High-precision estimation of split PS-wave time delays and polarization directions

R. Ross Haacke

ABSTRACT

The measurement of split PS-wave time delays and polarization directions is notoriously difficult in field data, partly because the signals are small and overprinted by competing mechanisms. This contribution describes a new processing method that suppresses many of the overprinted traveltime and amplitude anomalies, allowing depth-averaged split PS-wave time delays and polarization directions to be measured simply and precisely. These depth-averaged properties are then inverted using a simple forward model, allowing an earth model of split PS-wave time delays and polarization azimuths to be estimated without the need for layer stripping. In the field data used as an example, the processing and inversion methods are used to estimate split PS-wave time delays and polarization directions for ten layers spanning about 500 m depth from the seabed downward. Inversions using data-error covariances estimated from prestack data show model uncertainties less than 0.3 ms of time delay and 3° of polarization azimuth. However, it is clear that if the data-error covariances cannot be estimated from prestack data, due to low fold for example, model uncertainties would rise considerably. Repeating the inversions using data-error covariances of a postulated form leads to a range of maximum-likelihood models. When the data-error covariances cannot be estimated from prestack data, it seems reasonable to report precision levels implied by the spread of maximum-likelihood models, which in this case is up to 0.5 ms of time delay and 20° of polarization azimuth. The principal achievement of this processing and inversion scheme is to constrain a relatively large number of depth layers with similar levels of model uncertainty. The depth resolution available to this new method may have important implications for the development of tight-gas and shale-gas plays, in which variations of stress, strain, and fracture properties in discrete layers are important.

INTRODUCTION

Multicomponent instruments deployed on the seabed record useful information from P- and S-wavefields. Incident S-waves (produced by mode-conversion from downgoing P-waves) are highly sensitive to elastic anisotropy caused by the effects of stress, pore-fluid pressure, and strain in the earth (e.g., Crampin and Peacock [2008] and citations therein). Such anisotropy leads to S-wave polarization and traveltime anomalies that depend on offset and azimuth. In anisotropic rocks, S-waves often split into independent components that propagate with different polarizations and velocities, accumulating a time delay between them (Crampin, 1981).

Split S-wave time delays and polarizations are sensitive to preferential alignment of minerals, pores and fractures. Consequently, S- or PS-wave data can be used to detect and quantify subtle properties of subsurface rocks that are of interest to geoscientists and engineers. For example, Mueller (1991) reports spatial changes in crack-related hydraulic conductivity detected through processing of S-waves, Angerer et al. (2002) discuss the S-wave expression of zones with increased pore-fluid pressure, and Olofsson et al. (2003) use PS-wave data to map strain around a zone of seabed subsidence. It has also been shown in numerous examples that long shear wavetrains generated by interference of split PS-waves can dramatically degrade the quality of PS-wave images. This motivates the measurement of PS-wave splitting and compensation for it by layer-stripping methods applied prior to migration. A clear discussion of this, with some compelling case studies, is given by Bale et al. (2009).

Given the sensitivity of PS-wave polarization and traveltime anomalies to the physical properties of earth’s rocks, it makes sense...
to measure such anomalies as precisely as possible. Many targets of interest occur in thin horizons far beneath the shots and receivers (e.g., detection of overpressured petroleum reservoirs). In these cases, the target layers generate only very small split PS-wave time delays such that the ratio of desired signal to noise is low (signal being the time delay between split PS-waves). The robust estimation of such signals is difficult for a variety of reasons. In this contribution, we look at a new method of measuring split PS-wave polarization directions and time delays that copes well with noise and instability of measurement to provide precise estimates of these parameters with high levels of depth resolution.

Methods for measuring split S-wave time delays and polarization directions have been extensively explored, particularly for land data sets with controlled shear sources that excite S-waves with different azimuths of polarization. These 2C × 2C data (two source components recorded on two receiver components) can be computationally combined to isolate the split shear polarizations, allowing the time delay between them and their polarization directions to be determined (Li and Crampin, 1993; Zeng and MacBeth, 1993; Shieh, 1997; Thomsen et al., 1999). Nearly all the cited methods are generalizations of Alford (1986) rotation in one way or another (refer to the overview in Dellinger et al. (2001) and also to their robust formulation of the procedure), and have been shown to work well with layer stripping approaches when the orientation of the anisotropic fabric varies with depth (Winterstein and Meadows, 1991; Winterstein et al., 2001).

The Alford-rotation and layer stripping methods noted above must be modified to work with PS-wave data because the S-wave energy is sourced on reflection of the downgoing P-wavefield that, under weak anisotropy, vibrates near the sagittal plane (note that the sagittal plane is the vertical plane containing the shot and receiver). Consequently, P-wave shots at different locations are needed to reliably generate the split PS-wave polarizations with high enough amplitude for their properties to be measured. PS-wave data have been processed in this way using orthogonal source-receiver pairs in a direct application of the 2C × 2C methodology (e.g., Gaiser, 1997, 1999). However, this approach suffers from overprinted signals caused by heterogeneity along the different PS raypaths as well as the anisotropy of the rocks. For example, the time taken for the downgoing P-wave to undergo mode conversion is directly affected by P-wave velocity anisotropy, and by velocity or structural heterogeneity (including reflector dip). There are ways to mitigate this problem, mainly by migration of the PS-wave data, which, if accurate enough, will account for subsurface heterogeneity, structure, and anisotropy encountered on the P (downgoing) and S (upgoing) legs of the raypath. However, all migrations of field data are imperfect, leading to errors in the timing of PS events that depend on offset and azimuth. This is particularly acute when the data contain interfering split PS-waves, which hinder the analysis of PS gathers for tomographic purposes. The main timing errors of concern arise from the P-wave leg of the raypath, which are referred to here as conversion-time differences.

Further methods for measuring the properties of split PS-waves, with varying degrees of sensitivity to conversion-time differences, have been described. These include anisotropic forward modeling and inversion of PS-waves (Gratacos, 2006; Simmons, 2009), measurement of polarity changes on the transverse component (Bale et al., 2005; Mattocks et al., 2005), or similar approaches that look for transverse amplitude nulls (Garotta and Granger, 1988; Granger et al., 2001). It is also possible to modify the Alford-type layer-stripping algorithm to explicitly correct for conversion-time differences by introducing an additional degree of freedom to the problem (e.g., Gorskalev et al., 2004; Haacke et al., 2009, 2010). One further method of interest is described by Thomsen (2001), in which problems of conversion-time difference are compensated by first reducing, in a least-squares sense, a 3D gather of PS-wave data to remove differences caused by the variation in shot position and raypath of the downgoing P-waves. The analysis proceeds by use of Alford’s rotation on the reduced data, and subsequently by layer-stripping to resolve vertical changes in split PS-wave polarization directions and time delays.

Although the problems of conversion-time difference can be decreased by better PS-wave migration or data reduction, errors persist and their accumulation during Alford-type layer stripping leads directly to a loss of vertical resolution in measured PS time delays and polarizations. This reduces the sensitivity to subtle changes in pore-fluid pressure, crack density, or applied stress that might be of interest. Experience with multicomponent data sets shows that errors in layer stripping PS data with an Alford-type method quickly accumulate to unmanageable levels when the number of layers increases beyond a handful. As knowledge of subsurface stress and strain becomes more desirable with the rise of unconventional resource plays such as tight-gas and shale-gas, advanced methods of extracting small PS time delays from seismic data are needed that can provide high levels of depth resolution. In the current contribution, the approach is to avoid layer stripping altogether, instead choosing to invert data processed in a manner that simplifies the seismic record and removes the need for accurate modeling of seismic wave propagation.

The remainder of this paper outlines a high-precision method for robustly measuring polarization directions and time delays in PS-wave data. In the method section the field data are first introduced, along with their preprocessing, then used as an example data set for application of the main processing technique. The processing produces attributes sensitive to the depth-averaged properties of split PS-waves (the split PS time delays and polarization directions). Interpretation of these attributes is described using results from the example data set. We then turn to the problem of simulating the key attributes of the processed field data by developing a forward model that is suitable for eventual inversion. Following this, the inverse scheme is described then applied assuming independent data errors. The model-parameter uncertainties arising from this scheme are investigated first by linearizing the system around its maximum-likelihood solution, then again using more expensive Monte-Carlo methods. We then turn to the issue of data errors, how to estimate them, and what form their covariances should take. In the final results and discussion section, the inversions are repeated using different data-error covariances. We conclude with a summary of precision levels achieved by the inverse problem, and discussion of accuracy of the earth model versus precision of its estimated parameters.

**METHOD AND INITIAL RESULTS**

The data processing described in this section transforms PS-wave seismic data into a type of semblance-based lag attribute that enhances depth-averaged split PS time delays and polarization directions while suppressing competing signals caused by reflector dip, heterogeneity, P-wave velocity, anisotropy, etc. These
depth-averaged attributes are then inverted for an earth model of split PS-wave time delays and polarization azimuths. The processing and inversion methods are developed and described by receiver-domain application to unmigrated field data from an ocean-bottom seismic (OBS) survey offshore northwest Svalbard. The algorithm requires no significant changes for application to data after migration, however.

**Introduction to the field data**

The Svalbard OBS data were acquired by the EC-funded Hydratech project in 2001 with autonomous 4C nodes deployed by free-fall from the sea surface. The data are of high quality, with excellent ground coupling and vector fidelity. Shooting took place with a pair of 40 cubic-inch sleeve guns that provided power at frequencies in excess of 150 Hz (4.5 m tow depth), and dominant frequency of about 30–60 Hz for the strong and abundant PS-waves. The survey placed shots at a variety of offsets up to 5 km and used the full range of azimuths, Figure 1a. An example of data from shotline five is provided in Figure 2a, in which it is clear that the PS-wave arrivals are strong at a wide range of offsets. Readers are referred to Westbrook et al. (2008) for a detailed description of data acquisition and an overview of the study area. In general, Figure 2b, the seabed dips by about 2° to the southwest and the stratigraphy dips the same way but to a lesser degree. Some steep faulting is present that is related to the nearby Molloy transform zone. Despite this, however, the reflector structure is minimal in the area and we proceed with unmigrated data in the receiver domain.

**Preprocessing**

For brevity, the navigation processing, timing corrections, receiver repositioning, tilt correction, and vector fidelity tests applied to the field data are not described here. The reader is referred to Haacke and Westbrook (2006) and Exley et al. (2010) for details. After these steps, the data were prepared for analysis using a workflow similar to that described in Haacke et al. (2009). In brief, once the data had been preprocessed, they were band-limited to 20–90 Hz and binned in azimuth sectors of 3°. The data were move-out-corrected using PS-wave velocities with a fourth-order term applied to ensure reflections are flat across the range 0–2000 m of offset, Figure 2a. It will become apparent in later sections that errors in moveout, particularly errors that depend on azimuth, are decoupled from the measured time delay between split PS-waves, and thus a simple velocity function is adequate for this analysis.

The shot pattern used in the survey provides data redundancy by stacking in offset–azimuth bins, thus simulating a ring of shots around the receiver at constant offset and at all azimuths, Figure 1b. However, the goal of the analysis is to measure time delays between split PS-waves with a high degree of precision, and it is important that stacking does not corrupt the signal being measured. Hence, the offset range in the stack must be limited because the velocity difference between split PS-wave polarizations depends on the incidence angle. In Figure 3, we see predicted S-wave velocity plotted as incidence angle varies in the vertical plane normal to a set of parallel, vertical, penny-shaped cracks, and at incidence angles of 5° and 10° from the vertical as azimuth varies. The prediction is based on the model of Hudson (1981) using liquid-filled cracks of aspect ratio 0.001 and crack density 0.02 in a matrix of soft hemipelagic sediment representative of that at the West Svalbard study site (the isotropic matrix has \( V_p \) of 1854 m/s and \( V_s \) of 394 m/s). The predicted velocities are for polarizations parallel and perpendicular to the vertical plane containing the crack normal, and are numerical predictions that are independent of the seismic acquisition parameters. Although the details are specific to the model, in general, the same trends in shear velocity occur with angles around a set of aligned liquid-saturated cracks. Importantly, at finite offset the velocity difference between split S-waves depends on the azimuth of propagation. For our application, we therefore require the
PS-waves to propagate close to the vertical (assuming that the anisotropy is caused by vertical cracks) to reduce the variation of split PS-wave time delay with sagittal azimuth and allow it to be accurately preserved in the eventual stack. This holds true for all analyses that assume horizontal transverse isotropy, HTI, in the subsurface. Although the west Svalbard data are not migrated, the points raised here about offset-limited stacking would apply equally to the preparation of data by PS-wave migration. To ensure the fidelity of signals in the current analysis is high, a subset of data are used for which the range of offsets is limited to 250–1000 m (for a target depth of 1500–2000 m). The data are prepared for stacking early in the process flow, but tests show that the most stable results come by postponing the actual stack until the last stage of the analysis described in the next section.

The key difference in preprocessing to that described by Haacke et al. (2009) is in the rotation stage, which for this analysis transforms the horizontal-component data to a 45° geometry (see Figure 1c). After rotation, the horizontal geophones are bisected by the sagittal plane. The rotation is done shot-by-shot using the direct wave to optimize the process and ensure that the effect of errors in tilt and positioning are minimized. This preprocessing flow provides two final data sets, one from each rotated horizontal geophone, that show enhanced PS-wave arrivals with limited offsets but a full range of azimuths (Figure 4a). As one would expect from the 45° rotation geometry, the two data sets look largely the same on first inspection; there are, however, important differences in the detail of the PS-waveforms when the waves have passed through anisotropic rocks. The following sections describe a way to extract accurate and robust measurements of these second-order differences in the waveforms.

Data analysis and interpretation

The significance of the 45° geometry accomplished by rotation is easiest to visualize for eight evenly spaced source locations around the receiver. These eight salient points, representations of which are illustrated in Figure 4b, provide a useful way to introduce the data processing described in the next section and serve as reference points for interpreting the results. In the following, we shall assume a horizontal transversely isotropic (HTI) or orthorhombic elastic symmetry, which are indistinguishable with a limited offset range. Furthermore, we assume that the anisotropy is weak, such that the seismic record contains split PS-waves with nearly orthogonal polarizations and that the downgoing P-wave vibrates near the sagittal plane. The errors that arise from these approximations are

Figure 2. (a) Four-component data from part of Line 5 recorded on Svalbard OBS n639. The arrivals are zero-phase filtered and flattened with a fourth-order normal-moveout correction. (b) Part of line five as recorded on a short-offset single-channel streamer. The projection of the receiver position is marked on the seabed, which dips by about 2° to the southwest.

Figure 3. Predicted shear velocities in soft hemipelagic sediments containing aligned, vertical penny-shaped cracks filled with liquid (using the model of Hudson (1981) with crack density 0.02 and aspect ratio 0.001). For S-wave polarizations parallel and perpendicular to the crack normal, we see the variation in shear velocity with propagation angle in the vertical plane containing the crack normal, and with azimuth at constant incidence angles of 5° and 10°.
small under weak anisotropy and can be neglected when dealing with field data. For example, the crack models described above (using the theory of Hudson, 1981) predict nonorthogonality of arriving split PS-waves of no more than 2° in the west Svalbard study area (Haacke et al., 2009). Finally, for simplicity, we assume a 1D flat earth, with the usual caveats that apply to this, particularly about migration and selection of short offsets.

Four salient points occur at 90° intervals when the sagittal plane coincides with a plane of elastic symmetry in the anisotropic fabric (for example, parallel or perpendicular to a set of aligned vertical cracks). In these cases, the downgoing P-wave undergoes mode-conversion but produces an upgoing S-wave that vibrates only in that plane of symmetry, such that no shear-wave splitting occurs. Consequently, the geophones show identical PS-wave signals on both horizontal components. The four remaining salient points occur when the source is midway between the planes of symmetry. In these cases, the downgoing P-wave (which vibrates close to the sagittal plane if the anisotropy is weak) produces a shear wave in each of the elastic symmetry planes that differ in their phase and group velocities. There is an attenuation and reflection-coefficient difference also, but these differences can be neglected for reasons that will become apparent in the next few paragraphs.

For shots midway between the planes of symmetry, we can think of PS-waveforms vibrating in the orthogonal symmetry planes with almost the same shape and character. The horizontal geophones thus show approximately the same PS-wave records but with a relative time shift between them. Between the salient points coinciding with the planes of symmetry and those bisecting them, the seismic record varies smoothly and continuously. It is this variation between generation of only the fast split PS-wave (e.g., shooting parallel to cracks) and only the slow split PS-wave (e.g., shooting perpendicular to cracks) that gives rise to the traveltine undulations visible for deeper reflections in the seismic data; e.g., at 2.5 s in Figure 4a.

As noted above, Figure 3, the velocity difference between split PS-waves depends on azimuth for shots at finite offset. Consequently, the time delay between split PS-waves, $\Delta t_s$, is also azimuthally varying, Figure 4c. Most published methods for measuring split S-wave polarization azimuths and time delays neglect this azimuthal dependence of $\Delta t_s$. For the common assumption of azimuthally constant $\Delta t_s$ to be a good one, we see that it is important to limit the range of offset being used in the PS-wave splitting analysis.

Processing of the Svalbard data proceeds by independently rotating each source point in the shot circle to the 45° geometry, then time-shifting the horizontal-components through a range of relative lags (typically for 25 ms) and calculating a semblance trace at each lag. The semblance traces computed for each shot generate an azimuth slice from a semblance cube built in lag-azimuth-traveltime space (Figure 5) as the source point moves around the shot circle. This cube shows maximum semblance on a smooth and continuous surface through the volume, the lag-azimuth-traveltime coordinates of which are auto-picked to form the output lag surface. So far, we have worked with irregularly spaced prestack data. Once the lag surface has been formed from the semblance cube, the next step is to stack lag traces in azimuth bins (we have already selected a subset of offsets), reducing the data to a set of evenly spaced lag traces with common offset and traveltime equivalent to zero-offset PS-wave arrivals. This stacked lag surface, Figure 6a and 6b, is output as the basic form of data used to determine the depth-averaged time delays and polarization directions of the split PS-waves.

The lag surface shows broad corrugations in lag that oscillate with increasing azimuth. When the sagittal plane is parallel to a plane of elastic symmetry (e.g., vertical planes containing crack-normal and crack-parallel axes) the PS-wavelets on each geophone are essentially the same; they have the same phase and maximize semblance at zero applied lag. When the sagittal plane is midway between two planes of elastic symmetry, the PS-wavelets on each geophone are similar but with an overall time delay between them that corresponds to the amount of PS-wave splitting, which is the zero-to-peak amplitude of the lag surface corrugations. This measurement of time delay between wavelets is only weakly sensitive to amplitude differences caused by anisotropic attenuation or reflection coefficient as long as the bandwidth of each wavelet is at least...
roughly the same (see Gratacos et al. [2009] for the consequences of strong differential attenuation). Thus, the lag-surface corrugations have zero-to-peak amplitudes sensitive to the depth-averaged time delay between split PS-waves, and a phase that is sensitive to their depth-averaged polarization azimuths.

Anomalies of conversion-time difference are still present in the lag surface (as are errors in moveout, etc.) but they are expressed as traveltime crinkles in the lag surface. These traveltime crinkles (varying in the traveltime-azimuth plane for a single PS arrival) are decoupled from the lag corrugations (varying in the lag-azimuth plane for a single PS arrival) that provide the measurement of split PS-wave properties. As long as the traveltime crinkles are not excessively large, their effects are considerably reduced by smoothing in the semblance calculation. Thus, by processing the seismic data in the manner described above, we have somewhat decoupled the PS-wave splitting from the effects of structural or velocity heterogeneity, P-wave velocity anisotropy, and from moveout errors, all of which generate conversion-time differences. Furthermore, provided it is not excessively strong, anisotropy of attenuation and coefficients of reflection do not strongly alter the lag at which semblance is maximized in the lag cube, and thus this method of measuring the PS-wave splitting has reduced sensitivity to these effects.

The split PS-wave properties output by the lag-surface method are for depth-averaged quantities; it now remains to determine how the split PS-wave time delays and polarization azimuths change with depth. The next section describes a simple but powerful algorithm that simulates the measured lag surface and thus allows a layered earth model of polarization directions and PS-wave time delays to be estimated by forward modeling and inversion.

**Forward modeling**

The shapes seen in the lag surfaces derived from field data, Figure 7a, show coherent and systematic patterns that increasingly

![Figure 5](image)

Figure 5. (a) Semblance cube in traveltime-azimuth-lag space. (b) Time slices illustrating the semblance peak smoothly varying through the volume.

![Figure 6](image)

Figure 6. (a) Mean stack of lag traces in 3° azimuth bins. (b) Median stack. (c) Residuals from an L2 inversion using independent Gaussian-distributed data errors. (d) Residuals from an L1 inversion using independent Laplacian-distributed data errors. (e) Measure of suitability of a Gaussian data-error model for the prestack traces (blue is most suitable), see text. (f) Interquartile range of the prestack traces. (g) Prestack data deviations from a Gaussian probability model. (h) Prestack data deviations from a Laplacian probability model.
deviate from sinusoidal behavior as the time delay between split PS-waves increases. For any forward model of the lag surface, it is important that the algorithm can quickly and robustly simulate these nonsinusoidal shapes as well as the overall amplitude and phase of the main corrugations. Speed of computation is a critical factor if the lag surface is eventually to be inverted, because very rapid forward modeling allows directed Monte-Carlo techniques (such as simulated annealing, Ingber, 1989) to be used with what may turn out to be a strongly nonlinear problem (depending on the nonrandom noise present in the lag surface). This section outlines a very simple algorithm that, although crude, captures all the main features of the observed lag surfaces. Critically, this forward model is able to generate the range of sinusoidal and nonsinusoidal shapes derived from field data (Figure 7 column b: note these synthetics are not optimized to fit the field-data) and does so with very little computation time.

The forward model is solely kinematic, in that it perturbs a pre-existing synthetic seismogram to introduce traveltime and polarization anomalies that mimic the presence of split PS-waves. This is done with a simple geometric approach that assumes incident plane waves can be propagated upward from the conversion point to the receiver by projecting each wave onto orthogonal vectors parallel and perpendicular to the symmetry axis of the anisotropic medium and then introducing a time difference between the projections. This is repeated for each layer in the earth model, with a final projection onto geophone axes at 45° to the sagittal plane. The perturbed synthetic seismogram is put through the same semblance optimization process as described above for field data.

The synthetic seismogram being perturbed by the forward model has isotropic PS-wave events placed at the traveltimes of the main arrivals visible in the field data. In this way, the number and spacing of layers is controlled by the number and spacing of key events present in the input seismic data. It is only the kinematic perturbations due to the split PS-wave polarization azimuths ($\phi$) and time delays ($\Delta t_i$) in each layer that may be changed to produce the synthetic lag surface. Thus, an inversion of the lag surface must deal with only $M = 2n$ model parameters for $n$ layers in the earth model $m$. Generally speaking, the value of $\Delta t_i$ is constrained by the overall amplitude of the lag-surface corrugations, whereas the value of $\phi$ is constrained by the phase and the nonsinusoidal shape of the lag-surface corrugations. The algorithm is implemented by computing a series of synthetic traces at azimuth intervals matching those in the field data, which must be appropriately spaced to capture variations in the lag surface shape.

Anisotropy of attenuation and reflection or transmission coefficients is neglected in the forward model on the basis that the semblance optimization process is not sensitive to these effects. For the same reason, the shape of the source wavelet has no significant effect on the lag surface and can be modeled with an analytic function such as a Ricker wavelet. Furthermore, the overall scaling of the source wavelet has no effect on the final lag surface and so the synthetic need not consider the amplitude reduction due to intrinsic attenuation of the background isotropic rock or to geometric spreading. We also neglect the nonorthogonality of PS-waves and the deviation of the downgoing P-wave vibration from the sagittal plane because these assumptions are implicit in the data processing and interpretation. The model is very basic, therefore, and assumes:

1. that the anisotropy is weak, such that (a) the downgoing P-wave is polarized in the sagittal plane, (b) the recorded S-waves are polarized in orthogonal directions, (c) anisotropy of attenuation is negligible, (d) anisotropy of conversion, transmission, or reflection coefficients is negligible; (2) that the upcoming S-waves propagate vertically, such that the time delay between split PS-waves does not depend on azimuth and there is no significant change in this property within the restricted offset range used (this requires short offsets in the analysis); (3) that there is no significant out-of-plane energy produced by subsurface heterogeneity or structure (the validity of this assumption is improved after PS-wave migration).

Given these assumptions, it is quite remarkable that the synthetics are able to generate the lag-surface shapes seen in field data (Figure 7). This adds weight to the idea that the lag-surface corrugations are an expression only of the PS-wave splitting present in the data, with complications due to reflector structure, P-wave anisotropy, etc., being represented by traveltime crinkles that are decoupled from the signal we are interested in (and not included in the forward model). One direct consequence of this is that the inversion is not sensitive to these physical properties, and they do not interfere with the estimation of split PS-wave time delays and polarization azimuths.

Speed of computation is a key factor when inverting a lag surface. One way to reduce the number of azimuths being computed in the forward model is to generate synthetic traces positioned such that a

Figure 7. (a) Lag-surface time slices from field data (Svalbard survey, OBS n639 and n634) showing typical shapes displayed by the lag corrugations as they oscillate with azimuth. (b) Comparable shapes synthesized by the forward model. The synthetics are not optimized to fit the field data in these examples.
Chebyshev polynomial (Press et al., 2002) can be used to approximate the lag at any azimuth. Tests show that the Chebyshev polynomial approach using only a few tens of azimuths produces a lag surface equivalent to that synthesized with 180 regularly spaced azimuths to a very high degree of accuracy (errors <0.1 ms). As the lag surface becomes more nonsinusoidal in shape, the number of terms in the Chebyshev polynomial needed to achieve a good approximation also increases. With the range of shapes observed in the field data (Figure 7) use of 70 or more coefficients seems warranted to approximate the synthetic lag surface with an accuracy better than 0.1 ms.

Inversion

In this section, the inverse procedure is described with particular attention to the data errors, which map to model errors via the model response and therefore define the precision of the result. The inverse problem is solved with fast simulated annealing (Ingber, 1989) using a Cauchy-distributed perturbation function,

\[ m' = m + \Delta \frac{T}{T_0} \tan(\pi[\eta - 0.5]), \]

(1)
to modify the model vector \( m \) (producing the perturbed model \( m' \)) using a vector of a priori limits \( \Delta \) and a random number \( \eta \) on \([0,1]\). The perturbations become smaller as the annealing temperature, \( T \), decreases from the initial value, \( T_0 \), using an exponential scheme, \( T_{i+1} = T_i e^{-\frac{E_i}{\epsilon}} \) for decay constant \( 0 < \epsilon < 1 \). The model perturbation leads to a change in the object function, \( \Delta E \), via the computed model response. The perturbation is accepted probabilistically when

\[ \eta < \exp(-\Delta E/T) \]

(2)

(for a newly drawn random number, \( \eta \)), which is the usual Metropolis algorithm. The annealing schedule is chosen carefully to ensure that at the initial temperature 65% or more of the model perturbations are accepted and that the rate of cooling is slow enough to allow movement into and out of local minima as the system cools with typical values of \( \epsilon = 0.9 \) and 2000 model perturbations drawn at the \( i \)th step. Upward of 40,000 model perturbations are typically tested in one inversion using this scheme for a vector of 20 model parameters.

The object function \( E \), minimized by reannealing, is

\[ E(m) = r(m)^tC^{-1}r(m), \]

(3)

(where superscript \( t \) indicates transpose), for data covariance matrix \( C_d \) and vector of residuals \( r \). Note that there is no model regularization used. The form of the residuals vector depends on the probability distribution of the data errors in the lag surface. Because the semblance traces that comprise the lag surface are stacked in azimuth bins prior to use as a data vector in the inversion, we may derive some information about the statistics of these poststack data from their prestack values. In Figure 6e, we see a measure of the degree to which the prestack lag-trace data are Gaussian-distributed. This measure compares the standard score to prestack data values at quartiles 1 and 3 (respectively \( Q_1 \) and \( Q_3 \)) to quantify the departure from a Gaussian distribution according to

\[ g = |Q_1 - \mu_G + 0.67\sigma_G| + |Q_3 - \mu_G - 0.67\sigma_G|, \]

(4)

for mean \( \mu_G \) and standard deviation \( \sigma_G \). The \( g \)-values in this plot indicate that much of the poststack data can be modeled with Gaussian-distributed errors (the blue colors). However, by the same test about 20% of the data are not Gaussian-distributed (the red colors: note that the large red patches at the latest times begin after the arrival of the water multiple at 2.8 s). The caveat to this is that \( \mu_G \) and \( \sigma_G \) are estimated values from a sample set with variable fold in the azimuth bins typically not exceeding 30. The conclusion is that most of the data may be inverted using a Gaussian data-error probability distribution, but that parts of the data set do not have Gaussian-distributed errors and thus a more robust formulation might be needed.

The inversion now has two possibilities. The first is to use an L2-norm with an associated model of Gaussian-distributed data errors. This requires the vector of \( N \) data residuals to take the usual form

\[ r(m) = d - d(m) = [r_1, \ldots, r_N]^t, \]

(5)
such that the minimum of \( E(m) \) is a least-squares maximum likelihood solution. In this equation, \( d \) is a vector of \( N \) measured data (the lag surface) whereas \( d(m) \) is the vector model response. Alternatively, we can also use an L1-norm with an associated model of Laplacian-distributed data errors. The Laplacian distribution is a double exponential with expected value \( \mu_L \) given by the median of the \( N' \) data in each azimuth bin, and with variance \( 2\sigma_L^2 \) for deviation

\[ \sigma_L = \frac{1}{N'} \sum_{i=1}^{N'} |d_i - \mu_L|. \]

(6)
The univariate Laplacian distribution has a probability density function

\[ P(d) = \frac{1}{2\sigma_L} \exp\left(-\frac{|d - \mu_L|}{\sigma_L}\right), \]

(7)

which allows more robust inversion in the presence of data outliers than the equivalent Gaussian distribution.

Data errors represented by prestack deviations under assumptions of Gaussian and Laplacian distributions (Figure 6g and 6h), show correlations across regions of the data set. Consequently, it seems unsatisfactory to use a cost function derived from the sum of exponents in the univariate Laplacian distribution given above, because this implies independent data errors. To represent correlations in the data errors while satisfying the usual L1-norm in the limit that the data errors become uncorrelated, the object function used here takes a modified residuals vector

\[ r'(m) = \left[ \frac{r_1}{|r_1|^{1/2}}, \ldots, \frac{r_N}{|r_N|^{1/2}} \right]^t, \]

(8)

which has elements normalized by the square-root of their absolute value. In a similar manner, the data errors used to derive the data covariance matrix are normalized by the square-root of their magnitudes. In this case, the object function given in equation 3 now yields the usual L1-norm for uncorrelated data errors (equivalent to a sum of exponents in equation 7), yet also respects the sign
of the residuals in the event of correlated data errors. For the L2-norm inversion, the residuals and the elements of the data-error covariance matrix are not normalized.

The data used in the inversion are plotted in Figure 6, with a mean stack (a) and Gaussian standard deviations (g) used for the L2-norm inversion, and a median stack (b) with Laplacian standard deviations (h) used for the L1-norm inversion. Initially, the inversions are conducted with L1 and L2 norms using independent data errors expressed as prestack data variances on the main diagonal of \( \mathbf{C}_d \) (normalized as described above for the L1 case). The maximum likelihood models from these inversions are plotted in Figure 8a and 8b, with residuals in Figure 6c and 6d. The residuals are low in magnitude and mostly random, although with some correlation near the center of the lag surface (excluding data after the water multiple at approximately 2.7 s). The nature of these residuals indicate that the forward model has generally described the data well, justifying the deliberate omission of overprinting mechanisms such as anisotropic attenuation or layer dip from the forward model. The low magnitude of these correlated residuals shows some nonrandom effect manifest in the data that is not captured in the forward model. Given the low magnitude of these correlated residuals, and the simplicity of the forward model, however, the overall character of the residuals is quite impressive.

We now look at model uncertainties using an assumption of independent data errors, then turn to the issue of data-error correlations and how to model them before finally discussing the model precision level.

**Model uncertainties**

The maximum likelihood models from L1- and L2-norm inversions with independent data errors are surprisingly similar (Figure 8a and 8b). The similarity of these two results is a good sign of the repeatability of the directed Monte-Carlo inversion. To keep the proliferation of results under control, we now look at model uncertainties using only the L2-norm formulation.

The first approach to quantifying model uncertainty is to linearize the system around the maximum likelihood model obtained from simulated annealing, \( \mathbf{m} \), then to compute the model covariance matrix from

\[
\mathbf{C}_m = \left( \mathbf{J}(\langle \mathbf{m} \rangle)^\top \mathbf{C}_d^{-1} \mathbf{J}(\langle \mathbf{m} \rangle) + \mathbf{C}_m^{-1} \right)^{-1},
\]

where \( \mathbf{J}(\langle \mathbf{m} \rangle) \) is the Jacobian evaluated at the maximum-likelihood model, and the elements of \( \mathbf{C}_d \) are the Gaussian covariances (not normalized). The a priori model covariance matrix, \( \mathbf{C}_m \), stabilizes this calculation in the event that the Jacobian or the data covariances are singular. In this well-posed problem, neither of these quantities are singular and \( \mathbf{C}_m \) is an uninformative prior containing constants on the main diagonal only. Consequently, the calculation proceeds by dropping \( \mathbf{C}_m \) from equation 9. The results of this are shown in Figure 9, where for increasing depth of the earth layer the model parameters are numbered 1–10 for polarization azimuth and 11–20 for split PS time delay. Hence, the top-left quadrant of this plot has model variances for polarization azimuths on its main diagonal and similarly for split PS-wave time delays in the bottom-right quadrant. The largest values on the main diagonal indicate model deviations of about 2° of polarization azimuth and 0.3 ms of PS time delay.

The second approach to quantifying model uncertainty uses a Monte-Carlo method described as a Fast Gibbs Sampler by Molnar et al. (2010), following on from the discussions in Dosso (2002) and Dosso and Wilmut (2002), to which the reader is heartily referred. The Fast Gibbs Sampler is a Monte-Carlo Markov-Chain analysis that in practical usage is very closely related to simulated annealing, although its theoretical basis is less so. The method stems from an incorporation of the Metropolis Algorithm into Bayes’ rule, which expresses the conditional probability density function of a model given the data, \( P(\mathbf{m}|\mathbf{d}) \), in terms of the conditional probability density function of the data given the model \( P(\mathbf{d}|\mathbf{m}) \), the probability density function of the model \( P(\mathbf{m}) \), and the probability density function of the data \( P(\mathbf{d}) \), via

\[
P(\mathbf{m}|\mathbf{d}) \propto P(\mathbf{d}|\mathbf{m}) P(\mathbf{m}),
\]

\[
P(\mathbf{d}|\mathbf{m}) = \frac{1}{\mathcal{Z}(\mathbf{m})} \prod_{i=1}^{n} P(d_i|m_i),
\]

\[
P(\mathbf{m}) = \prod_{i=1}^{m} P(m_i|m_i),
\]

\[
P(\mathbf{d}) = \frac{1}{\mathcal{Z}(\mathbf{d})} \prod_{i=1}^{n} P(d_i),
\]

\[
\mathcal{Z}(\mathbf{m}) = \int P(\mathbf{m}) \prod_{i=1}^{n} P(d_i|m_i) \, d\mathbf{m},
\]

\[
\mathcal{Z}(\mathbf{d}) = \int P(\mathbf{d}) \prod_{i=1}^{n} P(d_i|m_i) \, d\mathbf{d},
\]

\[
\text{subject to } \sum_{\mathbf{m}} P(\mathbf{m}|\mathbf{d}) = 1.
\]

**Figure 8.** (a) Polarization azimuth of fast split PS-wave in maximum likelihood models from independent data errors with the L1 formulation (circles) and the L2 formulation (squares). (b) Similarly for the time delay between split PS-waves. (c and d) Maximum likelihood models inverted using data-error covariances from Toeplitz data residuals (blue), Toeplitz data deviations (red), and unstructured data deviations (black).
(Ulych et al., 2001; Tarantola, 2005). The model probability density, \( P(m) \), is an a priori function, which in this case remains fixed at some uninformative constant over the allowed model space and zero elsewhere. The probability density of the data is also constant, because the data measurements are unchanging. The conditional probability density function of the data given the model can be expressed as a likelihood function that depends on the data misfit, and hence on the object function defined above (equation 3).

The marginal probability density functions for the model parameters are given by

\[
P(m_i|d) = \int_{\Omega} \delta(m_i' - m_i) P(m|d) dm_i',
\]

(Aster et al., 2005; Dekking et al., 2005) where \( \Omega \) denotes integration over the entire model space, subscript \( i \) or \( j \) indicates a component of the model vector \((i, j = 1, \ldots, M)\) and in this case \( j \neq i \). The Dirac delta function, \( \delta(m_i' - m_i) = 1 \) if \( m_i' = m_i \) and 0 otherwise, causes the integration to be evaluated at point \( m_i \) of the \( i \)th dimension of the allowed model space. By integrating with respect to \( dm_i' \), equation 11 represents an integration with respect to all dimensions of the model space except for the \( i \)th. Gaussian and Laplacian probability models have likelihood functions \( P(d|m) \propto \exp[-E(m)] \).

Substitution, with equation 10, into equation 11 thus gives

\[
P(m_i|d) \approx \beta \sum_{u=1}^{Q} \delta(m_i' - m_i) \alpha \frac{\exp(-E(m_u'))}{s(m_u')} \sum_{v=1}^{\infty} \exp(-E(m_v')) P(m_v').
\]

(12)

for model ensembles \( u = 1, \ldots, Q \) and \( v = 1, \ldots, \infty \). Here, \( P(m) \) is the prior probability density of the model, which we can deem to be an uninformative constant. Also, \( \beta \) is a constant which has absorbed \( P(d) \). The denominator contains the probability density of a sampling function \( s(m) \), which is used to draw the set of \( Q \) models needed for summation (e.g., a uniform distribution, or, better, the Gibbs distribution in equation 13). Note that \( Q \) must be large for the summations to approximate integrations accurately. Ensemble \( v \) is drawn using a uniform sampling function, with constant probability density \( \alpha \). Unfortunately, it is not practical to evaluate the denominator of equation 12 prior to the calculation because ensemble \( v \) spans the entire model space. Interestingly, however, Dosso (2002) (citing Gilks et al. [1996] and Mackay [2003]) concludes that the Metropolis algorithm, when used to test models at a constant temperature, draws samples from the Gibbs distribution without having to know the denominator. Hence, for a Gibbs sampling distribution with probability density function

\[
s(m^u) = \frac{\exp[-E(m^u)/T]}{\sum_{v=1}^{\infty} \exp(-E(m^v)/T)},
\]

and with temperature held fixed at \( T = 1 \), we end up with

\[
P(m_i|d) \approx \beta \sum_{u=1}^{Q} \delta(m_i'^u - m_i).
\]

As long as the number of tested models \( Q \) is large (here, this means many hundreds of thousands of models), then the marginal probability density functions are proportional to histograms of the \( Q \) models drawn with a Metropolis algorithm.

The width of the marginal probability density functions described above provide an indication of uncertainty in the model parameters. These functions are plotted in Figure 10 for the L2-norm with independent data errors and evaluated with \( Q = 10^6 \). This calculation aims to sample the model space with no a priori bounds. To increase sampling efficiency, however, it is necessary to start the sampling process with bounds centered on and near the maximum likelihood solution, then to grow these bounds as the region of accepted models widens. The sampling bounds are grown to maintain a width that is a factor of two larger than the width of the accepted sample set. Convergence is monitored by periodically computing the mean model. From the results of this calculation (Figure 10) we immediately see that the probability density function is broad and flat for the polarization azimuth in layer 1 (which is centered on the direct arrival), indicating that all azimuths have similar low likelihood. This is because the PS-wave time delay is very tightly grouped around zero, which is what we expect from the direct wave. It is interesting to see that the PS time delays are actually very tightly defined, in terms of their marginal probability density functions, to one or two histogram bin widths for all the layers, with the exception of layer 10 which lies in the noisy part of the lag surface in the region of the water multiple. These histogram bins are 0.1 ms wide.

---

Figure 9. Model covariance matrix estimated by linearizing the L2 maximum likelihood model obtained using independent data errors. The model parameters are numbered 1–10 for polarization azimuths, and 11–20 for split PS time delays (increasing number corresponds with increasing depth in the earth model).
which implies an astonishing level of precision in the model parameters estimated by inversion. Similarly, the bin widths for polarization azimuths are of 0.5°, and again the Fast Gibbs Sampler indicates extremely high levels of precision in the inversion results, with model uncertainties generally <1° of azimuth and <0.1 ms of time delay. These values seem too good to be true, and probably are so when one remembers that these statistics have been derived using an assumption of independent data errors. Under this assumption, the implied model precision levels are a consequence of the massive data redundancy present in the lag surface, which contains more than 20,000 data points for only 20 model parameters. This means that even a tiny movement of the model response from its maximum likelihood position results in a small increase in the residuals and a change in probability given by the product of more than 20,000 negative exponentials. Hence, the likelihood of the model perturbation given the data becomes tiny. The way out of this problem is to allow data-error correlations, which have the net effect of increasing the uncertainty in the model parameters. We now turn to the issue of data-error covariance and how to quantify what it should be.

Data-error covariance

The effect of correlated data-errors on the inversion are explored in three ways. The first approach is to postulate a correlation based on an auto-regressive model, and to calculate a Toeplitz data-error covariance matrix from the data deviations according to

$$\mathbf{C}_{d_{i,j}} = \frac{1}{N-M} \sum_{k=1}^{N-|j-i|} (\sigma_i' - \langle \sigma' \rangle)(\sigma_{i+j}' - \langle \sigma' \rangle),$$

for $i = 1, \ldots, N$ and $j \geq i$, and where

$$\mathbf{C}_{d_{i,j}} = \mathbf{C}_{d_{j,i}},$$

for $j < i$ (after Molnar et al., 2010). Here, we allow $\sigma' = \sigma_l/|\sigma_l|^{1/2}$ for a Laplacian model and $\sigma' = \sigma_G$ for a Gaussian model, as previously. The second approach is to do the same using the residuals from the maximum likelihood model, which also provides some measure of error arising from the failure of the forward model to capture the nonrandom variations in the data. In this case, the above equation is modified to take the average residual $\langle r \rangle$, and a residual $r^0$ that directly substitutes for $\langle \sigma' \rangle$ and $\sigma'$ with normalization as necessary. The final method for estimating the data-error covariance is to use the prestack lag traces to compute

$$\mathbf{C}_{d_{i,j}} = E\left\{ \left[ \frac{d_i - E\{d_i\}}{|d_i - E\{d_i\}|^{1/2}} \right] \left[ \frac{d_j - E\{d_j\}}{|d_j - E\{d_j\}|^{1/2}} \right] \right\},$$

where in this case $E\{\}$ indicates the expected-value operator and once again the normalization (the denominators) is used only for the Laplacian case. We already have $E\{d_i\}$ from the stack. Hence, the above is the median (Laplacian) or mean (Gaussian) of products of prestack lag-trace deviations in bin $i$ with those in bin $j$. This data-error covariance is referred to here as an unstructured covariance, in contrast to the first two covariance matrices which have Toeplitz structures. The data-error covariance matrices output by these three methods are plotted for a Gaussian model (i.e., without normalization) in Figure 11.
straightforwardly invertible, however, due to the presence of near-zero values on the main diagonal that arise from prestack data around the direct wave. This is where the seismic data contain a strong PP-wave, and all lag traces find the same maximum semblance at zero lag, thus having zero deviation and preventing the data-error covariance matrix from being full rank. To proceed with the unstructured covariance matrix, a pseudoinverse is produced, where singular values with running condition number larger than $2 \times 10^{12}$ have their reciprocals zeroed out. The running condition number is defined here for the $i$th data point as the ratio of largest singular value to the $i$th singular value.

**FINAL RESULTS**

The inversions are repeated by reannealing the problem using the two Toeplitz data-error covariance matrices and the unstructured matrix defined above. We proceed using the L2 approach, the maximum likelihood models from which are plotted in Figure 8c and 8d. It is immediately apparent that between layers 4 and 9, data-error correlations do not dramatically alter the maximum likelihood models from those obtained under the independent data-error assumption. Layer 4 is where the cumulative time delay between split PS-waves starts to exceed 0.5 ms, and layer 9 is where the water multiple arrives (which has not been addressed in processing). Between layers 4 and 9, the maximum likelihood values obtained under the different assumptions of data-error covariance have a spread of about 0.5 ms or less in PS time delay, and about 20° of polarization azimuth. Notably, despite absorbing a great deal of effort, the final inversion conducted using the pseudoinverse of unstructured data-error covariances provides almost the same solution as that obtained with independent data errors, particularly in the parts of the model deeper than layer 4.

The size of the matrix multiplications needed to evaluate the object function using correlated data errors prevents the use of a Fast Gibbs sampler to determine model uncertainties, and so the linearized approach is used instead. Using the pseudoinverse of the unstructured covariance matrix built from Gaussian-distributed data errors, the linearized model uncertainties are about 0.3 ms of PS time delay and 3° of polarization azimuth.

**DISCUSSION**

Interestingly, the inversions did not struggle with the number of layers in the earth model, which in this example was ten layers spanning about 500 m of depth from the seabed downward. The layers were defined by the arrival times of strong and coherent PS arrivals present in the seismic data. One might expect an overparameterized model (i.e., too many layers) to produce oscillations in the maximum likelihood models, which are not observed to any great degree in Figure 8. This is a fortunate aspect of the processing and inversion strategy outlined in this paper because it means a high level of depth resolution can be achieved without having to use a model regularization to stabilize the inversion. Such regularizations (e.g., model flatness or smoothness) reduce the sensitivity of the inversion to anomalous layers which might be the specific target of interest.

As a validation of the processing and inversion method, all ten layers of the L2-norm result with independent data errors are used to layer-strip the input data using a method similar to that described in Bale et al. (2009). The input data (Figure 12b) show clear traveltime undulations in the radial component and an azimuthal pattern of four evenly spaced amplitude highs and lows in the transverse. After layer-stripping (Figure 12c), these characteristic patterns in the input data are no longer present, confirming that the model of split PS time delays and polarization azimuths describes the data adequately. This layer-stripping result does not, however, demonstrate the extent to which simpler earth models may produce an equivalent, or better, reduction of the data.

The ability of the processing and inversion scheme outlined above to constrain split PS-wave properties over a large number of depth layers with relatively high levels of precision has some important implications for the exploration of earth’s geologic
systems. Perhaps most important is the potential use of this method in the exploration of tight-gas and shale-gas plays, for which the presence and density of fractures (and their degree of alignment) are important factors in their exploration and development (e.g., Gale et al., 2007). The polarization azimuths and time-delays of split S-waves are thought to be considerably more sensitive to the properties of rock fractures than are other forms of seismic data (Crampin, 1999). Clearly, if this is the case, then it is important that these seismic characteristics are measured with as much depth resolution as possible.

Also of significance for this method is its reduced sensitivity to layer dip and velocity heterogeneity, which affect the traveltime to the point of mode conversion. The processing scheme outlined in this paper appears to suppress much of the competing signal arising from these mechanisms in the data being inverted, leaving only the split PS time delay and polarization azimuth to strongly influence the estimated earth model. Although this is a potential strength of this method, it should be noted that the dip and heterogeneity in the west Svalbard study area are not large (the seabed dips by 2° to the southwest), and despite using data that have not been migrated, the ability of the method to cope with dip and heterogeneity has not been properly tested. Nonetheless, as a first example, the results presented here are encouraging. To reiterate an argument from the introduction, the problems of layer dip and heterogeneity can always be addressed by more accurate PS-wave migration prior to analysis. Conversely, the ability to measure split PS-wave properties robustly in the presence of migration inaccuracies is very much a desirable trait, and is an area in which the method outlined here may prove to be strong.

**CONCLUSIONS**

The method of measuring split PS-wave time delays and polarization azimuths described in this paper falls into two parts. First is the processing of data in a sensible domain, such as the receiver domain for investigation of shallow layers, or the postmigration image-point domain for deeper targets. The processing gives a semblance cube in lag-azimuth-traveltime space, which considerably simplifies the PS data while preserving their most important characteristics for the measurement of split PS-wave properties. The lag cube provides a robust means of extracting depth-averaged split PS-wave properties via a lag surface. The processing scheme appears to work well with noise-bearing field data, and is able to separate the important split PS-wave properties from competing signals caused by heterogeneity of structure or velocity as well as P-wave azimuthal anisotropy. The latter, if not excessively large, produce traveltime crinkles in the lag surface that are decoupled from the lag corrugations representing the properties of the PS-wave splitting.

The second part of this method is an inversion of the lag-surface to estimate the properties of split PS-waves with depth. Experience with the example field data indicates that the earth model can be remarkably simple, containing only traveltimes of key PS-wave arrivals and the polarization azimuths and time-delays for each overlying layer. Because the earth model is simple (although the earth itself may not be) and the lag-surface data traces are well sampled in time and in azimuth, we arrive at a massively overdetermined problem, which leads to very high precision levels in the maximum likelihood models. Using independent data errors estimated from prestack lag traces, model-parameter uncertainty levels from a linearization around the maximum-likelihood model are typically 0.3 ms of time delay and 2° of polarization azimuth. A more exhaustive Monte-Carlo method applied to the same model and data errors indicates these linearized model uncertainties are too high by

Figure 12. (a) Maximum-likelihood model inverted with an L2 norm and independent data errors given by prestack deviations. Also present are an indication of model parameter uncertainties for the worst-case scenario of unknown data-error covariances. The model uncertainties here are the spread of results obtained with all data-error models, or the values on the main diagonal of the linearized model covariances computed with independent data errors, whichever is the larger. (b) Radial and transverse data in a shot circle formed by NMO correction and stacking in offset-azimuth bins of 250–1000 m and 3°. Data are shown before (b) and after (c) layer stripping conducted with the polarization azimuths and time delays plotted in panel (a).
at least a factor of two, giving its own model uncertainties of less than 0.1 ms and 1°. The model uncertainties rise when data-error correlations are introduced, however. Using data-error covariances estimated from prestack data, model uncertainties increase to about 0.3 ms and 3° when estimated by the linearization approach (the Monte-Carlo method is not practical with the large matrix multiplications required by correlated data-errors). This seems the most appropriate precision level to cite in this case, where the fold characteristics of the field data allow the data-error covariances to be estimated from the prestack data. Finally, however, by repeating the inversions using postulated forms of data-error covariance (for example, Toeplitz forms of auto-regressive type) it is clear that uncertainty in the form of the data-error covariances leads to a significant increase in model-parameter uncertainty. Therefore, if the fold of the field data is too low to allow the data-error covariances to be estimated from prestack data, then it is reasonable to report precision levels implied by the spread of maximum-likelihood models found with different data-error covariances, which in this case would be up to 0.5 ms of time delay and 20° of polarization azimuth.

In the above discussion, the focus has been on precision levels with no discussion of accuracy (the degree to which the earth model represents the true earth properties rather than just describing the data). Clearly, it is not possible to measure the accuracy of the earth models estimated from the west Svalbard seismic data because the true properties of earth in that region are not known. To understand the limitations of accuracy when processing and inverting seismic data as described in this paper, it is necessary to study synthetic data for which the real earth properties are known a priori. In this way, the simplifications of the forward model and also the effect of the processing scheme applied to the seismic data could be assessed quantitatively. Unfortunately, though, the synthetic earth may take a wide range of possible states, exhibiting all the forms of noise that the processing scheme seeks to suppress and the forward model seeks to avoid (e.g., layer dip, heterogeneity, P-wave anisotropy, and so on). Consequently, inverting a single synthetic data set tells us nothing in general about accuracy levels because the applicability of the forward model varies as the earth varies. It is beyond the scope of the present paper to invert synthetic data sets spanning the range of possible earth systems. Instead, we can qualify the methods reported here by pointing out that the processing and modeling do not incorporate the effects of steep dips and strong lateral heterogeneity, so study areas that exhibit these characteristics only weakly should correspond with higher levels of model accuracy than those that exhibit them strongly.

ACKNOWLEDGMENTS

This work began at the University of Birmingham, continued at the University of Victoria and the Geological Survey of Canada, at Royal Holloway University of London, and finally at CGGVeritas. I thank management and colleagues at all the above institutions for their support. I particularly thank Graham Westbrook, Russell Exley, Stan Dosso, and Sergio Grion for their discussions and encouragement. I thank also Richard Bale, Leon Thomsen, and Clement Kostov for their constructive reviews. The inversion aspect of this work was funded by the U. K. Engineering and Physical Sciences Research Council (EPSRC), grant EP/H018573/1. None of the work reported here would have been possible without Seismic Unix and Generic Mapping Tools, GMT.

REFERENCES


Gratacos, B., 2006, A robust algorithm and associated QC for finding the anisotropy directions for converted wave data: 68th Annual International Conference and Exhibition, EAGE, Extended Abstracts, 68, P188.


Measurement of PS-wave splitting


Downloaded 02/25/13 to 212.118.224.156. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/