JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 107, NO. B5, 2094, 10.1029/2001JB000498, 2002

Constraints on the *S* wave velocity structure in a continental shield from surface wave data: Comparing linearized least squares inversion and the direct search Neighbourhood Algorithm

J. Arthur Snoke

Department of Geological Sciences, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA

Malcolm Sambridge

Research School of Earth Sciences, Australian National University, Canberra ACT, Australia

Received 28 February 2001; revised 13 December 2001; accepted 18 December 2001; published 18 May 2002.

[1] In their study of upper mantle structure beneath the Paraná Basin of SE Brazil, *Snoke and James* [1997] concluded, on the basis of a linearized least squares inversion (LLSI) of surface wave dispersion data, that a strong (5% contrast) low-velocity zone (LVZ) beginning at a depth less than ~150 km was not required to fit the data. They were unable to establish a quantitative estimate, however, on the maximum depth at which such a LVZ could be resolved by their data. *Sambridge* [1999a, 1999b] has introduced the Neighbourhood Algorithm (NA), a direct search method for nonlinear inversion which can be tuned to extract information from an ensemble of models in addition to finding a single best fit model. Applying NA to the Brazilian dispersion data quantifies the statistics of the ensemble of models classified as ''acceptable'' based on a data misfit criterion and a smoothness constraint. The NA best fit model is not significantly different from the LLSI best fit model, but the analysis of the ensemble of models provides new insights regarding how well constrained the model is. Synthetics runs show that for this data set, our modeling procedures could resolve a strong LVZ that began at a depth of 120 km but could not rule out such an LVZ beginning at a depth of 180 km. *INDEX TERMS:* 7218 Seismology: Lithosphere and upper mantle; 3260 Mathematical Geophysics: Inverse theory; 8120 Tectonophysics: Dynamics of lithosphere and mantle—general; 8162 Tectonophysics: Rheology—mantle; 7255 Seismology: Surface waves and free oscillations; *KEYWORDS:* direct search methods for nonlinear inversion, surface waves, continental lithosphere-asthenosphere rheology, Neighbourhood Algorithm

1. Introduction

[2] For oceanic paths, using linearized least squares inversion (LLSI), the depth to the low-velocity zone (LVZ) which characterizes the asthenosphere, can be found if observed surface wave dispersion is well constrained to periods of at least 80 s [e.g., *Woods and Okal*, 1996; *Priestley and Tilmann*, 1999]. In continental shield regions the lithosphere is far thicker than under oceans, perhaps in excess of 250 km [e.g., *VanDecar et al.*, 1995; *Ritsema and van Heijst*, 2000; *James et al.*, 2001].

[3] Even for oceanic paths, LLSI provides little information about the range of physically acceptable models which fit the data, and because of the heavy damping required to stabilize the iteration process the model resolution and variances produced by LLSI are not easily interpreted. The conclusion drawn by *Snoke and James* [1997] in their LLSI for the *S* wave structure beneath the eastern Paraná Basin in SE Brazil is that no strong (5% velocity contrast) LVZ beginning at a depth shallower than ~150 km was required to fit their surface wave dispersion data. The objective of this paper is to see if further constraints on the velocity structure beneath the Paraná Basin can be obtained by using an alternative procedure to LLSI to analyze the interstation dispersion-velocity data set of *Snoke and James* [1997]. The procedure used is the Neighbourhood Algorithm (NA), developed by *Sambridge* [1999a, 1999b]. NA is a direct search (i.e., derivative-free) method for nonlinear inversion, which can be used to extract information from an ensemble of models in addition to finding a single best fit model.

[4] In sections 2 and 3 we review briefly the LLSI study of *Snoke and James* [1997] and give an outline of the NA. We then present results from the application of NA to the dispersion data set, and we follow this with synthetics tests in which both NA and LLSI are applied to synthetic dispersion data sets generated from velocity structures with well-defined LVZs at depths of 120 and 180 km.

2. Data

[5] The surface wave study presented by *Snoke and James* [1997] is based on data from the Brazilian Lithosphere Seismic Project: an array of portable broadband seismic stations operated in south central Brazil between 1992 and 1995, with the objective of mapping heterogeneity in the lithosphere and upper mantle in the region and correlating it with the principal tectonic provinces. The entire region, including the Andean basement, is composed of what is loosely termed the Brazilian shield, a patchwork of cratonic nuclei of varying ages, most of which were amalgamated during late Proterozoic time (Brasiliano/Pan-African, circa 600 Ma), but

0148-0227/02/2001JB000498$09.00

some of which in the west and south were accreted through mid-Paleozoic time during the continuing consolidation of the Gondwana supercontinent [see *Brito Neves and Cordani*, 1991; *Ramos*, 1988]. The interstation paths of interest here traverse the eastern part of the intracratonic Paraná Basin (Figure 1). The present study is based on four events of the seven used by *Snoke and James* [1997]. Three events were removed because it was decided that they were not pure paths [*Assumpção et al.*, 2002]: the eastern stations for those paths were in either the Brasília Belt (a mobile belt of Proterozoic terranes mobilized during collision of the Paraná Basin and São Craton in Brasiliano time circa 600 Ma) or the São Francisco craton (composed of Archean and Paleoproterozoic rocks).

[6] Two other studies using data from the Brazilian Lithosphere Seismic Project have some overlap with our study region: a three-dimensional (3-D) body wave tomography analysis [*VanDecar et al.*, 1995], and a 3-D surface wave tomography study [*van der Lee et al.*, 2001]. The data used in our studies have better vertical resolution than either of the tomography studies for depths shallower than ~200 km depth. A receiver function analysis for the crustal structure near several Brazilian stations [*Assumpção et al.*, 2002] helped constrain the crustal structure for our surface wave studies.

[7] The data in the Brazilian Lithosphere Seismic Project were recorded on broadband STS-2 seismometers which have a flat velocity response from 0.008 to 50 Hz with no clipping for the

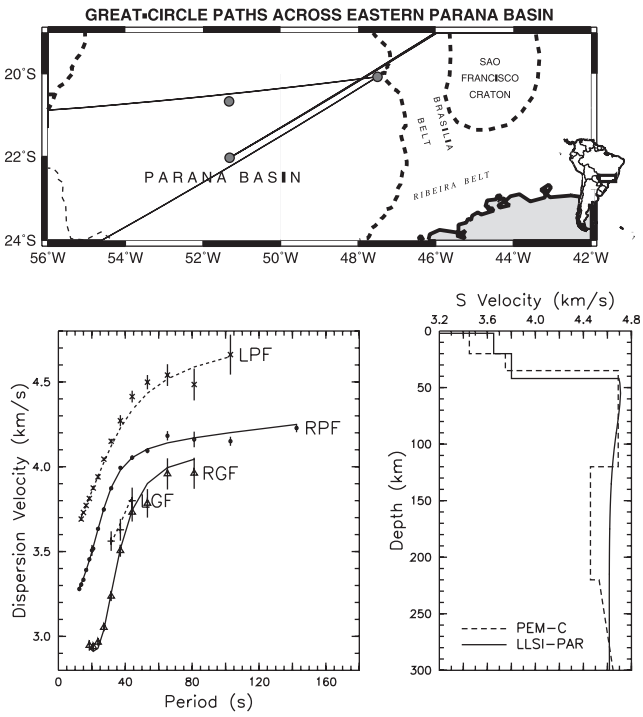

Figure 1. (top) Four interstation paths in the Paraná Basin used for surface wave analysis. (Three are between the same two stations.) (bottom left) Fundamental mode Rayleigh and Love phase and group velocities. Symbols are the data (averaged over up to four events for each period and mode), and lines are calculated from the best fit LLSI velocity model. Each dispersion velocity datum is a composite from up to four events, weighted by the inverse of the estimated variances. (bottom right) Best fit *S* wave model (LLSI-PAR) calculated using LLSI. The LLSI-PAR velocity model has 43 constant velocity layers, and the damping used is 0.1 times the the maximum eigenvalue of the data kernel matrix. Also shown for reference is the continental PEM model (PEM-C) [*Dziewonski et al.*, 1975].

events used. The results shown in Figure 1 are found by applying *Herrmann*'s [1987] SURF package first to find interstation phase and group velocities for both Love and Rayleigh waves and then to invert these for the best fitting *S* wave velocity structure using LLSI. Because of the high damping required to stabilize the inversion, the error estimates calculated using LLSI for model LLSI-PAR are <0.004 km/s for all depths and decrease for depths >150 km, both inappropriate results.

[8] There is one difference in the data preparation between this study and that of *Snoke and James* [1997]. In the earlier work, each interstation velocity set was culled to the same set of discrete periods but no averaging was done over events. Thus, for our four events, there could be as many as four dispersion values for a given mode and period. Here we use the direct search NA to generate a total of 10,000 velocity models in each run. To reduce the computational cost of forward modeling, we decrease the number of dispersion values by averaging over events for each mode and period. The dispersion data used in the present study are as shown in Figure 1. We ran LLSI with the original (74) and also with the modified (44) dispersion data, and results did not differ significantly. (There was also no significant change in the final model after reducing the number of events from seven to four.) We refer the reader to *Snoke and James* [1997] for further details on the data preparation and LLSI inversion.

3. The Neighbourhood Algorithm

[9] The Neighbourhood Algorithm (NA), introduced by *Sambridge* [1999a, 1999b], is a direct search method for nonlinear inversion. This approach is applicable to a wide range of inversion problems, particularly those where the relationship between the observables (data) and the unknowns (a finite set of model parameters) is rather complex and derivatives may be expensive or cumbersome to calculate, e.g., fitting of seismic waveforms for earth structure or source parameters.

[10] The approach is divided into two stages. In the first, known as the search stage, one samples a multidimensional parameter space for combinations of parameters (models) which provide satisfactory fit to observed data. In the second, known as the appraisal stage, one extracts information from the complete ensemble of models collected in the search stage, e.g., to provide estimates of resolution and variance. The search algorithm is in the same class of technique as genetic algorithms (GA) and simulated annealing (SA) [*Davis*, 1987], in that it uses randomized decisions to drive the search and avoids the need for calculation of derivatives of the data misfit function. These techniques are often associated with global optimization problems. The NA differs from previous techniques in that it requires just two control parameters to be tuned, and the search process is directly driven by only the rank of models with respect to the data misfit criterion, and not the misfit itself. This allows considerable flexibility because any combination of data fit criteria, or other information, can be used to rank models. Recently, the NA has been applied to hypocenter location [*Sambridge and Kennett*, 2001] and seismic source characterization [*Kennett et al.*, 2000; *Marson-Pidgeon et al.*, 2000].

[11] The NA makes use of simple geometrical concepts to search a parameter space. The basic idea is illustrated with a simple example: Figure 2 shows results from a two-parameter problem where the NA has been used to maximize a multi-peaked fitness function in a plane. Figure 2 (top left) shows an initial set of 10 models distributed quasi-randomly. At each stage the entire parameter space is partitioned into a set of Voronoi cells (nearest-neighbor regions), one about each previously sampled model. In this example the distance metric is a simple L2 norm. The Voronoi cells are used to guide subsequent sampling in a randomized fashion. As iterations proceed, the algorithm concentrates sampling in promising regions. Figure 2 (top right) shows the resulting Voronoi cells after 100 new model realizations, and Figure 2

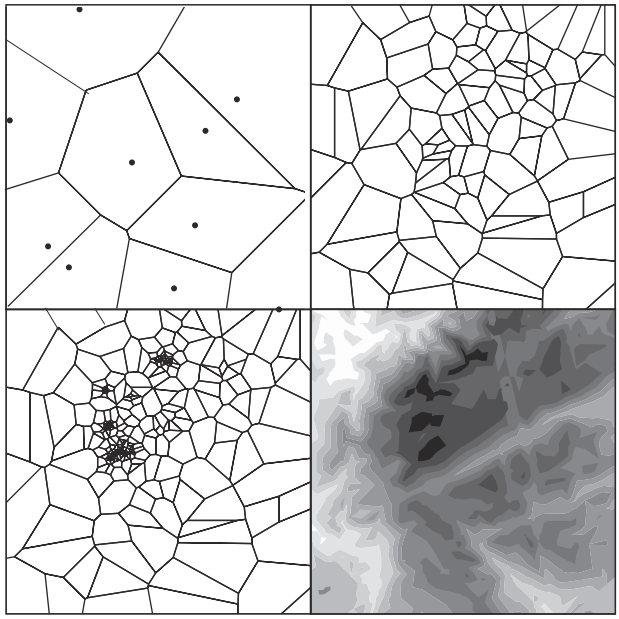

Figure 2. The first three panels show three stages of a Neighbourhood Algorithm search. (top left) Initial 10 uniformly randomly distributed model realizations and their corresponding Voronoi cells. (top right) Voronoi cells of the first 100 models generated by the NA plus the original 10. (bottom left) Voronoi cells after the generation of 500 new models. (bottom right) True fitness landscape. Darker shading indicates higher fitness. With increased sampling in NA, the concentration is much higher in the regions of higher fitness, all four maxima are found by NA.

(bottom left) after 500 new models have been added. Clearly, all four prominent maxima are well sampled. The global maximum was located after 442 samples.

[12] Voronoi cells form a piecewise continuous approximation to the data misfit function at each iteration which is based on all previous sampling. (The data misfit is treated as constant inside each cell.) This neighborhood approximation surface gradually drives the sampling into one or more better (lower misfit) regions simultaneously. *Sambridge* [1999a] has demonstrated that even though the NA is based on simple geometrical principles, it results in a highly self-adaptive search and remains computationally practical even in much higher dimensional spaces (e.g., 10–100). It has also been demonstrated that it is capable of handling misfit surfaces containing multiple minima [*Sambridge*, 1999a, 2001].

4. Analysis

[13] NA requires a parameterization of model space, a data set, a method of forward modeling to calculate data for a given model, a definition of the data misfit, two tuning parameters (n_s, and n_r), and N, the number of iterations. The procedure is as follows:

1. In the initialization stage, an initial set of n_s models are generated uniformly randomly and a data misfit measure is calculated for each model, i.e., forward modeling is performed n_s times.

2. In the generation stage, Voronoi cells (see above) are defined about each of the n_r models with lowest misfit, and a uniform random walk is performed inside each Voronoi cell to generate a total of n_s new models; that is, n_s/n_r models are generated in each cell.

3. In the forward modeling stage, the data misfit is calculated for the n_s new models generated in step 2 and the procedure returns to step 2. As more models are introduced, the size and shape of the Voronoi cells automatically adapt to the previously sampled models.

[14] Steps 2 and 3 constitute an iteration which is then repeated N times resulting in a total of $(N + 1) \times n_s$ model evaluations. The character of the parameter space search will be influenced by the choice of n_s and n_r. In general, larger n_s and/or n_r, the more exploratory the search will be, skipping over local minima but only slowly increasing the concentration of sampling. As the tuning parameters are reduced, the search will be quicker to converge but more akin to a series of local searches and more likely to be entrapped in local minima of the data misfit function. *Sambridge* [1999a] offers some advice on tuning of the two parameters. A slight variation of the above procedure is employed here: We allow the number of new models to vary between the initialization and iteration stages; that is, we generate more models during initialization than at each subsequent iteration. This was generally found to improve performance because the initial set of Voronoi cells are constructed around a larger (uniformly distributed) initial ensemble. The tuning parameters for the three runs discussed here are n_{si} = 500, n_s = 100, n_r = 50, and N = 95, giving a total of 10,000 model evaluations. (Here we define n_{si} to be the value of n_s in stage 1.)

[15] The model parameters used in this study are eight, overlapping, weighted averages over the velocity-depth model (Figure 3) which represent perturbations of a base model (here the LLSI-PAR model shown in Figure 1). The dispersion values do not go to low enough periods to resolve fine structure in the crust, so the crust is represented by a single "box car" which gives uniform weighting. The crustal structure, including the Moho depth (42 km), is felt to be well constrained by the LLSI analysis and the receiver function study mentioned above [*Assumpção et al.*, 2002]. Test runs were made with a model parameterization allowing for a varying Moho, but the resulting models did not differ significantly from those presented here. The overlapping box car and triangles over the mantle depths allow for smooth perturbations of the base model; first-order discontinuities within the mantle cannot be modeled with this parameterization. (Even though the top and bottom of the LVZ in model PEM-C shown in Figure 1 are shown as first-order discontinuities, no evidence for such sharp changes has ever been presented, and because the periods of surface waves considered here are all above 10 s, such sharp discontinuities could not be resolved even if they were there.) The perturbations go to zero at 400 km depth, which is well beyond the resolvable depth for this data set. Ranges for these parameters increase from ±0.6 km/s for the crustal velocities (parameter 1) to ±1.75 km/s for parameter 8, which has a maximum perturbation at 320 km depth. *Herrmann*'s [1987] SURF routines are used for the forward modeling.

[16] The data misfit (ϕ) for each model realization is the length of the error vector with each element weighted by the inverse of that datum's variance. Dispersion velocities calculated from surface wave velocities are constrained by the averaged velocity over depth ranges, rather than the detailed shape, so models with an unphysical "S" shape of velocity versus depth can often be found with low misfits. To down-weight such models, we add a penalty ($\mathcal{P}$) to the calculated misfit if successive parameters from among parameters 2–8 have opposite signs and differ by more than a preset value. The total misfit is then

$$\phi = \sqrt{\frac{\sum_{j=1}^{N} \left[\frac{o_j - c_j}{\sigma_j}\right]^2}{N \sum_{j=1}^{N} \left[\frac{1}{\sigma_j^2}\right]}} + \mathcal{P},$$

where N is the number of dispersions (44 in these runs), o_j and c_j are the observed and calculated values, respectively, for the j^{th} dispersion, σ_j is the estimated standard deviation for o_j, and the penalty $\mathcal{P}$ is zero unless any of the successive parameters in the

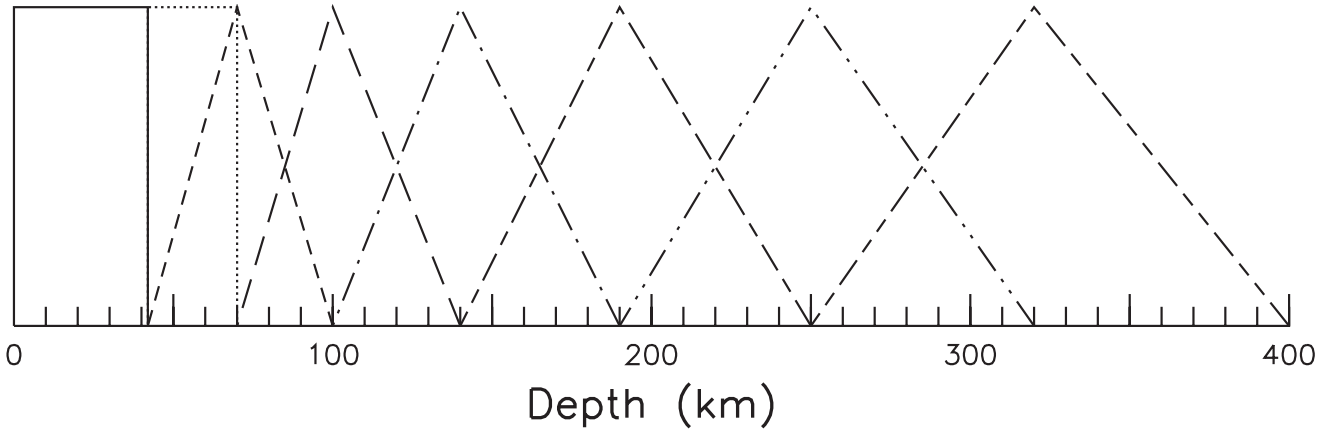

Figure 3. Interpolation model parameters (termed basis functions by *Nolet* [1990]) used in the NA analysis.

range 2–8 for that model realization is $\geq$0.175 for that run. In that case, $\mathcal{P}$ = 5.

5. Results

[17] Figure 4a shows the full ensemble of 10,000 models produced by NA for the Paraná dispersion data set projected onto the first two parameters: the perturbation in the crustal structure and the perturbation in the uppermost mantle velocity. The concentration of sampling increases in the regions of better data fit, but the full range of values between the hard limits is sampled.

[18] Figure 4b is an enlargement of the central portion of Figure 4a. Here the shading scale has been adjusted to cover the range of data misfits in this subregion. The correlations of data fit with position indicate that the V_s crust parameter is more tightly constrained than is the velocity of the uppermost mantle.

[19] Figure 5(top) shows the 71 models and their predicted dispersion curves for all models from the NA run which have misfits <0.01235. This value for the cutoff misfit was chosen so that all models would fit within the estimated errors for the Rayleigh phase velocities, which are the best constrained for this data set. (See Figure 5(top left).) Figure 5(bottom) shows the calculated dispersions and the ''average'' velocity-depth model (NA-PAR) calculated from the acceptable models along with calculated standard deviations at several depths for the 71-model ensemble. Even though the dispersion kernels have little resolving power for depths greater than ~200 km for this data set, the envelope of velocities for acceptable models does not broaden systematically with depth as found in other studies [e.g., *Lomax and Snieder*, 1994]. This indicates that given the smoothness constraint and the assumption that the model perturbation goes to zero at depths $\geq$400 km, the range

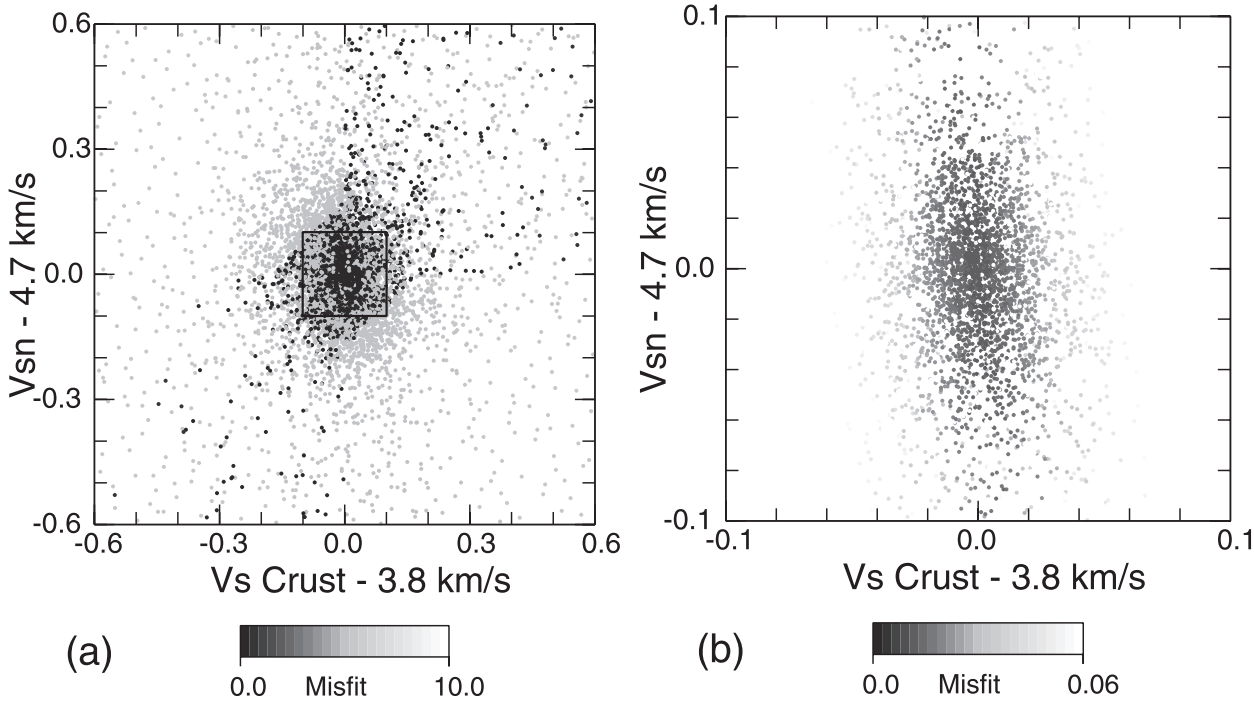

Figure 4. The full ensemble of models produced by the NA for the Paraná dispersion data set projected onto a pair of parameter axes (parameters 1 and 2). Each model is represented by a dot grey scale coded by data misfit. (a) Parameter ranges shown are the prescribed hard limits and the full range of calculated misfits. (b) As in Figure 4a but for a restricted range in parameter values (box in Figure 4a) and in misfits.

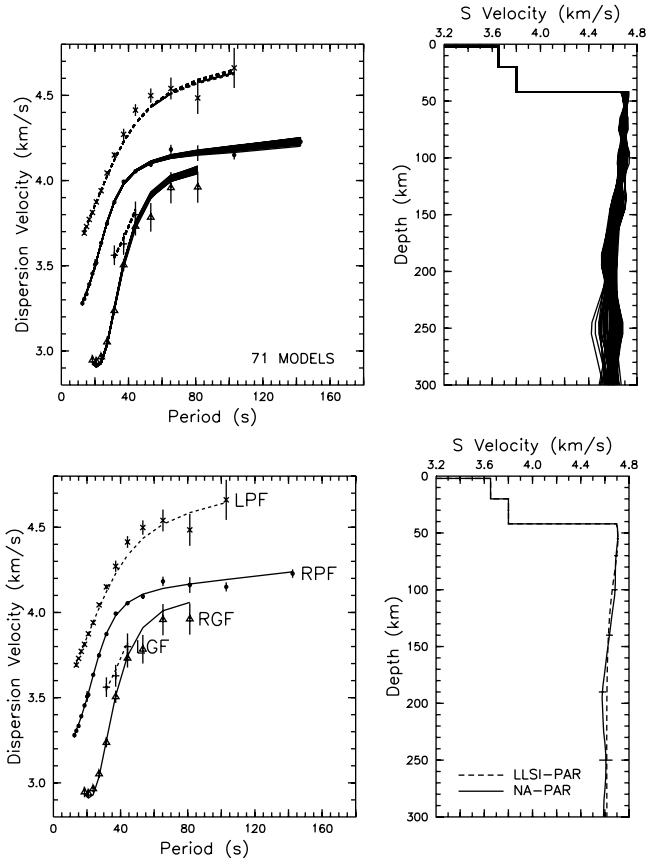

Figure 5. (top) Predicted dispersions and models for the ensemble of all models from the NA run total of 10,000 evaluations which have a misfit <0.01235. Notation as in Figure 1. (bottom) Average model (NA-PAR) and its dispersion calculated from the ensemble shown above. The misfit is 0.0121, just slightly higher than the 0.0120 misfit which is the smallest for the full ensemble. The LLSI-based model, LLSI-PAR, is included for comparison.

in velocities is constrained at all depths for the ensemble of acceptable models.

[20] The LLSI-based model, LLSI-PAR, is included in Figure 5(bottom right) for comparison. The misfits are the same. On the basis of the estimated errors for NA-PAR, the two velocity models are not significantly different.

6. Synthetics Runs

[21] To check our methodology and our interpretation of results, we applied both LLSI and NA to dispersion-velocity data sets which are exact fits for two models with well-defined LVZs. The first one is model PEM-C42, based on the continental PEM-C model of *Dziewonski et al.* [1975] (included in Figure 1) which has a 100-km-thick LVZ starting at a first-order discontinuity at 120 km depth with a 5% velocity contrast. We replaced the two-layer 35-km-thick crust in PEM-C with the 42-km-thick Paranná 2-layer crust overlain by a sedimentary layer, and we replaced the PEM-C velocities below the bottom of the LVZ (220 km depth) with the velocities for those depths from LLSI-PAR. The second model, PEM-C42R, is the same as PEM-C42 except the LVZ is shifted to run from 180 to 280 km depths, and the same procedure was followed as for PEM-C42. The synthetic sets of dispersion values are for the same periods as the Paraná data and have the same estimated errors, but the average values give an exact fit for PEM-C42 or PEM-C42R.

[22] We began the synthetics runs by running LLSI for synthetic data sets, starting in both cases with the LLSI-PAR velocity structure (shown above in Figures 1 and 5) as the initial model. The inversion was significantly more stable than for the real data case, so damping could be reduced to 0.0001 times the maximum eigenvalue compared with 0.1 for the real data. The LLSI best fit models (LLSI-PEM and LLSI-PEMR) were used as the reference models for the NA runs. The same model parameterization and tuning parameters are used as in the NA application to the real data, and, as before, 10,000 models were generated for each run.

[23] Figure 6 shows results for the PEMC-42 and PEMC-42R synthetics runs. As with the real data runs, our ensemble of acceptable models was a misfit cutoff that included all from among the 10,000 models with Rayleigh phase velocity dispersions which fit within the range of the plotted estimated error bars.

[24] For both synthetics runs the models with the lowest misfit are the same as the best fit LLSI models and are indistinguishable from the ensemble-averaged models. Because of the small damping required for the LLSI runs on the synthetic data the estimated errors for the mantle *S*-wave velocities are ~10 times as large as for the LLSI runs for the real data, and the estimated errors are similar in size to those calculated for the NA ensemble-averaged models. Figure 7 shows the ensemble-averaged models compared with the "true" models and NA-PAR, the NA ensemble-averaged model for the data run.

[25] From Figure 7, we deduce the following:

1. NA and LLSI can image equally well both the top and bottom of the LVZ for the PEM-C42 data set (Figure 7a) and, with poorer resolution, the velocity decrease at 180 km depth for the PEM-C42R data set (Figure 7b).
2. On the basis of the shapes of the velocity-depth curves and the estimated errors in the velocities, the Paraná data derived NA-PAR model is significantly different from models representing an LVZ starting at 120 km depth (Figure 7a). Although the data do not require an LVZ starting at 180 km depth, the velocity structure at this depth is at the limit of resolution for this data set (Figure 7b).

7. Discussion

[26] Because of the way they were implemented in our analyses, neither LLSI nor NA can resolve a first-order discontinuity within the mantle. As mentioned above, this may be of little practical importance, as there is no evidence that the lithosphere-asthenosphere boundary is a sharp interface. For the period range in our data set a first-order discontinuity cannot be resolved, but a model with such a discontinuity could be found to be consistent with the data by an appropriate choice of the reference model and the data parameterization.

[27] Our inversions do not find models which give a good fit to the middle periods for the Love phase velocities. Separate NA runs inverting for only the Rayleigh dispersion, produced effectively the same model as shown in Figure 5. Inversions for only the Love dispersion produced relatively poorly constrained models, which follows from the fact that the estimated errors for the Love dispersion velocities are larger than for the Rayleigh dispersion velocities. The ensemble-averaged model for the Love-only inversion has a significantly higher velocity lid in the mantle from the Moho down to ~70 km depth. That model produces an extremely poor fit for the Rayleigh group velocities above 35 s period and a poor fit for the midperiod Rayleigh phase velocities.

[28] Point measurements of azimuthal anisotropy based on *SKS* splitting have been made at the three stations used in this study, but nowhere else along the paths [*James and Assumpção*, 1996]. On the basis of their results we estimate path effects from azimuthal anisotropy to be at most a few percent. In addition, azimuthal

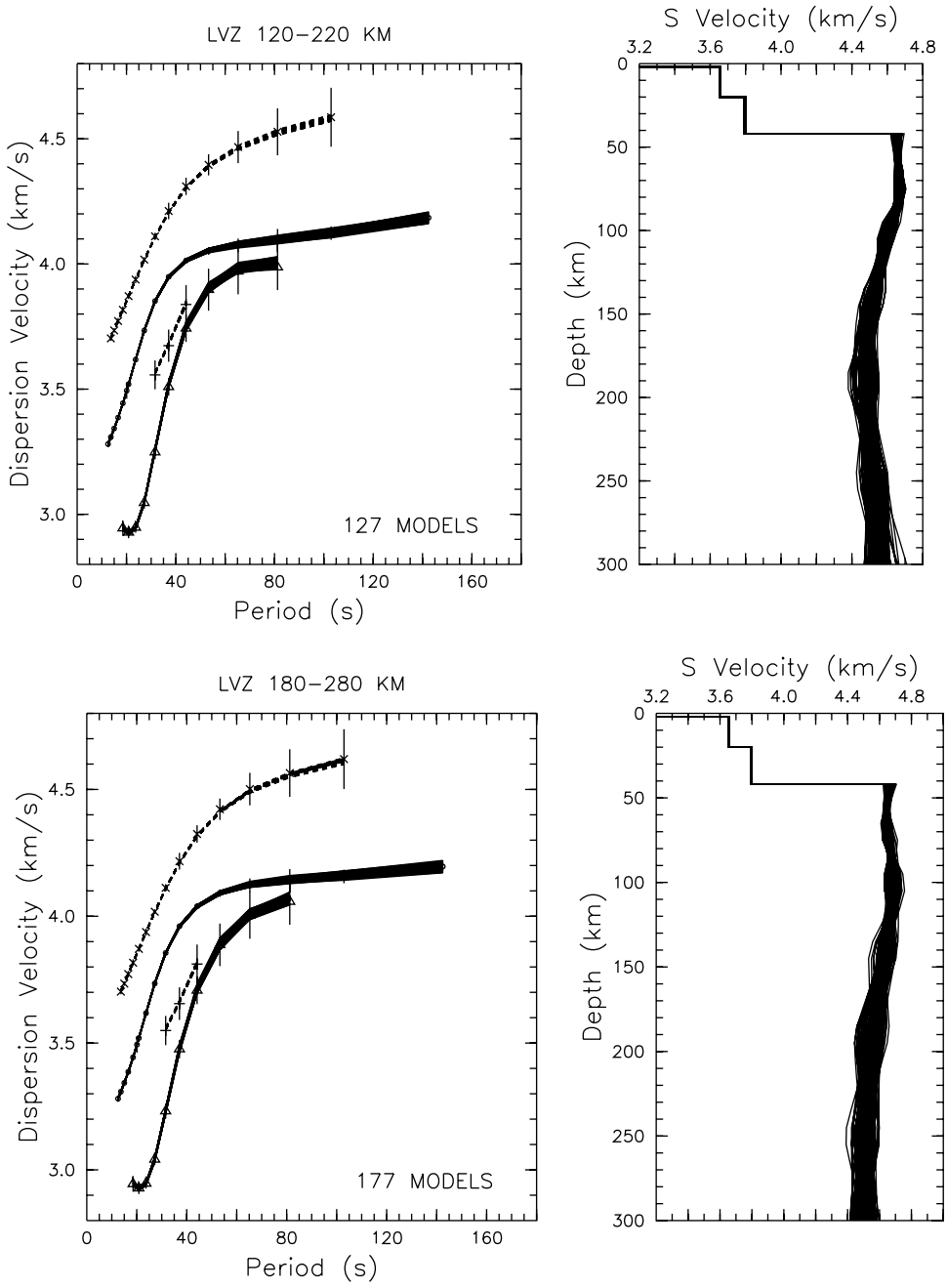

Figure 6. As in Figure 5(top) except the dispersion velocities "data" are generated by (top) the PEM-C42 model and (bottom) the PEM-C42R model.

anisotropy cannot explain differences between Rayleigh and Love models. The inability to model precisely both Love and Rayleigh dispersion may very well be produced by radial anisotropy [e.g., *Freybourger et al.*, 2001], but there are no data at present to constrain the degree of radial anisotropy.

[29] Other effects not taken into account in our modeling of the data include lateral heterogeneity, as well as variations in depth of anelasticity, density, and Poisson's ratio. While none of these is very well constrained for this region, nothing among published results indicates that such effects could have more than a second-order effect on the velocity structure in the upper mantle.

[30] It is easy to quantify the relative importance from among the data by doing NA runs with data subsets, as is discussed above for comparisons of Love-only and Rayleigh-only inversions. By varying the parameterization scheme or the definition of misfit, one can concentrate the analysis on different parts of the model space and interpret the results more easily with NA than with LLSI.

[31] The calculated standard deviations for "average" models in the NA runs give only relative estimates of the velocity spread as a function of depth. Even though NA searches the full parameter space (e.g., Figure 4a), if the number of iterations exceeds a certain value, there may still be "saturation" resulting in an ever increasing number of models near one which are indistinguishable within sets of model parameterizations. This is discussed by *Sambridge* [2001]. Another approach, not done in this study, is to get estimates of uncertainty using Bayesian appraisal [*Sambridge*, 1999b].

[32] Although beyond the scope of this study, it is relevant to conjecture about how one can explain the decoupling between the

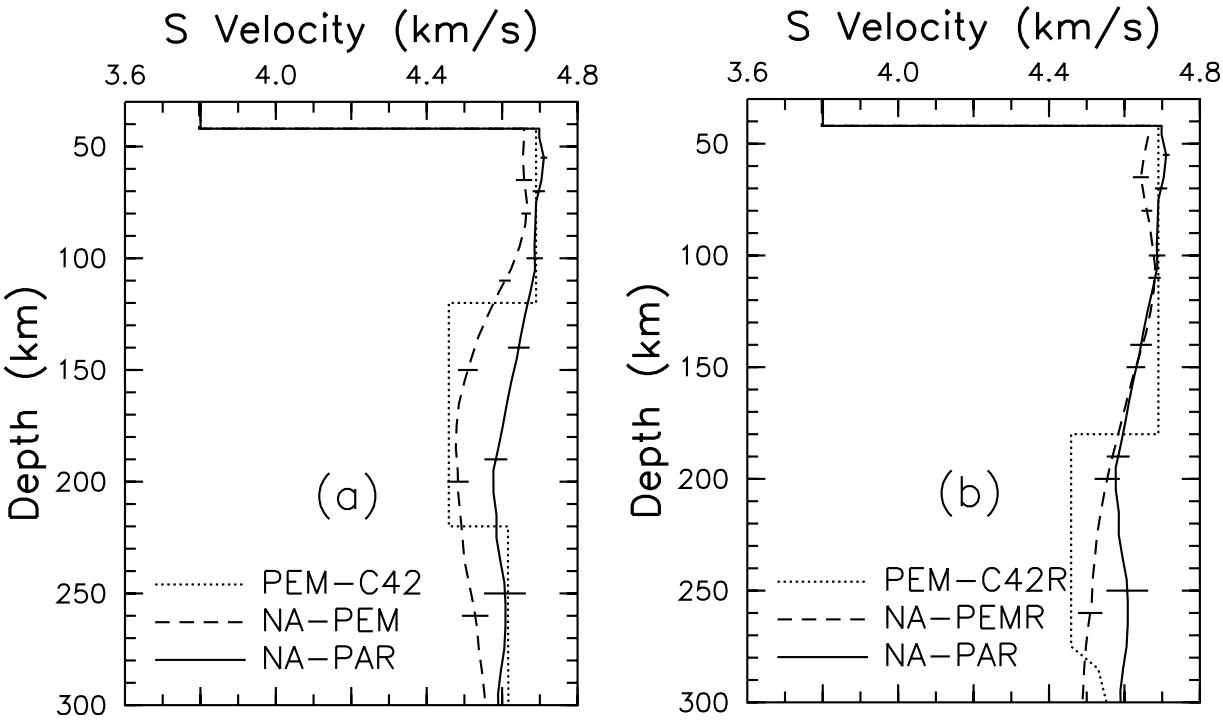

Figure 7. Velocity models generated from synthetic data sets for (a) PEM-C42 with an LVZ starting at 120 km depth and (b) PEM-C42R with an LVZ starting at 180 km depth. Shown are the models generated using NA for each case. Also included in both plots is model NA-PAR, generated by NA with the Paraná data.

continental plate and the deeper mantle, if there is no well-defined LVZ in the upper mantle,. We list here three possible explanations:

1. The rheology at depth beneath continental shields may be such that the low viscosities required to accommodate plate motion may not be correlated with low velocities.

2. The differential shearing may take place over an extended depth range with either small contrasts in viscosity between the lowest part of the continental root and the underlying mantle or a sequence of very narrow LVZs (which would not be resolvable).

3. As suggested by *VanDecar et al.* [1995], the whole upper mantle may move with the plate, so that any decoupling takes place at a greater depth, perhaps in the transition zone.

8. Concluding Remarks

[33] Earlier we had applied genetic algorithms (GA), another fully nonlinear global optimization technique, to this data set [*Zhang et al.*, 1998]. As found by *Sambridge* [1999a, 1999b], in his comparisons between NA and GA for a receiver function data set, NA is more easily tuned, the distribution of final solutions in NA is less dependent on the starting solution, and NA produces a more uniform sampling of the ensemble of acceptable models.

[34] For this Brazilian dispersion-velocity data set we have shown that NA confirms the conclusion drawn by *Snoke and James* [1997] that the data are best fit by velocity models which do not have a well-defined LVZ starting at depths shallower than ~150 km. The ensemble of acceptable models determined through NA provide insights about the deeper velocity structure not easily discernible using LLSI for less-than-perfect data. In particular, the standard deviations calculated for the ensemble-averaged models from NA are easier to interpret than those calculated using LLSI. To constrain the velocities for greater depths requires data with small estimated errors at longer periods than were available for these studies.

[35] **Acknowledgments.** We extend our thanks to David James for useful discussions and to the two reviewers along with the Associate Editor for helpful suggestions.

References

Assumpção, M., D. E. James, and J. A. Snoke, Crustal thicknesses in SE Brazilian shield by receiver functions analysis: Implications for isostatic compensation, *J. Geophys. Res.* 10.1029/2001jb00422, 2002.

Brito Neves, B., and U. Cordani, Tectonic evolution of South America during the late Proterozoic, *Precambrian Res.*, *53*, 23–40, 1991.

Davis, L., *Genetic Algorithms and Simulated Annealing (Research Notes in Artificial Intelligence)*, Pitman, London, 1987.

Dziewonski, A. M., A. L. Hales, and E. R. Lapwood, Parametrically simple Earth models consistent with geophysical data, *Phys. Earth Planet. Inter.*, *10*, 12–48, 1975.

Freybourger, M., J. B. Gaherty, and T. H. Jordan, Structure of the Kapvaal craton from surface waves, *Geophys. Res. Lett.*, *28*, 2489–2492, 2001.

Herrmann, R. B., Computer programs in seismology, St. Louis Univ., St. Louis, Mo., 1987.

James, D. E., and M. Assumpção, Tectonic implications of *S*-wave anisotropy beneath SE Brazil, *Geophys. J. Int.*, *126*, 1–10, 1996.

James, D. E., M. J. Fouch, and J. C. VanDecar, Tectonic structure beneath southern Africa, *Geophys. Res. Lett.*, *28*, 2485–2488, 2001.

Kennett, B. L. N., K. Marson-Pidgeon, and M. Sambridge, Seismic source characterization using a neighbourhood algorithm, *Geophys. Res. Lett.*, *27*, 3401–3404, 2000.

Lomax, A., and R. Snieder, Finding sets of acceptable solutions with a genetic algorithm with application to surface wave group dispersion in Europe, *Geophys. Res. Lett.*, *21*, 2617–2620, 1994.

Marson-Pidgeon, K., B. L. N. Kennett, and M. Sambridge, Source depth and mechanism inversion at teleseismic distances, using a neighbourhood algorithm, *Bull. Seismol. Soc. Am.*, *90*, 1369–1383, 2000.

Nolet, G., Partitioned waveform inversion and two-dimensional structure under the network of autonomously recording seismographs, *J. Geophys. Res.*, *95*, 8499–8512, 1990.

Priestley, K., and F. Tilmann, Shear-wave structure above the lithosphere above the Hawaiian hot spot from two-station Rayleigh-wave phase-velocity measurements, *Geophys. Res. Lett.*, *26*, 1493–1496, 1999.

Ramos, V. A., Late Proterozoic-Early Paleozoic of South America—A collisional history, *Episodes*, *11*, 168–174, 1988.

Ritsema, J., and H. van Heijst, New seismic model of the upper mantle beneath Africa, *Geology*, *28*, 63–66, 2000.

Sambridge, M., Geophysical inversion with a neighbourhood algorithm, I, Searching a parameter space, *Geophys. J. Int.*, *138*, 479–494, 1999a.

Sambridge, M., Geophysical inversion with a neighbourhood algorithm, II, Appraising the ensemble, *Geophys. J. Int.*, *138*, 727–746, 1999b.

Sambridge, M., Finding acceptable models in nonlinear inverse problems using a neighbourhood algorithm, *Inverse Problems*, *17*, 287–403, 2001.

Sambridge, M., and B. L. N. Kennett, Seismic event location: Nonlinear inversion using a neighbourhood algorithm, *Pure Appl. Geophys.*, *158*, 241–257, 2001.

Snoke, J. A., and E. E. James, Lithospheric structure of the Chaco and Paraná Basins of South America from surface wave inversion, *J. Geophys. Res.*, *102*, 2939–2951, 1997.

VanDecar, J. C., D. E. James, and M. Assumpcão, Seismic evidence for a fossil mantle plume beneath South America and implications for plate driving forces, *Nature*, *378*, 25–31, 1995.

van der Lee, S., D. E. James, and P. D. Silver, Upper-mantle *S* velocity structure of central and western South America, *J. Geophys. Res.*, *106*, 30,821–30,834, 2001.

Woods, M., and E. Okal, Rayleigh-wave dispersion along the Hawaiian swell: A test of lithospheric thinning by thermal rejuvenation at a hot spot, *Geophys. J. Int.*, *125*, 323–329, 1996.

Zhang, K., J. A. Snoke, and D. E. James, Lithospheric structure of the eastern Paraná Basin of central Brazil from surface save inversion: Comparing genetic algorithms and linearized least-squares inversion, paper presented at IRIS Workshop, Santa Cruz, Calif., July 1998.

M. Sambridge, Research School of Earth Sciences, Australian National University, Canberra ACT 0200, Australia. (malcolm@rses.anu.edu.au)

J. A. Snoke, Department of Geological Sciences, Virginia Tech, Blacksburg, VA 24061, USA. (snoke@vt.edu)