Geof. Int., Vol. 26-3, 1987, pp. 393 - 406

A NEW GENERALIZED LEAST MEAN-SQUARE ALGORITHM FOR PROCESSING NON-STATIONARY SEISMIC DATA

A. H. COMINGUEZ*
(Recibido: 14 de junio, 1985)
(Aceptado: 2 de julio, 1986)

RESUMEN

Se presenta un algoritmo adaptable apropiado para deconvolver trazas, el cual está basado sobre una expresión generalizada de la técnica de mínimo error cuadrático medio. El uso del nuevo proceso se recomienda especialmente para elaborar sismogramas de reflexión sísmica que contengan reverberaciones variables en el tiempo.

Mediante la aplicación del sistema adaptable los coeficientes del operador se recalculan para cada tiempo de la señal de entrada.

Tanto las características de convergencia del algoritmo como sus propiedades de estabilidad se analizan y comparan con las del algoritmo tradicional LMS. Para tal efecto se presentan ilustraciones con sismogramas sintéticos.

La aplicabilidad del método expuesto parece promisoria para pruebas sísmicas en aguas poco profundas.

ABSTRACT

An adaptive deconvolution algorithm based upon a generalized expression of the Least Mean-Square (LMS) error technique is presented. The use of this process is recommended for reflection seismic data which contain time-varying reverberations.

Filter coefficients are designed for each sample of the input trace using the proposed method.

Convergence characteristics of the new algorithm, and its stability properties, are analyzed and compared to the simple LMS algorithm. Illustrations using synthetic seismic data are presented.

Future possibilities of application to real shallow-water seismic data are found promising.

* *Instituto del Petróleo, Facultad de Ingeniería, Universidad de Buenos Aires.*

INTRODUCTION

In order to meet the needs of data processing in geophysical prospecting, a number of high-speed data methods have been developed in recent years for deconvolving seismic information. The purpose of this paper is to present a generalized expression of the LMS Adaptive Filter (Widrow and Hoff, 1960), which is appropriate for deconvolution of non-periodic multiple waves from shallow water seismograms (marine prospecting).

The application to seismic traces of the simple LMS algorithm was discussed by Griffiths *et al.* (1977).

Early papers discussing expressions of generalized techniques were written by Mantey and Griffiths (1969), and by Mueller (1975).

The New Algorithm requires a greater number of arithmetic operations than the conventional LMS system, but it presents the following advantages:

1) Accurate estimation of time-varying parameters is not required.

2) Instability intervals are not generated.

3) The algorithm shows efficiency in the elimination of some additional noises, for example noises that could be discriminated from the main signal by their frequency content.

The specific application of the New Algorithm in the elimination of multiple signals is achieved by adapting the filter coefficients to minimize the mean-square difference between the filter output and the data values in the seismogram occurring an appropriated distance later in time (Peacock and Treitel, 1969). Thus, the method renews the coefficients of an L-dimension prediction operator as the filter moves along the input trace.

GENERAL ADAPTIVE DECONVOLUTION

Let us assume an N-length input vector $\bar{x}$, and a prediction operator $\bar{c}$ with L samples.

The output y_k at time k is given by

$$y_k = \bar{x}_k^T \bar{C} \tag{1}$$

where

$$\bar{x}_k = \begin{bmatrix} x_k \\ x_{k-1} \\ \cdot \\ \cdot \\ \cdot \\ x_{k-L+1} \end{bmatrix} \text{ and } \bar{C} = \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_L \end{bmatrix}$$

y_k will not, in general, be exactly equal to the desired output d_k, so the error at time k will be

$$e_k = y_k - d_k \ . \tag{2}$$

The mean square error at time k will thus be

$$\epsilon^2 = E \left\{ e_k^2 \right\} . \tag{3}$$

Now $\bar{g}$, the gradient of $1/2\ \epsilon^2$ with respect to $\bar{c}$, can be calculated by combining (1), (2) and (3)

$$\bar{g} = 1/2 \frac{\partial \epsilon^2}{\partial \bar{c}} = E \left\{ e_k \bar{x}_k \right\} = A\bar{c} - \bar{v} \ , \tag{4}$$

where A is the signal autocorrelation matrix,

$$A = E \left\{ \bar{x}_k \bar{x}_k^T \right\} , \tag{5}$$

and $\bar{v}$ is the correlation vector between the input and the desired output,

$$\bar{v} = E \left\{ \bar{x}_k d_k \right\} \tag{6}$$

The best filter $\bar{c}$, in the mean square sense, is such that it minimizes ϵ^2, and is obtained by setting $\bar{g}$ to zero (Lee, 1960),

$$\bar{c}_{opt} = A^{-1} \bar{v} \ . \tag{7}$$

We will use the recursive algorithm proposed by Mueller (1972),

$$\bar{c}_{m+1} = \bar{c}_m - Q_m \bar{g}_m \ , \tag{8}$$

where $\bar{g}_m$ is the gradient vector calculated with the filter-vector after the mth iteration, and Q_m is a nonsingular matrix. The generalized algorithm (8) will stop updating when $c = c_{opt}$. Obviously, if the algorithm converges at all, it will converge to the optimum solution given by (7).

If we consider the simplified case $Q_m = \beta I$, equation (8) will reduce to

$$\bar{c}_{m+1} = \bar{c}_m - \beta \bar{g}_m \quad , \tag{9}$$

this is the expression of the steepest descent gradient algorithm (Widrow and Hoof, 1960).

From (4), expression (8) can be written in the following form

$$\bar{c}_{m+1} = (1 - Q_m A)\,\bar{c}_m + Q_m\,\bar{v} \quad , \tag{10}$$

and introducing the filter error vector

$$\Delta\,\bar{c}_m = \bar{c}_m - \bar{c}_{opt} \quad , \tag{11}$$

it is possible to arrive to the following expression

$$\Delta\,\bar{c}_{m+1} = (I - Q_m A)\,\Delta\bar{c}_m = \left[\prod_{n=1}^{m} (I - Q_n\,A)\right] \Delta\bar{c}_1 \quad . \tag{12}$$

We conclude from equation (12) that the term $(I - Q_n A)$ is a critical factor which controls the convergence characteristics of the algorithm given by (8).

CHOICE OF THE Q MATRIX

If the change in the statistical properties along the input signal is not extreme, the algorithm can be used to update the coefficients of the vector $\bar{c}$ as the filter moves along the trace. We could initialize algorithm (8) by selecting

$$\bar{c}_1 = \hat{A}^{-1}\,\bar{v}' \quad , \tag{13}$$

where $\hat{A}$ = average $\{\bar{x}_k\,\bar{x}_k^T\}$ along the input trace,

and $\bar{v}'$ = average $\{\bar{x}_k\,d_k^T\}$ along the same trace.

Equation (13) is equivalent to the least-squares filter suggested by Treitel and Robinson (1966).

From equation (12) it is observed that the estimate $\hat{A}$ could be used to estimate Q,

$$Q_k = \beta_k\,\hat{A}^{-1} \quad . \tag{14}$$

If we define

$$\Delta_k = (I - \hat{A}^{-1}\,A) \quad , \tag{15}$$

the coefficients a_{ij} of the matrix Δ_k have the following properties

$$E\left\{a_{ij}\right\} = 0, \text{ and } E\left\{a_{ij}^2\right\} \ll 1, \tag{16}$$

because the input signal was described as exhibiting small oscillatory fluctuations around its mean characteristics.

Selecting $\beta_k = 1$ in (12), the filter error will take the form

$$\Delta\bar{c}_{m+1} = \left[\prod_{k=1}^{m}(\Delta_k)\right]\Delta C_1 \tag{17}$$

Therefore, taking $Q = \hat{A}^{-1}$, a rapid convergence of the filter coefficients must be expected.

THE GLMS ALGORITHM

As shown by Peacock and Treitel (1969), the elimination of stationary events from a trace can be achieved by designing a Wiener filter $\bar{c}_{opt}$ appropriate for generating a trace estimate $\hat{x}$, γ samples later,

$$\hat{x}_{k+\gamma} = \bar{x}_k^T\bar{c}_{opt} \quad , \tag{18}$$

and calculating the deconvolved trace z_k as the difference between the input x_k and the predicted value $\hat{x}_k$:

$$z_k = x_k - \hat{x}_k \quad . \tag{19}$$

Using algorithm (8) with $Q_m = \mu A^{-1}$ as proposed in (14), we obtain

$$\bar{c}_{k+1} = \bar{c}_k - \mu\hat{A}^{-1}\bar{g}_k \quad . \tag{20}$$

Widrow and Hoff (1960) suggested that the gradient $\bar{g}_k$ could be estimated by replacing the autocorrelation terms A and $\bar{v}$ by their instantaneous values:

$$A \rightarrow \bar{x}_k\bar{x}_k^T \quad ,$$

$$\bar{v} \rightarrow x_{k+\gamma}\bar{x}_k \quad ,$$

following this idea, the algorithm in equation (20) becomes

$$\bar{c}_{k+1} = \bar{c}_k + \mu\hat{A}^{-1}(x_{k+\gamma} - \hat{x}_{k+\gamma})\bar{x}_k \quad , \tag{21}$$

or

$$\bar{c}_{k+1} = \bar{c}_k + \mu\hat{A}^{-1}z_{k+\gamma}\bar{x}_k \quad , \tag{22}$$

where $\hat{A}^{-1}$ is the inverse autocorrelation matrix of the input trace and z_k is the deconvolved trace at time k.

Equation (22) will be called GLMS algorithm (Generalized expression of LMS algorithm).

SELECTION OF ALGORITHM PARAMETERS

Griffiths *et al.* (1977) presented the application of the LMS algorithm

$$\bar{c}_{k+1} = \bar{c}_k + \mu' z_{k+\gamma} \bar{x}_k \quad , \tag{23}$$

for use in processing reflection seismic data which included multiples with time-varying periods. They selected μ' according to Widrow's suggestion (1976):

$$\mu' = \frac{\alpha}{L \bar{v} \sigma_x^2} \quad , \tag{24}$$

where

$$0 < \alpha < 2 \quad , \tag{25}$$

and $\bar{v}\ \sigma_x^2$ is the average power level of the input trace. L, as defined in (1), is the dimension of the prediction operator $\bar{c}$.

From the above results, it is not difficult to derive the GLMS algorithm parameters, which would be used in expressions (21) and (22). The $\bar{v}\ \sigma_x^2$ parameter must not be included in the expression of μ because it is included in the diagonal terms of the $\hat{A}^{-1}$ matrix. On the other hand, as the gradient estimate $z_{k+\gamma} \bar{x}_k$ is multiplied by the $\hat{A}^{-1}$ matrix, which is an LxL dimension matrix, the new μ value must be proportional to $1/L^2$.

Therefore, for the GLMS algorithm, the following expression will be adequate:

$$\bar{c}_{k+1} = \bar{c}_k + \mu \hat{A}^{-1} z_{k+\gamma} \bar{x}_k \quad , \tag{26}$$

where

$$\mu = \frac{\alpha}{L^2} \quad , \tag{27}$$

and

$$0 < \alpha < 2 \quad . \tag{28}$$

Equations (27) and (28) will be empirically ratified in the applications discussed in the final part of this paper.

NUMERICAL MODEL APPLICATIONS

The adaptive deconvolution procedures discussed in the first part of this paper are designed for use in processing seismic traces having non-stationary statistical behavior

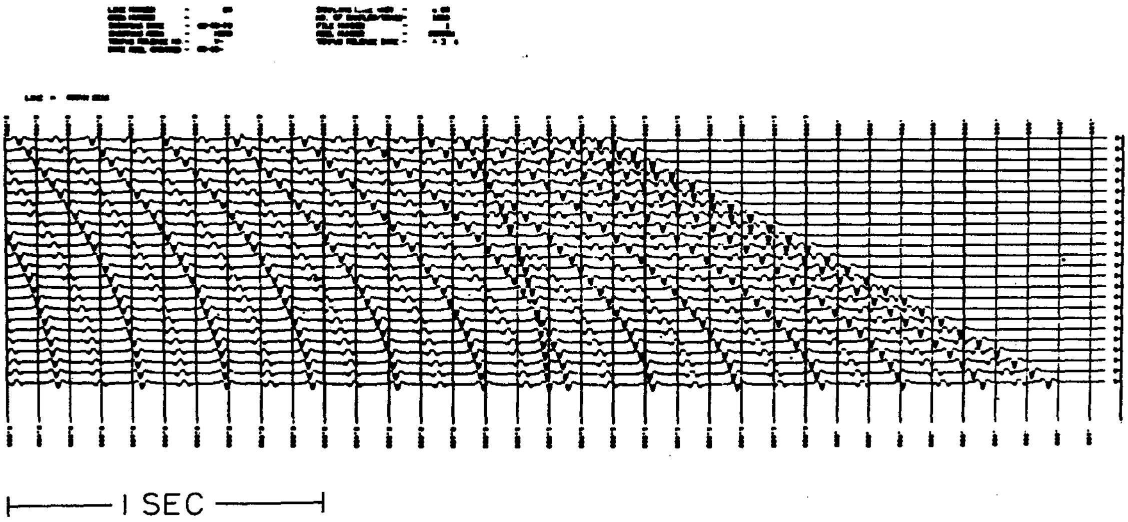

Fig. 1. Unprocessed 24 trace synthetic seismogram showing the water-bottom reflection with its multiples and a primary reflection. Near-offset distance = 290 m, group interval = 91.4 m, water depth = 100 m, water velocity = 1,500 m/sec, stack velocity (primary reflection) = 2,500 m/sec, t_o (primary reflection) = 1,800 sec.

over relatively short-time intervals. A synthetic seismogram (4 msec sample period) representing a marine experiment is presented in Figure 1. The traces contain the water-bottom reflection and its multiples, like a primary reflection. The multiples have constant amplitude, and the nonperiodic spacing intervals represent a critical example of shallow-water record. The distant receiver traces (upper part of Figure 1) show overlap between successive wavelets in the zone near the water-bottom. This detail represents a very difficult example for testing the stability properties of time-varying algorithms.

Figure 2 illustrates the use of LMS adaptive deconvolution (equation 23) on a seismogram generated by filtering the traces of Figure 1 with a 0-40 Hz low-pass filter. The values of the parameters used to deconvolve the seismogram were: filter length = 49 samples (196 msec), prediction coefficient = 23 samples (92 msec), and $\alpha = 0.20$. Experiments using values of α greater than 0.20 presented details of instability* on the output traces. However, a tendency to instability was observed increasing the frequency band of the input synthetic seismogram. In general, the tests demonstrated that the LMS deconvolution required a very accurate selection of the algorithm parameters.

Figure 3 shows the output calculated using the GLMS adaptive deconvolution (name of the new algorithm proposed in this paper) on the multi-channel seismogram presented in Figure 1. The operator was characterized by filter length = 19 samples (76 msec), prediction coefficient = 23 samples (92 msec), and $\alpha = 1.00$. Application of a previous low-pass filtering was not necessary when the GLMS algorithm was used. Inspection of the results shows that events are more readily distinguishable in Figure 3 than in Figure 2. However, good trace deconvolutions were obtained applying the GLMS filter with α values in the range of 0.10 to 1.20, and operator lengths in the range of 9 to 19 samples (36 to 76 msec).

An experiment was made adding 30% white noise to the input seismogram (Fig. 4) and applying GLMS deconvolution. Figure 5 shows the filtered output when the traces were deconvolved by using the same filter as in Figure 3. It is observed a robustness of the output interpretation under additive white noise conditions.

---

* The character of the unstable operator response is typically a rapid oscillating of the polarity of the output trace.

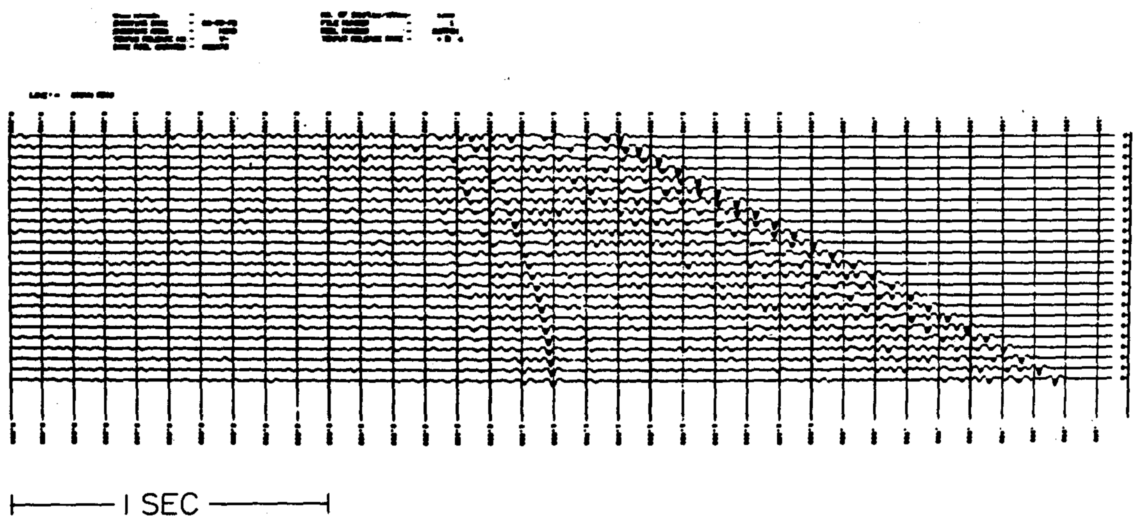

Fig. 2. The effects of LMS adaptive deconvolution on filtered (0-40 Hz) synthetic seismogram presented in Figure 1. The parameters used were: filter length = 49 samples, prediction coefficient = 23 samples, α = 0.20. The adaptive deconvolved seismogram was obtained using time-reverse processing.

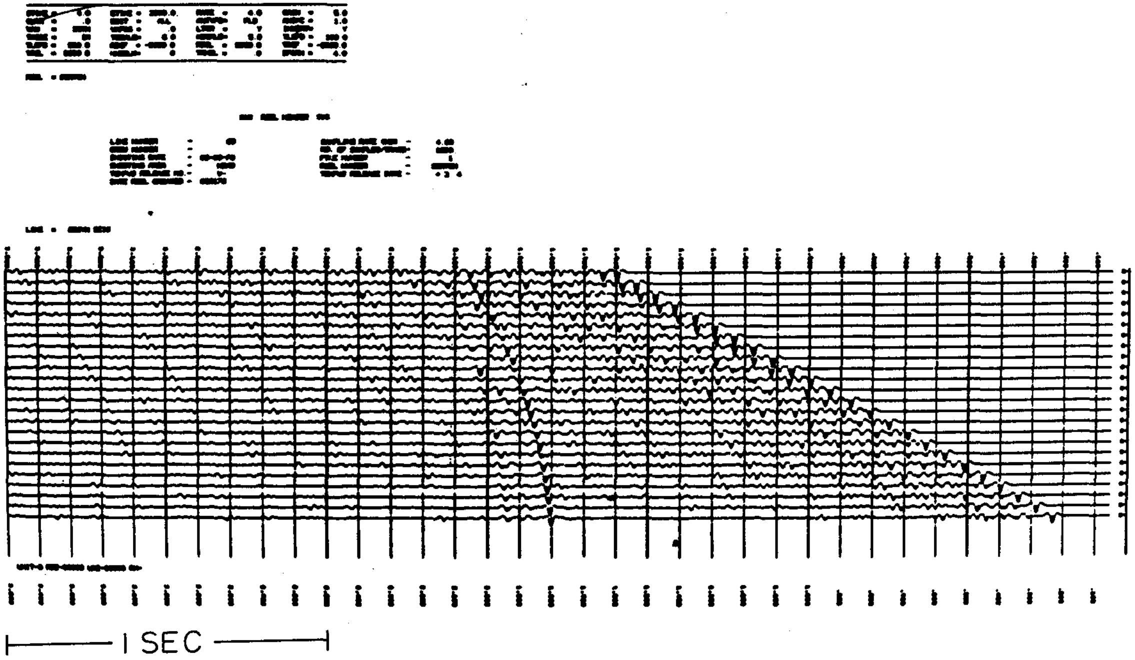

Fig. 3. The effects of GLMS adaptive deconvolution on "unfiltered" synthetic seismogram presented in Figure 1. The parameters used were: filter length = 19 samples, prediction coefficient = 23 samples, α= 1.00. The adaptive deconvolved seismogram was obtained using time reverse-processing.

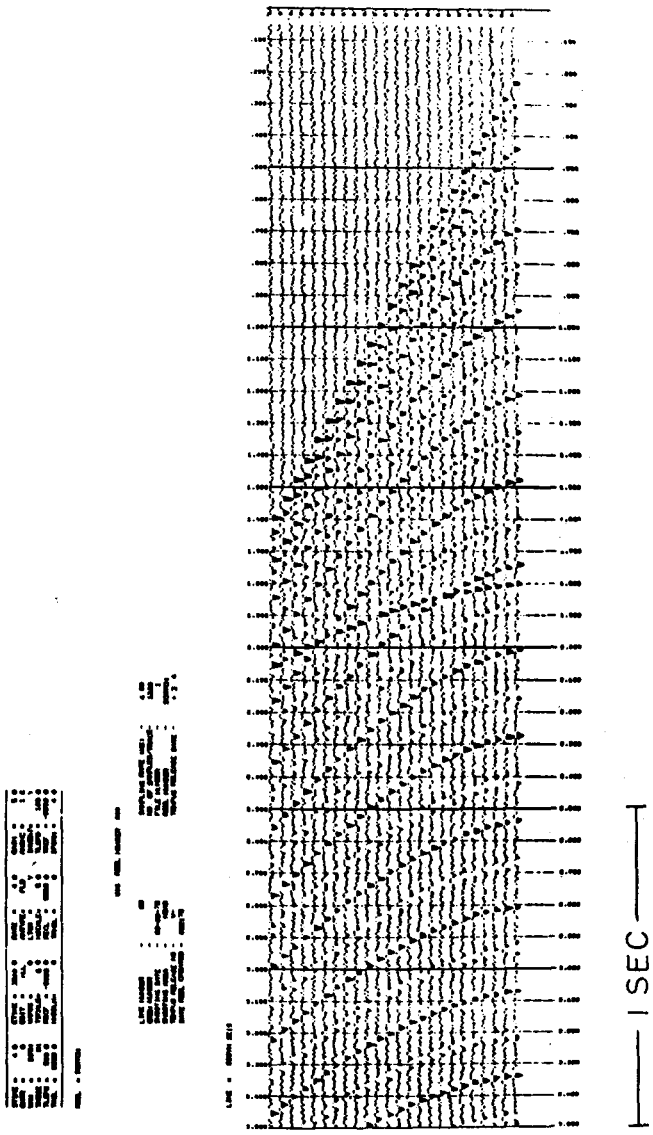

Fig. 4. Multi-Channel synthetic seismogram obtained adding white noise to the traces on Figure 1. The rms signal-to-noise ratio was 3.

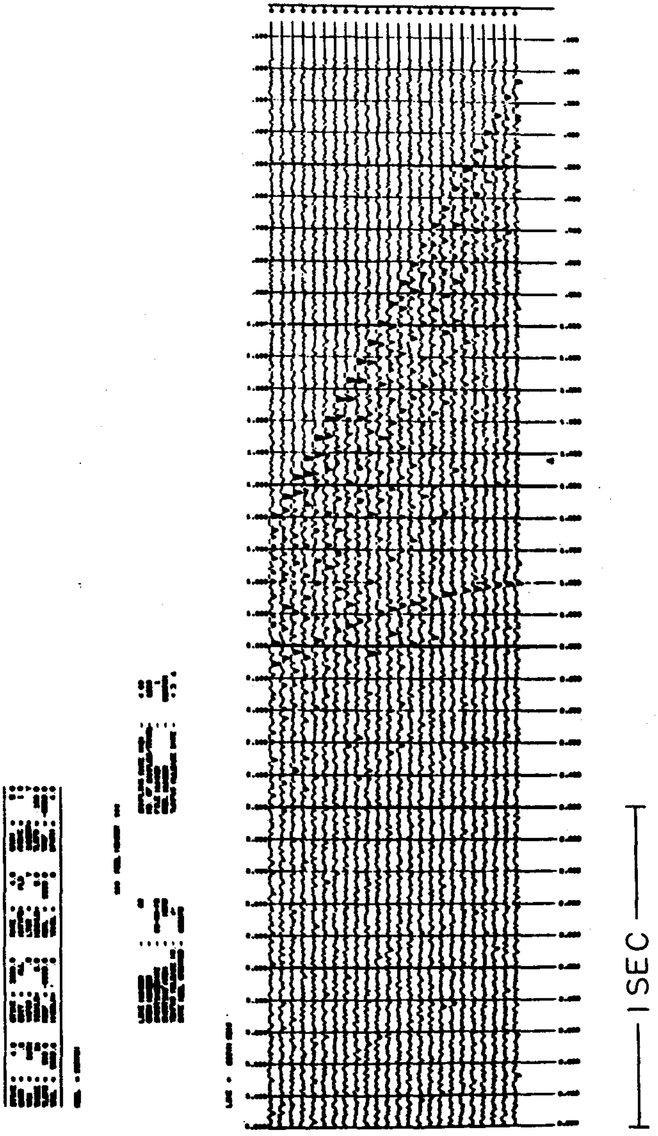

Fig. 5. The effects of GLMS deconvolution on the Multi-Channel synthetic seismogram of Figure 4.

CONCLUSIONS

A comparison between (22) and (23) shows that in the GLMS algorithm all the filter coefficients converge at the same rate, while in the LMS process the coefficients converge with different rates. Widrow *et al.* (1976) estimated that dissimilar convergence of the LMS filter coefficients produced critical results when the eigenvalues of the input correlation matrix were highly disparate ($\lambda_{max} / \lambda_{min} \gg 10$).

Practical applications to synthetic seismograms indicate that the performance of the GLMS algorithm in removing multiple wavelets is not a critical function of the process parameters (previous low-pass filtering is not necessary).

The effectiveness of the GLMS algorithm does not decrease when white noise is added to the input traces.

The good results obtained applying the GLMS algorithm to synthetic data suggest that this method could be used with advantage on real shallow-water seismic data.

BIBLIOGRAPHY

GRIFFITHS, T. L., 1969. A simple adaptive algorithm for real-time processing in antenna arrays. *Proc. IEEE, 57,* 1696-1704.

GRIFFITHS, T. L., F. R. SMOLKA and L. D. TREMBLY, 1977. Adaptive deconvolution: a new technique for time-varying seismic data. *Geophysics, 42,* 742-759.

LEE, Y. W., 1960. Statistical theory of communication. New York, John Wiley and Sons.

MANTEY, P. E. and L. T. GRIFFITHS, 1969. Iterative least-squares algorithm for signal extraction. Proc. Second Hawaii Int. Conf. System Sciences, Western Periodicals Co., 767-770.

MUELLER, K. H., 1975. A new fast-converging mean-square algorithm for adaptive equalizers with partial-response signaling. *Bell. Syst. Tech. J., 54,* 143-153.

PEACOCK, K. L. and S. TREITEL, 1969. Predictive deconvolution. Theory and Practice. *Geophysics, 34,* 155-169.

TREITEL, S. and E. A. ROBINSON, 1966. The design of high-resolution digital filters. *IEEE Trans. on Geosci. Electronics, GE-4,* 25-38.

WIDROW, B. and M. E. HOFF, 1960. Adaptive switching circuits. *WESCON Conv. Rec., p.t.4,* 96-140.

WIDROW, B., T. M. McCOOL, M. G. LARIMORE and C. R. JOHNSON, 1976. Stationary and nonstationary learning characteristics of the LMS adaptive filter. *Proc. IEEE, 64,* 1151-1162.