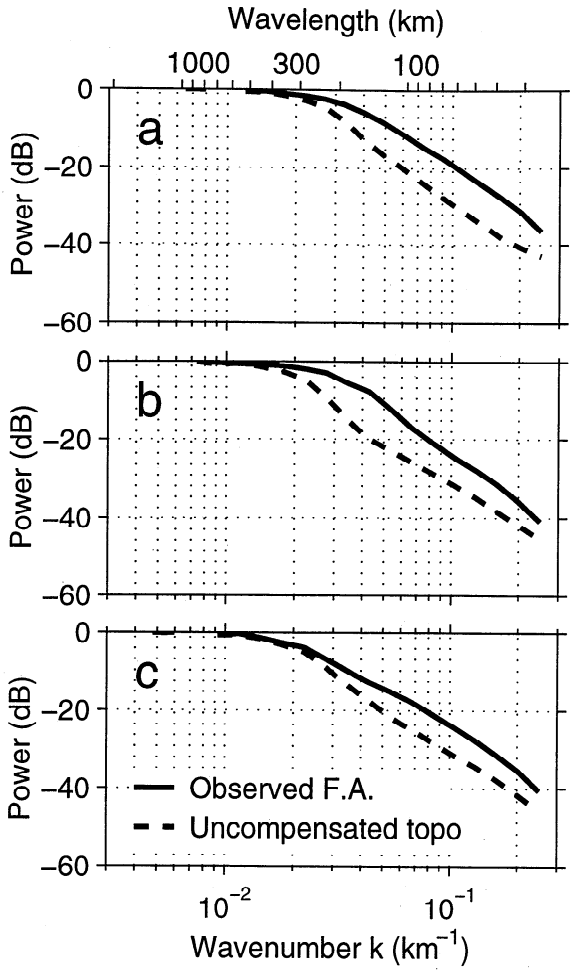

Figure 7. Power in the free-air gravity field compared to the power of the gravitational attraction due to uncompensated topography. The letters a, b, and c correspond to the labeled boxes in Plate 2.

lated 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 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.

6.2.2. Influence of the degree of the subtracted geoidal field. In Plate 3 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 section 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 paper are minor.

6.2.3. Influence of resolution parameter *NW*. For a large part this paper 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 averaging. The synthetic tests presented in section 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 win-

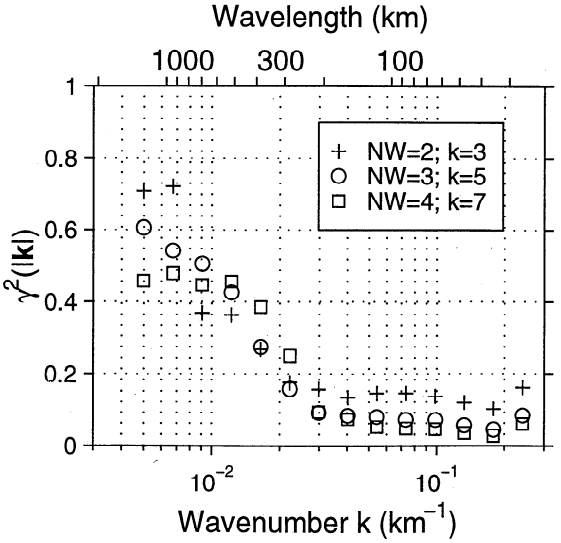

Figure 8. Influence of resolution parameter NW on the central Australian coherence data (box b in Plate 2). The physical resolution $W = j/(N\Delta t)$ for $j = 2, 3, 4$ (quoted as $NW = 2, 3, 4$) depicted in this figure corresponds for this data set to a resolution in wavenumber space of 0.01, 0.015,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

dows $k = 2NW - 1$. In Figure 8 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 Figure 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.

6.2.4. 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.

6.2.5. 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.

6.2.6. 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.* [2000]) 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 Plate 2 are given in Figures 9a–9c. 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 Figure 8 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) (Figure 9b). We have kept the value of that minimum error in choosing the optimal T_e for both western (Figure 9a) and eastern Australia (Figure 9c). 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].

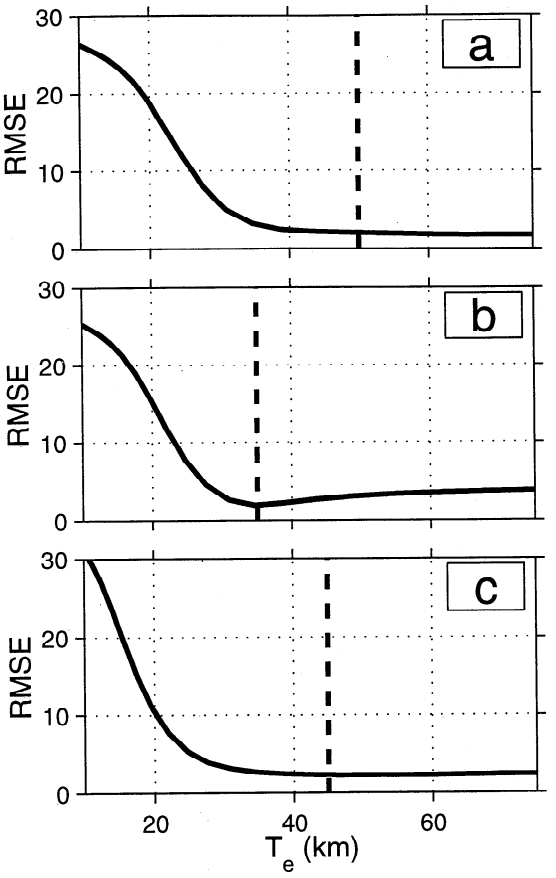

Figure 9. Model misfit in the inversion of coherence for effective elastic thickness for the three analysis regions corresponding to the lettered boxes in Plate 2. RMSE, root-mean-square error.

6.2.7. Elastic thickness values. Figure 10, 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. Figures 10a–10c give the multitaper results (solid), for the three regions labeled a through c in Plate 2. Figures 10d–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.* [2000], were fixed at 50 km (Figures 10a and 10d), 40 km (Figures 10b and 10e) and 35 km (Figures 10c and 10f).

As we have seen in Plate 3, 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

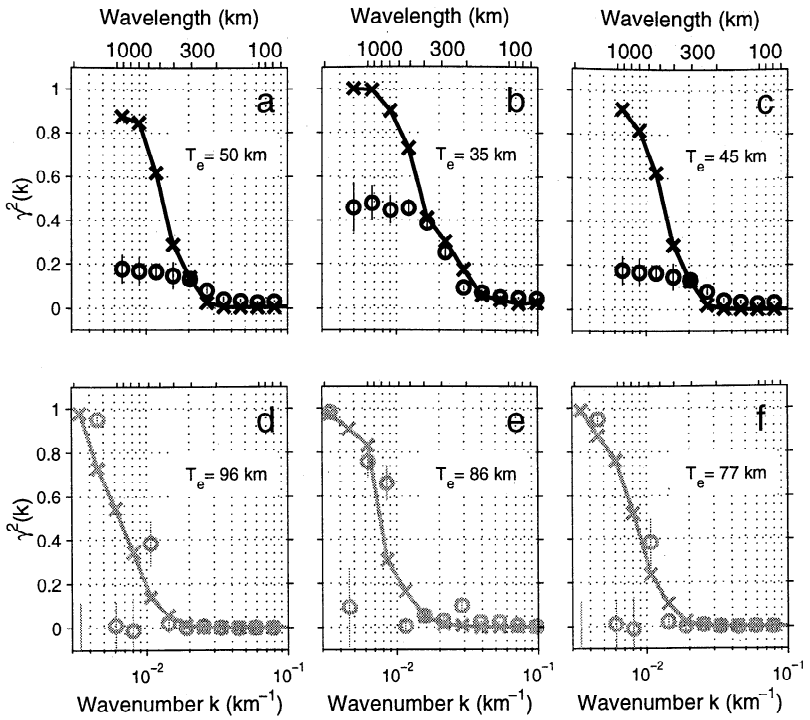

Figure 10. Isotropic T_e inversions of the three boxes lettered a through c in Plate 2. (a–c) Inversion with multitaper coherence estimates. Circles, measurements with 2σ error bar; crosses, prediction. (d–f) Same for periodogram measurements. The inversion for both coherence measurement methods is identical: 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 Figures 10a and 10d, 45 km for Figures 10b and 10e and 35 km for Figures 10c and 10f.

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].

7. Discussion

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.

7.1.1. 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.

7.1.2. 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].

7.1.3. 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.

7.1.4. 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.

7.1.5. 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].

7.1.6. 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.*, 1999a], as does our ability to detect it beneath thick cratons.

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.

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; *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.*, 1999b] (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.*, 1999b] and its azimuthal anisotropy [*Simons et al.*, 1999a] and the long-term mechanical isostatic response.

8. Conclusions

The focus of this paper 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.

Appendix: A. 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}:

$$\gamma^2_{fg}(\mathbf{k}) = \frac{|S_{fg}(\mathbf{k})|^2}{S_{ff}(\mathbf{k})S_{gg}(\mathbf{k})}. \tag{A1}$$

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

$$\hat{S}(f) = \Delta t \left| \sum_{t=1}^{N} h_t x_t e^{-i2\pi f t \Delta t} \right|^2 = \Delta t\, Y(f) Y^*(f), \tag{A2}$$

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

$$E\{\hat{S}(f)\} = \int_{-1/(2\Delta t)}^{1/(2\Delta t)} \mathcal{H}(f - f') S(f')\, df'. \tag{A3}$$

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

$$\mathcal{H}(f) = \frac{1}{\Delta t} |H(f)|^2 = \Delta t \left| \sum_{t=1}^{N} h_t e^{-i2\pi f t \Delta t} \right|^2. \tag{A4}$$

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 $\hat{S}_l, l = 1...k(f)$

$$\hat{S}^{\mathrm{MTM}}(f) = \frac{1}{k} \sum_{l=0}^{k-1} \hat{S}_l(f), \tag{A5}$$

where each $\hat{S}_l$ is given by an equation of the kind of (A2). From (A3) it then follows that the first moment of this multitaper spectral estimate is given by

$$E\{\hat{S}^{\mathrm{MTM}}(f)\} = \int_{-f_N}^{f_N} \bar{\mathcal{H}}(f - f') S(f')\, df', \tag{A6}$$

with

$$\bar{\mathcal{H}}(f) = \frac{1}{k} \sum_{l=0}^{k-1} \mathcal{H}_l(f). \tag{A7}$$

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 (A4), the frequency concentration $\beta^2(W)$ of the data tapers is measured as the fraction of energy in this frequency interval, by

$$\beta^2(W) = \int_{-W}^{W} \mathcal{H}_l(f)\, df \Big/ \int_{-f_N}^{f_N} \mathcal{H}_l(f)\, df\,, \tag{A8}$$

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

$$\alpha^2(N) = \sum_{t=0}^{N-1} |h_{t,l}|^2 \Big/ \sum_{t=-\infty}^{\infty} |h_{t,l}|^2\,, \tag{A9}$$

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 $|f| \le W < f_N$. The concentration problem amounts to an optimization. It is straightforward to rewrite the concentration measure of (A9) in the frequency domain as

$$\alpha^2(N) = \left(\int_{-W}^{W} |H_l(f)|^2\, df \right)^{-1} \int_{-W}^{W} \int_{-W}^{W} H_l(f) N \mathcal{D}(f, f'; N) e^{i\pi(N-1)(f-f')} H_l^*(f')\, df\, df', \tag{A10}$$

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

$$\mathcal{D}(f, f'; N) = \frac{1}{N} \frac{\sin[N\pi(f' - f)]}{\sin[\pi(f' - f)]}. \tag{A11}$$

We can recognize (A10) 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 (A10) will be extremized if the following equation is satisfied:

$$\int_{-W}^{W} N \mathcal{D}(f, f'; N) e^{i\pi(N-1)(f-f')} H_l(f)\, df = \alpha_l^2 H_l(f'), \tag{A12}$$

for $|f'| \le W$. It is convenient to reformat (A12) to the time domain. After some manipulation (as given by *Park et al.* [1987a] and *Lindberg and Park* [1987]), we can rewrite it as

$$\mathbf{C} \cdot \mathbf{h} = \alpha^2(N)\mathbf{h}, \tag{A13}$$

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

$$\mathbf{C}_{tt'} = \frac{\sin[2\pi W(t-t')]}{\pi(t-t')} \quad ; \quad t, t' = 0, 1, ..., N-1, \tag{A14}$$

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 Figure 2).

The data tapers are orthogonal. This causes the individual eigenspectra of (A3) 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 (A14)), (2) windowing the data, (3) taking the Fourier transform of the windowed data and calculating eigenspectra as in (A2), and (4) performing the average as in (A5), weighted with the eigenvalues. The calculation of the cross-spectral density function in order to arrive at the coherence function of (A1) is as in (A2), 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

$$\sigma^2 = \frac{k-1}{k(k+1)}\left[\frac{1}{k} + 2\frac{k-2}{k+2}\sqrt{\hat{\gamma}^2_{fg}}\right]. \tag{A15}$$

This measure gives the variance of the 2-D coherence estimation. In case the coherence is averaged azimuthally, this 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 Figure 3). So for n dimensions, the number of tapers is $K = k^n$. The associated spectral windows have been given in Figure 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 A3). 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.

Acknowledgments. We acknowledge numerous discussions with Sébastien Chevrot, Brad Hager, Tom Jordan, Anthony Lowry, Philip Tracadas, Rob van der Hilst, and Tony Watts. The insightful and constructive reviews by Gareth Armstrong, Cindy Ebinger, and Chris Swain are greatly appreciated. This research was supported by NSF grant EAR-9614341 and NASA grant NAG5-3264.

References

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., 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.

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

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.

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.

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.

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

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.

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.

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

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.

Djomani, Y. H. P., 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.

Dooley, J. C., Implications of Australian seismic and gravity measurements for the structure and composition of the upper mantle, *BMR J. Aust. Geol. Geophys.*, pp. 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.

Drummond, B. (Ed.), *The Australian Lithosphere*, Geol. Soc. Aust., Sydney, 1991.

Ebinger, C., and N. Hayward, Soft plates and hot spots: Views from Afar, *J. Geophys. Res.*, *101*, 21,859–21,876, 1996.

Finlayson, D. M., Geophysical differences in the lithosphere between Phanerozoic and Precambrian Australia, *Tectonophysics*, *84*, 287–312, 1982.

Forsyth, D. W., Subsurface loading and estimates of the flexural rigidity of continental lithosphere, *J. Geophys. Res.*, *90*, 12,623–12,632, 1985.

Fowler, C., *The Solid Earth*, Cambridge Univ. Press, New York, 1990.

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.

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.

Gough, D. I., Electrical conductivity under western North America in relation to heat flow, seismology and structure, *J. Geomagn. Geoelectr.*, *26*, 105–123, 1974.

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.

Hanssen, A., Multidimensional multitaper spectral estimation, *Signal Process.*, *58*, 327–332, 1997.

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.

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.

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 thermomechanical properties of the continental lithosphere, *Nature*, *304*, 250–253, 1983.

Kay, S. M., and S. L. Marple, Spectrum analysis: A modern perspective, *Proc. IEEE*, *69*, 1380–1419, 1981.

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.

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.

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.

Lees, J. M., and J. Park, Multiple-taper spectral analysis: A standalone 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.

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.

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, 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.

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.

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.

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.

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.

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.

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.

Percival, D. B., and A. T. Walden, *Spectral Analysis for Physical Applications, Multitaper and Conventional Univariate Techniques*, Cambridge Univ. Press, New York, 1993.

Phillips, R. J., and K. Lambeck, Gravity fields of the terrestrial planets: Long-wavelength anomalies and tectonics, *Rev. Geophys.*, *18*, 27–76, 1980.

Polet, J., and D. L. Anderson, Depth extent of cratons as inferred from tomographic studies, *Geology*, *23*, 205–208, 1995.

Priestley, M., *Spectral Analysis and Time Series*, Academic Press, San Diego, Calif., 1981.

Riedel, K. S., and A. Sidorenko, Minimum bias multiple taper spectral estimation, *IEEE Trans. Signal Process.*, *43*, 188–195, 1995.

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.

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.

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.

Simons, F. J., R. D. van der Hilst, and J.-P. Montagner, Azimuthal anisotropy of the Australian upper mantle (abstract), *Eos Trans. AGU*, *80(46)*, Fall Meet. Suppl., F729, 1999a.

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, 1999b.

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.

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.

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.

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.

Timoshenko, S., and S. Woinowsky-Krieger, *Theory of Plates and Shells*, 2nd ed., McGraw-Hill, New York, 1959.

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.

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.

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.

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.

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.

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., 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.

J. Korenaga, F. J. Simons, and M. T. Zuber, Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Rm 54-517A, Cambridge, MA 02139.
(fjsimons@mit.edu)

(Received November 17, 1999; revised March 16, 2000;
accepted April 26, 2000.)