MODELING, IDENTIFICATION AND CONTROL, 1985, VOL. 6, NO. 2, 57–73

Maximum-likelihood estimation of seismic impulse response†

B. URSIN and O. HOLBERG‡

Keywords: *Maximum-likelihood estimation, Seismic data processing, Deconvolution*

A seismic trace is assumed to consist of a known signal pulse convolved with a reflection coefficient series plus a moving average noise process (colored noise). Multiple reflections and reverberations are assumed to be removed from the trace by conventional means. The method of maximum likelihood (ML) is used to estimate the reflection coefficients and the unknown noise parameters. If the reflection coefficients are known from well logs, the seismic pulse and the noise parameters can be estimated.

The maximum likelihood estimation problem is reduced to a non-linear least-squares problem. When the further assumption is made that the noise is white, the method of maximum likelihood is equivalent to the method of least squares (LS). In that case the sampling rate should be chosen approximately equal to the Nyquist rate of the trace. Statistical and numerical properties of the ML- and the LS-estimates are discussed briefly. Synthetic data examples demonstrate that the ML-method gives better resolution and improved numerical stability compared to the LS-method.

A real data example shows the ML- and LS-method applied to stacked seismic data. The results are compared with reflection coefficients obtained from well log data.

1. Introduction

A seismic section is a mapping of geological structure. The ultimate aim of seismic data processing and interpretation is to solve the inverse problem: i.e., to infer the geological structure from noisy seismic data. Particularly in field development accurate and detailed geologic information is required. Most reservoirs are small in the vertical dimensions and even those that are not may thin out to zero thickness at the edges. For the delineation of such structures high vertical resolution is required.

Due to earth filtering, the seismic wavelets are distorted during their travel through solid earth. Deconvolution methods (Robinson and Treitel 1980) are used to compensate for this filtering effect and improve resolution. They have been successful for seismic events which do not closely overlap. The constraint limiting the applicability of such inverse filters is the limited resolution due to lack of bandwidth. It has been shown that with inverse filtering it is impossible to retrieve information outside the data bandwidth (Van Riel 1982; Van Riel and Berkhout 1983).

Classical interpretation techniques fail to resolve events with time separation less than approximately $1/(1{\cdot}4B)$ where B is the bandwidth of a broadband seismic wavelet (Kallweit and Wood 1982). This resolution limit is the best one can achieve in a noise-free case and, in practice, a larger time separation is required to resolve

Received 1 February 1985.
† Also published in *Geophysical Prospecting*, April 1985.
‡ Petroleum Technology Research Institute, 7034 Trondheim-NTH, Norway.

overlapping events. The state-of-the-art resolving power of reflection seismology is at best in the order of 20 m. This is not sufficient for the current aim of reservoir seismology.

The classical techniques are mostly based on least-squares (LS) inverse filtering which is optimal only if the noise is white. When the noise is colored, a weighted least-squares procedure should be used with a weighting matrix equal to the inverse of the noise covariance matrix. In practice, it is a problem to determine the noise covariance matrix. This problem can be solved by assuming that the noise is a moving-average process (white noise which has passed through a linear filter) and then estimate the coefficients in the noise filter. This corresponds to the method of maximum likelihood (ML) which is well-known in system identification (Åstrøm and Bohlin 1966; Åstrøm and Eykhoff 1971; Åstrøm 1980; Ljung and Søderstrøm 1983; Mendel 1983). A first attempt at using this technique to estimate seismic reflection coefficients was made by Özdemir (1982). Our results differ significantly from his, because we have used a different algorithm for the non-linear minimization which is required to compute the noise filter coefficients. Chi, Mendel and Hampson (1984) have developed a method for fast maximum-likelihood deconvolution in which the wavelet is estimated at the same time as the strength and location of the significant reflectors.

Our model does not include multiple reflections. These could be included by introducing auto-regressive parameters in the seismic model (Robinson 1978).

2. The seismic impulse response model

A seismic trace is assumed to consist of a known signal pulse convolved with a reflection coefficient series plus a moving average noise process (colored noise). Multiple reflections and reverberations are assumed to be removed from the trace by conventional means. If there are multiple reflections present, they will appear as part of the signal and thus give rise to false reflection coefficients. It is also assumed that the data have been corrected for geometrical spreading, and that identical scaling has been applied to the seismic traces.

An observed seismic trace may then be modeled as

$$y_k = s_k + w_k, \qquad k = 0, 1, \ldots, n_y \tag{1}$$

where s_k is the noise-free signal and w_k accounts for noise and model errors. $n_y + 1$ is the number of samples in the part of the trace that is being considered.

The signal is given by the convolution of the seismic pulse p_k, $k = 0, \ldots, n_p$, with a reflection coefficient series r_k, $k = 0, \ldots, n_R$:

$$S_k = \sum_{j=0}^{n_R} r_j p_{k-j}, \tag{2}$$

where $n_R + 1$ is the number of reflection coefficients. The reflection coefficients represent a discretized reflectivity model in which the reflectors are distributed in depth with two-way travel times between the reflectors being equal to the sampling interval. The seismic pulse is assumed to be constant in the time-gate considered. The shape of the pulse would then possibly include the effect of anelastic absorption. When p_{n_P} is the last sample of the seismic pulse which is significantly different from zero, n_y should be taken as $n_R + n_P$. We have tried to illustrate the assumed build-up of a noise-free seismic trace graphically in Fig. 1.

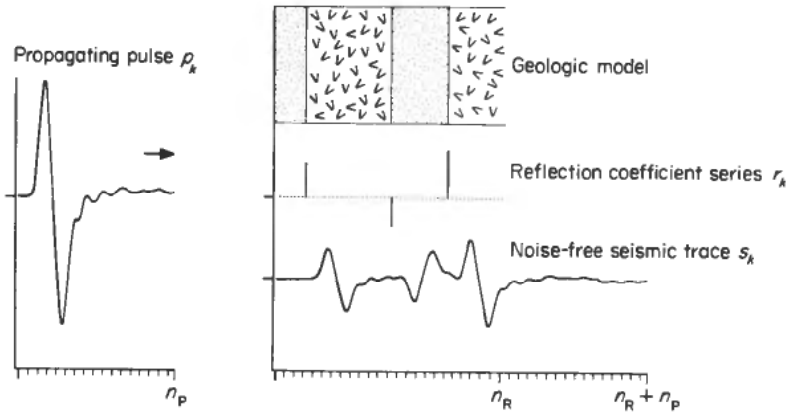

Figure 1. The assumed build-up of a noise-free seismic trace.

The noise is a low-order moving average process:

$$w_k = \sum_{j=0}^{nC} c_j e_{k-j}, \tag{3}$$

where $c_0 = 1$ and $c_j, j = 1, \ldots, n_C$, are the unknown filter coefficients. $\{e_k\}$ is a zero-mean Gaussian white noise series with variance σ^2.

The seismic impulse response model consists of a moving average model both for the signal and the noise. The signal is assumed to be deterministic while the noise is a stochastic process, as illustrated in Fig. 2.

From the observed seismic trace we may estimate the reflection coefficient series and the noise filter coefficients provided that the seismic pulse is known. Alternatively, if the reflection coefficients are known from well log data, the seismic pulse and the noise filter coefficients may be estimated.

The input wavelet is never known exactly. In offshore applications a recorded pulse may be used. For two reasons the recorded pulse is usually a poor estimate of the wavelet. Firstly, the effect of the surface reflection ('ghost') at the receiver must be taken into account by filtering the recorded pulse in the computer. This includes uncertainty in the surface reflection coefficient (usually assumed to be -1) and in the depth of the receiver array. The second source of error is more fundamental; the recorded pulse does not include the filtering effects of the sub-surface geology. An improved estimate of the seismic pulse from deep reflections may be obtained by

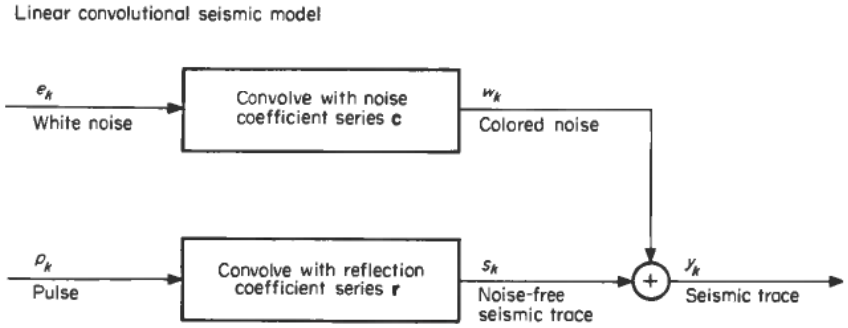

Figure 2. The seismic model used.

using reflection coefficients from well logs as input to our seismic model. This estimated pulse can in turn be used to estimate the reflection coefficient series in a corresponding time gate for the seismic traces in surrounding areas.

Generally it is difficult to recover the true values of the reflection coefficients. The relative amplitudes of the reflection coefficients are, however, preserved if the estimation is applied within a short time gate of a seismic trace with constant scaling within that time gate.

We have assumed that the reflectivity function is zero n_p samples before and n_p samples after the estimation time window. Therefore, we must expect edge effects when only a limited part of a trace is considered.

3. Maximum-likelihood estimation

Given the seismic data

$$y_k = \sum_{j=0}^{n_R} r_j p_{k-j} + \sum_{j=0}^{n_C} c_j e_{k-j}, \qquad k = 0, 1, \ldots, n_y \tag{4}$$

and the seismic pulse p_k, $k = 0, \ldots, n_p$, we want to compute estimates of the reflection coefficients r_k, $k = 0, \ldots, n_R$, the noise filter coefficeints c_k, $k = 1, \ldots, n_C$ ($c_0 = 1$) and the noise variance σ^2. The principle of maximum likelihood consists in choosing the values of **r**, **c** and σ^2 which make the observed data **y** the most probable for the given pulse **p**.

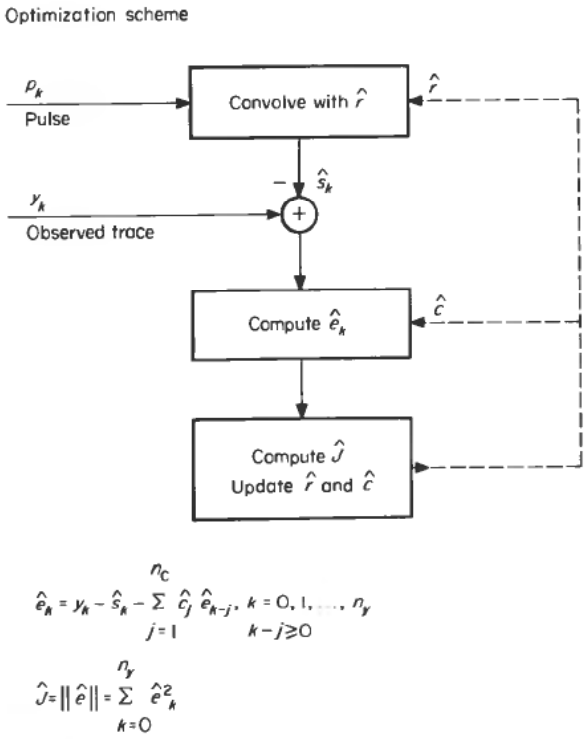

Figure 3. Iterative estimation of the reflection coefficient series **r** and the noise coefficeint series **c**. The model is perturbed until the trace is fitted: i.e. **r** and **c** are perturbed until no further minimization of J is feasible.

The sequence **e** is assumed to be Gaussian and white. Therefore, because our model is linear, the sequence **y** is Gaussian. The conditional ML-estimates of the vectors **r** and **c** are obtained by maximizing the conditional probability distribution of **y** for fixed **p**. It is easy to show (Åstrøm 1981) that this is equivalent to minimizing

$$J(\mathbf{r}, \mathbf{c} \mid \mathbf{y}, \mathbf{p}) = \sum_{k=0}^{n_y} e_k^2 \tag{5}$$

with respect to **r** and **c**. This is a non-linear least-squares problem. The ML-estimate for the variance of the white noise sequence **e** is:

$$\sigma^2 = \min (J)/(n_y + 1). \tag{6}$$

Estimates of the white noise process, needed in (5), are given by the recursive formula

$$e_k = y_k - s_k - \sum_{j=1}^{n_C} c_j e_{k-j}, \qquad \begin{cases} k = 0, 1, \ldots, n_y \\ k - j \geqslant 0 \end{cases}. \tag{7}$$

Clearly, for $n_C \neq 0$, J is strongly non-linear in **c**. We have tried to illustrate the maximum likelihood estimation of **r** and **c** graphically in fig. 3.

In seismic applications it is commonly assumed that the noise **w** is white. This is equivalent to choosing $n_C = 0$ in the equations above. The method of maximum likelihood then reduces to the method of least squares.

When searching for a minimum, we need the partial derivatives of e_k with respect to r_j and c_j. A recursive formula is obtained by differentiating (7) which gives (Åstrøm and Bohlin 1966):

$$\partial e_k/\partial r_j = -p_{k-j} - \sum_{i=0}^{n_C} c_i \partial e_{k-i}/\partial r_j, \quad \begin{cases} k = 0, 1, \ldots, n_y \\ j = 0, 1, \ldots, n_R \\ k - j \geqslant 0 \\ k - i \geqslant 0 \end{cases}; \tag{8a}$$

$$\partial e_k/\partial c_j = -e_{k-j} - \sum_{i=1}^{n_C} c_i \partial e_{k-i}/\partial c_j, \quad \begin{cases} k = 0, 1, \ldots, n_y \\ j = 1, 2, \ldots, n_R \\ k - j \geqslant 0 \\ k - i \geqslant 0 \end{cases}. \tag{8b}$$

These are currently termed the first order sensitivity functions. From (8) it is seen that

$$\partial e_k/\partial r_j = \partial e_{k-1}/\partial r_{j-1}, \tag{9a}$$

$$\partial e_k/\partial c_j = \partial e_{k-1}/\partial c_{j-1}. \tag{9b}$$

Therefore, (8) need only be computed for the first index j. This is utilized to reduce the computational effort.

Differentiation of (8) yields the second order sensitivity functions. As these are not required by our optimization algorithm (Moore, Garbow and Hillstrøm 1980) they are not included in this text.

The statistical properties of the ML-estimates are well-known. It can be shown (Goodwin and Payne 1977) that the ML-estimates obtained from (4) have the following asymptotic properties.

(i) The estimates converge to their true values as the number of independent observations goes to infinity (asymptotic consistency).
(ii) The estimates converge to normal random variables (asymptotic normality).
(iii) There exists no other unbiased estimator which gives smaller variance (asymptotic efficiency).

It should, however, be emphasized that very little is known about the statistical properties of the estimates obtained from small data sets (Åstrøm 1981).

4. Choice of sampling interval

An inherent property of the discrete system (4) is that no estimates can be obtained between sampling instants. Therefore, we cannot expect to resolve events with separation smaller than the sampling interval. How should the sampling interval be chosen to achieve the best possible resolution? We have not been able to fully answer this important question on a theoretical basis.

In least squares estimation it is assumed that the noise spectrum is white. If the noise spectrum is flat within the bandlimits of the signal, the noise will appear as white noise when the Nyquist frequency is equal to the bandwidth of the noise. If a higher sampling rate is used, the noise spectrum does not occupy its full bandwidth and the spectrum is definitely not white.

On the other hand, to avoid aliasing the Nyquist frequency cannot be chosen lower than the signal bandwidth. This leaves the upper cut-off frequency of the signal as the only allowed Nyquist frequency in least-squares estimation. Experimental observations with synthetic data support this conclusion, which has also been obtained by Özdemir (1982). If the Nyquist frequency is chosen larger than the signal bandwidth the least-squares procedure will become unstable. It is possible to stabilize the least-squares procedure by using the singular value decomposition cut-off method or ridge regression (Ursin and Zheng 1983), but this decreases the resolution. A spectral characterization of the ill-conditioning in numerical deconvolution is given by Ekstrom (1973).

When we consider the maximum-likelihood method, the answer is no longer trivial. The noise spectrum is now allowed to have any shape, and it does not put any restrictions on the choice of sampling rate. When a Nyquist frequency higher than the bandwidth is used, one cannot assume the observations (samples) to be independent. This problem is taken care of by the noise coefficients provided that the number of noise coefficient is large enough to approximate the spectrum reasonably well. A finite-impulse-response low-pass filter cannot be implemented with fewer coefficients than the ratio between the sampling frequency and the cut-off frequency. On the other hand, because our optimization problem is non-linear in $\mathbf{c}$, a high value for n_c is not recommended. If n_c is chosen carefully with these two limits in mind, we are able to perform estimation with a Nyquist frequency higher than the bandwidth of the signal. Nevertheless, we have a strong intuitive feeling that—as the sampling rate is increased—we will sooner or later face a non-uniqueness problem. This is easily demonstrated by convolving a band-limited pulse with a series of spikes of opposite polarity. The output tends to zero as the interval between the spikes of opposite polarity is decreasing. From a practical point of view, a non-uniqueness situation may not always represent a serious problem. If good initial estimates are provided, one can use a high sampling rate and hope that the

optimization algorithm will converge to the correct solution, irrespective of whether the solution is unique or not. This is a possible approach in stratigraphic extrapolation.

Using linear algebra Van Riel and Berkhout (1982) have shown that $2BT$, i.e., twice the product of bandwidth B and time gate T, set an upper limit to the number of independent parameters which can be uniquely retrieved from a seismic trace. For the least-squares method this gives the empirical result stated above. The data are n_y samples with a sampling period Δt, giving $T = n_y \Delta t$. The signal bandwidth should be $B = 1/2\Delta t$, and we see that the number of independent parameters is

$$2BT = 2n_y \Delta t/2\Delta t = n_y.$$

Keeping the statement above in mind we are led to conclude that, if we successfully have retrieved more than $2BT$ parameters from a seismic trace, these parameters are not independent. The estimated parameter set then contains no more information than a corresponding set with only $2BT$ parameters. However, the result with the smallest sampling interval may be easier to interpret.

In general, for maximum-likelihood estimation of the parameters in the model (4), we expect the resolution bound to depend on bandwidth, time gate, pulse length, signal-to-noise ratio and *a priori* information. We have not been able to establish a theoretical resolution limit, and we leave this question open.

5. Numerical optimization

For the minimization of the function J in (5) we have used the MINPACK library (Moore *et al.* 1978). This is a program package for the solution of systems of non-linear equations and for the solution of non-linear least-squares problems.

The MINPACK-implemented general non-linear least-squares solver, LMDER1, uses a modification of the Levenberg–Marquardt algorithm (Moore *et al.* 1980; Wolfe 1978). This algorithm generally guarantees global convergence, even from starting points far from the solution.

Our experience with the computer program is that the optimization algorithm generally converges, even from zero starting values. However, the convergence is very slow for $n_c \neq 0$.

During testing of the program we encountered numerical problems in a few cases. The problem seemed to be that the MINPACK optimization subroutine LMDER1 produced estimates that made the inverse of the filter $\mathbf{c}$ unstable. This resulted in overflow during the recursive calculations of $\mathbf{e}$ in (7). However, this problem was always easily solved by shifting the estimation time window a few samples or by choosing a lower order of the noise process. The problem did not occur when good initial estimated for the noise coefficients were provided.

6. Numerical results

6.1. Synthetic data

All our synthetic data are generated using a delayed pulse model without multiple reflections. Each trace is built up by summing the primary reflected pulses.

$$y(t) = \sum_{i=1}^{n} r_i p(t - \tau_i) + w(t), \tag{10}$$

where r_i is the reflection coefficient and τ_i is the two-way traveltime for reflecting interface number i. $p(t)$, the seismic pulse, is an airgun array signature. This pulse and its amplitude spectrum are displayed in Fig. 4. $w(t)$ is a low-pass filtered white noise sequence which is different for each trace. The bandwidth of the noise is chosen equal to the bandwidth of the pulse (125 Hz). The signal-to-noise ratio (SNR) is here defined as the ratio of the signal energy and the noise energy in the part of the trace under consideration. The displayed SNR-value on top of each synthetic data example is the average SNR for all traces in the ensemble. The delayed pulse model in (10) differs slightly from the model (4) used in the estimation: the two-way traveltimes for the reflecting interfaces are not restricted to the sampling instants, and the number of non-zero reflection coefficients is generally much less than the number of estimated reflection coefficients.

The first example is a simple pinchout where the reflection coefficients have opposite sign. The reflectivity model, the synthetic traces, and their frequency spectra are displayed in Fig. 5. For plotting, all traces are scaled by the highest amplitude in the ensemble. The sampling interval is 1·0 ms, and the time gradient is zero for the first interface and 1·0 ms/trace for the second. The wavelet used to generate the synthetic data is also used to obtain the estimates. All our assumptions are thus satisfied.

Two sets of reflection coefficient estimates are displayed in Fig. 6, where all reflection coefficients are scaled by the highest amplitude in each set. The first set is the least-squares result ($n_C = 0$). To meet the white-noise requirement we had to re-sample the traces to their Nyquist interval (4·0 ms). Where signal events arrive at the sampling instants or close to them, the estimates are good. Events arriving between sampling instants result in two reflection coefficients with smaller amplitude than the true amplitude. In attempting to increase the resolution by using a Nyquist frequency above the bandwidth, the LS-method becomes unstable. The estimates are then dominated by spurious events with large amplitudes and alternating polarity.

The next set in Fig. 6 is the result of maximum-likelihood estimation with the number of noise coefficients equal to twelve. We have obtained 1·0 ms resolution

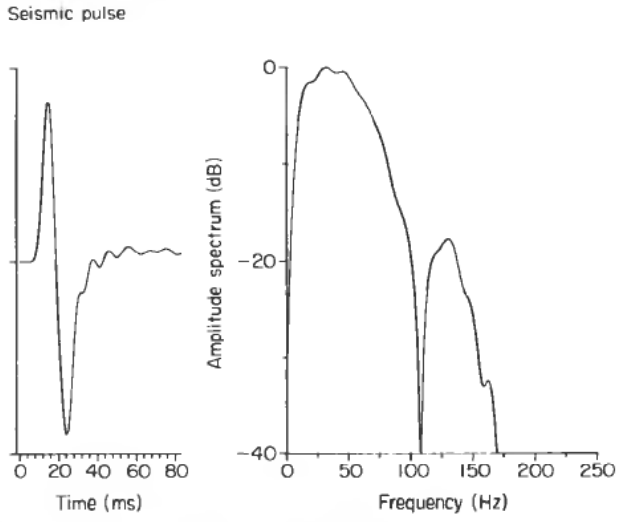

Figure 4. The recorded airgun signature used in synthetic data generation.

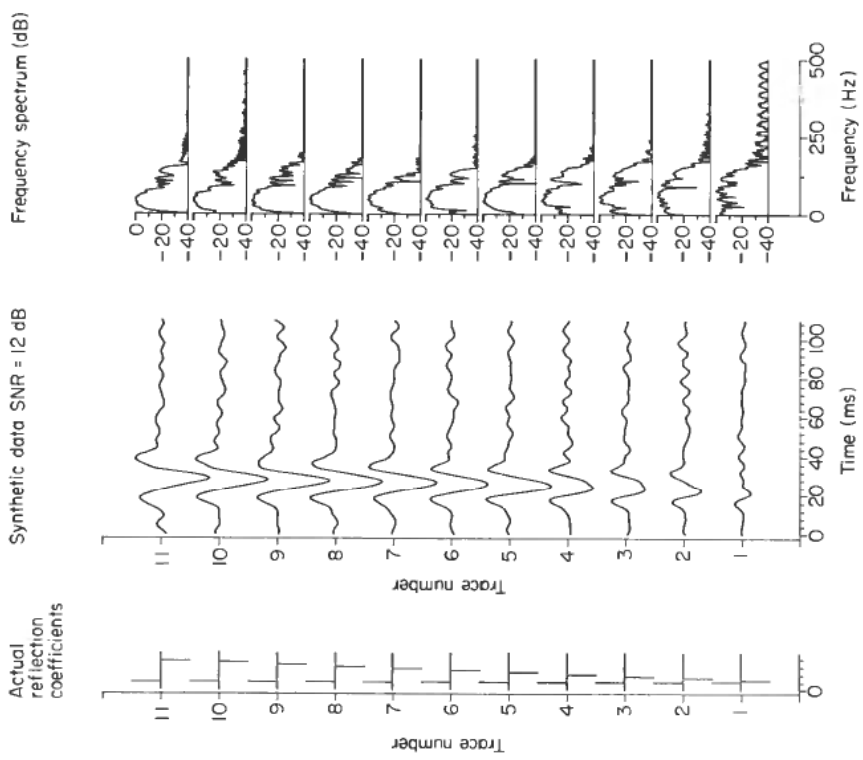

Figure 5. Thin layer-reflectivity model with corresponding synthetic traces and their frequency spectra.

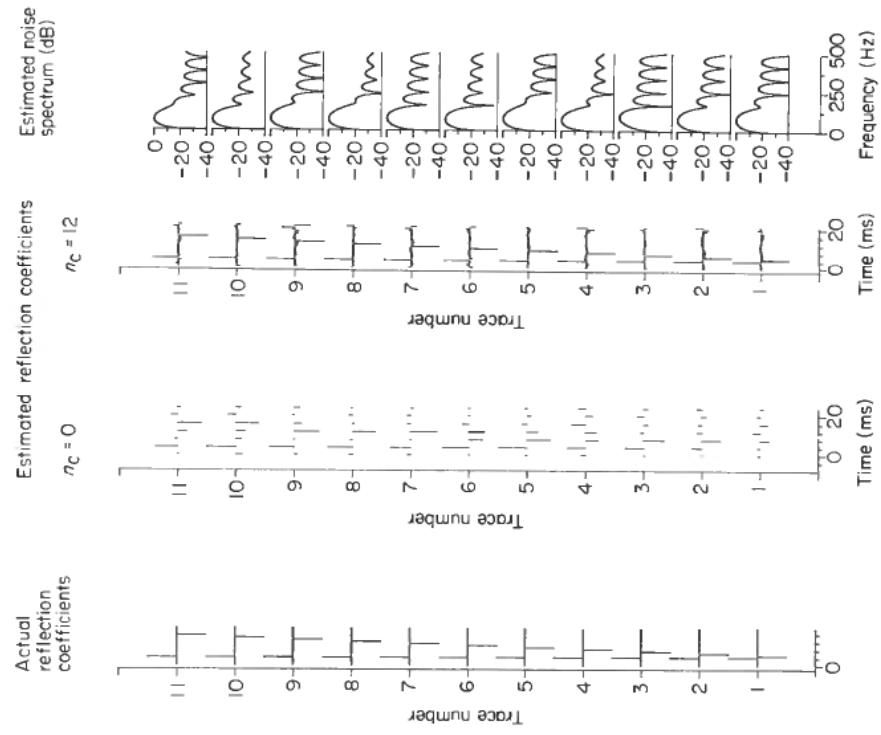

Figure 6. Least-squares and maximum-likelihood estimates obtained from the traces in Fig. 5.

from a signal with 125 Hz bandwidth. The number of estimated parameters is approximately 2·6 times the product of the bandwidth and the time gate. The numerical values of the estimated reflection coefficients do not deviate more than 5% from their theoretical values. It can be seen that the spurious reflection coefficient are correlated. The spurious events generally have small amplitudes, and they are not continuous from trace to trace. This is probably due to the fact that the noise in our model is spatially incoherent. We cannot normally expect the same to hold for real data. By trial and error we found that the ML-estimates were slightly better with increasing order of the noise model and in this case the optimum value of n_C seemed to be about twelve. Further increase in n_C to over twenty produced spurious events with larger amplitudes. The displayed noise spectra are obtained by Fourier-transforming the estimated noise coefficients. These spectra indicate a noise spectrum with bandwidth of approximately 125 Hz. Figures 7 and 8 are equivalent to Figs. 5 and 6, but now the average signal-to-noise ratio is 2 dB. Some of the spurious estimates for $n_C = 12$ are large in amplitude and therefore the reflection coefficient series are scaled individually.

The results shown were obtained using zero starting values for **r** and **c**. If the true values are used as starting points, the algorithm still converges to the displayed values. It has been verified that the minima of the loss function in (5) corresponding to the displayed solutions are smaller than the values of the loss function corresponding to the true parameter values.

We also did experiments with different locations and lengths of the estimation time window. A narrower estimation window containing the interesting events did not produce significantly better results. One property that may be important in the interpretation of the estimated reflection coefficients, is that the spurious events are changing when n_R is changed.

The next example is a simple reservoir model. The reflectivity model and the resulting traces are displayed in Fig. 9. We here simulated two problems that we face in real data applications. The first problem is that we cannot expect the reflectivity function to be zero n_P-samples before and n_P-samples after the estimation time window. This edge effect will probably be most serious for short time gates. The second problem is that we do not know the pulse exactly. Results obtained with the original pulse are displayed in Fig. 10. Inversion is successful only when the time window contains all events. In this case the number of estimated parameters is approximately 4·5 times the product of bandwidth and time gate. Again, the spurious reflection coefficients are correlated. We shall now assume that the reflectivity function for trace number six is known from well logs. This reflection coefficient series together with trace number six is used to provide pulse estimates. In Fig. 11 two estimated pulses are shown. The first pulse is obtained by considering the reflection coefficients in the time interval between 10 and 100 ms. Visual comparison with the original pulse in Fig. 4 shows that the shape of the pulse is recovered. The main difference lies in the tail. The next pulse is obtained by considering only the reflection coefficients in the time interval between 38 and 68 ms. Although we chose the desired pulse length to be only 40 ms in order to reduce the edge effect, this pulse is seriously distorted. The two sets of estimated reflection coefficients shown in Fig. 12 are obtained by using the two estimated pulses. With the exception of a few spurious events for traveltimes around 10 ms, the first set may be compared with the result obtained using the original pulse. The second set shows a result that is very poor even for trace number six. We also did tests with other and nicer looking short pulses, but we were not able to reduce the edge effect in this way.

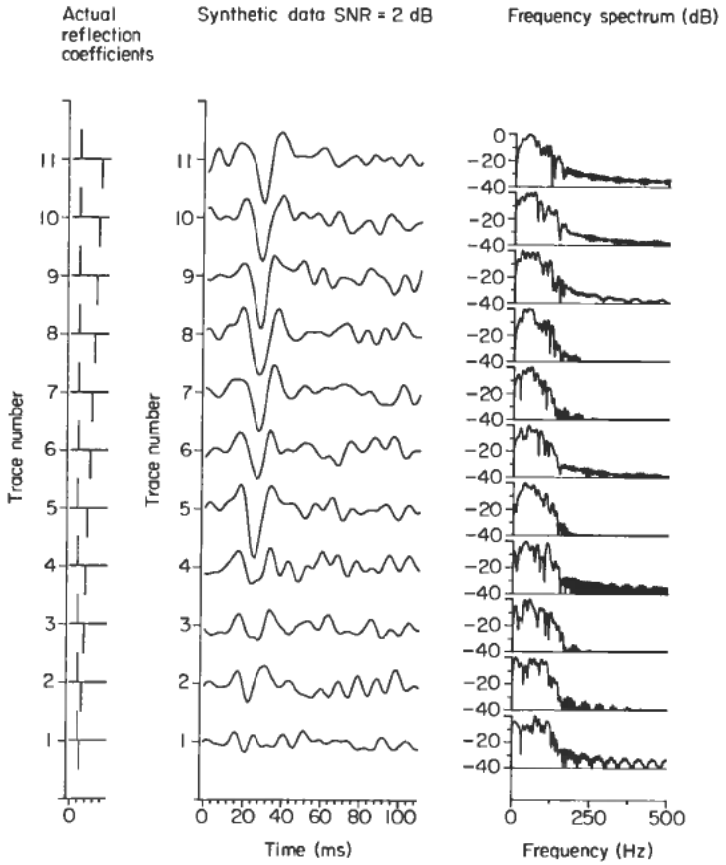

Figure 7. Thin-layer reflectivity model with corresponding synthetic traces and their frequency spectra.

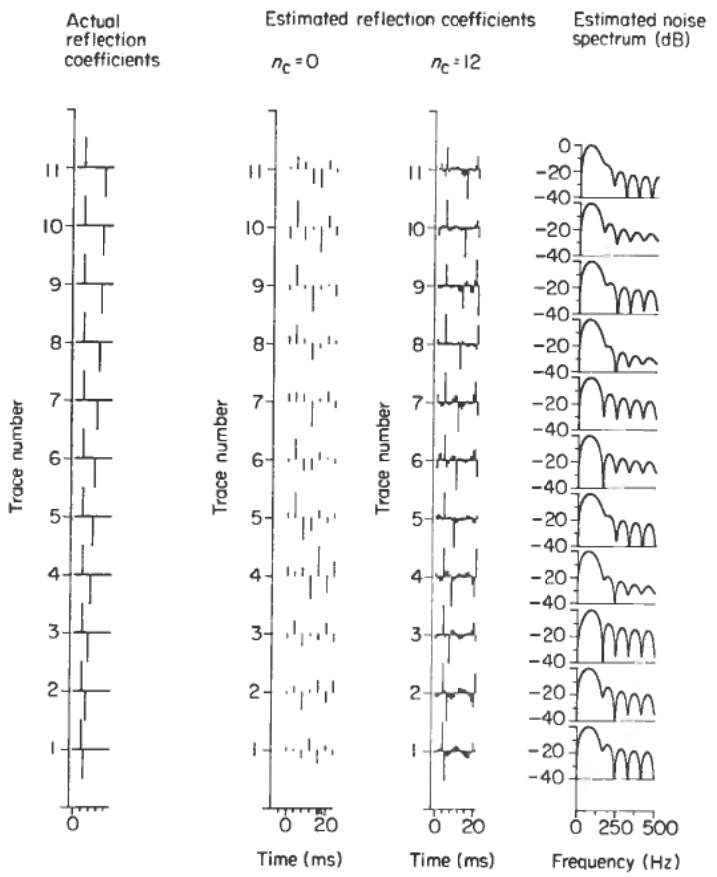

Figure 8. Least-squares and maximum-likelihood estimates obtained from the traces in Fig. 7.

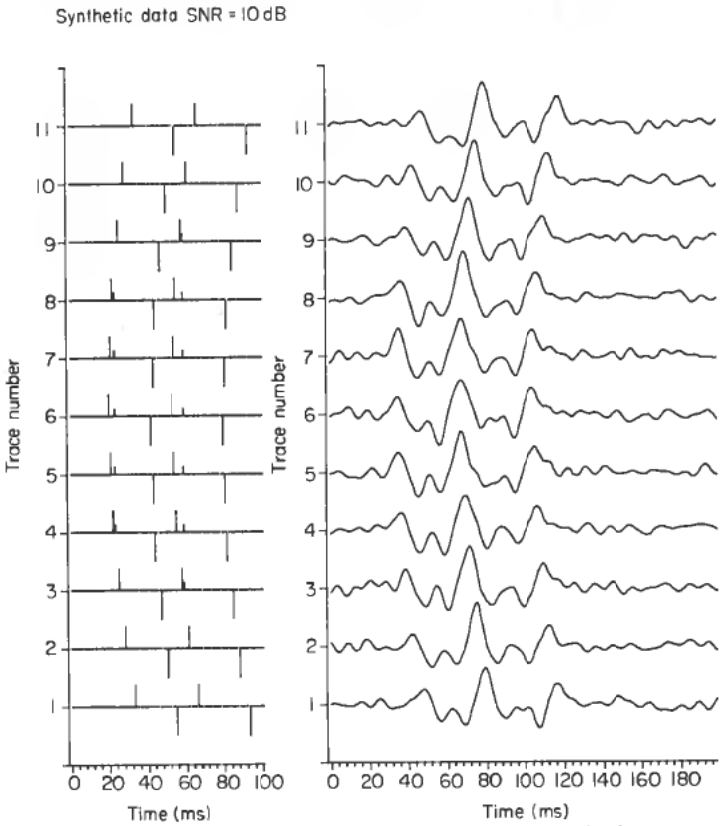

Figure 9. Reflectivity model and corresponding synthetic traces.

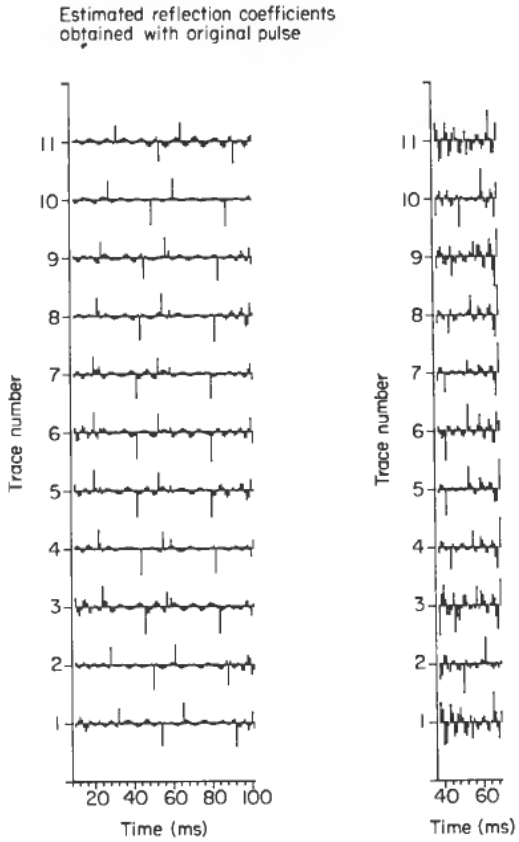

Figure 10. Effect of considering only a limited part of a trace with many events (edge effect). Left: The result of ML-estimation when all events are within the estimation time window. Right: The result of an attempt to resolve only the events within a short time window.

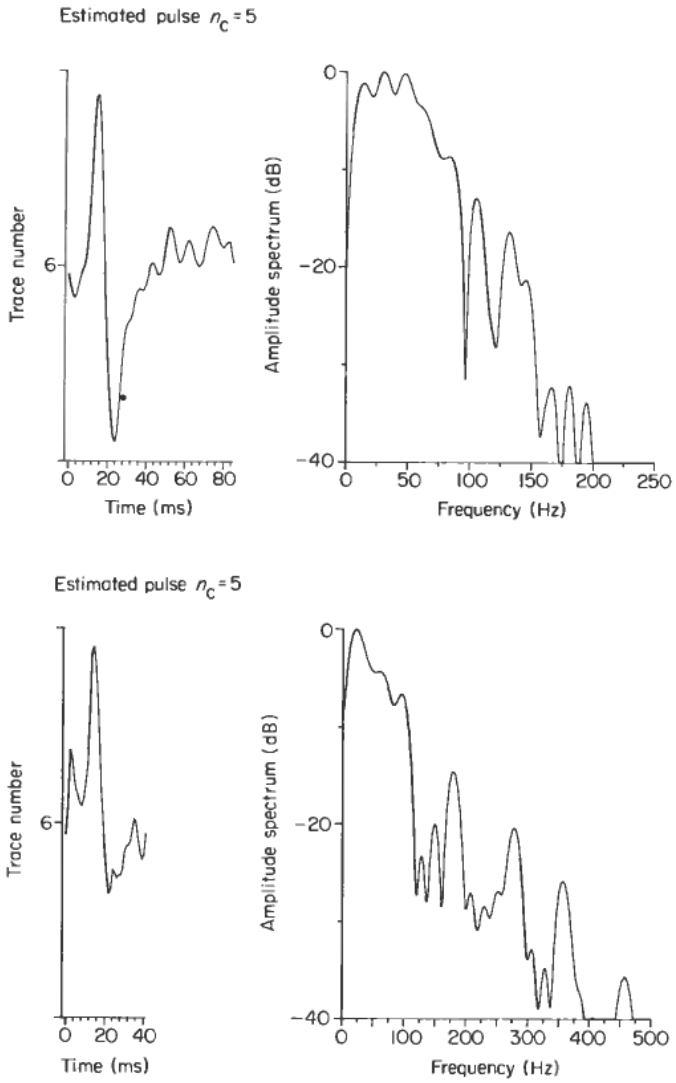

Figure 11. Pulses estimated from trace number six in Fig. 7. Upper: The estimation time window contained all events. Lower: Only the events between 38 and 68 ms were considered.

The simple examples shown above do not complete the investigation of the properties of the ML-method. They are merely intended to indicate how the algorithm can be applied and what kind of problems we must expect in real data applications.

6.2. Real data

Figure 13 shows a part of a stacked unprocessed seismic section. The sampling interval is 4 ms. A reflection coefficient series obtained from well data at CDP 278 is also displayed. The sampling interval of the reflection coefficient series is 2·0 ms. We shall focus our attention on the area between 2082 and 2222 ms two-way traveltime.

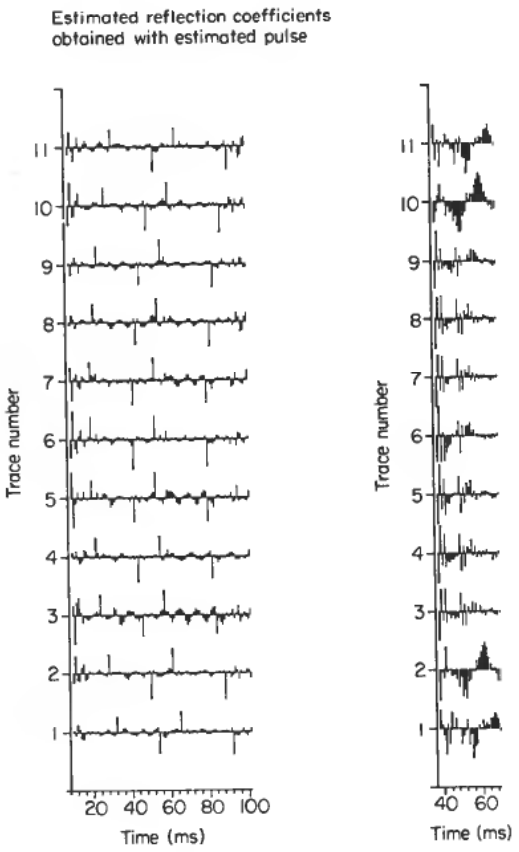

Figure 12. Effect of not knowing the input wavelet exactly. Left: ML-estimates obtained with the upper pulse in Fig. 11. Right: ML-estimates obtained with the lower pulse in Fig. 11.

Reflection coefficients in this interval together with the stacked trace for CDP 278 are used to estimate the pulse displayed in Fig. 14. As can be seen from its tail, the estimated pulse is distorted by multiple reflections and/or edge effects. These effects cannot be expected to be spatially invariant. Hence the estimated pulse will only be suited for inversion of traces close to the well.

The results displayed in Fig. 15 are obtained using the estimated pulse and with zero starting values for **r** and **c**. For least squares estimation, the trace had to be resampled to 8·0 ms sampling interval. The event at approximately 2100 ms two-way traveltime is resolved but the rest of the trace is corrupted by spurious events with alternating polarity. Maximum-likelihood estimation was performed with 4·0 ms sampling interval and n_C equal to five. The result of ML-estimation appears to be a smoothed version of the reflection coefficient series obtained from the well log. The event at approximately 2100 ms two-way time is nicely resolved, but the negative peak at approximately 2150 ms is not properly recovered. The polarity changes from 2150 to 2220 ms vary consistently along the traces.

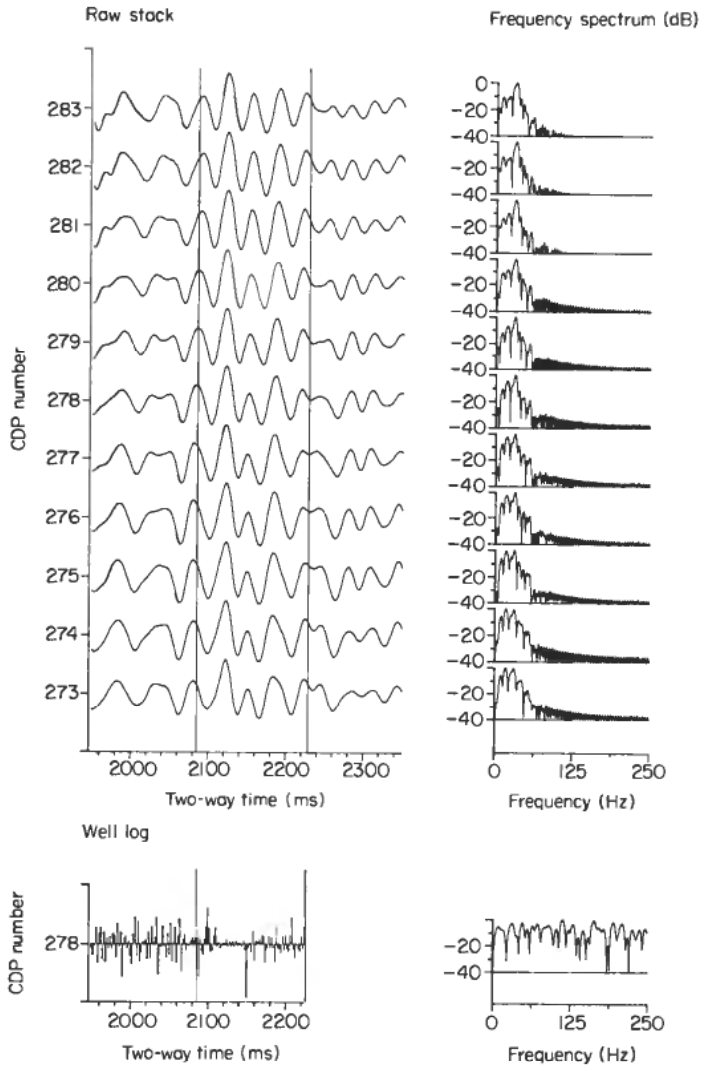

Figure 13. Stacked and unprocessed data with corresponding frequency spectra. Reflectivity function obtained from well logs. The well is located at CDP 278.

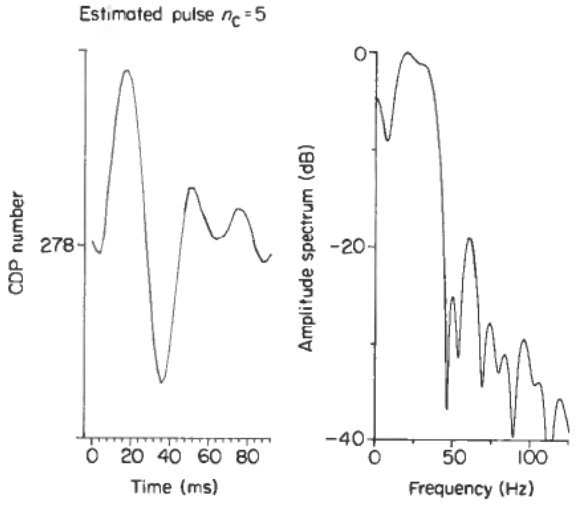

Figure 14. Pulse estimated from the reflectivity function and the stacked trace at CDP 278.

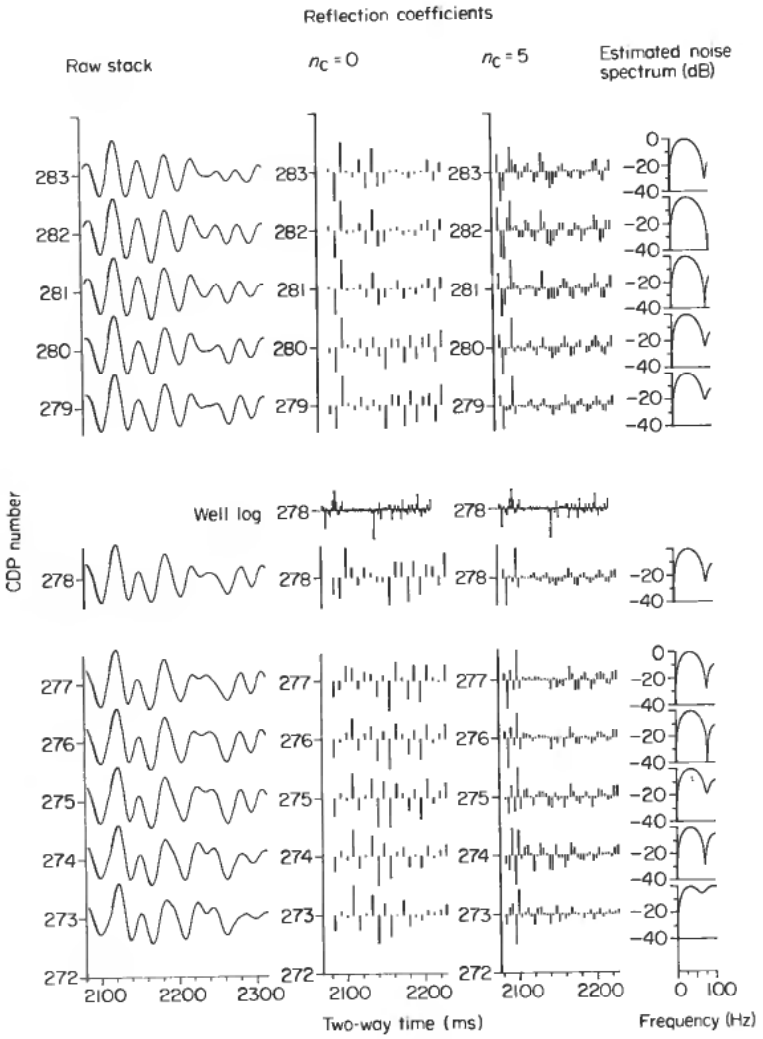

Figure 15. The stacked traces and the estimates obtained from them. The reflectivity function obtained from the well at CDP 278 is included for comparison.

7. Conclusions and discussion

Synthetic and real data examples have demonstrated the effectiveness of the maximum-likelihood method.

When least-squares estimation is applied, the sampling rate must be chosen approximately equal to the Nyquist rate of the trace. Maximum-likelihood estimation yields resolution superior to that provided by least-squares estimation. The ML-method also appears to be numerically more stable.

In real data applications the inversion is distorted by multiple reflections and also by edge effects when only a limited part of the seismic trace is considered.

The synthetic and real data examples indicate where to focus attention as far as possible improvements are considered. The edge effect may be reduced by using different boundary conditions in the model. Because the optimization problem is nonlinear in **c**, faster convergence may be achieved by using a fixed noise filter **c**.

Acknowledgments

We wish to thank S. Haug and Y. Zheng for producing the computer programs required for the synthetic data generation and for plotting, and we wish to thank H. Øzdemir, S. Vaage and E. Berg for valuable discussions related to this work.

This work has received financial support from Saga Petroleum a.s and from Den norske stats oljeselskap a.s.

We wish to thank Saga Petroleum a.s for allowing us to present the data used in this paper.

References

Åstrøm, K. J. (1980), Maximum likelihood and prediction error methods, *Automatica*, **16,** 551–574.

Åstrøm, K. J. and Bohlin, T. (1966). Numerical identification of linear dynamic systems from normal operation records, *in* P. H. Hammond (ed.), *Theory of Self-adaptive Control Systems* (Plenum Press, New York).

Åstrøm, K. J. and Eykhoff, P. (1971), System identification—a survey, *Automatica* **7,** 123–162.

Chi, C. Y., Mendel, J. M. and Hampson, D. (1984). A computationally fast approach to maximum-likelihood deconvolution, *Geophysics*, **49,** 550–565.

Ekstrom, M. P. (1973). A spectral characterization of the ill-conditioning in numerical deconvolution, *IEEE Trans. on Audio and Electroacoustics*, **AU-21,** no. 4, 344–348.

Kallweit, R. S. and Wood, L. C. (1982). The limits of resolution of zero-phase wavelets, *Geophysics*, **47,** 1035–1047.

Goodwin, G. C. and Payne, R. L. (1977). *Dynamic System Identification: Experiment Design and Data Analysis* (Academic Press, London).

Ljung, J. and Søderstrøm, T. (1983). *Theory and Practice of Recursive Identification* (MIT Press, Cambridge).

Mendel, J. M. (1983). *Optimal Seismic Deconvolution: An Estimation-based Approach* (Academic Press, London).

Moore, J. J., Garbow, B. S. and Hillstrøm, K. E. (1978). User guide for MINPACK-1. Argonne National Laboratory ANL-80-74.

Özdemir, H. (1982). Seismic impulse response estimation by the maximum-likelihood and the least-squares methods, SINTEF Report STF28 A82033, Trondheim.

Robinson, E. A. (1978). Iterative identification of noninvertible autoregressive moving average systems with seismic applications. *Geoexploration*, **16,** 1–19.

Robinson, E. A. and Treitel, S. (1980). *Geophysical Signal Analysis* (Prentice-Hall, Englewood Cliffs).

Ursin, B. and Zheng, Y. (1984). Identification of seismic reflections using singular value decomposition, technical report, Petroleum Technology Research Institute, Trondheim.

Van Riel, P. (1982). Seismic trace inversion, MSc Thesis, Delft University of Technology.

Van Riel, P. and Berkhout, A. J. (1983). Resolution in seismic trace inversion by parameter estimation, submitted for publication in Geophysics.

Wolfe, M. A. (1978). *Numerical Methods for Unconstrained Optimization* (Van Nostraind Reinhold, New York).