Array Analysis of Two-Dimensional Variations in Surface Wave Phase Velocity and Azimuthal Anisotropy in the Presence of Multipathing Interference

Donald W. Forsyth

Department of Geological Sciences, Brown University, Providence, Rhode Island

Aibing Li

Department of Geosciences, University of Houston, Houston, Texas

Multipath propagation of surface waves introduces distortions in waveforms that can bias array measurements of phase velocities. We present a method for array analysis of laterally and azimuthally varying phase velocities that represents the incoming wavefield from each earthquake as the sum of two interfering plane waves. This simple approximation successfully represents the amplitude and phase variations for most earthquakes recorded in the MELT Experiment on the East Pacific Rise in the period range from 16 to 67 s. The inversion for velocities automatically reduces the importance of data from earthquakes or periods that are not described well by this approximation. Each iteration in the inversion involves two stages: a simulated annealing inversion for the best wave parameter description of each event, and a linearized inversion for velocities and changes in the wave parameters. At 29 s period, the two-plane-wave solutions indicate that nearly every signal is significantly affected by multipathing. The larger of the two plane waves typically has an apparent azimuth of propagation that is within a few degrees of the great circle path. The smaller wave is more scattered, differing in apparent azimuth from the larger wave by an average of about 13° at 29 s. Both lateral and azimuthal variations in Rayleigh wave phase velocity in the study area are significant, although it is possible to trade off azimuthal anisotropy with rapid and probably unrealistic lateral variations in velocity. Apparent azimuthal anisotropy reaches 5 to 6%, with the fast direction approximately perpendicular to the ridge.

1. INTRODUCTION

The measurement of phase velocities of surface waves crossing an array of seismometers is a powerful way to detect variations in lithospheric and asthenospheric structure. Traditional methods of array-analysis solve for slowness and apparent direction of propagation assuming the incoming wave is planar. Surface waves, however, encounter many heterogeneities between the source and receivers that distort the waves, causing deviations from great-circle azimuths [*Evernden*, 1953; *Alsina, et al.*, 1993; *Laske*, 1995], wave interference effects sometimes described as multipathing [*Capon*, 1970], and scattering [*Snieder and Nolet*, 1987]. Non-planar energy can dominate the signals for Rayleigh waves at periods less than 50 s [*Friederich et al.*, 1994]. Neglecting wave interference by selecting only those records from an event that have a "clean" appearance can lead to a systematic bias in the apparent phase velocity [*Wielandt*,

1993] if the clean appearance stems from constructive interference of waves that have traveled different paths.

To overcome the inaccuracies in phase velocity determinations caused by non-planar waves, Friederich and Wielandt [1995] have developed a method that simultaneously solves for phase velocity variations within an array and for the structure of the incoming wavefield from each event. The wavefields are represented by a set of basis functions in the form of Hermite-Gaussian functions that describe the perturbations from a plane wave. This elegant approach has been successfully applied in studies of regional structure in Germany [*Friederich*, 1998] and California [*Pollitz*, 1999]. There are, however, some practical difficulties that arise. Given a finite number of stations, the wavefield solution is non-unique [*Wielandt*, 1993] and must be constrained by an additional condition, such as that the model wavefield have the same total energy as is estimated from the observations at the existing stations [*Friederich et al.*, 1994]. Even with this condition, the inversions tend to put the maximum predicted amplitude variations in gaps between stations or outside the array (see examples in Figure 11 of *Friederich et al.*, 1994); a typical problem with any orthogonal basis function expansion with incomplete coverage.

Another problem is that many parameters may be needed to represent relatively simple interference patterns of large amplitude but unknown wavelength. With Hermite-Gaussian expansion up to order twenty, 44 parameters are required to describe the wavefield [*Friederich et al.*, 1994; *Friederich*, 1998] at any given frequency. With this large number of wavefield parameters for each source event, it is not surprising that it is difficult to resolve the relatively small changes in phase of the waves associated with variations in phase velocity within the study area. The spatial pattern of phase velocity variations seems to be robust in the joint inversions for structure and wavefields, but the amplitude of the velocity variations is highly dependent on the relative damping of the wavefield and the velocity parameters [*Friederich*, 1998; *Pollitz*, 1999].

Motivated by the relatively simple interference patterns revealed by wavefield modeling with orthogonal basis functions and by the need to find a more stable inversion with fewer parameters for arrays with gaps between stations, we have developed a procedure for joint inversion of velocity and wavefield structure in which each incoming wavefield is represented by the interference between two plane waves. This approach reduces the number of wavefield parameters for each event to six. In most cases, this representation provides an adequate description of the amplitude variations across an array. The procedure automatically reduces the influence of events with more complex wavefields that are not well-represented by this model of two-plane-wave inter-

ference. This approach is illustrated with an example of Rayleigh waves propagating across the array of ocean-bottom seismometers deployed in the MELT Experiment [*MELT Seismic Team*, 1998; *Forsyth et al.*, 1998]. Shear wave splitting [*Wolfe and Solomon*, 1998], seismic refraction experiments [*Raitt et al.*, 1969], ophiolite samples [*Christensen and Salisbury*, 1979], and propagation of Rayleigh and Love waves [*Forsyth*, 1975] all indicate that the oceanic lithosphere should be anisotropic in the MELT study area. Consequently, the two-dimensional structure in the inversion includes terms for azimuthal anisotropy of phase velocity. Other papers [*Forsyth et al.*, 1998; *MELT Seismic Team*, 1998] describe more details of the experiment and discuss the tectonic significance of the results. This paper focuses on the two-plane-wave method of array analysis, which has also been successfully applied in continental settings [*Li et al.*, 2002, 2003; *Weeraratne et al.*, 2003].

2. DATA AND DATA ANALYSIS

To illustrate the method, we use Rayleigh wave data from 21 earthquakes recorded at ocean-bottom stations of the MELT Experiment on the East Pacific Rise between November 1995 and May 1996. Sensors were either three-component, 1-Hz seismometers (model L4C, Mark Products) or Cox-Webb [*Cox et al.*, 1984] differential pressure gauges (DPGs). The instruments came from four different instrument pools with different recording and filtering operations, so before analysis, all responses for seismometers and pressure gauges were equalized to duplicate the response of the seismometer type that produced the most abundant data [*Forsyth et al.*, 1998]. Despite the short natural period of the seismometers, in the stable thermal environment of the seafloor, they yield reliable responses out to periods on the order of 100 s. The DPGs were designed initially for oceanographic purposes and have flat response to acceleration out to periods of thousands of seconds [*Cox et al.*, 1984], although at periods greater than 35 s, pressure variations from infra-gravity waves [*Webb*, 1998] in the 2700 to 3500 m water depths of the MELT Experiment increasingly interfere with the recording of earthquake signals. The phase responses of all the DPGs were well matched, but the absolute gain varied somewhat from instrument to instrument, so in the inversions described in later sections, terms were added to find the best gain factor for each station. Stations were deployed primarily in two lines across the axis (Plate 1a), an arrangement intended to provide optimum resolution for small features beneath the ridge for body wave tomography, rather than optimum coverage for surface waves. An average of 30 stations recorded each event with good signal-
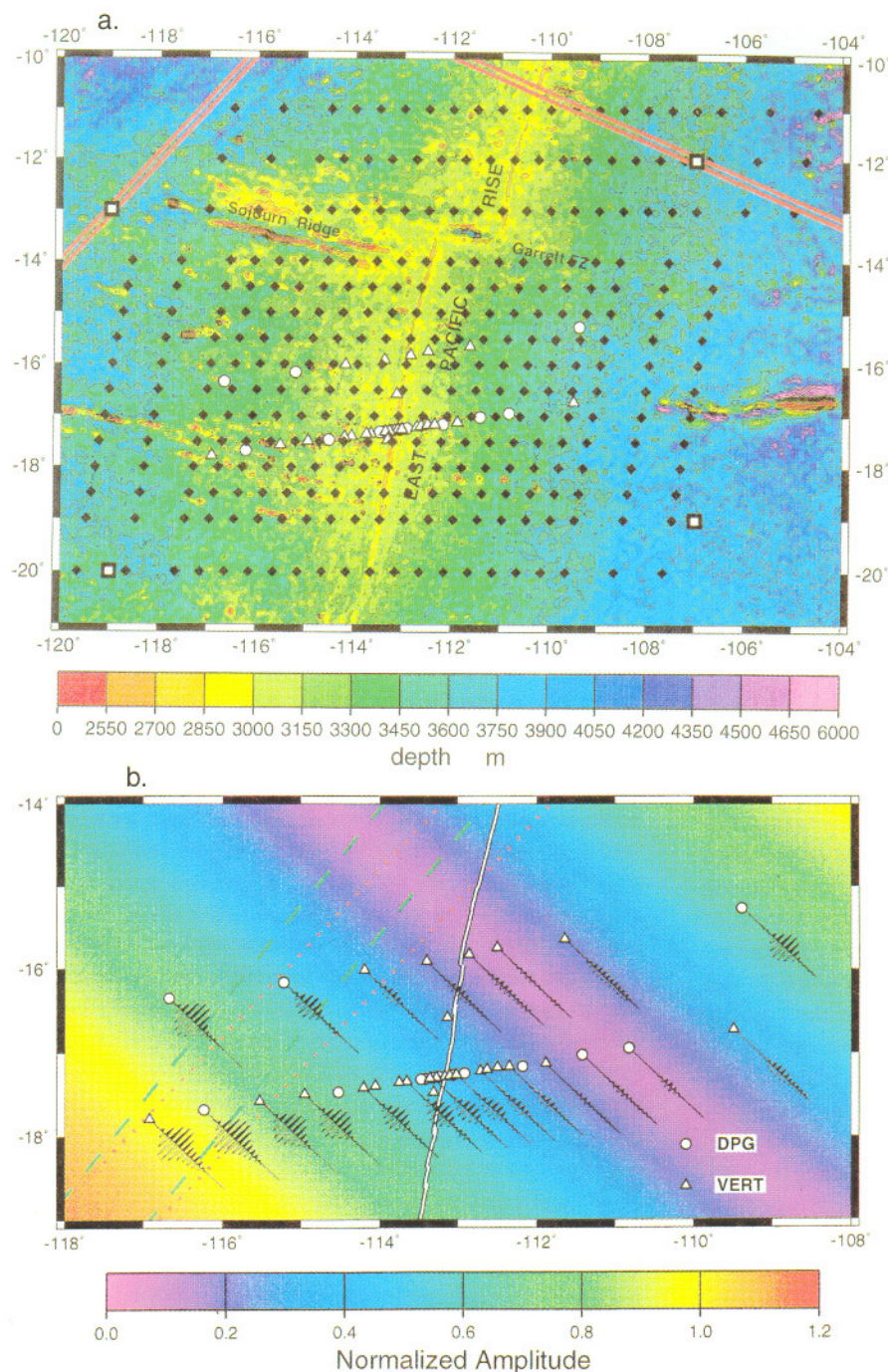
Plate 1. (a) Location of MELT Experiment seismic array on the East Pacific Rise. Bathymetry is from Smith and Sandwell [1997]. The spreading center is marked by the narrow, axial high region shallower than 2850 m, trending about N12°E. Open circles mark locations of differential pressure gauges, open triangles are three-component seismometers. Small diamonds indicate grid of nodal points used in the velocity models. Large, open, squares indicate corners defining edges of the region of interest; travel times (text equation 5) are calculated from the point where a wavefront, assumed to be perpendicular to the great circle path for this purpose, first encounters one of these corners. Double red lines mark these wavefront intersections with the corners for an earthquake along the Kurile trench (Plate 1b and Figure 5a and b) approaching from the northwest and an earthquake off the coast of Oaxaca, Mexico (Figure 5c and d) approaching from the north-northeast. (b) Multipath interference effects for the Kurile earthquake (Ms 7.1 at 2136 on 7 Feb. 1996) at 0.035 Hz. At selected stations, chosen so that the displayed seismograms do not overlap, 500 s records are plotted along a line in the direction of great circle propagation. Time increases away from the station symbol. These records are generated by narrow-passband filtering the original, observed seismograms with corner frequencies at 0.030 and 0.040 Hz. Colors indicate predicted amplitudes for two-plane wave representation found in inversion for velocity model 4, Table 1. Amplitudes are normalized to the maximum, observed amplitude for each event, hence the predicted amplitude can exceed 1.0. To illustrate the two plane waves found in the solution, two of the peaks of the model, sinusoidal wavefields for the two plane waves are shown as green, dashed and red, dotted lines, representing a spatial snapshot at the reference time for the event. One entire wavelength is shown for each wave, although in the idealized, Fourier representation of a single frequency, the sinusoidal oscillations extend across the entire area. Constructive interference occurs wherever the waves are in phase, as in lower left, and destructive interference wherever the waves are out of phase, as along violet band striking diagonally across center of map. The strike of the interference pattern is along the bisector of the two wavefronts, which in this case is nearly, but not exactly, along the great-circle direction.

to-noise ratio, ranging from a minimum of 20 to a maximum of 38 stations.

The amplitudes and phases of Rayleigh waves at selected frequencies were determined by Fourier analysis after instrument correction, filtering and windowing the records. Frequencies analyzed ranged from 15 to 60 mHz. A narrow bandpass filter (4th order Butterworth, zero-phase shift) centered at the frequency of interest with corner frequencies separated by 10 mHz was applied to each record. The filtered records were then windowed (boxcar with 50-s cosine tapers on each end) to isolate the fundamental mode Rayleigh wave from other phases. The width of the window varied depending on the dispersion of the incoming wave around that frequency, but the same frequency-dependent window length was applied to all stations of the array. Amplitudes were corrected for geometrical spreading on a spherical earth and differential attenuation. These corrections are relatively minor, but significant across an array spanning 800 km. To avoid an emphasis on earthquake size in the inversion, the measured spectral amplitudes were normalized to unit RMS amplitude for each event.

Earthquakes from around the Pacific rim were used as sources (Figure 1). To facilitate separation of fundamental mode Rayleigh waves from S phases, a minimum epicentral distance of 3000 km was employed. The earthquakes were chosen to be as well distributed in azimuth as possible, but there is a substantial azimuthal gap, because no earthquakes of sufficient size occurred to the south of the array during the recording period. Nevertheless, the distribution of sources and receivers yielded an excellent set of crossing paths within and to the north of the array (Figure 2) that provide the basis for tomographic imaging of the velocity variations. Paths that traveled across ocean-continent boundaries or along island arcs were avoided as much as possible to reduce the severity of multi-pathing or scattering [*Friederich et al., 1994*]. The period range used for each event depended on the magnitude of the earthquake and the nature of the path. At short periods for more complex paths or paths traveling across the deep water of the northwest Pacific, Rayleigh waves were highly scattered and incoherent from station to station; those periods for particular events were discarded. Scattering in the northwest Pacific has previously been described by Lerner-Lam and Park [1989]. Despite the careful selection of events and period ranges, significant interference and wavefield complexity remains that cannot be ignored in phase velocity analysis. The interference can be recognized in individual records in the form of nodes or beats in the dispersed waveforms, or by systematic amplitude variations across the array (Plate1b). Some of the nodes or amplitude minima may be associated with the excitation function at the source; the effects of multipathing can be
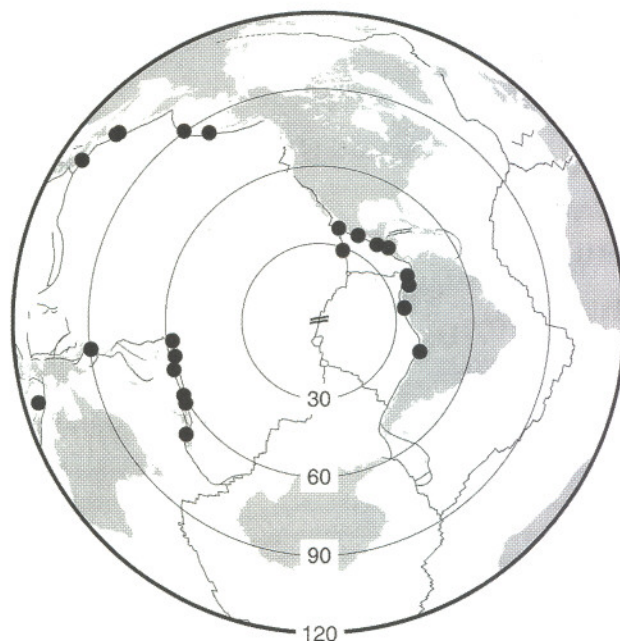
Figure 1. Epicenters of earthquakes used as Rayleigh wave sources in this study. Projection is azimuthal equidistant centered on middle of array, with distance in degrees noted by concentric circles. Geometry of the array indicated by two short, parallel lines. Thin lines mark plate boundaries. The two earthquakes far to the west of the array in the Solomon and Banda Seas were useful only at periods greater than 30 s, at which multi-pathing interference effects are less severe.

extreme when the waves leave the source in a near nodal direction where significant phase shifts in the interfering waves can be introduced by small changes in azimuth.

3. FORWARD PROBLEM

3.1 Coordinate Systems

A local cartesian coordinate system is set up for each earthquake source based on the epicenter and the location of a reference station (Figure 3). The reference station is established as the origin of the coordinate system, with the +x direction in the direction of propagation along a great circle path from the epicenter and the +y direction 90° counterclockwise from the x direction. The x-coordinate is equal to the difference between the epicentral distance at each point and the epicentral distance to the reference station, thus removing the curvature of the expected wavefront on a uniform, spherical earth. The y-coordinate is defined as the distance along a small-circle about the epicenter from the point to the x-axis (the great-circle path through the reference station). This earth-flattening coordinate system
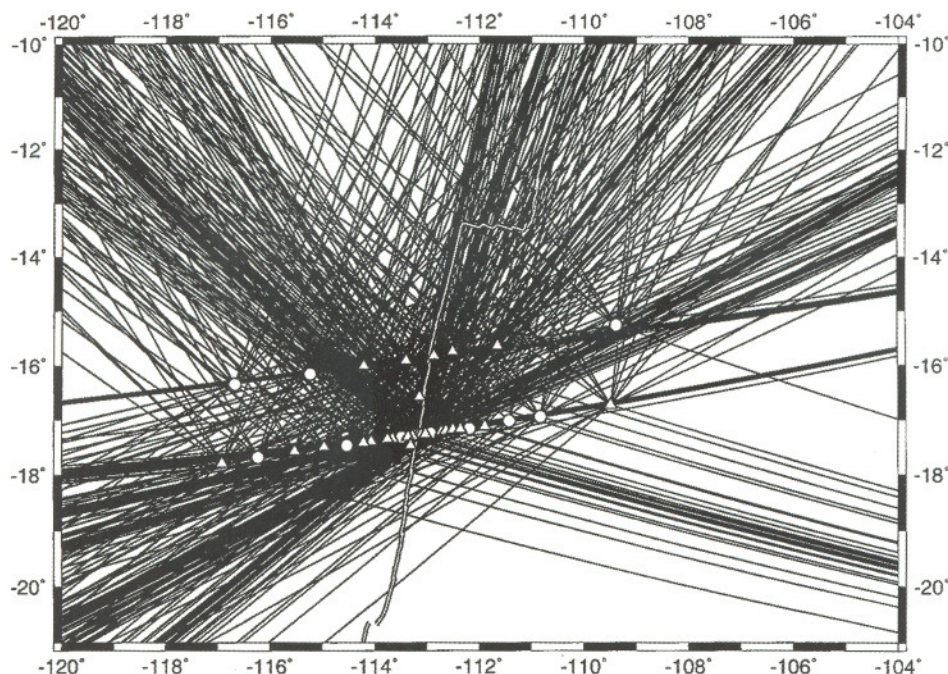
Figure 2. Great circle paths in the vicinity of the array used in the analysis at 0.035 Hz. Double line marks location of plate boundary, circles and triangles the locations of instruments. Some tomographic control is possible wherever there are crossing paths.

represents a compromise between (1) a mercator projection centered about an axis about 90° from the study area, and (2) a mercator projection centered on the epicenter. Option (1) preserves the relative geometry of points within the study area, but leaves the wavefront curved. Option (2) flattens the wavefront and makes all great circle paths lie along the x-direction, but distorts distances between points in the y-direction. The coordinate system used in this paper flattens the wavefront, but preserves distances between points better than option (2), which is important because the waves may not be approaching the array along a great circle path and the interpolation scheme we use for phase velocity is based on distances to fixed grid points.

Rather than describe variations in phase velocity in terms of a set of orthogonal basis functions or a set of cells of constant velocity, we use a continuous function that is a weighted average of velocities at neighboring grid points. The characteristic velocity at the jth grid point is azimuthally anisotropic and given by

$$_iV_j = B_{0,j} + B_{1,j} \cos 2\,_i\theta_j + B_{2,j} \sin 2\,_i\theta_j \quad (1)$$

where $_i\theta_j$ is the backazimuth from the jth grid point to the ith event in the original geographic coordinate system. Higher order azimuthal terms are neglected as they are expected to be small for Rayleigh waves [*Smith and Dahlen*,

1973]. We use a 2-D Gaussian weighting function that has a characteristic scale L_w. The phase slowness at every point (x,y) in the medium, including at the grid points themselves, is given by

$$_iS = \sum_{j=1}^{N} \frac{_iw_j}{_iV_j} \Big/ \sum_{j=1}^{N} {_iw_j} \quad (2)$$

where N is the number of grid points, the weights are given by

$$_iw_j = \exp\left[-\frac{(x-_ix_j)^2 + (y-_iy_j)^2}{L_w^2}\right]$$

and $(_ix_j,_iy_j)$ is the location of the jth gridpoint in the coordinate system of the ith event. The weighting function acts as an averaging or smoothing function, with L_w controlling the length scale for variations in velocity of the medium. With this approach, velocities are described everywhere, even outside the actual grid of points employed, and will return to the average velocity of the grid at infinite distance.

There is no absolute criterion for selecting L_w. As is usual for this type of inverse problem, there is a tradeoff between model resolution and model variance or between data misfit and model length. Decreasing L_w increases the resolution,
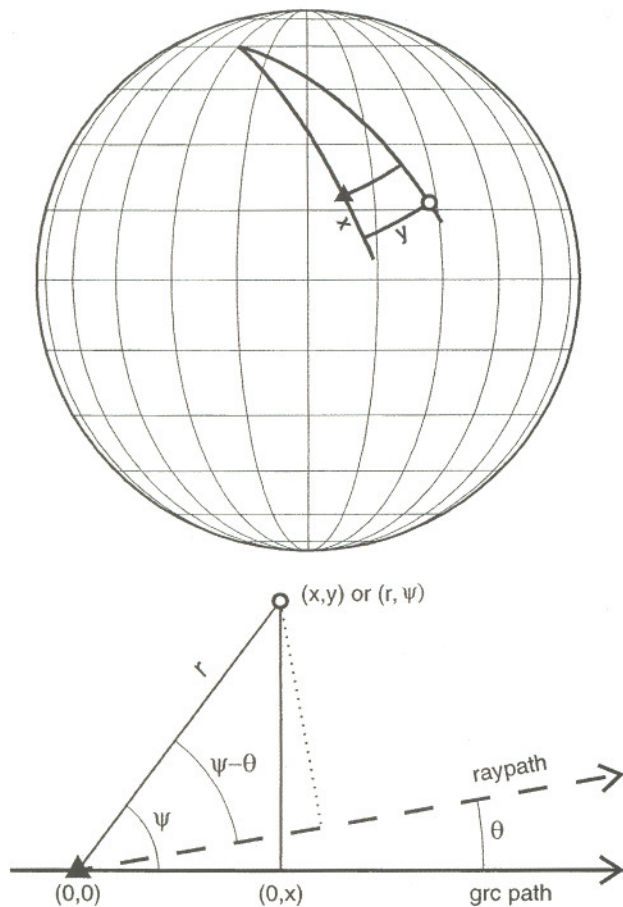
Figure 3. A different coordinate system is employed for each event. Upper diagram shows two great-circle paths from the epicenter to stations in the array. x-coordinate is distance along great circle path relative to distance to reference station (triangle). The y-coordinate of any other point (circle) is distance along the perpendicular, small-circle passing through that point to the great circle path through the reference station. These coordinates are used as Cartesian system for local analysis of phase velocities, converted to local polar coordinates (r, ψ) for convenience in describing local direction of propagation θ of an incoming plane wave (lower diagram).

allowing more details of velocity variations to be resolved or represented, usually resulting in smaller data residuals, but at the expense of greater amplitude, shorter wavelength, perhaps unrealistic, oscillations in the velocity model that have greater uncertainty. It is a choice between resolving more with greater uncertainty and less with greater certainty that must be made in the light of the particular tectonic question that is being addressed by the tomographic study.

The actual grid of points employed in this study is shown in Plate 1a. There is no requirement that the points be regularly spaced on a rectangular grid. Here we base our grid on the natural coordinate system provided by the ridge, as it is expected that the dominant variations in velocity will occur as a function of the age of the seafloor. Points are regularly spaced in latitude, but at each latitude, the center of the grid is tied to the ridge axis, which makes it easy during the inversion process to impose requirements such as that velocity vary only as a function of distance from the ridge. It is important that the grid of points extend well-outside the area of interest that will have well-resolved velocities constrained by crossing paths; in the tomographic inversion, these outer points "absorb" travel-time variations that represent additional deviations from the idealized two-plane wave representation of the incoming wavefield. Describing the resolution of the tomographic inversion is easier if the grid points are regularly spaced in the region of interest, but to reduce the total number of variables in the inversion, we employ a sparser grid in the outer, surrounding region. The total number of grid points in this model is 315.

3.2 Two-Plane-Wave Representation

At each frequency ω for each event, the incoming wave is regarded as the sum of two horizontally propagating plane waves so that the vertical displacement is of the form

$$U_z(\omega) = A_1(\omega) \exp[-i(\mathbf{k}_1 \bullet \mathbf{x} - \omega t)] + A_2(\omega) \exp[-i(\mathbf{k}_2 \bullet \mathbf{x} - \omega t)] \quad (3)$$

where $\mathbf{k}$ is the horizontal wavenumber of each wave and $\mathbf{x}$ is the position vector. In practice, we describe the position within the array relative to the reference station and use a common reference origin time for all records of a single event, so that the predicted displacement at the kth station for the ith event at each frequency is simply

$$_i^k U = {_iA_1} \exp(-i{_i^k\phi_1}) + {_iA_2} \exp(-i{_i^k\phi_2}) \quad (4)$$

where

$$_i^k\phi_1 = {_i^0\phi_1} + \overline{{_i^kS}}\,\omega\{{_i^kr}\cos({_i^k\psi} - {_i\vartheta_1}) - {_i^kx}\} + \omega({_i^k\tau} - {_i^0\tau})$$

and

$$_i^k\phi_2 = {_i^0\phi_2} + \overline{{_i^kS}}\,\omega\{{_i^kr}\cos({_i^k\psi} - {_i\vartheta_2}) - {_i^kx}\} + \omega({_i^k\tau} - {_i^0\tau})$$

$_i^0\phi_1$ and $_i^0\phi_2$ are the phases of the first and second plane waves at the reference station, $_i^k\tau$ and $_i^0\tau$ are the travel times along the great circle path from the edge of the study area to the kth and reference stations, $_i\vartheta_1$ and $_i\vartheta_2$ are the angular

deviations from the great circle path of the first and second waves, $\overline{{}_i^k S}$ is the average slowness, and the (x, y) coordinates of the stations are described also in terms of polar coordinates (r, ψ) centered on the reference station and oriented in terms of the great circle path. The incoming wavefield at each frequency is thus described by six parameters; the amplitude, reference phase and direction of each of two waves. The directions and relative phases of the two waves are illustrated in Plate 1b for one of the earthquakes of this study by plotting peaks of the two sinusoidal oscillations for one frequency at one instant, the reference time. Wherever the waves are in phase, they will add constructively, and where they are out of phase, they will interfere destructively, producing a characteristic interference pattern that varies slowly in space.

The travel times are found by integrating along the great circle path (x-direction) from the edge of the study area to each station. The edge is defined by the planar wavefront perpendicular to the great circle path, i.e. constant value $_i x_{edge}$ in the coordinate system for each event, at the point where the wavefront intersects one of the specified corners of the study area (Plate 1a). Thus,

$$_i^k \tau = \int_{_i x_{edge}}^{_i^k x} {}_i S dx \qquad (5)$$

In the correction for deviations from a great circle path (second terms in the expressions for phase in equation 4), the average slowness is defined as

$$\overline{{}_i^k S} = \left[{}_i^k \tau / ({}_i^k x - {}_i x_{edge}) + {}_i^0 \tau / ({}_i^0 x - {}_i x_{edge}) \right]/2 \qquad (6)$$

Both the average slowness and the travel times for each path are functions of the velocity coefficients at all grid points as defined by the Gaussian averaging functions, although of course the points nearest the great circle ray path are most heavily weighted. The expressions for phase are exact in Cartesian coordinates if the waves lie along the great circle path or if the velocity in the medium is uniform, but if the waves deviate from a great circle path in a heterogeneous medium, the second term in these expressions for phase provides only a first order correction. This approximation is justified because, as will be shown later, deviations from a great circle path are small and typical velocity variations are only a few percent.

4. INVERSE PROBLEM

Solution of the inverse problem for the wave parameters and velocity parameters is accomplished through two sets of iterations with two stages of inversion in each iteration. Rather than trying to match the observations in terms of normalized amplitude and phase, we choose to minimize misfit to the real and imaginary components. This approach has the advantage that misfits in phase have little importance at those stations where destructive interference causes low amplitudes and rapid phase fluctuations; a satisfactory fit is found if the predicted real and imaginary components are both small, regardless of phase. To begin the inversion, the station with the largest amplitude is chosen as the reference station, because the two plane waves are most likely to be in phase at that station and the amplitude should be representative of the sum of the amplitude of the two waves. Observed amplitudes at other stations are normalized by the reference station amplitude and phase, and the phases of the two plane waves in the starting model are set to zero at this point (but are not fixed during the inversion). With 6 wave parameters required to describe the incoming wavefield for each event and 2 pieces of information per station (amplitude and phase or real and imaginary component), a minimum of four stations per event is required to have any information about the velocity structure. In practice, a local array of 5 or 6 stations yields reasonable results with a measure of uncertainty, but clearly better results are obtained with more stations.

In the first stage of each iteration, the velocity model is fixed either with starting values or with the current values found in the previous iteration. In this stage, we find the best fitting wave parameters in a least squares sense using a downhill simplex method of simulated annealing [p. 444, *Press et al., 1992*]. We found that this stage was necessary, because the combination of the periodic non-linearity of the problem and the ambiguity in the solution if the two plane waves have similar azimuth makes it very difficult to achieve global minimization of the objective function with standard linearized inversion techniques. With the velocity fixed, each event is independent, so we perform a separate simulated annealing search for the 6 wave parameters for each event. To insure (almost always) that a global minimum is found, we start the simulated annealing in each iteration for each event with several different wave-parameter models, one of which is the best model from the previous iteration.

In the second stage of each iteration, we use a standard linearized inversion technique [*Tarantola and Valette*, 1982] to solve simultaneously for corrections to the current velocity model (values of velocity and azimuthal anisotropy at the

grid points) and wave parameters for each event. The solution to the general non-linear least-squares problem is

$$\Delta\mathbf{m} = (\mathbf{G}^T\mathbf{C}_{nn}^{-1}\mathbf{G} + \mathbf{C}_{mm}^{-1})^{-1}(\mathbf{G}^T\mathbf{C}_{nn}^{-1}\Delta\mathbf{d} - \mathbf{C}_{nm}^{-1}[\mathbf{m} - \mathbf{m}_0]) \quad (7)$$

where $\mathbf{m}$ is the current model, $\mathbf{m}_0$ is the original starting model, $\Delta\mathbf{m}$ is the change to the model, $\Delta\mathbf{d}$ is the difference between the observed and predicted data for the current model, $\mathbf{G}$ is the partial derivative or sensitivity matrix relating predicted changes in $\mathbf{d}$ to perturbations in $\mathbf{m}$, and $\mathbf{C}_{nn}$ and $\mathbf{C}_{mm}$ are the *a priori* data and model covariance matrices, respectively. The observed data are the real and imaginary components at a single frequency (we invert each frequency independently) for each of the filtered and windowed seismograms. In the examples discussed later at 0.035 Hz, we have 629 seismograms from 21 earthquakes, or 1258 observations (real and imaginary from each record). The predicted data are functions of the wave parameters and the velocity parameters at the nodes or grid points, which translate to integrals of the weighting functions along the great circle paths. The velocity and wave parameters together make up $\mathbf{m}$. With B_0, B_1 and B_2 at each of 315 nodes and 126 wavefield parameters there are potentially 1071 parameters, leading to non-unique or underdetermined model solutions. In addition, there may be other parameters such as the DPG gains or an attenuation coefficient. In the terminology of Menke [1984], it is a mixed determined problem. It is overdetermined in the sense that there are more observations than model parameters, but it is underdetermined in the sense that not all the model parameters are well resolved.

Regularization or damping of underdetermined solutions is achieved by introducing diagonal and off-diagonal terms into $\mathbf{C}_{mm}^{-1}$. We use either a minimum length criterion employing diagonal terms only [*Marquardt*, 1970], or both minimum length and a smoothing criterion using off-diagonal terms also [*Constable et al.*, 1987]. For the minimum length criterion, we employ uniform variance for points in the interior of the grid. We increase the variance for points on the edge of the grid by a factor of 10 so that these essentially undamped values can "absorb" travel-time deviations from the two-plane-wave model without introducing unnecessary velocity variations into the interior area of interest. Only very light damping is applied to the wave parameter terms so that analysis of the uncertainty or resolution of the velocity model will include full trade-offs with the wave parameter terms. Table 1 indicates what we mean by light damping; the difference between the total rank and the velocity rank is the number of independent combinations of wavefield parameters that are resolved, which range from 94 to 97% of the total of 126 wavefield parameters. Starting mod-

els for the wave parameters are from the simulated annealing inversions in the previous stage and the $\mathbf{m}$ - $\mathbf{m}_0$ terms (equation 7) penalizing changes from the original starting model are dropped for the wave parameters.

When the smoothing criterion is applied, the inversion minimizes the 2D second derivative for velocities [*Maurer et al.*, 1998] in the interior of the grid and the first derivative at the edges. For simplicity, we use the same variance (or weight) for the diagonal component of $\mathbf{C}_{mm}$ in the smoothing criterion as for the minimum length criterion, although clearly one could choose to vary the relative importance assigned to these regularization criteria. No smoothing is used for the wave parameter terms.

In the first set of iterations, all the observations are assigned equal variance. The data covariance matrix is assumed to be diagonal. Experience shows that a typical misfit to the normalized real and imaginary terms is on the order of 0.1, which we chose as the initial, *a priori* estimate of standard deviation in the first set. Typically, after 4 or 5 iterations, no further significant changes in velocity occur, although minor fluctuations in the model continue with further iterations due to the small random variations in wave parameters introduced in the simulated annealing steps. We usually terminate the inversion after 10 iterations. At that point, we estimate the *a posteriori* standard deviation of the data from each individual earthquake. Because some earthquakes have better signal-to-noise ratios than others and the wavefield from some earthquakes at some periods may not be well-described by the two-plane-wave model, there is a wide variation in how well the observations from different earthquakes are fit. The dominant source of variance in the original data is the deviation from the ideal plane wave character. If this deviation is not well-described by our simple representation of the incoming wavefield, we do not want the poor wavefield model to bias the velocity model. Therefore, in the second set of iterations, we assign *a priori* standard deviations to the observations based on the *a posteriori* standard deviations found for each earthquake after the first set of iterations. Our experiments indicate that with this scheme, eliminating one or two events with large residuals has little effect on the velocity model.

Our inversions are based on least squares minimization which implicitly assumes a Gaussian distribution of errors. We find no *a posterior* evidence in the form of an excess number of outliers in the residuals to individual events that would indicate a violation of this assumption, provided that the data have been carefully cleaned to avoid processing errors, glitches in the data, and noisy records. After the initial inversions, we routinely reexamine the data for any event with large residuals. If we identify clear problems, we either correct the error or eliminate that record if we had

Table 1. Velocity Models Used in Inversions at 0.035 Hz.

Model	Velocity Variations	Averaging Length	Damping	Total Rank	Velocity Rank	RMS Misfit Real&Imag	RMS Phase Misfit	Median RMS Misfit
1	uniform isotropic	65 km		123.6	1.0	.119	.95 s	.76 s
2	uniform anisotropic	65 km		123.8	3.0	.114	.90 s	.74 s
3	1-D isotropic	65 km	min length smoothing	138.4	16.8	.087	.65 s	.57 s
4	2-D & 1-D aniso	65 km	min length smoothing	219.2	101.0	.075	.51 s	.42 s
5	2-D isotropic	50 km	min length	255.5	133.2	.075	.51 s	.42 s

inadvertently accepted data with poor signal-to-noise ratio. If there is no obvious cause for the large residual, we leave the poorly fitting data in the analysis.

The renormalization of standard deviations changes the weight assigned to each earthquake based on the result of the first inversion. As a result, there is no one measure available to straightforwardly compare the quality of fit of different velocity models; every inversion of the same data with different model parameters or smoothing length or damping parameters will have somewhat different weights assigned to the data. Because the weights or standard deviations are different, we cannot directly point to the reduction in variance to describe how much better one model is than another in fitting the data. Because renormalization reassigns standard deviations based on how well each model fits the data, each model will tend to fit equally well in terms of minimizing the final objective function, i.e., the square of the ratios of the misfits to the standard deviations renormalized after the first set of iterations. This is a small price to pay for an inversion method that is insensitive to bad data or poor fits to one or two earthquakes. We can provide approximate descriptions of the quality of fit by removing the renormalization in a variety of ways at the end of the process.

We employ three approximate descriptions of quality of fit to the data (Table 1): the root-mean-square (rms) misfit in terms of real and imaginary components, unnormalized by multiplying by the standard deviations for each event as used in the second set of iterations; the rms phase misfit converted to equivalent seconds, based on all stations and events; and the median rms misfit in phase for an individual earthquake source. The first of these measures is closest to representing our actual minimization that is performed on real and imaginary components rather than phase and amplitude. The second represents the misfits that are most directly related to travel times and the velocity structure. Both of these measures, however, are sensitive to one or two poorly fit events that are given little weight in the renormalized, final set of iterations. The median rms phase misfit is more representative of the effect of renormalization, as it characterizes the "typical" event that is given intermediate weight in the inversion.

5. TWO-PLANE-WAVE SOLUTIONS

The solutions for relative amplitudes of the two plane waves for each earthquake demonstrate that wavefield complexity or interference is the dominant source of variance in travel times or phase from the simplest model of a plane wave propagating across a region of uniform velocity. If the ratio of the amplitude of the second, smaller plane wave to the first, larger wave is R_w, then the maximum possible phase shift γ introduced by the second wave is $\sin^{-1} R_w$. If the phase shift throughout the study area is approximately random, then the expected root-mean-square (rms) time deviation is approximately $T\gamma/(2^{3/2}\pi)$, where T is the period of the wave. Figure 4 shows the distribution of R_w for two frequencies. These ratios are taken from solutions in which the phase velocity is 2-D and anisotropic, model 4 in Table 1. At 0.035 Hz or 28.6 s, roughly in the middle of our useful frequency range, seven of 21 events have $R_w > 0.5$. The median ratio is 0.28, which corresponds to an expected rms scatter of 0.91 s in travel-time at 0.035 Hz. In contrast, in the simultaneous solution for wave parameters and velocity variations, the median rms misfit in phase at this frequency corresponds to 0.42 s. If the velocity is assumed to be uniform, the median misfit is 0.76 s after solving for the best velocity and wavefield parameters (Table 1). Thus, wavefield complexity is expected to be a greater contributor to variance than either velocity variations or noise (given our selection of high-quality signals).
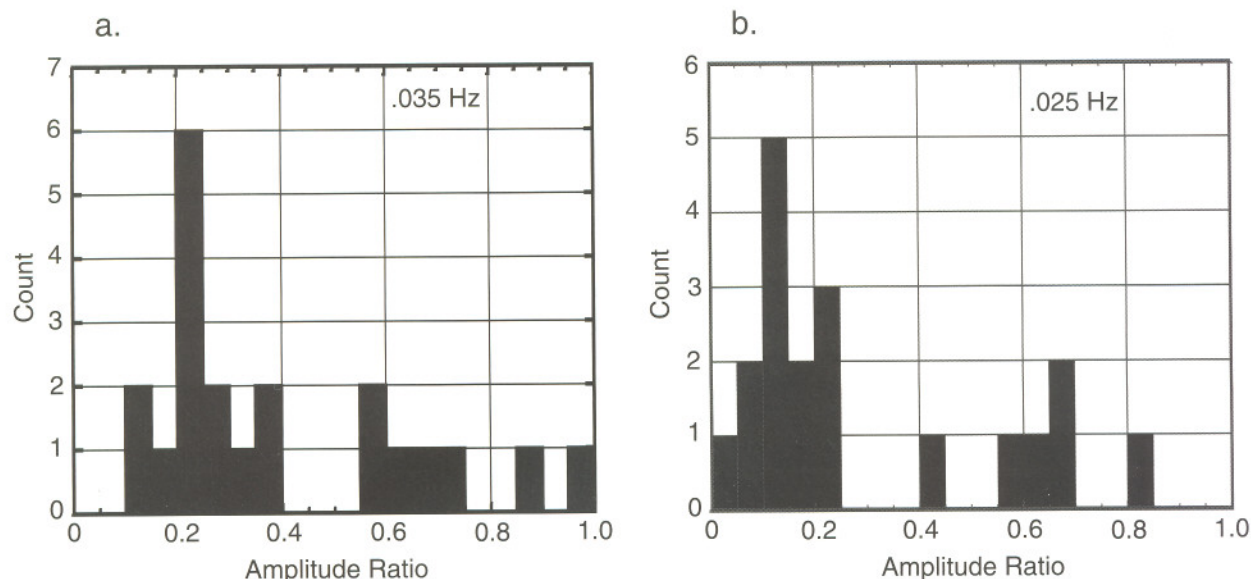
Figure 4. Amplitude ratios of the smaller of the two plane waves to the larger. Solutions from model 4, Table 1. (a) 0.035 Hz and (b) 0.025 Hz. Note that typical ratio decreases with decreasing frequency, indicating decreasing effect of multipathing.

At 0.035 Hz, virtually every event deviates significantly from a plane wave, despite the simplicity of the paths that are confined almost entirely to the Pacific basin. Figure 5 illustrates the details of fits for two events at this frequency. The Kurile earthquake has the largest amplitude variations, roughly an order of magnitude difference between the largest and smallest records (Plate 1b). The great-circle paths from this event traverse the deep water of the northwest Pacific and pass close to the Hawaiian hotspot. Despite the large variations, the interference pattern is relatively simple and the two-plane-wave model fits the observed amplitudes well with an R_w of 0.87 (Plate 1b and 5a). This event also has the largest rms deviations (1.09 s) between observed and predicted phase of any of the sources at this period. The phase misfits, however, stem largely from stations where destructive interference is occurring and the amplitudes are small (Figure 5b). At these stations, slight deviations from the model cause large phase shifts. It is important to keep these stations in the inversion, however, because their amplitudes constrain the wavefield parameters. No damage is done to the velocity model by inclusion of these stations, because there is no direct penalty in the least squares minimization for phase misfit; we minimize the real and imaginary component misfits, which are well satisfied by the modeled low amplitudes. The quality of overall fit for this event is average as measured by the rms misfit of real and imaginary components.

A more typical source at 0.035 Hz is the Oaxaca earth-quake. It shows amplitude variations of about a factor of 2 (Figure 5c) and is characterized by an R_w of 0.27, compared to the median value of 0.28. These amplitude variations are well-resolved and the phases are extremely well-fit at all stations (Figure 5d). The limit of resolution of R_w for this data set is about 0.1, corresponding to amplitude variations of +/− 10% and rms phase variations of about 0.3 s. At shorter periods, scattering becomes more severe, particularly for events from the western Pacific traversing regions where the short periods are sensitive to the thick water layer and are highly dispersed. The spectrum of R_w for events employed in the inversion, however, is not much different than at 0.035 Hz because we discard highly scattered arrivals. At 0.06 Hz, the median value of R_w is 0.31, but 6 of the 21 events were rejected. At longer periods, wavefield complexity is reduced [*Friederich et al.,* 1994]. At 40 s period, the median value of R_w is 0.19, which means that significant improvement in fit is obtained by using two plane waves for about half of the events. Thus, as expected, the degree of scattering or multipathing decreases with increasing period.

It is probably not realistic to interpret the two plane wave solutions literally as they are intended to be approximations to a more realistic and more complicated wavefield. Nevertheless, it is interesting to examine the character of the interference pattern implied by the solutions. The intensity of the interference is controlled by the amplitude ratio. The interference pattern consists of alternating maxima and minima aligned along a direction that is the bisector of the normals to the two wavefronts, typically nearly aligned in the
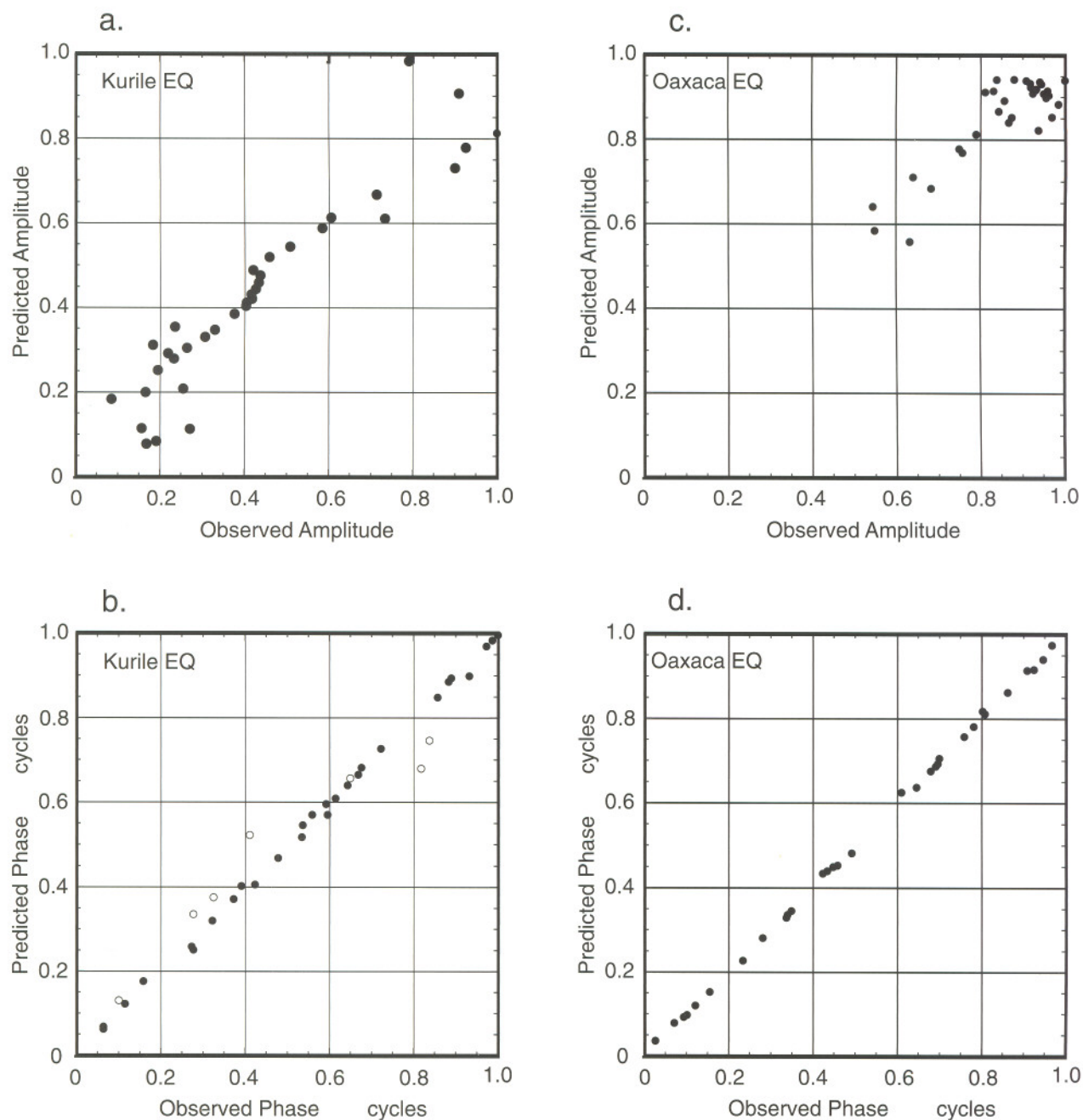
Figure 5. Comparison of observed and predicted (a) amplitudes and (b) phases for the Kurile event at 0.035 Hz and the observed and predicted (c) amplitudes and (d) phases for the Oaxaca earthquake (Ms 7.0 at 0308 on 25 Feb. 1996). Scatter appears to be least at intermediate amplitudes, because many of the intermediate values happen to come from the very closely spaced stations near the center of the southern line of the array (Plate 1b). In (b), phases from stations with observed amplitudes less than 0.2 are shown with open symbols. Note decreased accuracy of predicted phases for these stations with maximum, destructive interference.

great circle propagation direction (Plate 1b). The wavelength between maxima is controlled by the angle separating the two normals, β, and is given by

$$\lambda_I = \frac{cT \cos \beta/2}{\sin \beta} \qquad (8)$$

where c is the phase velocity of the plane waves (assumed to be equal for both). In the example given above for the Kurile earthquake at 0.035 Hz, the direction of propagation of the larger plane wave is 0.45° counter-clockwise (−) from the great circle path and the direction of the smaller wave is 3.98° clockwise (+). For the Oaxaca event, the primary wave is +3.85° and the smaller is −1.64°. All but one of the 21 primary waves at 0.035 Hz lie within 5° of the great circle path (Figure 6a). The secondary waves are more scattered and the scatter is greater the smaller the secondary wave (Figure 6b). All of the secondary waves for events with $R_w > 0.5$ lie within 12° of the great circle path. Of the six secondary waves that deviate by more than 12°, four have $R_w < 0.25$. In most cases, the interference pattern is aligned within a few degrees of the great circle path, similar to the trends found in Germany using the more complicated basis function expansions of wavefields [*Friederich et al.,* 1994; *Friederich,* 1998]. The mean difference in angle between the two waves is 12.8°, the median is 7.6°, and the maximum is 35.4°, corresponding to λ_I between interference maxima of about 490, 820, and 180 km, respectively. Thus,

although there are substantial variations, the typical scale of interference we detect is on the order of the maximum dimension of the array. This result is not surprising, because interference on much smaller wavelength scale would lead to rejection of the data as being incoherent from station to station and interference at much larger wavelengths would be difficult to resolve.

One of the common questions we have received from others who have used this method for their data sets is "How do you know when the two-plane-wave method will help?" We also have received skeptical comments such as "It works in the Pacific because the paths are so simple, but it won't work in my study area" and "What about other effects on amplitude, such as attenuation, radiation pattern and focusing/defocusing due to 3D structure?". Using the two-plane-wave method as opposed to the usual single plane wave method never hurts; if a single plane wave is an adequate representation, then the model finds a solution that is equivalent. The two-plane-wave method is most effective in the transition period range between a simple, planar wavefront and incoherent scattering. If, in examining the filtered, narrow-bandpass records from the array the amplitude is highly variable between adjacent stations or there are more than two beats in the bandpassed seismogram, then two planes waves will probably not provide a good representation of the scattering. The orthogonal polynomial approach of Friederich and Wielandt [1995] is effectively a multi-plane-wave representation of the incoming wavefield [*Pollitz,* 1999],
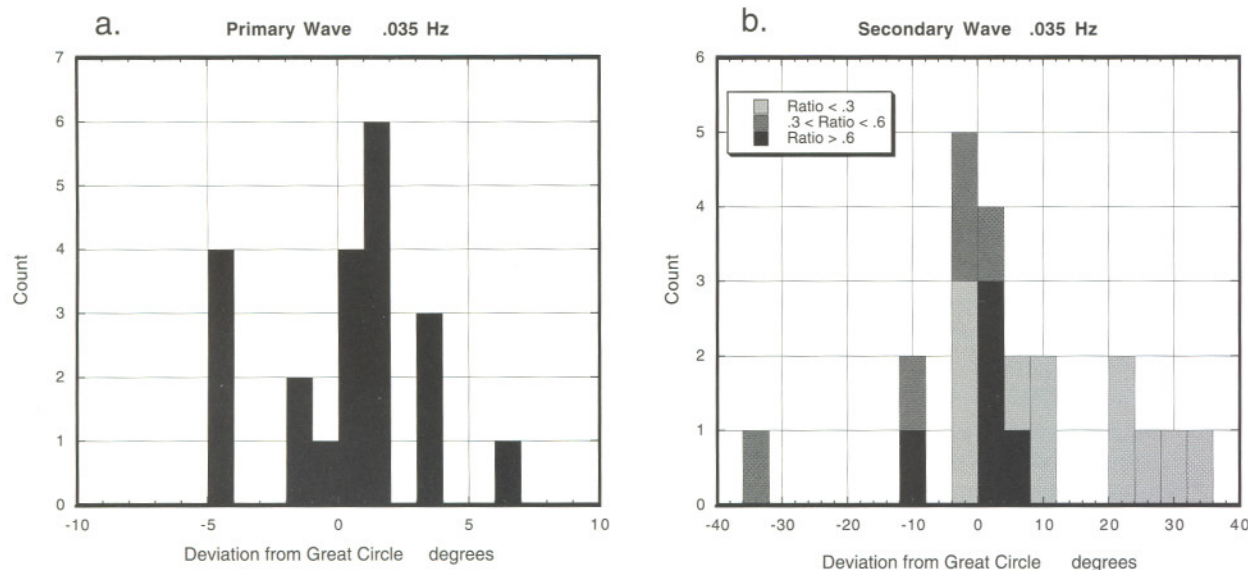
Figure 6. Distribution of differences between azimuths of the model plane waves and the great circle direction at 0.035 Hz. (a) Azimuths of the waves with the larger amplitudes of the pair for each event typically vary only a few degrees from the great circle direction. (b) Azimuths of the smaller waves are more scattered than the larger wave and scatter increases with decreasing amplitude ratio of the smaller to the larger. Note change in scale from lefthand plot.

but, unless there is a very high density of stations in a 2-D array, we find there is a rapid transition with decreasing period between the case where two-plane waves is an adequate description and highly incoherent scattering.

This approach has proved to be useful for arrays in a variety of settings, including Tanzania, New England, and Colorado. The complexity of paths to the array just changes the relative amplitudes of the two plane waves and the period of the transition to incoherent scattering. The radiation pattern is relatively unimportant as long as the dimensions of the array are much less than the distance to the source, so that range of azimuths from source to receiver is small. In this study, the range of azimuths is typically about 3 degrees. Attenuation is a minor effect across a small array, but we have performed inversions in which we solve for attenuation coefficients. As we will report in another paper, typical Q for Rayleigh waves in the MELT study area is about 80. Focusing and defocusing inside the array can have a bigger effect. We are incorporating into the method amplitude and phase response kernels for each of the plane waves to represent the internal scattering effects. Of course, the whole point of using two plane waves is to approximately represent the scattering effects of unknown 3-D structure outside the array.

6. PHASE VELOCITY SOLUTIONS

To explore the information contained in the data about the variations in phase velocity, we have performed a large set of inversions with different degrees of complexity allowed in the velocity model. Models range from constant velocity throughout the study area to 1-D variations that are a function only of distance from the ridge to 2-D variations in both isotropic velocity and azimuthal anisotropy. Within the models allowing lateral variations, resolution can be controlled by varying the averaging length L_w or by adjusting the *a priori* model covariance. For all the models summarized in Table 1, we employed an *a priori* model standard deviation of velocities at the grid points of 0.2 km/s for both B_0 and the anisotropy terms B_1 and B_2. This choice stabilizes the inversion, but is not very restrictive as the standard deviation is equal to or larger than the deviations from the mean velocity at each period within the well-resolved part of the grid inside the array. Increasing the standard deviation beyond this level makes little difference in the amplitude of the velocity variations for our particular problem and choices of L_w, but decreasing it dampens the velocity variations. With this choice, resolution is determined primarily by L_w.

Resolving power can be briefly summarized by the rank of the problem, equivalent to the number of pieces of information extracted from the data. The rank can be estimated by summing the diagonal elements of the resolution matrix $\mathbf{R}$, defined by

$$\mathbf{R} = (\mathbf{G}^T \mathbf{C}_{nn}^{-1} \mathbf{G} + \mathbf{C}_{mm}^{-1})^{-1} \, \mathbf{G}^T \mathbf{C}_{nn}^{-1} \mathbf{G} \qquad (9)$$

The contribution to the rank from the velocity terms can be summed separately from the wave parameters; both the total rank and the rank from the velocity terms are listed in Table 1. The rank from the wave parameters in this linearized inversion is nearly independent of the rank from velocity terms and typically about 97% of 6N, where N is the number of earthquakes. In addition, the wave parameters are unconstrained in the simulated annealing stage. Thus, the light damping of these parameters has essentially no effect on the velocity solution and the linearized descriptions of resolution and variance of the velocity solutions include nearly the complete tradeoff with wave parameters.

The inversions in Table 1 do not include the gains of the DPGs as variables. These inversions are for a single period, but the gain is best estimated from all periods. The approach we use is to solve for the gain factors in an initial inversion equivalent to model 1 with uniform velocity and wave parameters. We then find the weighted average of the gain factors at all periods for each station and apply them as station corrections before redoing all the inversions.

The simplest models are those with constant velocity throughout the area. If no azimuthal anisotropy is allowed, the average velocity B_0 at 0.035 Hz is $3.758 \pm .004$ km/s (model 1, Table 1). Adding 2θ terms reduces all three measures of misfit (model 2, Table 1), yielding estimates $B_0 = 3.736 \pm .006$, $B_1 = -0.067 \pm .007$, and $B_2 = -0.021 \pm .006$ km/s. The peak-to-peak variation in velocity with azimuth is nearly 4% with the fast direction orthogonal to the ridge axis. The direction is in agreement with that inferred for young sea floor on the East Pacific Rise in previous studies of Rayleigh wave anisotropy using long paths to land stations [*Forsyth*, 1975; *Nishimura and Forsyth*, 1988; *Montagner and Tanimoto*, 1990; *Laske and Masters*, 1998; *Larson et al.*, 1998], but the amplitude is about twice the estimate from those previous studies. It also agrees well with the direction found for shear wave splitting in the MELT Experiment [*Wolfe and Solomon*, 1998]. In tomographic studies, it is often assumed that neglecting azimuthal anisotropy will cause no significant bias in the isotropic model, because the paths are sufficiently distributed in azimuth to average out these effects. Here we find that the change in B_0 caused by introducing azimuthal terms is small, but significant at the 99% confidence level.

The uniform velocity model is used as a starting model for an inversion in which azimuthal anisotropy is neglected and B_0 is allowed to vary only as a function of distance

from the spreading center. Given the design of the grid, we find the velocities by simply collecting the partial derivatives from all points at common distances from the ridge axis and reducing the number of velocity variables to the 21 sets of these points (Plate 1a), but it could also be done by introducing off-diagonal terms into C_{mm} that force perfect correlation between grid points sharing a common distance. The improvement in fit achieved by introducing these lateral variations (model 3, Table 1) is much greater than that achieved by the introducing the uniform anisotropic terms. In this setting, we expected that the variation in velocity would be dominated by the effects of the cooling of the oceanic lithosphere as a function of the age of the seafloor. This expectation is largely confirmed by the large velocity variations (Plate 2a) and large reduction in data variance found for this model compared to model 2 (Table 1). The minimum in velocity is offset somewhat from the ridge axis and the gradients are asymmetric, leading to higher velocities to the east, on the Nazca plate side of the East Pacific Rise. These asymmetries are discussed more thoroughly in Forsyth et al. [1998]. The pattern is similar to that found at longer periods and greater distance scales by Ekstrom et al. [1997].

Maps of phase velocities are constructed using the same Gaussian weighting function of neighboring grid points (Equation 2) as employed in the inversion. Thus each point on the map represents a tapered average of phase velocities in the surrounding region, but this value is an accurate representation of velocities actually assumed in the inversion rather than a smoothed or filtered version of velocities in cells. Uncertainties in the maps are described by the variance of the weighted averages found by linear error propagation [*Clifford*, 1975] and given by

$$\sigma_v^2 = \mathbf{q} \mathbf{C}_{MM} \mathbf{q}^T \qquad (10)$$

where $\mathbf{q}$ is the vector of weights for all the grid points for a particular position and $\mathbf{C}_{MM}$ is the complete, *a posteriori*, model covariance matrix

$$\mathbf{C}_{MM} = (\mathbf{G}^T \mathbf{C}_{nn}^{-1} \mathbf{G} + \mathbf{C}_{mm}^{-1})^{-1} \qquad (11)$$

where we have extracted the elements dependent on velocity only. It is important to use the complete covariance matrix in the sense of including off-diagonal terms, because the weighted averages are much better known than would be inferred from the diagonal variance terms alone. Typically, there is a large, negative covariance between adjacent grid points; their individual values may not be well-constrained, but their sum is. When we combine grid points, as in model 3, we still use the Gaussian weighting functions to generate the maps and equation 10 to estimate the uncertainties, ex-

cept in this case there is a perfect positive correlation between combined points represented in $\mathbf{C}_{MM}$.

The isotropic solution with velocities dependent on distance from the axis is employed as the starting model for models with 2-D variations. Model 4 (Plate 2c and Table 1) allows 2-D variations in B_0, but the azimuthal anisotropy terms are allowed to vary only with distance from the axis. The characteristic averaging length L_w is 65 km and both minimum length and smoothing conditions are applied. As is well known, minimum length solutions tend to be oscillatory. The smoothing criterion alone tends to project or continue linear variations in well-constrained parts of the map into less well-constrained parts. Applying a combination of both criteria seems to minimize the appearance of these undesired, unresolved artifacts in the inversion. The map of variance in B_0 (Plate 2d) indicates that some of the along-axis variations in velocity are significant (exceed two standard deviations), but the dominant effect is still variations with distance from the axis. Note that there is some resolution outside the array, particularly to the north where there are many crossing paths (Figure 2). As expected, the best resolution is inside the array where there is a high density of stations and crossing paths. The fit to the data with this model is remarkably good, with typical misfits in phase on the order of 0.4s or about 1/70th of a cycle.

The average azimuthal anisotropy in model 4 is similar to that found in the uniform velocity model, but there is a minimum at or just to the east of the axis that is well resolved. Maximum anisotropy is about 5 to 6%. We can construct models in which anisotropy is allowed to vary along-axis as well, but we find that the uncertainty in the inversion is too large to be useful unless L_w is increased to the point where only one value is resolved across the along-axis dimension of the array.

A more critical issue than resolving 2-D variations in anisotropy is resolving whether azimuthal anisotropy is required at all. As has been pointed out many times previously, travel times or phase shifts caused by anisotropy can always be matched by a model of lateral heterogeneities if the variations are sufficiently strong and small-enough in scale. In this case, data from the MELT array forces a reduction in scale of the needed heterogeneity by at least an order of magnitude from previous regional or global studies, yet the fast direction is consistent with those studies and the apparent degree of azimuthal anisotropy is greater. In addition, the fast direction inferred from the Rayleigh waves agrees with shear wave splitting measurements at the same stations [*Wolfe and Solomon*, 1998] and the direction expected for simple geodynamic models of the spreading process [*Blackman et al.*, 1996]. Nevertheless, a laterally heterogeneous, isotropic model can satisfy the observations just as well as
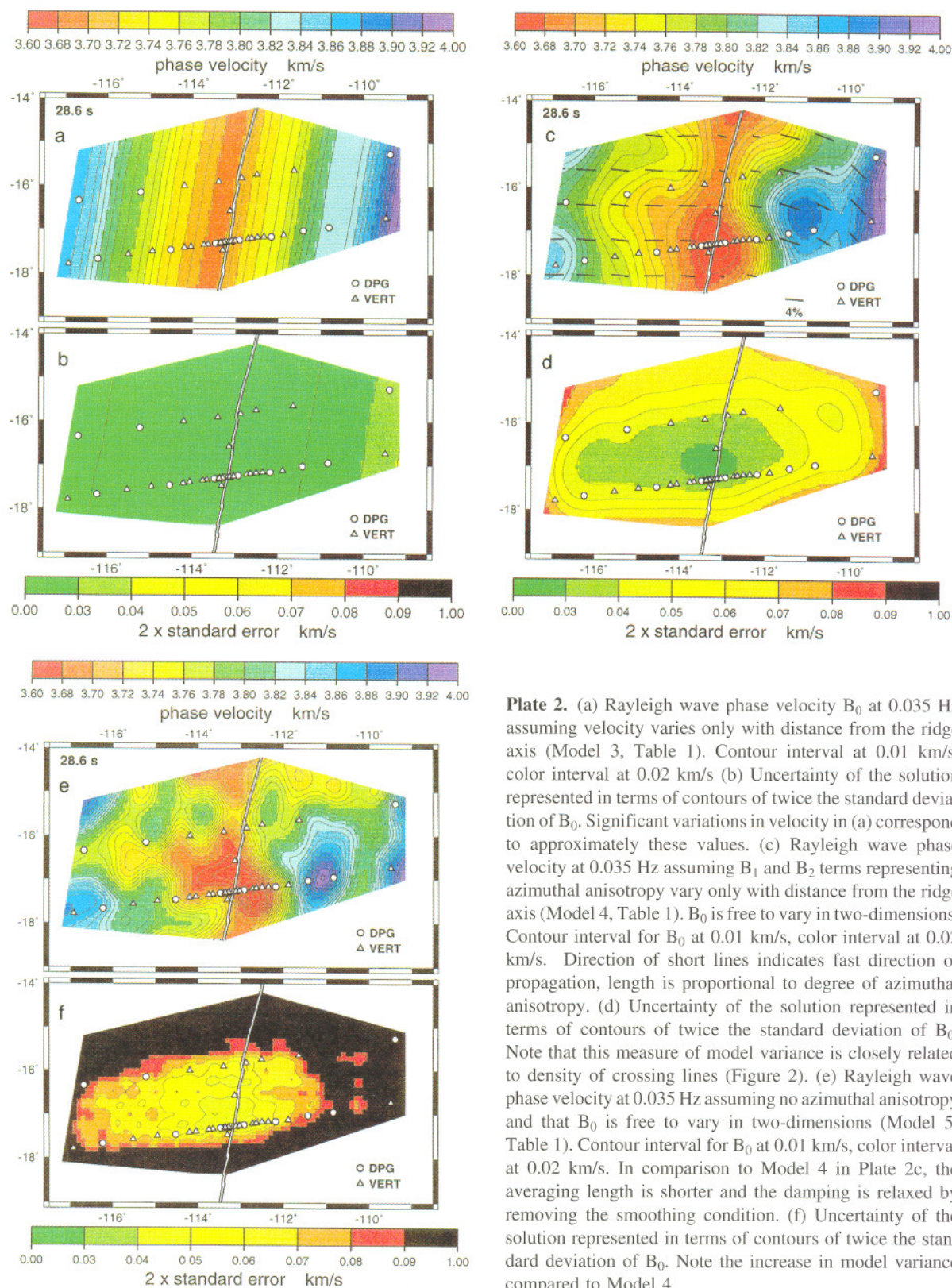
Plate 2. (a) Rayleigh wave phase velocity B_0 at 0.035 Hz assuming velocity varies only with distance from the ridge axis (Model 3, Table 1). Contour interval at 0.01 km/s, color interval at 0.02 km/s (b) Uncertainty of the solution represented in terms of contours of twice the standard deviation of B_0. Significant variations in velocity in (a) correspond to approximately these values. (c) Rayleigh wave phase velocity at 0.035 Hz assuming B_1 and B_2 terms representing azimuthal anisotropy vary only with distance from the ridge axis (Model 4, Table 1). B_0 is free to vary in two-dimensions. Contour interval for B_0 at 0.01 km/s, color interval at 0.02 km/s. Direction of short lines indicates fast direction of propagation, length is proportional to degree of azimuthal anisotropy. (d) Uncertainty of the solution represented in terms of contours of twice the standard deviation of B_0. Note that this measure of model variance is closely related to density of crossing lines (Figure 2). (e) Rayleigh wave phase velocity at 0.035 Hz assuming no azimuthal anisotropy and that B_0 is free to vary in two-dimensions (Model 5, Table 1). Contour interval for B_0 at 0.01 km/s, color interval at 0.02 km/s. In comparison to Model 4 in Plate 2c, the averaging length is shorter and the damping is relaxed by removing the smoothing condition. (f) Uncertainty of the solution represented in terms of contours of twice the standard deviation of B_0. Note the increase in model variance compared to Model 4.

model 4. Reducing L_w to 50 km and removing the smoothing constraint yields a 2-D model (Plate 2e) with nearly identical measures of misfit (model 5, Table 1). As is typical of such solutions, more parameters are required for the isotropic fit to match the anisotropic fit; the rank is increased by 32.

Averaging over smaller areas increases the variance of the mapped phase velocities (compare Plate 2f to Plate 2d) and introduces many short-wavelength oscillations between stations. If the array had been optimally designed for surface wave analysis with stations more randomly distributed in the study area, the scale of heterogeneity needed to match the anisotropic inversion could have been further reduced, but it would still have been possible to generate such a model. The scale of the heterogeneities required to mimic the effects of anisotropy in this study (Plate 2e) are substantially smaller than the typical width of the Fresnel zone for the wavelength of the Rayleigh waves and the scale of the experiment. Thus, it is highly unlikely that these short-wavelength fluctuations are real, and, applying Occam's razor, we strongly prefer the azimuthally anisotropic solution and conclude that there is an average of about 4% variation in phase velocity with direction in the vicinity of the East Pacific Rise.

7. CONCLUSIONS

Variations in surface wave amplitudes and phases caused by multipath propagation through heterogeneities often exhibit a relative simple pattern across an array of stations that can be satisfactorily described by the interference of two plane waves. We have developed a tomographic imaging method for phase velocities in the vicinity of an array of stations in which the velocity field is represented by a Gaussian weighting function of values at a set of grid points and the incoming wavefield from each source event is represented by the amplitude, phase and direction of two plane waves. For simple paths that lie primarily within the Pacific basin, the apparent azimuth of the larger of the two plane waves is typically within a few degrees of the great circle path. There is a wide variation in apparent relative amplitudes of the two plane waves, although usually interference effects are stronger and the amplitudes more nearly equal at shorter periods. Typically, the characteristic wavelength of the interference pattern is on the order of several hundred kilometers and the pattern is approximately aligned along the great circle path. Near the East Pacific Rise, the phase velocities of Rayleigh waves vary rapidly with distance from the ridge axis and azimuthal anisotropy reaches a maximum of 5 to 6%.

REFERENCES

Alsina, D., R. Snieder, and V. Maupin, A test of the great circle approximation in the analysis of surface waves, *Geophys. Res. Lett.*, 20, 915-918, 1993.

Blackman, D.K., J.-M. Kendall, P.R. Dawson, H.-R. Wenk, D. Boyce, and J. Phipps Morgan, Teleseismic imaging of subaxial flow at mid-ocean ridge: traveltime effects of anisotropic mineral texture in the mantle, *Geophys. J. Int.*, 127, 415-426, 1996.

Capon, J., Analysis of Rayleigh-wave multipath propagation at LASA, *Bull. Seism. Soc. Am.*, 60, 1701-1731, 1970.

Christensen, N.I., and M.H. Salisbury, Seismic anisotropy in the upper mantle: Evidence from the Bay of Islands ophiolite complex, *J. Geophys. Res.*, 84, 4601-4510, 1979.

Clifford, A.A., *Multivariate Error Analysis*, Wiley, New York, 1975.

Constable, S.C., R.L. Parker, C.G. Constable, Occam's inversion: a practical algorithm for generating smooth models from electromagnetic sounding data, *Geophys.*, 52, 279-288, 1987.

Cox, W.C., S.C. Webb, and T.K. Deaton, A deep sea differential pressure gauge, *J. Atmos. Oceanic Technol.*, 1, 237-246, 1984.

Ekstrom, G., J. Tromp, and E.W.F. Larson, Measurements and global models of surface wave propagation, *J. Geophys. Res.*, 102, 8137-8157, 1997.

Evernden, J.F., Direction of approach of Rayleigh waves and related problems, Part I, *Bull.Seism. Soc. Am.*, 43, 335-374, 1953.

Forsyth, D.W., The early structural evolution and anisotropy of the oceanic upper mantle, *Geophys.J. R. Astron. Soc.*, 43, 103-162, 1975.

Forsyth, D.W., S.C. Webb, L.M. Dorman, and Y. Shen, Phase velocities of Rayleigh waves in the MELT experiment on the East Pacific Rise, *Science*, 280, 1235-1238, 1998.

Friederich, W., and E. Wielandt, Intepretation of seismic surface waves in regional networks: joint estimation of wavefield geometry and local phase velocity - Method and tests, *Geophys. J. Int.*, 120, 731-744, 1995.

Friederich, W., E. Wielandt, and S. Stange, Non-plane geometries of seismic surface wavefields and their implications for regional surface-wave tomography, *Geophys. J. Int.*, 119, 931-948, 1994.

Friederich, W., Wave-theoretical inversion of teleseismic surface waves in a regional network: Phase velocity maps and a three-dimensional upper-mantle shear-wave velocity model for southern Germany, *Geophys. J. Int.*, 132, 203-225, 1998.

Larson, E.W.F., J. Tromp, and G. Ekstrom, Effects of slight anisotropy on surface waves, *Geophys. J. Int.*, 132, 654-666, 1998.

Laske, G., Global observation of off-great circle propagation of long-period surface waves, *Geophys. J. Int.*, 123, 245-259, 1995.

Laske, G. and G. Masters, Surface-wave polarization data and global anisotropic structure, *Geophys. J. Int.*, 132, 508-520, 1998.

Lerner-Lam, A.L., and J.J. Park, Frequency-dependent refraction and multipathing of 10-100 second surface waves in the western Pacific, *Geophys. Res. Lett.*, 16, 527-530, 1989.

Li, A., D.W. Forsyth, and K.M. Fischer, Evidence for shallow isostatic compensation of the southern Rocky Mountains from Rayleigh wave tomography, *Geology, 30*, 683-686, 2002.

Li, A., D.W. Forsyth, and K.M. Fischer, Shear wave structure and azimuthal anisotropy beneath eastern North America from Rayleigh wave tomography, *J. Geophys. Res.*, in press, 2003.

Marquardt, D.W., Generalized inverse, ridge regression, biased linear estimation, and nonlinear estimation, *Technometrics, 12*, 591-612, 1970.

Maurer, H., K. Holliger, and D.E. Boerner, Stochastic regularization: smoothness or similarity?, *Geophys. Res. Lett.*, 25, 2889-2892, 1998.

MELT Seismic Team, Imaging the deep seismic structure beneath a mid-ocean ridge: The MELT Experiment, *Science*, 280, 1215-1218, 1998.

Menke, W., *Geophysical data analysis: discrete inverse theory*, Acad. Press, San Diego, 1989.

Montagner, J.P., and T. Tanimoto, Global anisotropy in the upper mantle inferred from the regionalization of phase velocities, *J. Geophys. Res.*, 95, 4797-4819, 1990.

Nishimura, C.E., and D.W. Forsyth, Rayleigh wave phase velocities in the Pacific with implications for azimuthal anisotropy and lateral heterogeneities, *Geophys. J.*, 94, 479-501, 1988.

Pollitz, F.F., Regional velocity structure in northern California from inversion of scattered seismic surface waves, *J. Geophys. Res., 104*, 15043-15072, 1999.

Press, W.H., S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes in FORTRAN: The Art of Scientific Computing* (2nd edition), Cambridge Univ. Press., Cambridge, 963pp., 1992.

Raitt, R.W., G.G. Shor, T.J.G. Francis, and G.B. Morris, Anisotropy of the Pacific upper mantle, *J. Geophys. Res.*, 74, 3095-3109, 1969.

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

Smith, W.H.F. and D.T. Sandwell, Global seafloor topography from satellite altimetry and ship depth soundings, *Science*, 277, 1956-1962, 1997.

Snieder, R. and G. Nolet, Linearized scattering of surface waves on a spherical Earth, *J. Geophys.*, 61, 55-63, 1987.

Tarantola, A., and B. Valette, Generalized non-linear problems solved using the least-squares criterion, *Rev. Geophys. Sp. Phys.*, 20, 219-232, 1982.

Webb, S.C., Broadband seismology and noise under the ocean, *Rev. Geophysics*, 36, 105-142, 1998.

Wessel, P. and W.H.F. Smith, New version of the Generic Mapping Tools released, *Eos Trans. AGU*, 76, 329, 1995.

Wielandt, E., Propagation and structural interpretation of non-plane waves, *Geophys. J. Int.*, 113, 45-53, 1993.

Wolfe, C.J. and S.C. Solomon, Shear-wave splitting and implications for mantle flow beneath the MELT region of the East Pacific Rise, *Science*, 280, 1230-1232, 1998.

D.W. Forsyth, Dept. of Geological Sciences, Brown University, Providence, RI 02912-1846

A. Li, Dept. of Geosciences, University of Houston, 4800 Calhoun Road, Houston, TX 77204-5503