IMPROVED SURFACE WAVE DISPERSION MODELS AND AZIMUTH ESTIMATION TECHNIQUES

Jeffry L. Stevens, David A. Adams, and G. Eli Baker

Science Applications International Corporation

Sponsored by Defense Threat Reduction Agency

Contract No. DTRA01-01-C-0082

ABSTRACT

Surface waves are of primary importance for nuclear monitoring because the Ms:mb discriminant and its regional variants are among the most reliable means of determining whether an event is an earthquake or an explosion. When surface waves can be measured, in most cases they unambiguously identify an event as an earthquake or an explosion. The goal of this project is to reduce the magnitude threshold for which surface waves can be identified and measured reliably, and to improve the accuracy of surface wave measurement, using phase-matched filtering, development of global regionalized earth and dispersion models, and other techniques.

Global regionalized earth and dispersion models are being developed by inversion of a very large data set of phase and group velocity dispersion measurements. Two new data sets have recently been received and incorporated: a new data set of about 9,000 measurements from SLU, and an expanded data set from the University of Colorado. The complete data set now contains 725,000 dispersion data points. We are reviewing the current data set to find areas where new data is required to fill gaps or resolve ambiguity, and to identify incorrect or anomalous data. Because of the very large size of the data set, we are developing new methods to review and manage the data.

The dispersion measurements are inverted for earth structure, and the earth structure is then used to generate dispersion predictions. This is accomplished in the following way. The inversion is performed for approximately 600 distinct base structures, which were originally derived from the Crust 2.0 models over the AK135 mantle model. The Moho depth, bathymetry, and sediment depths vary on a one-degree grid. Moho depths are derived from Crust 2.0, sediment depths from the Laske and Masters sediment maps, and bathymetry from ETOPO5. Moho depth, bathymetry and sediment properties are fixed in the inversion, while crust and upper mantle velocities are allowed to vary in the base models. Phase and group velocity dispersion curves are calculated for each of 64,800 models on the one-degree grid. The phase velocity dispersion curves are used to calculate phase-matched filters to improve detection.

Automatic processing of surface waves at the International Data Center calculates a station to event azimuth estimate based on polarization analysis, but the algorithm currently in use has been found to be unreliable for a large fraction of surface wave arrivals, particularly for lower amplitude arrivals. Selby (2000) suggested that a different azimuth estimation algorithm, based on a technique due to Chael (1997) would be more robust than the algorithm currently used in the automatic processing system. Part of the problem is due to high noise levels on the horizontal components, particularly at stations using tri-axial instruments. This problem can only be resolved with better instrumentation; however better analysis techniques could improve results on the rest of the data set. We have implemented Selby's method for azimuth estimation within the automatic surface wave processing program Maxpmf and compared the results with the existing method for estimating azimuth from 3-Component data on data with known surface wave arrivals. The results show a reduction in the standard deviation of the azimuth residuals from 72º to 33º with the new method. Work is continuing to optimize the performance of both azimuth estimation techniques.

OBJECTIVE

The goal of this project is to reduce the magnitude threshold for which surface waves can be identified and measured reliably, and to improve the accuracy of surface wave measurement, using phase-matched filtering, development of global regionalized earth and dispersion models, and other techniques.

RESEARCH ACCOMPLISHED

During the first year of this project, we have focused on two areas: improvements to global earth models and dispersion maps, and implementation and testing of azimuth estimation techniques at three component stations based on polarization analysis. This project continues some of the work done under an earlier project (Stevens *et al*, 2001b). In that project, we started with a set of ~150 earth models defined on a five degree grid originally derived from Crust 5.0 (Stevens and McLaughlin, 2001), replaced this with a starting model derived from Crust 2.0, and developed a set of dispersion curves and earth models defined on a one degree grid. The technique for doing this is described in the following section. In the current project, we are developing improvements to the data set and the inversion technique. We are performing a comprehensive review of the data set to identify anomalous data and to determine whether these data are inconsistent with the model because of inaccuracies in the model, errors in the data, or anomalous paths that cannot be modeled with the current procedure. Because of the complexity of working with this very large data set, we have developed some interactive tools to help with this task. We have also greatly improved the performance of the inversion scheme, so that the time required for an inversion has been reduced from over 24 hours to about 7 hours on a Compaq ES40 computer with three 667 MHz Alpha processors.

Surface Wave Dispersion Data Set

We have continued to add new dispersion data to our global data set. Two new data sets have been received and incorporated: a new data set of about 9,000 measurements from St. Louis University (Perez, 2001), and an expanded data set from the University of Colorado (Levshin *et al*, 2002). The complete data set now contains 725,761 dispersion data points. Other data in this data set is described in Stevens *et al* (2001a,b).

One-Degree Earth Model

We have developed a one-degree Earth model made up of layers of ice, water, sediments, crust and upper mantle. Currently this model depends on 8886 free parameters that are adjusted by least squares fitting to group and phase velocity dispersion data of Rayleigh waves. Most of the data set consists of group velocity measurements. The free parameters are the S-wave velocities of layers of 570 different model types. Other constrained parameters in the model are P wave velocities, densities and Q. P wave velocities are constrained via a constant Poisson's ratio of 0.27, and density via Birch's Law of $\rho = 0.65\beta + 400$ where ρ is density and β is shear velocity in MKS units. Q corresponds to PREM in the mantle, and is set according to Swanger's law Q=β/10 (in MKS units) in the crust and upper mantle unless other information is available. The model types are based on the Crust 2.0 2x2 degree crustal types (Bassin *et al*., 2000 and Laske *et al*. 2001) and also on ocean ages (Stevens and Adams, 2000).

The procedure for constructing the inversion models is illustrated in Figure 1. The top few km of the model (consisting of water, ice and/or sediments) are fixed and match data from one degree bathymetry maps made by averaging Etopo5 five minute measurements of topography, and Laske and Masters (1997) 1 degree maps of sediments. There is an explicit discontinuity between the bottom of the sediments and the crust. There are three or more crustal layers. The Crust 2.0 models that were the starting point for these structures have three crustal layers, but we found it necessary to add more layers in regions of thick crust. There is another explicit discontinuity at the Crust/Mantle boundary. The Moho depth is derived from Crust 2.0 and varies on a 2-degree grid. The mantle starting model is derived from AK135 (Kennett, *et al*, 1995). With these constraints, the inversion is performed for the shear velocity of the crust and upper mantle to a depth of 300 km. Below 300 km the earth structure is fixed, and the inversion model is required to be continuous with the mantle structure at the base of the inversion.

In broad ocean areas, we replace the Crust 2.0 model with models distinct for each ocean and subdivided by ocean age. We also separate into distinct models Crust 2.0 models that are geographically separated. So, for example, if Crust 2.0 has the same model type in North America and in Asia, we use the same starting model for each, but treat them as separate models in the inversion. Figures 2 and 3 show the global distribution of starting model types.

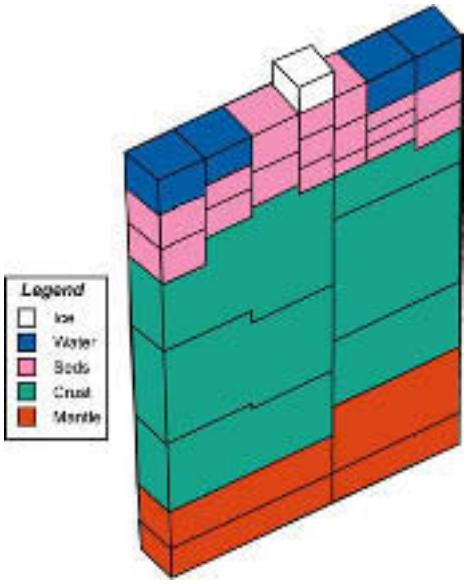

Figure 1. Surface wave inversions are performed for a set of earth models consisting of mantle, crust, sediments, and in some layers, water or ice. The properties of the ice and sediment layers are held fixed, as is the Moho depth, and inversion is performed for the shear velocity in the crust and upper mantle.

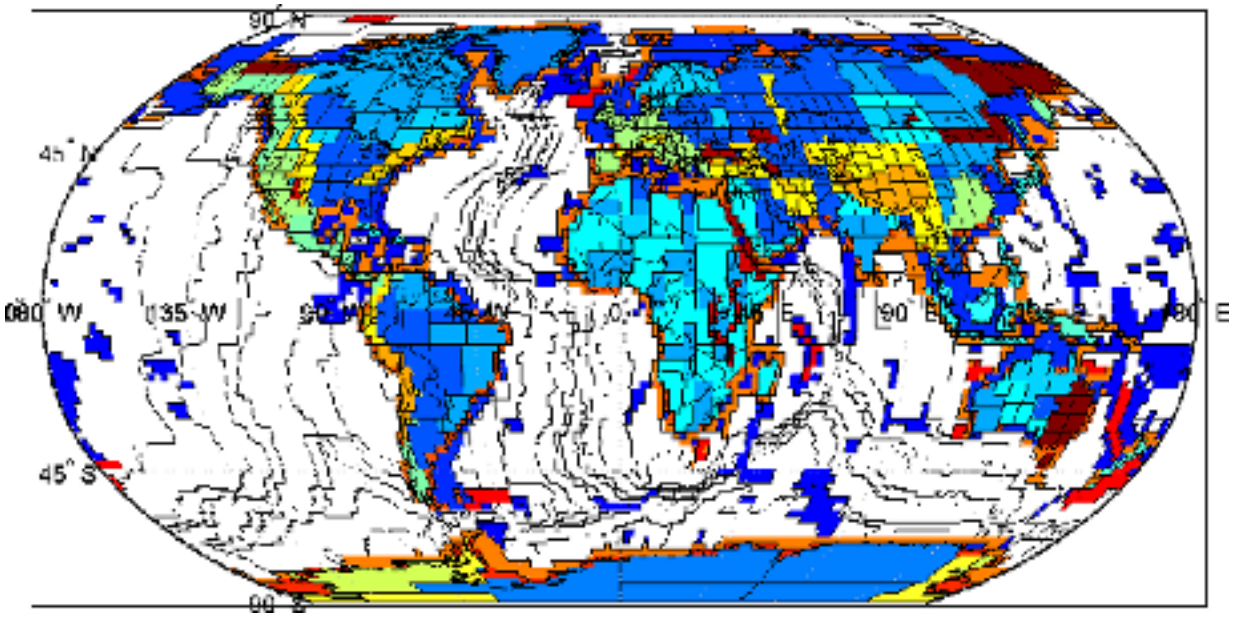

Figure 2. Starting models for inversion in continental areas. Each color is a different model type. The initial starting models for inversion are the Crust 2.0 model types. Separate models (with the same initial structure) are used for geographically separated Crust 2.0 models of the same type.

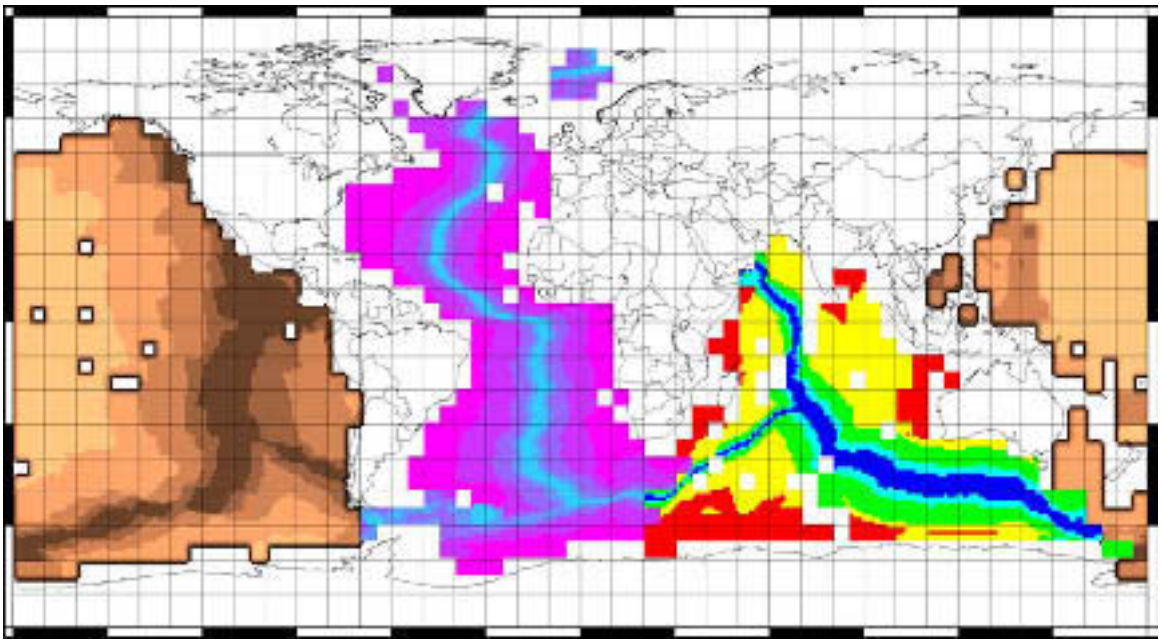

Figure 3. Starting models for broad ocean areas are separated by age. Distinct models are used for each ocean.

Tools for Model and Data Analysis

Because of the very large size of the data set, it has become increasingly difficult to evaluate the results, identify problem data, and determine where new models are needed. To assist with this effort, we have developed data and model analysis and visualization tools using Matlab and the Matlab Mapping Toolbox. These are interactive tools and include multiple ways of representing data and models with zoom adjustable scales. The tools are modular and can be easily modified and extended. These characteristics are important for finding quick answers to questions about the large and complex data set and model. The link between data appraisal and model parameter adjustment is strong, and the ability to examine data and model frequently and in different ways allows us to focus in on those aspects of the inversion that need to change. For example, we can examine spatial variations of data fit, find outliers, and find under or over regularized regions in the model. Such tasks can be very time consuming using non-interactive graphics tools.

Example of Interactive Tools Applied to Data Analysis

The following example demonstrates how the Matlab tools can be used to categorize data according to data fit and to find long wavelength variations in data misfit. We categorize all data into four sets.

1. Good data consistent or almost consistent with the model. These correspond to smooth dispersion curves with similar dispersion curves on neighboring paths.
2. Good data inconsistent with model. Typically these correspond to multiple dispersion curves along similar paths that consistently disagree with model predictions. These show that model improvement is needed.
3. Indeterminate data. These correspond to isolated paths where ray density is too low to determine whether differences are caused by data or model problems. In this case, more data along similar paths is needed.
4. Poor data points. These are identified by being inconsistent with dispersion curves of neighboring paths, and or with other frequencies in the same dispersion curve.

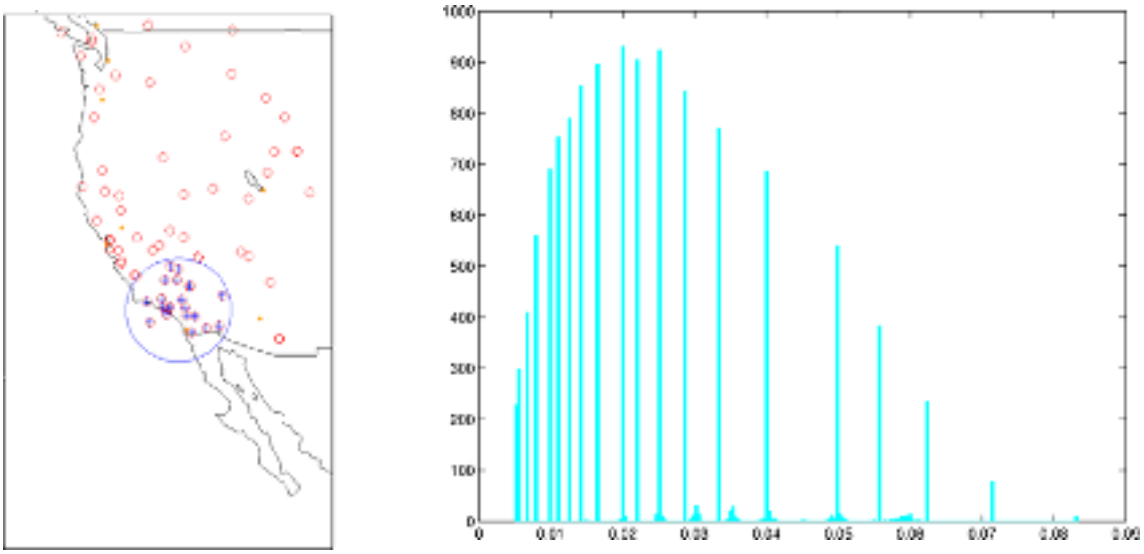

Figure 4. Locations of all stations in the western US appearing in our data set. The histogram to the right shows the distribution of group velocity measurements made at the stations in southern California marked with blue crosses.

Using the interactive tool, the user can select a region and find the characteristics of all rays starting or ending in the region. Figure 4 shows the locations of stations and sources selected within a circular region in and near Southern California appearing in the data set and a histogram of the distribution of corresponding group velocity measurements. The user can then select a subset of data, for example, by frequency band. Figure 5 shows all of the paths recorded at the Southern California stations at frequencies between .018 and .037 Hz. The paths are color coded according to the residual $1\text{-}s_o/s_p$, where s_o and s_p are the observed and predicted group slownesses. Blue paths indicate paths where observations are slower than predicted, red faster than predicted. The scale is truncated to make colors more sensitive to variations nearer zero. Long wavelength variations in data fit appear as smooth color variations in Figure 5. These can be used in the future to build station dependent correction surfaces or tables, or to drive model improvements through reparameterization by densification of model types, or layers, or by localized relaxation of regularizing parameters (smoothing and damping parameters). Dark red and blue paths are mostly associated with poor data points (category 4), however there are several nearby paths that show consistent low velocities between Southern California and the South Atlantic Ocean.

Geographic and model information can be displayed together with the paths. Figure 6 shows sedimentary thicknesses together with paths starting in the blue circle shown in Figure 4. Figure 7 shows color-coded group velocity dispersion curves for the paths shown in Figure 6 with crosses for predicted and circles for observed velocities. The colors are used for cross-referencing between Figures 6 and 7. Interactively, the user can select a dispersion curve on Figure 7 and find the path on Figure 6, or vice-versa. Comparing Figures 5-7 it can seen that the anomalous data points discussed above pass close to or within the sediments off of the northern coast of South America. These therefore correspond to a geophysical effect rather than data error, and most likely will require the introduction of a new model type.

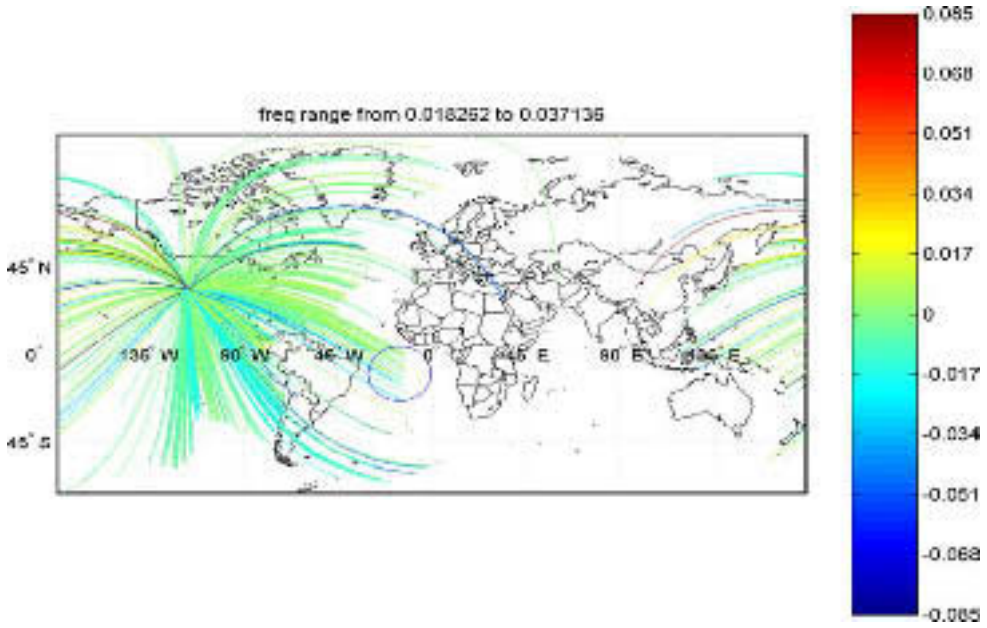

Figure 5. All ray paths for surface waves recorded at the southern California stations at frequencies between .018 and .037 Hz. The paths are color coded according to the residual $1\text{-}s_o/s_p$. Blue means the observations are slow relative to the model, red fast relative to the model.

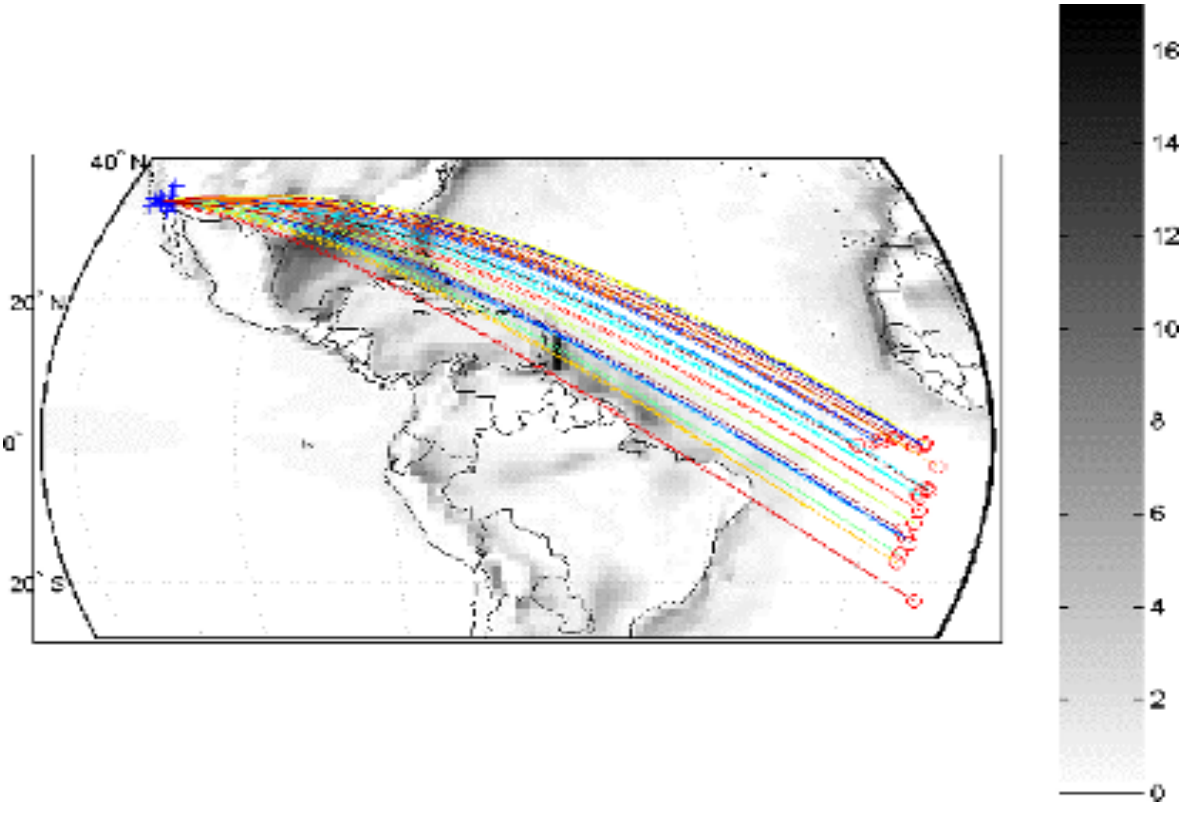

Figure 6. Sedimentary thickness in kilometers and all paths starting in the blue circle shown in Figure 5. The ray colors correspond to the dispersion curve colors in Figure 7.

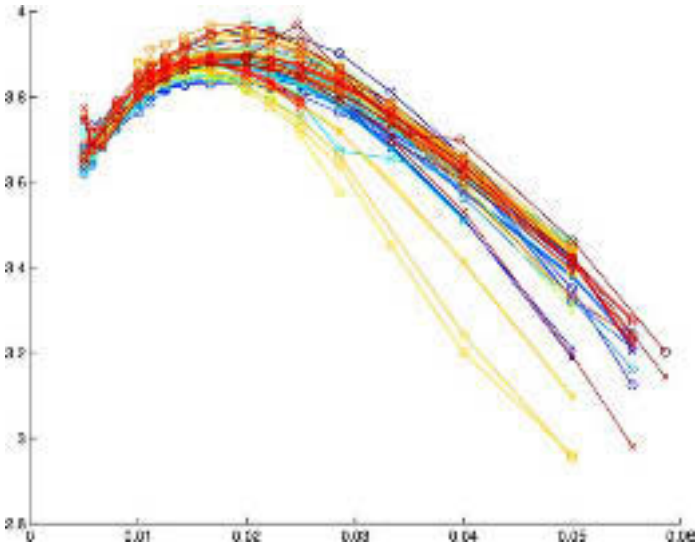

Figure 7. Color-coded group velocity dispersion curves for paths shown in Figure 6 with crosses for predicted and circles for observed velocities. The colors correspond to the path colors in Figure 6. The lowest velocity curves correspond to the anomalous curves discussed in reference to Figure 5.

The Selby/Chael Azimuth Estimation Algorithm

The azimuth estimation technique used by the automatic surface wave processing program Maxsurf at the IDC is based on an algorithm due to Smart (1977). When Maxsurf was first installed at the PIDC, it used this azimuth test in addition to a dispersion test to identify surface wave arrivals. That is, it used three-component or array data to estimate the azimuth to the source and compared this with the known azimuth. If it was not within a specified error, the arrival was rejected and not included in the database. However, review of processing results showed that a large number of arrivals, particularly at three component stations, were being rejected even though they showed a clear arrival on the vertical component. Further analysis showed that the azimuth errors were primarily due to high noise levels on the horizontal components. An example is shown in Figure 8, showing seismograms and spectra at ZAL from the Chinese nuclear test of July 28, 1996. Because of these problems, the azimuth test was removed and replaced with a post-processing algorithm to remove misassociated arrivals.

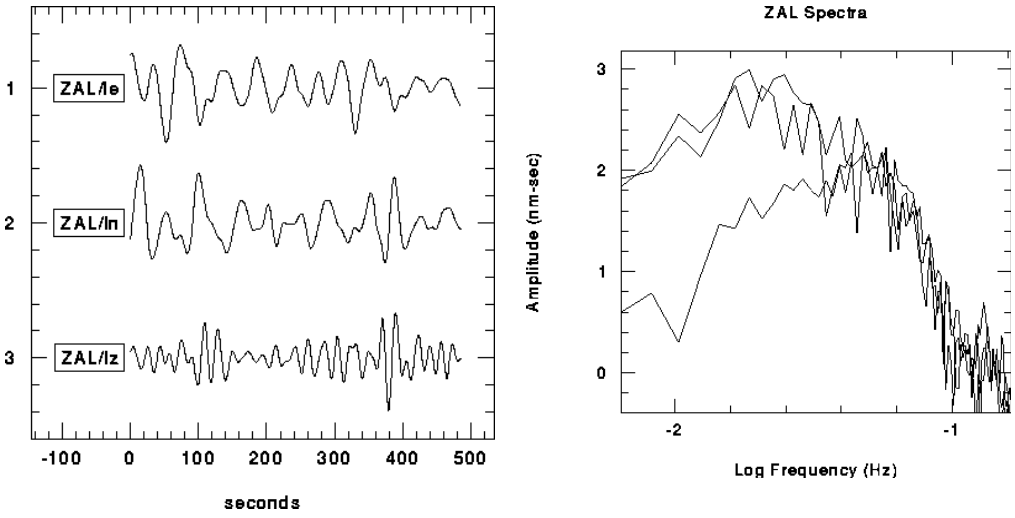

Figure 8. Seismograms (left) and corresponding spectra (right) from the Chinese nuclear test of July 28, 1996 recorded at ZAL at a distance of 12.7 degrees. Note the good signal on the vertical component, but high noise levels that obscure the signal on the horizontal components. This event had an M_s of 2.8 based on arrivals at ZAL and NRI.

Recently, Selby (2000) suggested that a different azimuth estimation algorithm, similar to that developed by Chael (1997) would be more robust that the algorithm currently used in Maxsurf. He processed a set of data using this algorithm and then compared it with the azimuth estimates in the Reviewed Event Bulletin (REB) and found significantly improved results. Although no algorithm is likely to work with very noisy data such as that shown in Figure 8, an improved algorithm could make the azimuth estimate more useful as a means of associating arrivals.

We have implemented Selby's method for azimuth estimation within Maxpmf and compared the results with the existing method for estimating azimuth from 3-component data. The algorithm takes advantage of the fact that the vertical and Hilbert transformed radial ground motion are highly correlated. Specifically, it finds the back-azimuth for which the correlation of the vertical and Hilbert transformed radial is a maximum. Selby presents the cross-correlation, normalized as follows,

$$C_{z\bar{r}} = \frac{S_{z\bar{r}}}{S_{zz} S_{\bar{r}\bar{r}}}, \tag{1}$$

where,

$$S_{jk} = \sum_{\tau=1}^{N} x_j(\tau) x_k(\tau). \tag{2}$$

The implementation uses the correlation of the Hilbert transformed vertical, as that only has to be calculated once, with the radial, for one degree increments of back-azimuth. While Selby uses the maximum positive correlation as the back-azimuth, Chael (1997) uses the central azimuth, determined by the circular mean.

Circular statistics are analogous, but not identical, to linear statistics. For example, while the mean of 359 and 1 is 180, the circular mean of 359° and 1° degree is 0°, which is the correct answer for averaging directions. We find the circular mean separately for the positive and negative values of the cross-correlation at each one-degree increment of back-azimuth, and average those results for greater accuracy. For synthetic Rayleigh waves with noise added, the circular mean provides more accurate estimates than does the maximum correlation. The poorer the S/N, the greater the advantage of using the circular mean. We have implemented both methods and are testing whether there is a significant difference for real data.

For noise-free data (i.e. synthetics), equation 1 will return a negative or positive constant. That is because cross-correlation of the Hilbert transformed vertical and radial and the autocorrelation of the radials change in synch as the algorithm steps through back-azimuths. Our implementation of the algorithm therefore provides two methods of estimating the similarity of the components. One uses the correlation coefficient, that is, nearly equation 1, but with the normalization using square root of the denominator, as follows,

$$C_{z\bar{r}} = \frac{S_{z\bar{r}}}{\sqrt{S_{zz} S_{\bar{r}\bar{r}}}}. \tag{3}$$

The second method is intended to avoid the problem described above when data are noise free. It does not normalize by $S_{\bar{r}\bar{r}}$, but instead uses,

$$C_{z\bar{r}} = \frac{S_{z\bar{r}}}{S_{zz}}. \tag{4}$$

The results of our initial test are shown in Figure 9. Work is continuing to determine the optimum time window, frequency bands, and other parameters for both methods, but at least this initial test shows a significant improvement, quantified by a reduction in standard deviation of the azimuth residual from 79 degrees to 45 degrees.

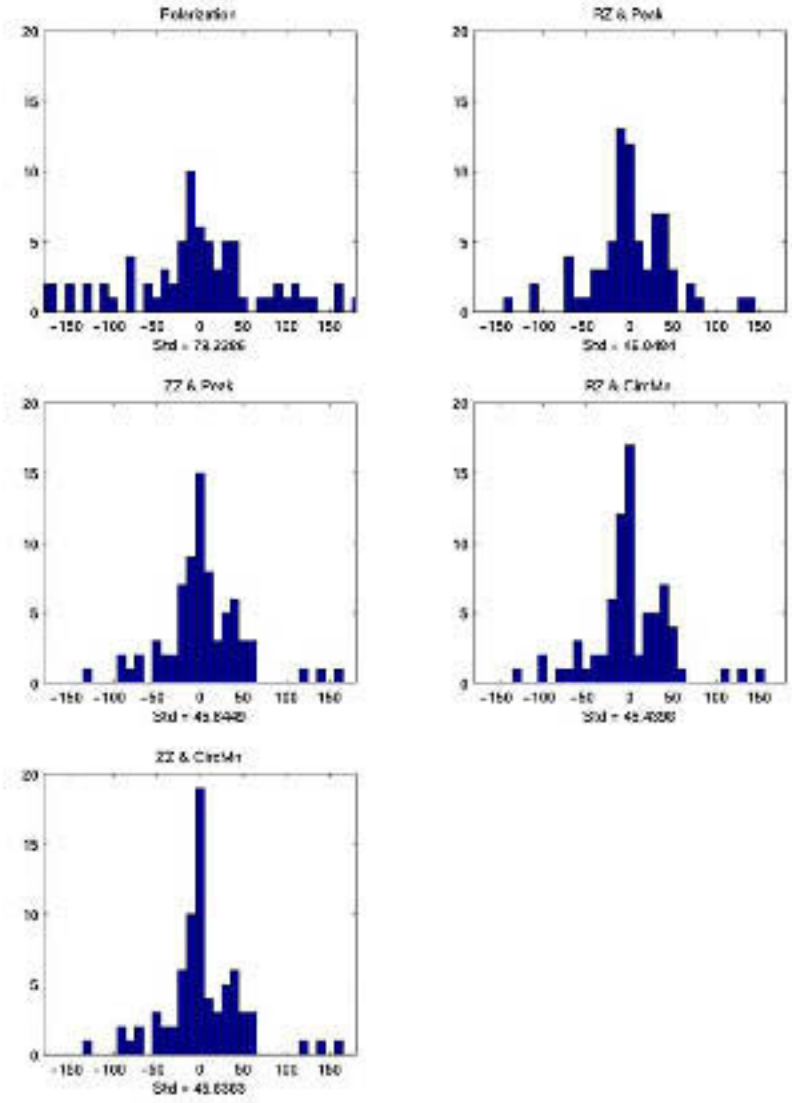

Figure 9. Distribution of azimuth residuals for known surface wave arrivals. Top left figure uses the algorithm that is currently implemented in Maxsurf. The other figures use the Selby/Chael method using different sets of parameters as discussed in the text. These preliminary results show that the error rate is reduced substantially by the new method.

CONCLUSIONS AND RECOMMENDATIONS

Surface wave research is continuing in the following areas:

- Surface wave dispersion models are being improved through:
  - A careful review of existing data, and continuing addition of new data.
  - Improvements to model types and constraints.
  - Development of interactive tools to improve speed and accuracy of analysis.
- An azimuth estimation technique for three-component data proposed by Selby is being implemented and tested. Initial results show substantial improvement.
- Surface wave measurement is being improved through
  - Phase-matched filtering enhanced by development of better dispersion models.
  - Development of path-corrected spectral magnitudes.
  - Development of methods to remove bad noise data.

ACKNOWLEDGEMENTS

We would like to thank Mike Ritzwoller, Anatoli Levshin and Nikolai Shapiro of the University of Colorado, Bob Herrmann of St. Louis University, and their coworkers for the use of their data in this project.

REFERENCES

Bassin, C., Laske, G. and Masters, G. (2000), The Current Limits of Resolution for Surface Wave Tomography in North America, *EOS Trans AGU*, **81**, F897.

Chael, E. P. (1997), "An automated Rayleigh wave detection algorithm," *Bull. Seism. Soc. Am*., **87**, 157-163.

Kennett, B.L.N., E.R. Engdahl, and R. Buland (1995), Constraints on seismic velocities in the Earth from travel times, *Geophys J Int*, **122,** 108-124.

Laske G., and G. Masters (1997), A Global Digital Map of Sediment Thickness, *EOS Trans. AGU,* **78**.

Laske, G., G. Masters, and C. Reif, Crust 2.0 (2001), A new global crustal model at 2*2 degrees, http://mahi.ucsd.edu/Gabi/rem.html.

Levshin, A., J. L. Stevens, M. H. Ritzwoller, and D. A. Adams (2002), "Short Period (7s-15s) Group Velocity Measurements and Maps in Central Asia," Proceedings of the 24th Annual DOD/DOE Seismic Research Symposium, 17-19 September 2002 (this volume).

Perez, F. M., (2001), Surface wave dispersion about the New Madrid Region, MS thesis, Saint Louis University.

Selby, N.D. (2000), "Identification, Association, and Analysis of Rayleigh Waves in a CTBT Monitoring Context," Proceedings of the 22nd Annual DOD/DOE Seismic Research Symposium, 12-15 September 2000.

Smart, E. (1977), "A Three-Component, Single-Station Maximum-Likelihood Surface Wave Processor", SDAC-TR-77-14, Teledyne Geotech, Alexandria, VA.

Stevens, J. L., and D.A. Adams, (2000), Improved Surface Wave Detection and Measurement Using Phase-Matched Filtering and Improved Regionalized Models, Proceedings of the 22nd Annual DOD/DOE Seismic Research Symposium, 12-15 September 2000.

Stevens, J. L., D. A. Adams, and G. E. Baker, (2001a), "Phase-matched filtering with a one degree dispersion model" Proceedings of the 23rd Annual DOD/DOE Seismic Research Symposium, 1-5 October 2001.

Stevens, J. L., D. A. Adams and G. E. Baker (2001b), "Surface wave detection and measurement using a one degree global dispersion grid," SAIC Final Report to Defense Threat Reduction Agency, SAIC-01/1084, December.

Stevens, J. L. and K. L. McLaughlin (2001), "Optimization of surface wave identification and measurement," *Pure and Applied Geophysics*, **158**, no. 7, pp 1547-1582.