

HOMOMORPHIC PROCESSING OF ACOUSTIC LOGGING DATA

by

K.J. Ellefsen and D.R. Burns

New England Research, Inc.
76 Olcott Drive
White River Junction, VT 05001

C.H. Cheng

Earth Resources Laboratory
Department of Earth, Atmospheric, and Planetary Sciences
Massachusetts Institute of Technology
Cambridge, MA 02139

ABSTRACT

A new processing method, which we developed for the guided waves generated during acoustic logging, accurately estimates the wavenumber when only a few seismograms are available or when the seismograms are irregularly spaced. The estimates of the attenuation coefficient are seemingly accurate when many seismograms are available but are inaccurate when only a few seismograms are available. The new method does not generate any spurious estimates as the Prony-based method does.

INTRODUCTION

The guided waves generated during acoustic logging are sometimes used to estimate formation properties (see e.g., Cheng et al., 1987; Hornby et al., 1989; and Ellefsen et al., 1991). Because the algorithms are often formulated in the frequency-wavenumber domain, the seismograms must be processed to calculate the parameters of the waves in this domain (viz., wavenumber, attenuation coefficient, and amplitude).

The most widely-known method for processing these guided waves, the Prony-based method (Parks et al., 1983; McClellan, 1986; Lang et al., 1987; Ellefsen et al., 1989), has four significant disadvantages. First, many spurious estimates are generated. Since no automated method currently exists for identifying and deleting these estimates, the task must be done manually. Second, many seismograms (i.e., at least 8) are required to obtain reasonably accurate results. Because many existing logging tools have only

4 receivers, this method cannot process the data collected by these tools. Third, the receivers must be regularly spaced. Although all tools are built in this manner, a data set in which one seismogram from the middle of the array is bad cannot be processed with this method. Fourth, the estimates of the attenuation coefficient are often inaccurate and biased. This problem is probably caused by the heterogeneity of the formation and by mismatched receivers. Another processing method, high resolution slant stacking (Hsu and Baggeroer, 1986; Block et al., 1986), can estimate the phase velocity but neither the attenuation coefficient nor amplitude.

In this paper, we describe a new processing method that overcomes many of the disadvantages of the Prony-based method. The method designed is to determine the parameters of only one wave, like the tube wave. First, we will explain the underlying mathematics. Then, using field data, we will demonstrate its performance and will also compare its performance to that of the Prony-based method.

METHOD

The mathematical basis of the processing method is derived from the solution to the wave equation in the frequency-wavenumber domain:

$$s(z, t) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} d\omega e^{-i\omega t} \int_{-\infty}^{\infty} dk X(k, \omega) e^{ikz} \quad (1)$$

(see e.g., Cheng et al., 1981). $s(z, t)$ is the response recorded by a receiver, z the distance between the receiver and the source, t the time, ω the frequency, and k the wavenumber. $X(k, \omega)$ is an expression which includes information about the geometry of the model, the source, the boundary conditions, etc., and we do not need to know its exact form to develop the processing method. In the frequency domain, this expression is

$$\tilde{s}(z, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk X(k, \omega) e^{ikz} \quad (2)$$

This integral, when it is evaluated using contour integration in the complex wavenumber domain, equals the sum of the residues enclosed by the contour (ignoring, for the moment, any contribution from the branch cuts):

$$\tilde{s}(z, \omega) = i \sum \text{Residues of } X(k_c, \omega) e^{ik_c z} \text{ enclosed by the contour} \quad (3)$$

(Peterson, 1974) where k_c is the complex wavenumber. In some frequency bands, the residue of one pole is the dominant contribution to $\tilde{s}(z, \omega)$. For example, between 0 and about 5 kHz, the residue of the pole associated with the tube wave is much larger than any of the other residues, and this fact is clearly demonstrated in the frequency-wavenumber plots by Schmitt and Bouchon (1985). In these bands, $\tilde{s}(z, \omega)$ is well

approximated by this one residue, which we will express as $R(k_p, \omega) e^{ik_p z}$ where k_p is the value of the complex wavenumber at the pole. What is important in the processing is the distance that the wave propagates from the first receiver in the array. This distance will be designated \hat{z} and equals $z - z_1$ where z_1 is the distance from the source to the first receiver. The receiver response due to the dominant pole is

$$\tilde{s}(z, \omega) = [iR(k_p, \omega) e^{ik_p z_1}] e^{-\alpha \hat{z}} e^{ik_r \hat{z}} \quad (4)$$

where k_r and α are the real and imaginary parts, respectively, of k_p . To simplify the expression within the brackets, let A and ϕ equal its modulus and argument, respectively:

$$\tilde{s}(z, \omega) = A e^{-\alpha \hat{z}} e^{i(\phi + k_r \hat{z})} \quad (5)$$

This expression has a very simple interpretation. The amplitude of the wave at the first receiver is A , and its phase ϕ . As the wave propagates along the receiver array, its amplitude changes by $e^{-\alpha \hat{z}}$, and its phase by $k_r \hat{z}$. Note that these four parameters are functions of frequency; this dependence has been omitted to simplify the notation.

The Prony-based method can be used to estimate A , ϕ , α , and k_r , but it has the problems that were already discussed. The method that we propose is a type of homomorphic signal processing (see e.g., Oppenheim and Schaffer, 1975, p. 480–531): it uses the natural logarithm of $\tilde{s}(z, \omega)$. The real and imaginary parts of the logarithm are

$$\Re[\ln \tilde{s}(z, \omega)] = \ln A - \alpha \hat{z} \quad (6)$$

and

$$\Im[\ln \tilde{s}(z, \omega)] = \text{principal value of } [\phi + k_r \hat{z}] \quad (7)$$

respectively. Note that the imaginary part is based on the principal value because the logarithm has a branch cut at π (Churchill et al., 1974, p. 62–64). Interestingly, the transformation completely separates the information about the amplitude and phase of the wave, putting it into equations 6 and 7, respectively. In terms of analytical geometry, the equation for the amplitude describes a straight line: the intercept is $\ln A$, and the slope $-\alpha$. Taking into account the effect of the principal value, the equation for the phase also describes a straight line: the intercept is ϕ , and the slope k_r . These two properties are important because the technique for fitting a line to data is well developed (see e.g., Draper and Smith, 1966, chapters 1–3; Daniel and Wood, 1971, chapters 2 and 3). An additional advantage of the logarithmic transformation is that the estimation problem is formulated with real numbers instead of complex numbers making the inversion simpler.

The inversions are based upon the difference called the residual, \mathbf{r} , between the observed data, \mathbf{d} , and the predictions, \mathbf{Zm} :

$$\mathbf{r} = \mathbf{d} - \mathbf{Zm} \quad (8)$$

\mathbf{Z} contains information about the receiver offsets. For the inversion of the phase data, \mathbf{m} consists of k_r and ϕ at one frequency; for the inversion of the amplitude data, \mathbf{m} consists of $\ln A$ and α at all frequencies. \mathbf{r} can also include data from successive depths in the borehole or from multiple sources. A detailed description of these terms is in the Appendix.

For the inversion of the phase data, we minimize a cost function that is the sum of the absolute values of the residuals:

$$\left(|\mathbf{r}|^{1/2}\right)^T \left(|\mathbf{r}|^{1/2}\right) \quad (9)$$

This criterion for minimization, called an l_1 norm, has the desirable feature that it is less affected by outliers in the data than other types of norms. The algorithm used for the minimization is iterative reweighted least squares (Scales and Gersztenkorn, 1988).

For the inversion of the amplitude data, a more sophisticated cost function must be minimized to obtain reasonable results:

$$\left(|\mathbf{r}|^{1/2}\right)^T \mathbf{C}_D^{-1} \left(|\mathbf{r}|^{1/2}\right) + c_1 \mathbf{m}^T \mathbf{D}^T \mathbf{D} \mathbf{m} + c_2 \Psi^T \Psi \quad (10)$$

The first term forces the predictions to fit the data. The residuals are weighted by \mathbf{C}_D whose elements, which are only on the main diagonal, are the reciprocals of the amplitude. When a wave has a low amplitude, the data tend to be noisy, and with this weighting the cost function is only slightly affected by the corresponding residuals. The second term forces the attenuation coefficient to vary smoothly as the frequency changes. The smoothness is measured with the second derivatives of the attenuation coefficients; the derivative operator is in matrix \mathbf{D} (Menke, 1984, p. 53). The importance given to the smoothness is controlled by c_1 . The third term forces the attenuation coefficient to be positive. If during the minimization of the cost function a given coefficient is positive, then the associated element of Ψ is zero. However, if this coefficient becomes negative, then the associated element of Ψ becomes large and positive (Bard, 1974, p. 141–145). To decrease the cost function the minimization algorithm forces the coefficient to be positive. The importance given to this constraint is controlled by c_2 .

To determine how well the predictions fits the data, the elements of the residual vector could be examined. However, this approach is not practical when a large amount of data is being processed. A good alternative is to examine the variance of the residuals, which is estimated using

$$\hat{\sigma}^2 = \frac{1}{N-2} \sum_{i=1}^N r_i^2 \quad (11)$$

where N is the number of elements in the residual vector (Devore, 1982, p. 435). The variance is valuable information when performing an inversion for formation properties because it indicates how much confidence should be given to the estimates. That is, the smaller the variance, the greater the confidence.

RESULTS

Very accurate estimates were obtained when this new method was used to process synthetic data, and these tests demonstrated that the mathematics are sound and the computer algorithm is properly written. However, the method must work well on field data to be useful, and consequently we will only present the tests with the field data here. The data were collected by a tool having two sources and twelve receivers (Figure 1). Estimates for the tube wave were obtained between 0 and 4000 Hz, the frequency range within which the amplitude of this wave is large. Since only the wavenumber, attenuation coefficient, and amplitude can be used in an inversion for formation properties, only these estimates will be discussed here.

First, the complete data set was processed with the homomorphic and Prony-based methods. The wavenumber and amplitude estimates obtained by both methods are similar (Figure 2a, b, e, and f). However, the attenuation coefficient estimates obtained by the homomorphic method increase smoothly with frequency whereas those obtained with the Prony-based method do not (Figure 2c and d). We believe, but cannot prove, that these smoother estimates are a more accurate characterization of the tube wave. The many spurious estimates generated by the Prony-based method are evident. The estimates obtained with the homomorphic method will be used as a standard against which the estimates from the subsequent tests will be judged.

The variances of the residuals (Figure 3) indicate that the estimates fit the data well from about 900 Hz to 3000 Hz. The fit degrades below 900 Hz probably because the wave has a low amplitude relative to the noise. The fit also degrades above 3000 Hz probably because a leaky mode with a low amplitude is present. Nonetheless the estimates in these two ranges are reasonable (Figure 2a, c, and e). Because the amplitude of the wave is an indication of the signal to noise ratio, it is another qualitative measure of the reliability of an estimate: the higher the amplitude, the greater the reliability. The variances and amplitudes for the subsequent tests will not be discussed because their basic features are similar to those in Figure 2e and Figure 3.

Second, to test the performance of the method when only a few receivers are available, four seismograms from the complete data set (Figure 1) were processed. The wavenumber estimates obtained with the homomorphic method for this small data set are like those obtained from the complete data set (Figure 4a); however, those obtained with the Prony-based method are not (Figure 4b). The attenuation estimates obtained with both methods are inaccurate (Figure 4c and d). The spurious estimates generated by Prony-based method are again evident (Figure 4b and d).

Third, six irregularly spaced seismograms from the complete data set (Figure 1) were processed. The wavenumber estimates from this data set are like those from the complete data set (Figure 5a). The attenuation estimates, however, are inaccurate

(Figure 5b).

DISCUSSION

The new processing method overcomes many, but not all, of the disadvantages of the Prony-based method. First, no spurious estimates are generated. Second, only a few seismograms are needed to obtain seemingly accurate wavenumber estimates. Third, the receivers can be irregularly spaced. Fourth, the estimates for the attenuation coefficient are apparently accurate when many seismograms are available. The only significant problem is that this new method cannot accurately estimate the attenuation coefficient when only a few seismograms are available. Since the amplitude of a wave is strongly affected by formation heterogeneity and matching the receivers is difficult, this problem will not be easily solved.

Although we have demonstrated this method with only one type of wave, the method is general: it can be used to process any guided or surface wave when the amplitude of that wave dominates within some frequency range. Perhaps this method could be used to process the flexural wave generated during shear wave logging and the Rayleigh and Love waves generated either from a surface seismic experiment or an earthquake.

An important issue is how the estimates are affected by random noise when a logarithmic transformation is applied to the data (Menke, 1984, p. 147). To gain some insight into this problem, consider the response of a noisy receiver:

$$\tilde{s}(\hat{z}, \omega) = (Ae^{-\alpha\hat{z}} + \epsilon) e^{i(\phi + k_r\hat{z} + \varepsilon)} \quad (12)$$

where ϵ and ε are independent random variables characterizing the noise. The natural logarithm is used to separate the amplitude and phase data. The variance of the phase data is

$$\text{Var}(\Im[\ln \tilde{s}(\hat{z}, \omega)]) = \text{Var}(\varepsilon) \quad (13)$$

This variance is unaffected by the transformation, and consequently the inversion for ϕ and k_r will also be unaffected. Using a series expansion, the variance of the amplitude data is approximately

$$\text{Var}(\Re[\ln \tilde{s}(\hat{z}, \omega)]) \approx \text{Var}(\epsilon) \frac{e^{\alpha\hat{z}}}{A} \quad (14)$$

Because $\alpha > 0$ always, this variance increases as a function of z , and the inversion for $\ln A$ and α is adversely affected. However, we found that in field data the increase in the variance is small. We believe that the benefits of the logarithmic transformation (viz., the simplicity of fitting a straight line to data) outweigh this detriment.

Another important issue is the effect that other waves with small amplitudes have on the accuracy of the estimates for the tube wave. To gain some intuition about this effect, imagine that two waves are present. The response of the receiver is the sum:

$$\bar{s}(\hat{z}, \omega) = A_1 e^{-\alpha_1 \hat{z}} e^{i(\phi_1 + k_{r1} \hat{z})} + A_2 e^{-\alpha_2 \hat{z}} e^{i(\phi_2 + k_{r2} \hat{z})} \quad (15)$$

where $A_1 \gg A_2$. Index 1 refers to the tube wave, and index 2 to either a pseudo-Rayleigh or leaky wave. Taking the natural logarithm and using a series expansion, the real part is

$$\Re[\ln \bar{s}(\hat{z}, \omega)] \approx \ln A_1 - \alpha_1 \hat{z} + \frac{A_2}{A_1} e^{-(\alpha_2 - \alpha_1) \hat{z}} \cos[(\phi_2 - \phi_1) + (k_{r2} - k_{r1}) \hat{z}] \quad , \quad (16)$$

and the imaginary part is

$$\Im[\ln \bar{s}(\hat{z}, \omega)] \approx \phi_1 + k_{r1} \hat{z} + \frac{A_2}{A_1} e^{-(\alpha_2 - \alpha_1) \hat{z}} \sin[(\phi_2 - \phi_1) + (k_{r2} - k_{r1}) \hat{z}] \quad . \quad (17)$$

Both parts of the response have small oscillatory components that decay exponentially with distance (assuming that $\alpha_2 > \alpha_1$). Consequently the presence of an additional wave with relatively low amplitude will have little effect on the estimates for the tube wave.

CONCLUSIONS

We have developed a new processing method based upon a transformation used in homomorphic signal processing. The method can estimate the parameters of one guided wave when the amplitude of that wave is much larger than the amplitude of any other waves. Although we have demonstrated this method by processing only a tube wave, the method could be applied to other types of guided and surface waves.

With the new processing method, accurate estimates for the wavenumber were obtained when only a few seismograms were available and when the seismograms were irregularly spaced. The estimates of the attenuation coefficient were seemingly accurate when many seismograms were available but were inaccurate when only a few seismograms were available. No spurious estimates were generated as they are with the Prony-based method. In general, the new processing method estimates the parameters of the guided wave better than the Prony-based method does.

ACKNOWLEDGEMENTS

We wish to thank ARCO Oil and Gas Company for the acoustic logging data. KJE and DRB were supported by a contract from Halliburton Logging Services. CHC was supported by the Full Waveform Acoustic Logging Consortium at M.I.T.

REFERENCES

- Bard, Y., *Nonlinear Parameter Estimation*, Academic Press Inc., 1974.
- Block, L.V., C.H. Cheng and G.L. Duckworth, Velocity analysis of multi-receiver full waveform acoustic logging data in open and cased holes, *56th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts*, Houston, Texas, 33–35, 1986.
- Cheng, C.H., and M. N. Toksöz, Elastic wave propagation in a fluid-filled borehole and synthetic acoustic logs, *Geophysics*, 46, 1042–1053, 1981.
- Cheng, C.H., J. Zhang, and D. R. Burns, Effects of in-situ permeability on the propagation of Stoneley (tube) waves in a borehole, *Geophysics*, 52, 1279–1289, 1987.
- Churchill, R.V., *Complex Variables and Applications*, McGraw-Hill Book Co., 1974.
- Daniel, C., and F.S. Wood, *Fitting Equations to Data*, John Wiley & Sons, Inc., 1971.
- Devore, J.L., *Probability and Statistics for Engineering and the Sciences*, Brooks/Cole Publishing Co., 1982.
- Draper, N.R., and H. Smith, *Applied Regression Analysis*, John Wiley & Sons, Inc., 1966.
- Ellefsen, K.J., C.H. Cheng, and K.M. Tubman, Estimating phase velocity and attenuation of guided waves in acoustic logging data, *Geophysics*, 54, 1054–1059, 1989.
- Ellefsen, K.J., M.N. Toksöz, K.M. Tubman, and C.H. Cheng, Estimating a shear modulus of a transversely isotropic formation, *Geophysics*, 1991, submitted for publication.
- Hornby, B.E., D.L. Johnson, K.W. Winkler, and R.A. Plumb, Fracture evaluation using reflected Stoneley-wave arrivals, *Geophysics*, 54, 1274–1288, 1989.
- Hsu, K., and A.B. Baggeroer, Application of the maximum-likelihood method (MLM) for sonic velocity logging, *Geophysics*, 51, 780–787, 1986.
- Lang, S.W., A.L. Kurkjian, J.H. McClellan, C.F. Morris, and T.W. Parks, Estimating slowness dispersion from arrays of sonic logging data, *Geophysics*, 52, 530–544, 1987.
- McClellan, J.H., Two-dimensional spectrum analysis in sonic logging, *Acoustics, Speech, Signal Processing Magazine*, 3, 12–18, 1986.
- Menke, W., *Geophysical Data Analysis: Discrete Inverse Theory*, Academic Press Inc., 1984.
- Oppenheim, A.V., and R.W. Schaffer, *Digital Signal Processing*, Prentice-Hall, Inc., 1975.
- Parks, T.W., J.H. McClellan, and C.F. Morris, Algorithms for full-waveform sonic logging, in *Proceedings of the Second Acoustics, Speech, and Signal Processing Workshop on Spectral Estimation*, pp. 186–191, Tampa, Florida, 1983.

Peterson, E.W., Acoustic wave propagation along a fluid-filled cylinder, *J. Appl. Phys.*, 45, 3340–3350, 1974.

Scales, J.A., and A. Gersztenkorn, Robust methods in inverse theory, *Inverse Problems*, 4, 1071–1091, 1988.

APPENDIX

In this appendix the terms in the residual vector, defined in equation 8, will be explained in detail.

For the inversion of the phase data, the residual vector is

$$\mathbf{r} = \begin{pmatrix} \Im [\ln \tilde{s}(z_1, \omega)] \\ \Im [\ln \tilde{s}(z_2, \omega)] \\ \vdots \\ \Im [\ln \tilde{s}(z_N, \omega)] \end{pmatrix} - \begin{pmatrix} \hat{z}_1 & 1 \\ \hat{z}_2 & 1 \\ \vdots & \vdots \\ \hat{z}_N & 1 \end{pmatrix} \begin{pmatrix} k_r \\ \phi \end{pmatrix} \quad (\text{A.1})$$

where each offset is indexed by a subscript ranging from 1 to N . If we compare the right hand side of this equation with that in equation 8, we see that the $N \times 1$ vector is \mathbf{d} , the $N \times 2$ matrix is \mathbf{Z} , and the 2×1 vector is \mathbf{m} .

When the inversion includes data from successive depths or from multiple sources, the residual vector must be slightly modified. The modification is based upon the assumption that k_r is the same for each data set but ϕ is different. When two sets of data are used, for example, the residual vector is

$$\mathbf{r} = \begin{pmatrix} \Im [\ln \tilde{s}^{(1)}(z_1, \omega)] \\ \vdots \\ \Im [\ln \tilde{s}^{(1)}(z_N, \omega)] \\ \Im [\ln \tilde{s}^{(2)}(z_1, \omega)] \\ \vdots \\ \Im [\ln \tilde{s}^{(2)}(z_N, \omega)] \end{pmatrix} - \begin{pmatrix} \hat{z}_1 & 1 & 0 \\ \vdots & \vdots & \vdots \\ \hat{z}_N & 1 & 0 \\ \hat{z}_1 & 0 & 1 \\ \vdots & \vdots & \vdots \\ \hat{z}_N & 0 & 1 \end{pmatrix} \begin{pmatrix} k_r \\ \phi^{(1)} \\ \phi^{(2)} \end{pmatrix} . \quad (\text{A.2})$$

The superscript denotes each set of data.

For the inversion of the amplitude data, the residual vector is constructed using all

frequencies of interest:

$$\mathbf{r} = \begin{pmatrix} \Re [\ln \bar{s}(z_1, \omega_1)] \\ \vdots \\ \Re [\ln \bar{s}(z_N, \omega_1)] \\ \vdots \\ \Re [\ln \bar{s}(z_1, \omega_M)] \\ \vdots \\ \Re [\ln \bar{s}(z_N, \omega_M)] \end{pmatrix} - \begin{pmatrix} \hat{z}_1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ \hat{z}_N & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & \hat{z}_1 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & \hat{z}_N & 1 \end{pmatrix} \begin{pmatrix} -\alpha(\omega_1) \\ \ln A(\omega_1) \\ \vdots \\ -\alpha(\omega_M) \\ \ln A(\omega_M) \end{pmatrix} \quad (\text{A.3})$$

where each frequency is indexed by a subscript ranging from 1 to M . If we compare the right hand side of this equation with that in equation 8, we see that the $NM \times 1$ vector is \mathbf{d} , the $NM \times 2M$ matrix is \mathbf{Z} , and the $2M \times 1$ vector is \mathbf{m} . To include data from successive depths or from multiple sources, the residual vector is modified in the same manner that the residual vector for the phase data is modified.

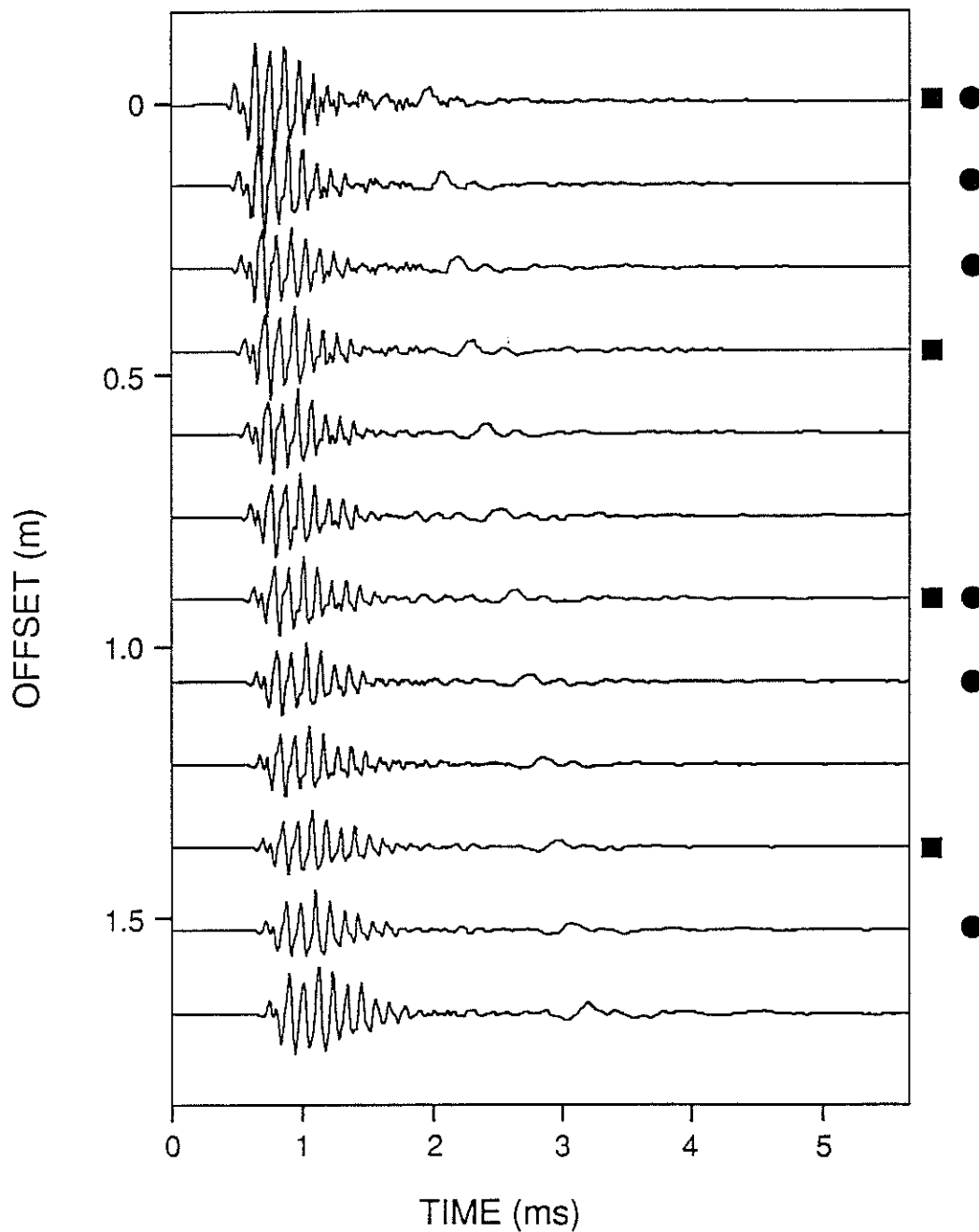
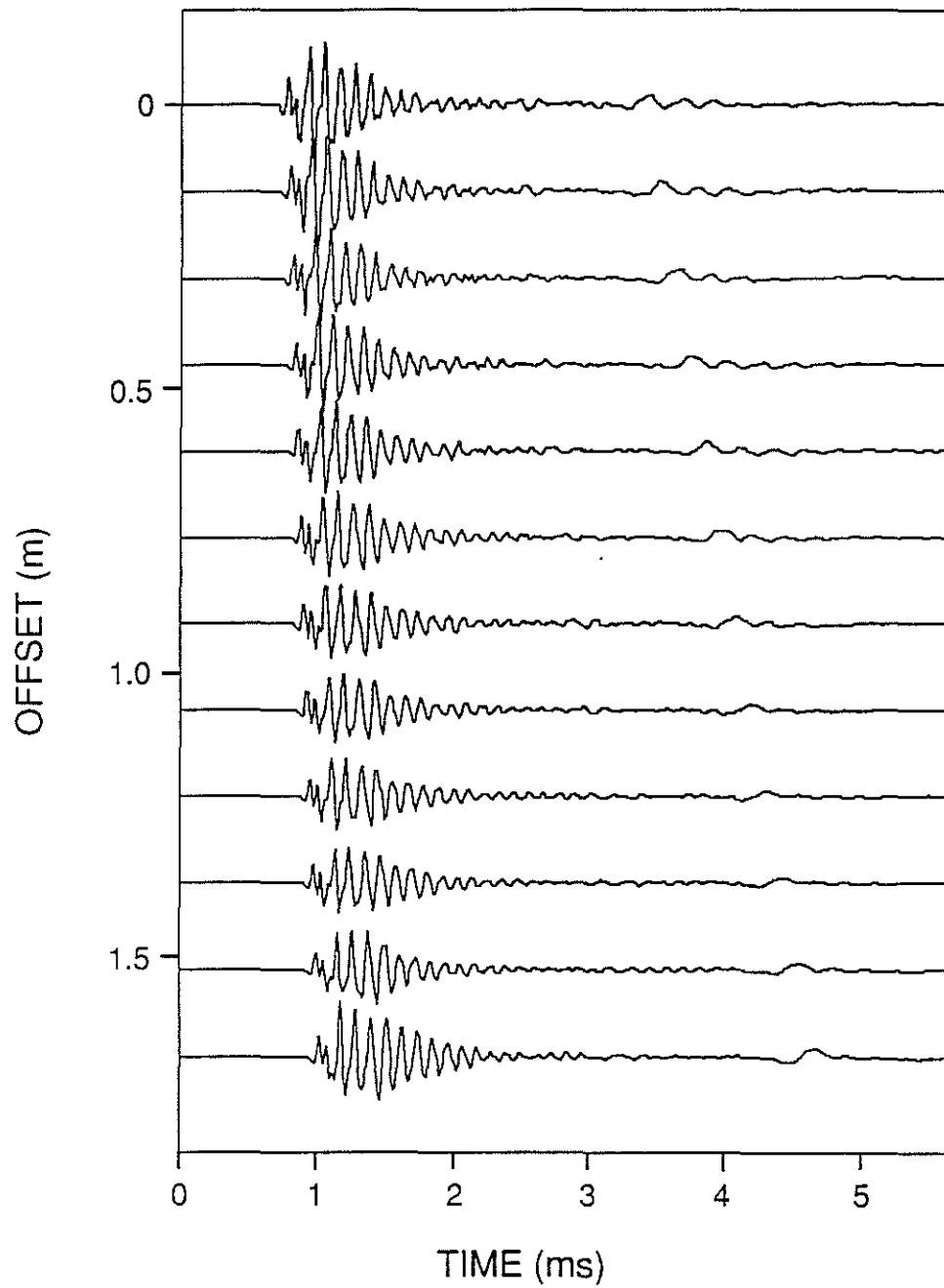


Figure 1: Data (above and overleaf) used to test the homomorphic processing method. The solid rectangles indicate those seismograms used to test its performance when only a few seismograms are available. The solid circles indicate those seismograms used to test its performance when only irregularly spaced seismograms are available.



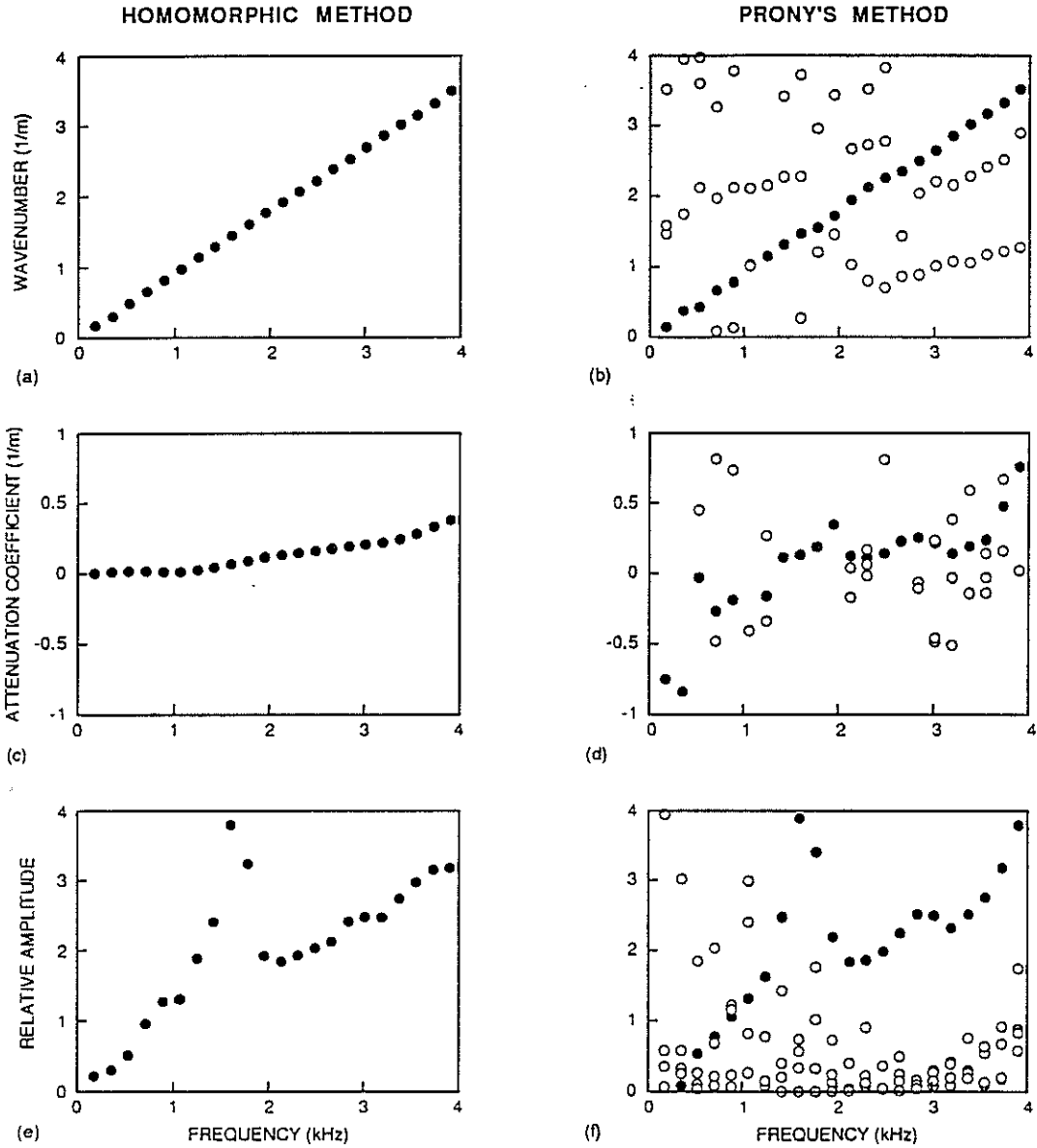


Figure 2: Parameter estimates for the tube wave (solid circles) obtained with the homomorphic and Prony-based methods using the complete data set (Figure 1). The open circles are spurious estimates generated by the Prony-based method.

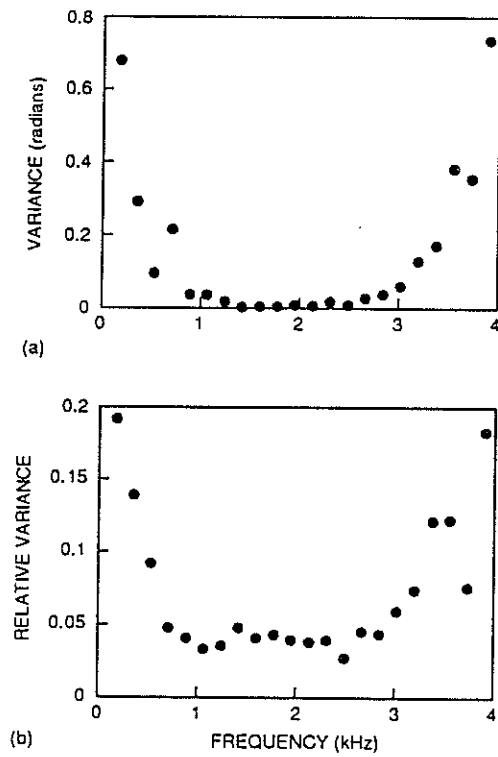


Figure 3: Variances of the residuals for (a) the phase data and (b) the amplitude data. The data were processed with the homomorphic method using the complete data set (Figure 1).

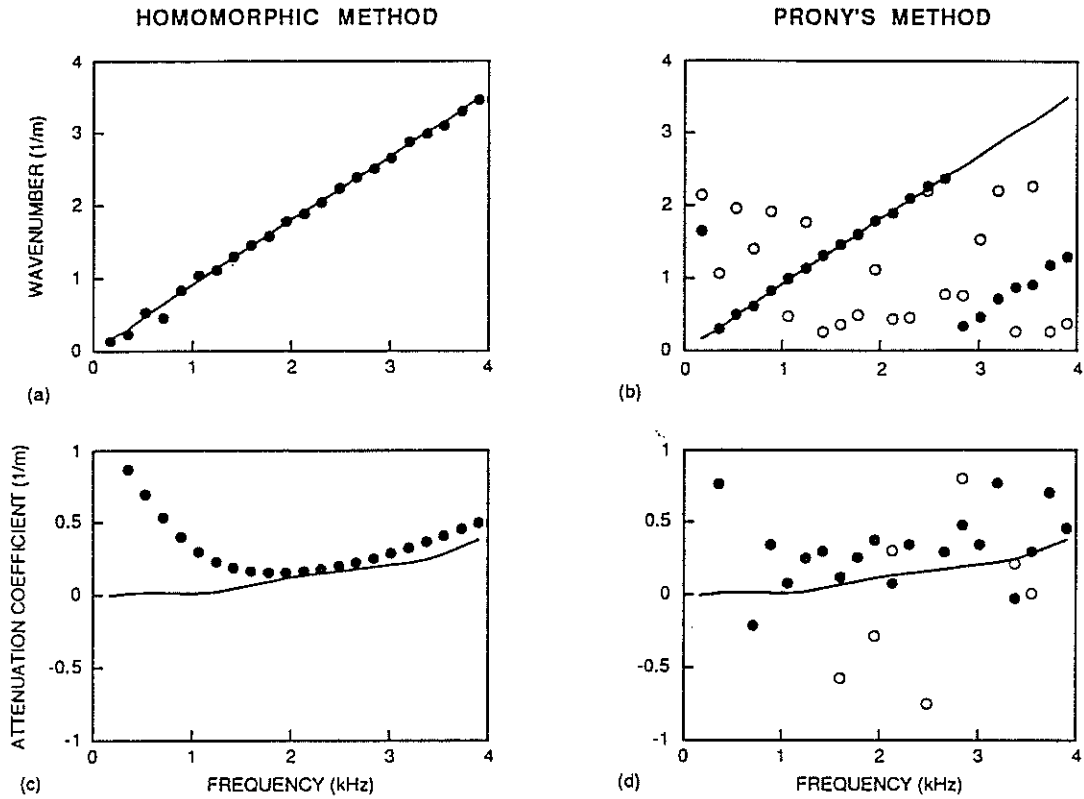


Figure 4: Parameter estimates for the tube wave (solid circles) obtained with the homomorphic and Prony-based methods using 4 seismograms from the complete data set (Figure 1). The lines are the estimates obtained with the homomorphic method from the complete data set (Figure 2). The open circles are spurious estimates.

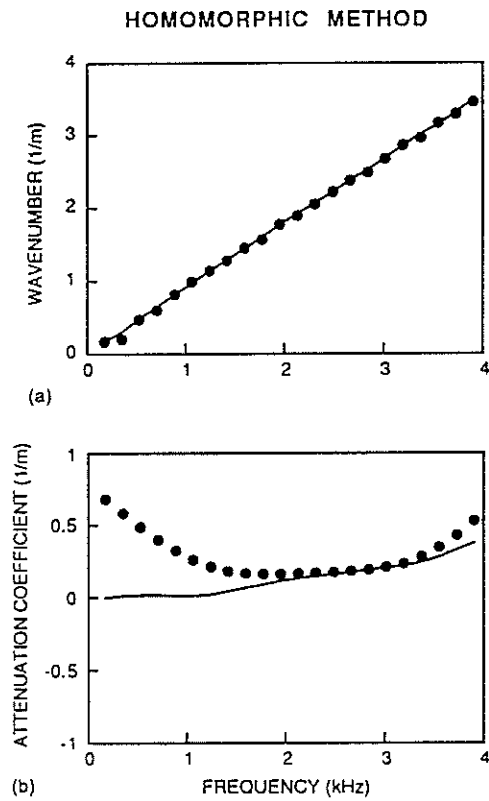


Figure 5: Parameter estimates for the tube wave (solid circles) obtained with the homomorphic method using 6 irregularly spaced seismograms from the complete data set (Figure 1). The lines are the estimates obtained with the homomorphic method from the complete data set (Figure 2).