AVO and velocity analysis

D. Sarkar*, J. P. Castagna‡, and W. J. Lamb‡

ABSTRACT

Velocity analysis using semblance estimates velocities based on a constant amplitude model for seismograms and does not take amplitude variation with offset (AVO) into account. In the presence of AVO, the constant amplitude model becomes inaccurate, particularly for events which exhibit polarity reversals.

An AVO sensitive velocity analysis procedure, which is a generalization of the traditional semblance method, can be devised by giving an offset dependence to the modeled seismograms. Incorporating AVO into velocity analysis requires additional parameters to describe the reflectivity. This results in reduced velocity precision. By introducing a regularization term which provides a controlled suppression of the contributions due to AVO effects, we describe an AVO sensitive velocity analysis algorithm that properly deals with events exhibiting polarity reversals or large amplitude variation with offset.

INTRODUCTION

Semblance-based coherency measures are commonly applied to perform velocity analysis (Taner and Koehler, 1969; Neidell and Taner, 1971) on seismic reflection data. Velocities are estimated by maximizing a coherence functional or, equivalently, by minimizing the difference between the data and an a priori amplitude distribution while varying the velocity. Thus, assuming a wrong a priori amplitude-versus-offset (AVO) function results in an incorrect velocity estimation. As a result, amplitude variation with offset and velocity analysis are intimately related. Just as the wrong velocity gives incorrect AVO parameters, the converse (estimating the wrong velocity by assuming wrong AVO parameters) is also true, although the latter exhibits itself in a more subtle way. This problem is particularly crucial for polarity reversals [such as may occur at the tops of some class 1 and class 2 sands (Rutherford and Williams, 1989)]. Taking AVO effects into account should improve normal moveout (NMO) corrections and determination of prestack attributes.

Semblance is today, as it has been for many years, the most widely used method of velocity analysis. It is robust and computationally efficient. Software is well developed and is readily available. However, because it implicitly assumes a constant amplitude model, it does not handle AVO properly.

More recently, other methods have been developed. The differential semblance method (e.g., Symes and Kern, 1994) works by comparing each trace with traces of similar offset (see Appendix A). It eliminates secondary maxima and produces a broad primary maximum. Because it compares only traces of similar offset, this method is less sensitive to AVO variations than traditional semblance analysis.

Eigenvalue methods (e.g., Biondi and Kostov, 1989; Key and Smithson, 1990) use the fact that the signal covariance matrix is of low rank in the absence of noise. These methods promise high resolution but in general are computationally more intensive than semblance. They easily incorporate AVO, even for complex variations (Appendix B). In principle, crossing and interfering events can be handled (Biondi and Kostov, 1989), but great simplification in computation results from assuming a single event (Key and Smithson, 1990). Less is known about the coherent noise immunity of these methods compared with the vast experience of the geophysical community with semblance. Kirlin (1992) proposed a method intermediate between eigenvalue methods and semblance. It has the same limitations as semblance with regard to AVO.

In this paper, we focus on generalizing the semblance method so that it incorporates AVO. To develop a procedure similar to the traditional semblance method and simultaneously account for amplitude changes with offset, we follow Corcoran and Seriff (1989) and define an objective function, \( \Theta(v_s, t_0) \), for a given stacking velocity \( v_s \) and zero offset time \( t_0 \), as...
\[ \Theta(r_v, v_s | t_o) = 1 - \frac{1}{\sum_{t=t_o-W\Delta T}^{t=t_o+W\Delta T} \sum_{i=1}^{N} (A(v_s | t) + B(v_s | t) \sin^2 \theta_i + C(v_s | t) \sin^2 \theta_i \tan^2 \theta_i - F_D(v_s | x_i, t))^2}{\sum_{t=t_o-W\Delta T}^{t=t_o+W\Delta T} \sum_{i=1}^{N} (F_D(v_s | x_i, t))^2}. \] (3)

was calculated over a time window of \((2W + 1)\) samples centered at a particular zero offset time \((t_o)\). This window should be such that its length is inversely proportional to the frequency content of the data and approximately equal to the length of the wavelet. \(\Delta T\) is the sample rate, \(r_v, v_s\) are the parameters that describe the amplitude variation at a stacking velocity \(v_s\). For example, in the Shuey (1985) approximation to the Zoeppritz equations, \(r_v\) represents the AVO intercept and gradient terms. It can be shown that when \(F_D\) is a least squares fit to the data, \(\Theta\) lies between 0 and 1. We refer to this function as the generalized semblance function.

Corcoran and Serif (1989) suggested giving an offset dependence to the amplitude in the forward model to account for AVO. The offset dependence is incorporated in the \(F_D\) vector. Traditional semblance can be shown to be a particular case of this generalized procedure. However, we show here that while this process is accurate, it has less velocity precision than traditional semblance when there is no amplitude variation with offset. The results can be improved by treating the problem as a mixed determined problem (Menke, 1984). We show that this provides a general procedure to move from traditional semblance to an AVO sensitive semblance by varying a regularization parameter. This parameter determines the balance between robustness and accuracy of the algorithm.

In the subsequent paragraphs, we outline the amplitude dependent semblance in detail and show some applications to synthetic and field data.

### AVO SENSITIVE SEMBLANCE

Based on Corcoran and Serif (1989), we describe an amplitude-dependent semblance procedure that is accurate in the presence of AVO. Using Shuey’s (1985) approximation for the reflection coefficient, the forward modeling operator at a particular zero offset time \(t\) for data moved out at velocity \(v_s\) can be written as

\[ F_D(v_s | x_i, t) \approx A(v_s | t) + B(v_s | t) \sin^2 \theta_i + C(v_s | t) \sin^2 \theta_i \tan^2 \theta_i, \] (2)

where \(\theta_i = \) average angle of incidence and refraction at the \(i\)th offset. It is a function of offset and depth. \(A\) is the AVO intercept, \(B\) is the AVO gradient, and \(C\) is the AVO curvature. The objective function can be written as

\[ \Theta(A, v_s | t_o) = \frac{1}{N} \sum_{t=t_o-W\Delta T}^{t=t_o+W\Delta T} \sum_{i=1}^{N} \left( \frac{F_D(v_s | x_i, t)}{F_D(v_s | x_i, t)} \right)^2, \] (6)
which is the familiar ratio of the average stacked energy to the total energy of its components—also called semblance. Traditional semblance ($A$ semblance) implicitly assumes a model with no change in amplitude with offset and where the constant amplitude $A$ is the normalized stacked trace for a particular trial velocity. This is accurate when AVO is negligible. As AVO variations increase, the value of the traditional semblance objective function gradually decreases. At the extreme case of a polarity reversal at the center of the offset range, the average stack energy tends to zero without affecting the total energy of its components. This gives rise to the misleading conclusion of a fitness value tending to zero at the correct velocity.

In an attempt to improve on this, the slope (gradient) term, $B$, and the curvature term, $C$, are introduced in the reflectivity functional. In doing so we obtain better fits in the presence of AVO and especially in the presence of polarity reversals. When $A$ (intercept) and $B$ (gradient) terms are used, we call the method “$AB$ semblance,” and when $A$, $B$, and $C$ (curvature) terms are used, we call it “$ABC$ semblance.” For our purpose the problem is simplified by (1) assuming hyperbolic moveouts and (2) using a straight ray approximation instead of the correct incidence angles. However, these simplifications can be relaxed for more complex problems.

INITIAL SYNTHETIC EXAMPLES

The first data set studied was a synthetic three-layer model as shown in Figure 1 and Table 1. The first event has a polarity reversal near the center of the offset range, whereas the second and third events have negligible AVO but opposite polarities. At the first interface, a decrease in Poisson’s ratio causes a negative $B$, whereas the $P$-wave velocity and density increase causes a positive $A$. At the second interface, an increase in Poisson’s ratio gives a positive $B$, whereas an increase in $P$-wave velocity and density gives a positive $A$. At the third interface, a decrease in Poisson’s ratio, $P$-wave velocity, and density gives a negative $B$ and a negative $A$. The fitness curves obtained for the different methods when applied to events 1 and 3 of Figure 1 are shown in Figures 2 and 3, respectively.

Of special importance is the failure of the traditional semblance ($A$ semblance) method to determine the correct velocity in the presence of a polarity reversal (Figure 2). This shows that using the wrong functional to describe the amplitude variation can cause poor velocity determination. The fitness function curves for the third event (Figure 3) show that, in the absence of strong AVO variations, traditional semblance ($A$ semblance) is more robust than $AB$ or $ABC$ semblance. Figures 2 and 3 show that there are basically two effects that couple together to determine the fitness function curve. The first effect is time dependent. It is the well-known effect of losing precision with depth because of the flattening of the moveout curve. The second effect is amplitude variation. When fewer parameters are used than necessary, the fitness function optimizes at a wrong velocity causing poor velocity estimation (Figure 2). Also, if more parameters than necessary are used, then precision is lost as indicated by a flattening of the fitness curves (Figure 3).

<table>
<thead>
<tr>
<th>Layer</th>
<th>$V_p$ ($P$-velocity) (m/s)</th>
<th>$V_s$ ($S$-velocity) (m/s)</th>
<th>$\rho$ (density) (g/cm$^3$)</th>
<th>$\sigma$ (Poisson’s ratio)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Layer 1</td>
<td>2500</td>
<td>982.759</td>
<td>2.192</td>
<td>0.4086</td>
</tr>
<tr>
<td>Layer 2</td>
<td>2780</td>
<td>1853.334</td>
<td>2.751</td>
<td>0.09999</td>
</tr>
<tr>
<td>Layer 3</td>
<td>3300</td>
<td>1672.414</td>
<td>2.349</td>
<td>0.3272</td>
</tr>
<tr>
<td>Layer 4</td>
<td>2780</td>
<td>1853.334</td>
<td>2.251</td>
<td>0.09999</td>
</tr>
</tbody>
</table>

Fig. 1. Synthetic gather consisting of three events. From top to bottom: event with a polarity reversal, and events with almost no AVO.

Fig. 2. Fitness plot for event 1 with polarity reversal, calculated at the center of the event. Curves are semblance ($A$ semblance), $AB$ and $ABC$ semblance, simple differential semblance (Appendix A), and modified Key and Smithson (1990) coherence (Appendix B).
In Figures 2 and 3 we have also included fitness curves for simple differential semblance (Appendix A) and a modified Key and Smithson (1990) coherence method (Appendix B). The essential point is that both have maxima at the correct velocity, regardless of the presence or absence of AVO. The differential semblance curve is much broader and the eigenvalue method much sharper than the semblance family of curves. Care should be used in relating the sharpness of these curves to velocity resolution. Because the normalization of the eigenvalue and semblance methods are different, visual comparisons can be misleading (Kirlin, 1992).

The increased sensitivity to AVO is illustrated with semblance contour plots (Figure 5) created from the synthetic data set shown in Figure 4. The plots show the reduction in velocity precision as the number of model parameters increases from one to three. This is especially so for events with negligible AVO, such as the events after 2 s.

To understand the effect of the loss in resolution on AVO analysis, crossplots of extracted As and Bs from time windows about the numbered events were made from the data set shown in Figure 1 (see Figures 6–9). The A and B values were automatically picked. Every pair corresponds to the velocity giving the highest fitness value at a zero offset time, sampled at the sample rate of the seismogram. Because of the presence of a wavelet instead of a spike, the A-B crossplot scatters in a straight line through the center for every event. The more aligned the points are for an event, the better is the reflectivity estimation. The A-B values determined analytically from Shuey’s (1985) equation are numbered and labeled on the crossplots for every event (Figures 6–9). Figure 6 shows how wrongly determined velocities using traditional semblance (A semblance) cause poorly determined A and B values for the event 1. Noteworthy also is the increase in A-B crossplot spread for events 2 and 3, and not for event 1, in Figures 7 and 8 with increasing number of reflectivity parameters. The cause is overfitting. Events 2 and 3 have small amplitude variation with offset, thus using higher degree polynomials causes tradeoff in parameter estimation and, hence, lower velocity resolution. These events however appear to be well resolved when traditional semblance (A semblance) and simple differential semblance (see Figures 6 and 9, respectively) were used.

The above results show that both moveout and amplitude variations are important in pinpointing the right velocity and both must be taken into account while doing velocity analysis.

The A-B crossplot (Figure 9) obtained from the simple differential semblance may be viewed as the control data set. A good A-B crossplot is obtained using differential semblance, in spite of a flat objective functional, because of its insensitivity to AVO and lack of parameters to overfit.

In the absence of AVO, we can conclude that traditional semblance has greater precision than AB semblance because any nonzero estimate of B causes error. However, in the presence of polarity reversals, traditional semblance fails and AB semblance is more accurate.

**IMPROVING THE RESOLUTION WITH A REGULARIZATION TERM**

The loss in precision occurs due to overparameterization that gives rise to parameter trade-off errors. Such a problem arises when one or more of the eigenvalues of the covariance matrix are small (Menke, 1984). To improve on this, we suggest determining the AVO parameters \( A(t) \) and \( B(t) \) by minimizing the following objective function at a zero offset time \( t \):

\[
\hat{\Theta}(A(t), B(t), \ldots | v_s, t) = \frac{1}{N} \| F_m(r[v_s]|v_s, t) - F_D(v_s, t) \|^2 + \sum_{i=2}^{p} \lambda_i k_i^2(t),
\]

where \( \lambda_i \) are the regularization factors, \( k_i \) are the parameters \((A, B, C \ldots)\) describing the reflectivity functional, and \( p \) is the total number of parameters required. This equation is only used to determine the reflectivity parameters, and the parameters thus obtained are substituted in equation (3) (refer to Appendix C for more details). Velocity and zero offset time are kept fixed and are written to the right of the vertical bar.

Thus, when two parameters are used to describe the reflectivity,\( A(t), B(t) | v_s, t \),

\[
\hat{\Theta}(A(t), B(t) | v_s, t) = \frac{1}{N} \| F_m(r[v_s]|v_s, t) - F_D(v_s, t) \|^2 + \lambda B^2(t),
\]

\[
\hat{\Theta}(B(t) | v_s, t) = \frac{1}{N} \| F_m(r[v_s]|v_s, t) - F_D(v_s, t) \|^2 + \lambda B^2(t),
\]

\[
\hat{\Theta}(B(t), C(t) | v_s, t) = \frac{1}{N} \| F_m(r[v_s]|v_s, t) - F_D(v_s, t) \|^2 + \lambda B^2(t) + \lambda C^2(t),
\]

\[
\hat{\Theta}(A(t), B(t), C(t) | v_s, t) = \frac{1}{N} \| F_m(r[v_s]|v_s, t) - F_D(v_s, t) \|^2 + \lambda B^2(t) + \lambda C^2(t),
\]

\[
\hat{\Theta}(A(t), B(t), C(t) | v_s, t) = \frac{1}{N} \| F_m(r[v_s]|v_s, t) - F_D(v_s, t) \|^2 + \lambda B^2(t) + \lambda C^2(t).
\]
where $\lambda$ is a parameter which determines which AVO parameter ($A$ or $B$) is to be given more importance. When $\lambda$ is 0, the parameters estimated from the objective function [equation (8)] equals that of $AB$ semblance, and when $\lambda$ goes to $\infty$ (in the examples considered, a $\lambda = 10$ approximates $\infty$), the parameters estimated from the objective function [equation (8)] becomes that of traditional semblance.

To find the effect of $\lambda$ for the data sets used, we experimented with a number of values of $\lambda$ ranging from 0 to 10 on data sets shown in Figures 1, 4, 13, and 14. The most significant four plots are shown. We found that, for the data sets considered here, a value of $\lambda(0.006)$ consistently stabilizes the results to a great extent and achieves the resolution comparable to traditional semblance for deeper events where AVO is weak, while at the same time it works well for some of the shallower events with considerable amplitude variation. This can be seen in the velocity fitness plots shown in Figures 10 and 11. Also, semblance contour plots of the synthetic data (Figure 4) shown in Figure 12 illustrate the influence of $\lambda$ on velocity analysis. Similar plots computed on processed field common midpoint (CMP) gathers are shown in Figures 13 and 14. The data has amplitudes compensated for spherical divergence and attenuation, multiples removed, and a bandpass filter applied. The gathers were NMO corrected with velocities that maximized the semblance curves at times 1.4, 1.6, and 1.8 s. The event around 1.6 s, shown in Figure 14, is significantly overcorrected, whereas Figure 13 shows that appropriate values of $\lambda$ flattens the event correctly. This again shows the failure of traditional semblance to estimate accurate velocities for such events.

In Figure 15, we show a crossplot of intercept and gradient terms extracted from the field gather using velocities that maximized $AB$ semblance ($\lambda = 0$) at every time sample between 1.4 and 1.8 s. This method may be used to estimate reflectivity parameters along with velocities simultaneously.

When velocity analysis is performed on CMP gathers, a reasonable value of $\lambda$ must be determined first. The factors that influence the estimate of $\lambda$ are (1) the data, which includes the signal and noise content, (2) number of offsets, which determines the number of data points, and (3) the geometry of the problem (offset to depth ratios). Since the geometry of the problem can be assumed to remain reasonably constant for large areas, a single value of $\lambda$, or at the most a few $\lambda$s, should be sufficient for an entire data set.
DISCUSSION AND CONCLUSIONS

Our study shows the limitations that arise when amplitude variation is ignored in velocity analysis, and also shows the problems of using more complicated amplitude dependence than is necessary. It shows the classic tradeoff between accuracy and robustness (or bias versus variance). Traditional semblance, though not quite accurate, is robust. However, the AVO-parameterized semblance, though not robust, can be accurate.

AVO-parameterized semblance (1) can determine $A$ and $B$ values simultaneously with velocity, (2) provides a coherence measure that is directly comparable to traditional semblance, and (3) requires a regularization term for routine application.

Because AVO parameters are sensitive to velocity errors ($B$ was seen to be more sensitive to velocity than $A$), care must be taken in their interpretation. Though $AB$ semblance may be used on single events, it is not as robust as traditional semblance for reflections with only small AVO variations. Treating the problem as a mixed determined problem allows one to get the best of both traditional semblance ($A$ semblance) and AVO sensitive semblance. AVO sensitive semblance with a regularization term has the potential of computing better stacking velocities in certain specific, but important, cases, thus improving the accuracy of the velocity analysis algorithm.

There are a few effects ignored in this study. The first is the effect of tuning. Tuning affects correct determination of $A$-$B$ parameters and hence can affect velocity determination. This is more of a problem for $AB$ semblance than for $A$ semblance (traditional semblance). For such cases, the eigenvalue method is particularly attractive (Biondi and Kostov, 1989). The second effect ignored is the angle of incidence used. For this analysis,
the angle was computed from the arc tangent of the depth to offset ratio. In most cases, this underestimates the true angle of incidence. Although the angle of incidence does not greatly affect the determination of the zero offset reflectivity ($A$), it has a serious effect on the gradient ($B$) term. From our limited analysis, the effect does not seem to be significant on the cases studied, but it may prove relevant in certain instances.

Fig. 11. Fitness versus velocity plots for different values of $\lambda$ when applied on event 3 (almost no AVO) in the data set shown in Figure 1.

Fig. 12. Semblance contour plots for different values of $\lambda$ when applied on the synthetic data shown in Figure 4. (a) $\lambda = 0.0$, (b) $\lambda = 0.006$, (c) $\lambda = 0.05$, (d) $\lambda = 10$.

Fig. 13. Processed CMP field gather after NMO correction. (a) $\lambda = 0.0$, (b) $\lambda = 0.006$.

Fig. 14. Processed CMP field gather after NMO correction. (a) $\lambda = 0.05$, (b) $\lambda = 10$. 
Now that semblance has been generalized to include AVO effects, an objective of future research will be to compare the performance of the different velocity analysis methods (AVO sensitive semblance, differential semblance, and eigenvalue methods) in the presence of various kinds of coherent noises and AVO variations.

ACKNOWLEDGMENTS

Thanks to Antonio Ramos and Bob Baumel (Conoco) for helpful comments throughout this study. Thanks to John Anderson (Exxon Mobil), who suggested our consideration of differential semblance. Thanks to Lynn Kirlin and two anonymous reviewers for many suggestions in improving the manuscript. We also thank the OU Geophysical Reservoir Characterization Consortium for financial support. Thanks to Conoco Inc. for providing the field data set.

REFERENCES


APPENDIX A

A SIMPLE DIFFERENTIAL SEMBLANCE

The concept of differential semblance exploits the similarity between two adjacent traces by calculating the difference between the two traces. It is based on the hypothesis that the right velocity will give the right moveout curve, causing the events to line up on adjacent traces and, hence, give the optimum of the objective function. We define it as the summation of the difference between pairs of traces. This concept is similar to the differential semblance optimization of Symes and Kern (1994). Following Siegfried and Castagna (1982), the differential semblance objective function between two adjacent traces $i$ and $i+1$ can be written as

$$\Theta_s(v_i | t) = \frac{||S_{i+1} - S_i||^2}{2(||S_{i+1}||^2 + ||S_i||^2)}$$  \hspace{1cm} (A-1)

where $S_i$ is the $i$th trace of a gather.

Equation (A-1) can be simplified to

$$\Theta_s(v_i | t) = \frac{1}{2} - \frac{||S_{i+1}|| ||S_i|| \cos \theta}{||S_{i+1}||^2 + ||S_i||^2},$$  \hspace{1cm} (A-2)

which is equivalent to

$$\Theta_s(v_i | t) = \frac{1}{2} - \left( \frac{r}{1 + r^2} \right) \cos \theta,$$  \hspace{1cm} (A-3)

where $r$ is the ratio of the norms of the two adjacent traces. If they have the same time dependence but have different amplitudes, $r$ is just the ratio of these amplitudes. $\theta$ is the angle between the two vectors $S_i$ and $S_{i+1}$. As the two waveforms become similar in shape and amplitude, the value of the objective function $\Theta_s$ decreases or, in other words, the value of $(r/1+r^2)\cos \theta$ increases. For a perfect match, this term equals 1/2, and $\Theta_s$ equals zero. Similarly, when there are $N$ traces, the differential semblance objective function is given by

$$\Theta_s(v_i | t) = 1 - \frac{2}{N-1} \sum_{i=1}^{N-1} \frac{r_i}{1 + r_i^2} \cos \theta_i,$$  \hspace{1cm} (A-4)

where the constant 2 is inserted to scale the function to a maximum fitness of 1; $r_i$ and $\theta_i$ are the amplitude ratio and the phase difference of the $i$th pair. The sum consists of two parts: an amplitude part $[r_i/(1 + r_i^2)]$ and a waveshape part $(\cos \theta_i)$. The waveshape part is the crosscorrelation between the traces. For a simple interface and reflections inside the precritical range, AVO does not affect the waveshape but it does affect the amplitude. Also, the amplitude part is significantly affected only if the trace separation is large. When trace separation is small, the overall amplitude variation of the event does not affect the differential semblance function. The relative importance of the amplitude and the waveshape can be selected by

Fig. 15. Crossplot of the intercept and gradient terms estimated from the field gather using velocities that maximized $AB$ semblance at every time sample within a time window of 1.4–1.8 s for $\lambda = 0.0$. 
varying the exponents \( m \) and \( n \) in the following objective function:

\[
\Theta_s(v_i | t) = 1 - \frac{1}{N-1} \sum_{i=1}^{N-1} \left( 2 - \frac{r_i}{1 + r_i^2} \right)^m \cos^n \theta_i. \tag{A-5}
\]

For comparison with semblance, we want an objective function which has a maximum when all traces are identical. Since \( \Theta_s \) is a minimum for this case, we work instead with \( 1 - \Theta_s \), which we refer to as simple differential semblance. Figures 2 and 3 show \( 1 - \Theta_s \), with \( n = m = 1 \).

**APPENDIX B**

**EIGENVALUE METHODS AND AVO**

Eigenvalue methods (e.g. Biondi and Kostov, 1989; Key and Smithson, 1990) attempt to provide higher resolution than semblance. The key to understanding eigenvalue methods is to recognize that the signal covariance matrix is of low rank (rank one for a single event) in the absence of noise. This is true regardless of AVO. To see this, consider equation (3) of Key and Smithson (1990),

\[
r_i(t) = s(t) + n_i(t), \tag{B-1}
\]

where \( r_i(t) \) is the measured reflection amplitude and the subscript \( i \) denotes the offset dependence. It consists of two parts: a signal, \( s(t) \), which is independent of offset, and noise, \( n_i(t) \), which depends on offset. We generalize it to arbitrary AVO:

\[
r_i(t) = s_i(t) + n_i(t), \tag{B-2}
\]

where \( s_i(t) = a_i s(t) \) and \( a_i \) is the amplitude dependence with offset. The signal part of the covariance matrix \( s's \) is still the outer product of a vector; hence, the rank of the matrix is still one. This fact is used to separate the covariance matrix into a subspace containing signal plus noise, and a subspace containing only noise. This separation is used in the design of coherency measures.

Key and Smithson (1990) proposed the coherence measure:

\[
\Theta_s(v_i | t) = M \ln \left[ \left( \frac{\sum_{j=1}^{N} \varepsilon_j}{\sum_{j=1}^{N} \varepsilon_j} \right)^N \right]
\]

\[
\cdot \frac{\varepsilon_1 - \frac{\sum_{j=2}^{N} \varepsilon_j}{(N-1)}}{\sum_{j=2}^{N} \varepsilon_j}, \tag{B-3}
\]

This function tends to one when signal to noise ratio tends to infinity, whereas it tends to zero when signal to noise ratio tends to zero. This is just the ratio of the estimates of signal to signal plus noise. Like semblance, it is bounded by one. It was also used by Gertzkenkorn and Marfurt (1999) as a coherency measure and deserves further study.

Figure 2 and Figure 3 were computed on (almost) noise free cases so the dynamic range of the Key and Smithson (1990) formula [equation (B-3)] is enormous. In order to accommodate this, we have omitted the first factor in brackets in equation (B-3) and normalized estimates to the peak value.

**APPENDIX C**

**REGULARIZATION OF THE COHERENCE MEASURE**

Solving for the reflectivity parameters is an inverse problem of the form \( \mathbf{d} = P \mathbf{m} \), where \( \mathbf{d} \) is the data vector consisting of data points along the moveout curve defined by a stacking velocity, \( \mathbf{m} \) is the model vector, and \( P \) is the data kernel. Overdetermined problems like ours can be solved by the least-squares method. Parameters are obtained that best fit data in a least squares sense (i.e., minimizes a \( L_2 \) norm). The problem can be written as

\[
\mathbf{m} = \mathbf{CD}, \tag{C-1}
\]

where \( \mathbf{C} = [P^T P]^{-1} \) and \( \mathbf{D} = P^T \mathbf{d} \). \( \mathbf{C} \) is called the covariance matrix. Its diagonal terms signify the variance of each parameter and thus give information about their accuracy.

For our problem of AB semblance, the normalized covariance matrix can be written as

\[
\mathbf{C} = \left[ \begin{array}{cc}
\frac{1}{\langle \sin^2 \theta \rangle} & \langle \sin^2 \theta \rangle \\
\langle \sin^2 \theta \rangle & \langle \sin^4 \theta \rangle
\end{array} \right]^{-1}.
\]
\[ C = \frac{1}{(\sin^2 \theta - \langle \sin^2 \theta \rangle)^2} \begin{bmatrix} \langle \sin^4 \theta \rangle & -\langle \sin^2 \theta \rangle \\ -\langle \sin^2 \theta \rangle & 1 \end{bmatrix}, \]  

(C-2)

where \( \langle \cdot \rangle \) implies average over number of offsets. Thus the variance of \( A \) is \( (\sin^4 \theta)/(\sin^4 \theta - \langle \sin^2 \theta \rangle^2) \) and of \( B \) is \( 1/(\sin^4 \theta - \langle \sin^2 \theta \rangle^2) \). The variance of \( A \) and \( B \) does not depend on the data but only on the angles of incidence. Since the range of angle of incidence decreases with depth, one would expect the variance of \( B \) to increase. Thus, reflectivity values determined at deeper horizons using the \( AB \) parameterization are less accurate than those obtained at shallower horizons. This, when coupled with the flattening of the moveout curve, causes an increased loss of resolution. This is primarily due to trade off between the estimated parameters. Using redundant parameterization in such cases causes high variances and highly inaccurate parameter estimation. This was observed in all the test gathers described above.

To improve results, one needs to find a way to stabilize the covariance matrix. The formulation in equations (7) and (8) does precisely that. Using \( \lambda \) as shown in equation (8) modifies the covariance matrix:

\[ \begin{bmatrix} 1 & \langle \sin^2 \theta \rangle \\ \langle \sin^2 \theta \rangle & \langle \sin^4 \theta \rangle \end{bmatrix}^{-1} \text{to} \begin{bmatrix} 1 & \langle \sin^2 \theta \rangle \\ \langle \sin^2 \theta \rangle & \langle \sin^4 \theta \rangle + \lambda \end{bmatrix}^{-1}. \]  

(C-3)

\[ m = \left[ P^T P \right]^{-1} P^T d \]  

In this case, the covariance matrix becomes

\[ \begin{bmatrix} 1 & \langle \sin^2 \theta \rangle \\ \langle \sin^2 \theta \rangle & \langle \sin^4 \theta \rangle + \lambda \end{bmatrix}^{-1} \begin{bmatrix} \langle F_D \rangle \\ \langle F_D \sin^2 \theta \rangle \end{bmatrix}. \]  

(C-4)

When \( \lambda \) is small the formulation goes to \( AB \) semblance. In other words, only the prediction error is minimized. This is desired when there is enough range in the angle of incidence. When \( \lambda \) is large, the formulation goes to \( A \) semblance (traditional semblance). This means only the model length is minimized, which is desired when there is not much AVO present, a case which arises when the aperture of incidence angles is small. The large value of \( \lambda \) stabilizes the covariance matrix by taking the formulation to a one-parameter estimation instead of a two-parameter estimation. Figures 12 and 14 indicate that this approach appears to improve the resolution of the semblance plots without loosing its ability to identify events with polarity reversals.