AD NUMBER
AD842560
LIMITATION CHANGES
TO: **Approved for public release; distribution is unlimited. Document partially illegible.**
FROM: **Distribution authorized to U.S. Gov't. agencies and their contractors; Critical Technology; SEP 1968. Other requests shall be referred to Air Force Technical Application Center, Washington, DC 20333. Document partially illegible. This document contains export-controlled technical data.**
AUTHORITY
usaf ltr, 25 jan 1972

PRINCIPLES OF WIENER AUTO-ADAPTIVE FILTERING

27 September 1968

Prepared For

AIR FORCE TECHNICAL APPLICATIONS CENTER
Washington, D.C.

By

P. R. Lintz
TELEDYNE, INC.

Under

Project VELA UNIFORM

Sponsored By

ADVANCED RESEARCH PROJECTS AGENCY
Nuclear Test Detection Office
ARPA Order No. 624

PRINCIPLES OF WIENER AUTO-ADAPTIVE FILTERING

SEISMIC DATA LABORATORY REPORT NO. 224

AFTAC Project No.:	VELA T/6702
Project Title:	Seismic Data Laboratory
ARPA Order No.:	624
ARPA Program Code No.:	8F10
Name of Contractor:	TELEDYNE INDUSTRIES, INC.
Contract No.:	F 33657-68-C-0945
Date of Contract:	2 March 1968
Amount of Contract:	$ 1,251,000
Contract Expiration Date:	1 March 1969
Project Manager:	Royal A. Hartenberger (703) 836-7647

P. O. Box 334, Alexandria, Virginia

AVAILABILITY

This document is subject to special export controls and each transmittal to foreign governments or foreign nationals may be made only with prior approval of Chief, AFTAC.
attn: VSC, wash. D.C. 20333

This research was supported by the Advanced Research Projects Agency, Nuclear Test Detection Office, under Project VELA-UNIFORM and accomplished under technical direction of the Air Force Technical Applications Center under Contract F 33657-68-C-0945.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

This report discusses various techniques of designing auto-adaptive (self-changing) filters, and the application of one particular technique to two teleseisms. This report also discusses the calculation of running correlation functions, using a moving time window, and iterative solutions to the multichannel normal equations. A listing is included of a program which will calculate an auto-adaptive multichannel filter by the use of running correlation functions and the method of steepest descent.

SECTION I. INTRODUCTION

In the analysis of seismic records, it is apparent to the analyst that there are both long-term and short-term variations in the noise statistics of seismic array data. An example of long-term variation in noise statistics is the seasonal change in the noise power level; an example of short-term variation is the short bursts of obviously coherent energy which appear on most seismic records which the analyst sees. The question arises as to whether significant rejection of time-varying noise can be achieved by multichannel filtering.

A filter whose coefficients depend upon the time-varying statistics of the noise is termed an adaptive filter. A filter which, as the noise statistics change, automatically updates itself without the need of human intervention or supervision is termed an auto-adaptive (self-changing) filter. The auto-adaptive filters described in this report were designed to be linear but time-varying multichannel filters.

SECTION II. THEORY OF MULTICHANNEL ADAPTIVE WIENER FILTERING

1. The Mean Square Error Criterion

The multichannel wave shaping problem as illustrated in Figure 1 may be stated as follows:

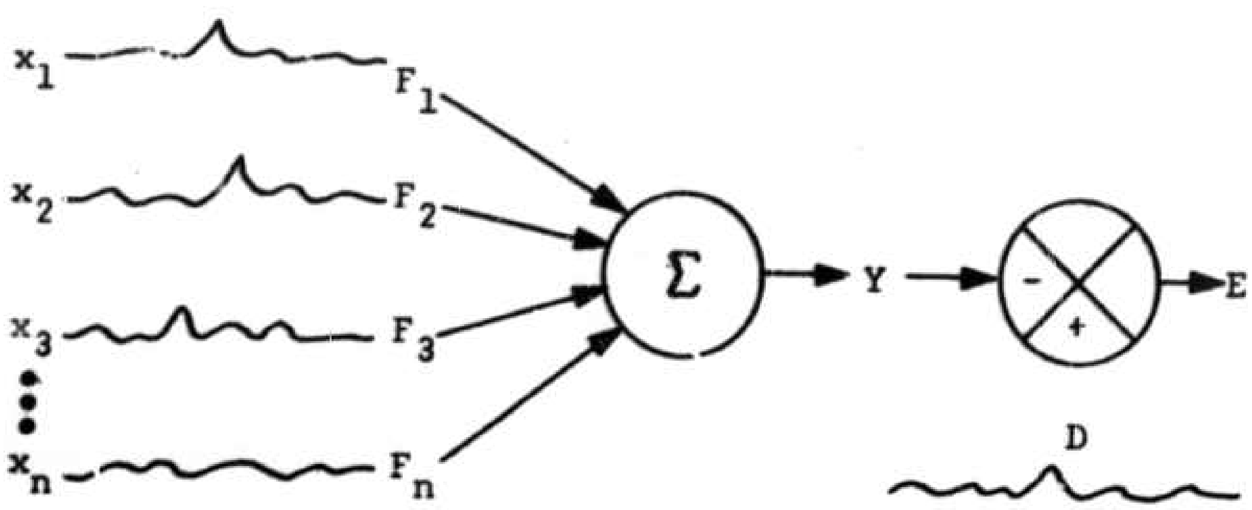


Figure 1. The multichannel wave shaping problem

Given a set of n input time-series data traces x_1 through x_n, and a set of m desired output traces d_1 through d_m, design a set of n filters to transform the n input data series into a set of approximations y_1 through y_m of the m output data series, subject to the constraint that the mean square error $\overline{E^2}$ (the time average of $\sum_{i=1}^{m} (d_i - y_i)^2$) be a minimum.

The problem may be succinctly stated in vector notation as follows.

The multichannel convolution matrix ($\underset{\sim}{X}$) of the set of input traces multipled by the filter vector (F) equals the output vector (Y):

In vector notation

$$\underset{\sim}{X} F = Y \tag{2.1}$$

where $\underset{\sim}{X}$ is the multichannel analog of the single channel rectangular Toeplitz convolution matrix.

The error vector equals the difference between the desired output vector and the actual output vector:

$$\xi = D - Y \tag{2.2}$$

The mean square error is the expected value of $\xi^t\xi$. The t superscript denotes matrix transposition.

$$E\ \{\xi^t\xi\} = MSE = E\ \{D^t - F^t\underset{\sim}{X}^t)\ (D - \underset{\sim}{X}F)\} \tag{2.3}$$

(2.3) can be expanded to give

$$E\ \{\xi^2\} = E\{D^tD\} + E\ \{-F^t\underset{\sim}{X}^tD - D^t\underset{\sim}{X}F\} + E\ \{F^t\underset{\sim}{X}^t\underset{\sim}{X}F\} \tag{2.4}$$

Since the transpose of a scalar is equal to itself,

$$(F^t\underset{\sim}{X}^tD) = D^t\underset{\sim}{X}F \tag{2.5}$$

and (2.4) can be written

$$E\ \{\xi^2\} = E\ \{D^2\} - 2\ E\ \{F^t\underset{\sim}{X}^tD\} + E\ \{F^t\underset{\sim}{X}^t\underset{\sim}{X}F\} \tag{2.6}$$

In terms of $R \equiv (E\ \{\underset{\sim}{X}^t\underset{\sim}{X}\})$, the multichannel correlation matrix of the input data, and $G \equiv (E\ \{\underset{\sim}{X}^tD\})$, the multichannel crosscorrelation vector between the input with the desired output, we have

$$E\{\xi^2\} = \overline{D^2} - 2\,F^tG + F^tRF \qquad 2.7$$

The criterion that the mean square error be a minimum leads to the equation

$$\frac{\partial \overline{\xi^2}}{\partial F^t} = -\,2G + 2RF \qquad 2.8$$

Setting $\frac{\partial \overline{\xi^2}}{\partial F^t} = 0$ gives the multichannel normal equations:

$RF = G$

where the expected value of the gradient of the mean square error is

$$-\frac{\partial \overline{\xi^2}}{\partial F^t} = -2(RF-G) \qquad 2.10$$

For the case N = 2, M = 2, L = 3, the multichannel normal equations take the form

$$\begin{bmatrix} r_{11}(0) & r_{12}(0) & r_{11}(1) & r_{21}(1) & r_{11}(2) & r_{21}(2) \\ r_{21}(0) & r_{22}(0) & r_{12}(1) & r_{22}(1) & r_{12}(2) & r_{22}(2) \\ r_{11}(1) & r_{12}(1) & r_{11}(0) & r_{12}(0) & r_{11}(1) & r_{21}(1) \\ r_{21}(1) & r_{22}(1) & r_{21}(0) & r_{22}(0) & r_{12}(1) & r_{22}(1) \\ r_{11}(2) & r_{12}(2) & r_{11}(1) & r_{12}(1) & r_{11}(0) & r_{12}(0) \\ r_{21}(2) & r_{22}(2) & r_{21}(1) & r_{22}(1) & r_{21}(0) & r_{22}(0) \end{bmatrix} \qquad 2.11$$

$$\times \begin{bmatrix} f_{11}(0) & f_{12}(0) \\ f_{21}(0) & f_{22}(0) \\ f_{11}(1) & f_{12}(1) \\ f_{21}(1) & f_{22}(2) \\ f_{11}(2) & f_{12}(2) \\ f_{21}(2) & f_{22}(2) \end{bmatrix} = \begin{bmatrix} g_{11}(0) & g_{12}(0) \\ g_{21}(0) & g_{22}(0) \\ g_{11}(1) & g_{12}(1) \\ g_{21}(1) & g_{22}(1) \\ g_{11}(2) & g_{12}(2) \\ g_{21}(2) & g_{22}(2) \end{bmatrix}$$

where

N = number of input channels

M = number of output channels

L = number of filter points (number of lags in the correlation functions)

It should be noted that the multichannel correlation matrix, R, in Equation (2.11) is a symmetric matrix, and moreover that writing it in terms of submatrices,

$$\begin{bmatrix} r_0 & r_1^t & r_2^t \\ r_1 & r_0 & r_1^t \\ r_2 & r_1 & r_0 \end{bmatrix} \times \begin{bmatrix} f_0 \\ f_1 \\ f_2 \end{bmatrix} :: \begin{bmatrix} g_0 \\ g_1 \\ g_2 \end{bmatrix} \qquad 2.12$$

we see that it is a block Toeplitz matrix whose cross-diagonal terms are transposes (i.e., $r_{ji}^t = r_{ij}$). This characteristic of the R matrix leads to some simplification in the solution of the equations.

2. The Principle of Orthogonality

The principle of orthogonality (Papoulis, 1965) states that for the optimum multichannel filter, the expected value of the convolution of the error trace and the input data channels is zero (for X or E having zero mean), i.e.,

$$E\,\{\underset{\sim}{X}^t \xi\} = 0 \qquad 2.13$$

We have shown (Equation(2.9)) that the optimum wave-shaping filter satisfies the equation

$$(RF - G) = 0 \qquad 2.14$$

Equation (2.14) can be factored:

$$E\,\{\underset{\sim}{X}^t\,XF - \underset{\sim}{X}^t D\} = 0 \qquad 2.15$$

or

$$E\,\{\underset{\sim}{X}^t\,(D - \underset{\sim}{X}F)\} = 0 \qquad 2.16$$

or

$$E\,\{\underset{\sim}{X}^t(\xi)\} = 0 \qquad 2.17$$

hence

$$E\,\{\underset{\sim}{X}^t\,\xi\} = G-RF \qquad 2.18$$

Equation (2.18) and (2.14) are thus equivalent. Therefore, the gradient of the mean square error, $-2(RF-G)$, is also given by $2E\,\{\underset{\sim}{X}^t\xi\}$. Widrow (1966) makes use of this equivalence to design adaptive filters. As a function of the iteration number n, the "method of steepest descent," or point-slope iterative procedure for the solution of the Wiener-Hopf equation in the time domain, may be written

$$F^{n+1} = F^n - k\,E\,\{\underset{\sim}{X}^t\xi\} \qquad 2.19$$

or

$$F^{n+1} = F^n - k(RF^n - G) \qquad 2.20$$

where k is the magnitude of the step taken in the direction opposite to the gradient of the mean square error.

Widrow does, however, make one further assumption, namely that the expected value of the convolution of the input data channels with the error trace is equal to the product of the input with the error trace at any one instant of time. This assumption leads to a very noisy, but unbiased, estimate of the gradient of the mean square error.

3. Adaptive Filtering Solutions

The solution to the multichannel Wiener-Hopf equation RF = G may be obtained by one of three methods:

1) direct inversion of the R matrix;
2) application of the Levinson-Robinson recursion formula (Levinson, 1949; Robinson, 1967);
3) application of an iterative technique.

The method of steepest descent (Equations(2.19) and (2.20) as well as other iterative solutions, depends on the fact that the mean-square error function is a quadratic function (Equation 2.6) of the filter weights. The mean-square error function has a unique minimum. Figure 2 illustrates the quadratic mean-square error surface as a function of a two point filter. The projection of lines of constant mean-square error on the MSE surface are conic sections in the filter space. The gradient of the scalar function in filter space is $-\frac{\partial \xi^2}{\partial F^t}$. As the noise statistics change, one can picture the mean-square-error surface translating, rotating, changing its ellipticity, opening and closing.

The taking of a step in the opposite direction of the gradient in filter space is equivalent to using Equation (2.19) or (2.20).

There are three other iterative solutions (Varga, 1965) under investigation at the Seismic Data Laboratory,

a) Iterative solution No. 1 is derived as follows:

$$RF = G \qquad 2.21$$

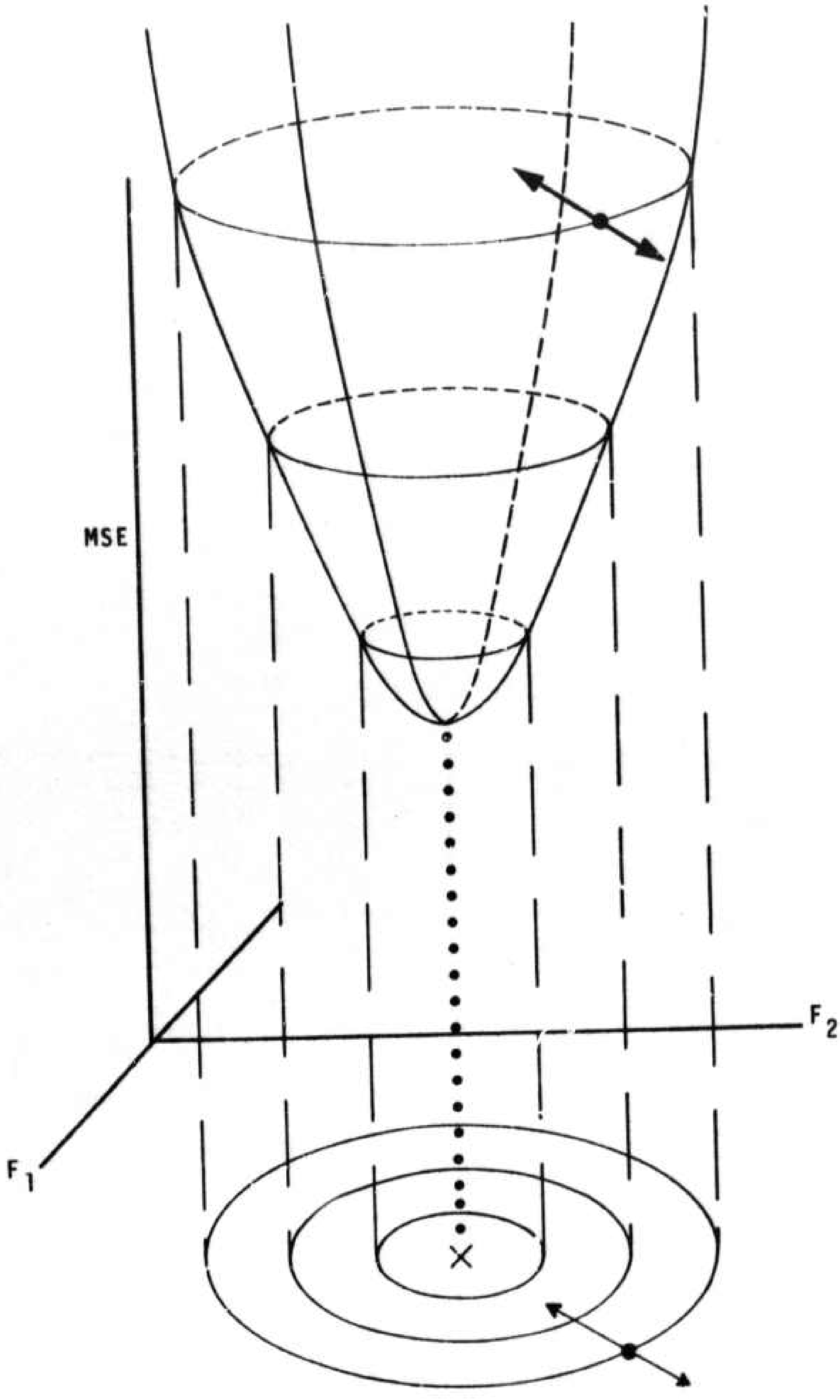


Figure 2. The mean-square-error surface as a function of a two point filter

$$(R+CI-CI)F = G \tag{2.22}$$

$$CIF = G-(R-CI)F \tag{2.23}$$

$$F = C^{-1}(G-RF) + F \tag{2.24}$$

$$F^{m+1} = C^{-1}(G-RF^{m}) + F^{m} \tag{2.25}$$

$$F^{m+1} = F^{m} + C^{-1}(G-RF^{m}) \tag{2.26}$$

where C is arbitrary. However, C should probably be r_0 in order that the filter should not be changed by values greater than unity.

It should be pointed out that this iterative solution is equivalent to Equation (2.20) since RF-G is the gradient of the mean square error.

b) Iterative solution No. 2, is a point-Jacobi technique in which R is split into a diagonal matrix, an upper triangular matrix, and a lower triangular matrix:

$$R = (r_0 I + U + L) \tag{2.27}$$

$$RF = G \tag{2.28}$$

$$(r_0 I + U + L)F = G \tag{2.29}$$

$$r_0 F = G - (U+L)F \tag{2.30}$$

$$F^{m+1} = r_0^{-1}\{G - (U+L)F^{m}\} \tag{2.31}$$

This method is quite simple to implement since r_0^{-1} can be easily stored in the computer.

c) Iterative solution No. 3 is the Gauss-Seidel method:

$$R = (r_0 I + U + L) \qquad 2.32$$

$$RF = G \qquad 2.33$$

$$(r_0 I + U)F = G - LF \qquad 2.34$$

$$F^{n+1} = (r_0 I + U)^{-1} (G - LF^n) \qquad 2.35$$

The inverse of an upper or lower Toeplitz triangular matrix was shown by Meserve (1968) to be an easily obtainable upper or lower Toeplitz triangular matrix. The derivation for a 4 x 4 case is:

$$\begin{bmatrix} r_0 & r_1^t & r_2^t & r_3^t \\ 0 & r_0 & r_1^t & r_2^t \\ 0 & 0 & r_0 & r_1^t \\ 0 & 0 & 0 & r_0 \end{bmatrix} \times \begin{bmatrix} x_0 & x_1 & x_2 & x_3 \\ 0 & x_0 & x_1 & x_2 \\ 0 & 0 & x_0 & x_1 \\ 0 & 0 & 0 & x_0 \end{bmatrix} = \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{bmatrix} \qquad 2.36$$

then

$$r_0 x_0 = I, \; x_0 = r_0^{-1} \qquad 2.37$$

$$r_1^t x_0 + r_0 x_1 = 0, \; x_1 = -r_0^{-1} r_1^t x_0 = -x_0 r_1^t x_0 \qquad 2.38$$

$$r_0 x_2 + r_1^t x_1 + r_2^t x_0 = 0, \; x_2 = + x_0 r_1^t x_0 r_1^t x_0 - x_0 r_2^t x_0 \text{ etc.} \qquad 2.39$$

Thus we have a simple recursive scheme for finding the inverse of any upper (or lower) Toeplitz triangular matrix.

We constructed a test case for the Gauss-Seidel method where the number of input channels was 2, the number of lags was 3 and the number of outputs was 2. The equation to be solved is

$$\begin{bmatrix} r_0 & r_1^t & r_2^t \\ r_1 & r_0 & r_1^t \\ r_2 & r_1 & r_0 \end{bmatrix} \times \begin{bmatrix} f_0 \\ f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} g_0 \\ g_1 \\ g_2 \end{bmatrix} \qquad 2.40$$

where

$$\begin{bmatrix} r_0 \\ --- \\ r_1 \\ --- \\ r_2 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 2 & 7 \\ \hline 1 & 3 \\ 3 & 4 \\ \hline 1 & 3 \\ 2 & 2 \end{bmatrix} \qquad 2.41$$

and

$$\begin{bmatrix} f_0 \\ \\ \\ f_1 \\ \\ \\ f_2 \end{bmatrix} = \begin{bmatrix} f_{11}(0) & f_{12}(0) \\ f_{21}(0) & f_{22}(0) \\ \\ f_{11}(1) & f_{12}(1) \\ f_{21}(1) & f_{22}(1) \\ \\ f_{11}(2) & f_{12}(2) \\ f_{21}(2) & f_{22}(2) \end{bmatrix} \qquad 2.42$$

and

$$\begin{bmatrix} g_0 \\ -- \\ g_1 \\ -- \\ g_2 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 3 & 4 \\ \hline 1 & 2 \\ 3 & 2 \\ \hline 1 & 0 \\ 1 & 2 \end{bmatrix} \qquad 2.43$$

The initial guess for the filters was

$$\begin{bmatrix} f_0 \\ \\ f_1 \\ \\ f_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad 2.44$$

The results for this test case are shown in Figure 3. The plotted points are the filter points corresponding the second column of Equation (2.42).

The Gauss-Seidel method is guaranteed to converge for diagonally dominant matrices (Varga, 1965). By diagonally dominant, we mean that

$$r_{ij}(k) < [r_{ii}(0)\ r_{jj}(0)]^{1/2};\ i \neq j \text{ for } k = 0.$$

The stipulation that a matrix be diagonally dominant is no real restriction in the analysis of time-series data since diagonal dominance implies that no input channel can be more highly correlated with another channel than with itself.

In order to use any of the three iterative solutions (which depend upon knowing R and G) in an adaptive filter program it is necessary to have some means of updating the multichannel correlation matricies with time. A method of updating the R matrix and the G matrix with time is presented below.

4. Running Correlation Functions

It is possible to update the correlation matrix of a multichannel time series with the arrival of each new data point. Given a multichannel time series $X_i(t)$, i = 1, 2, ...,n.

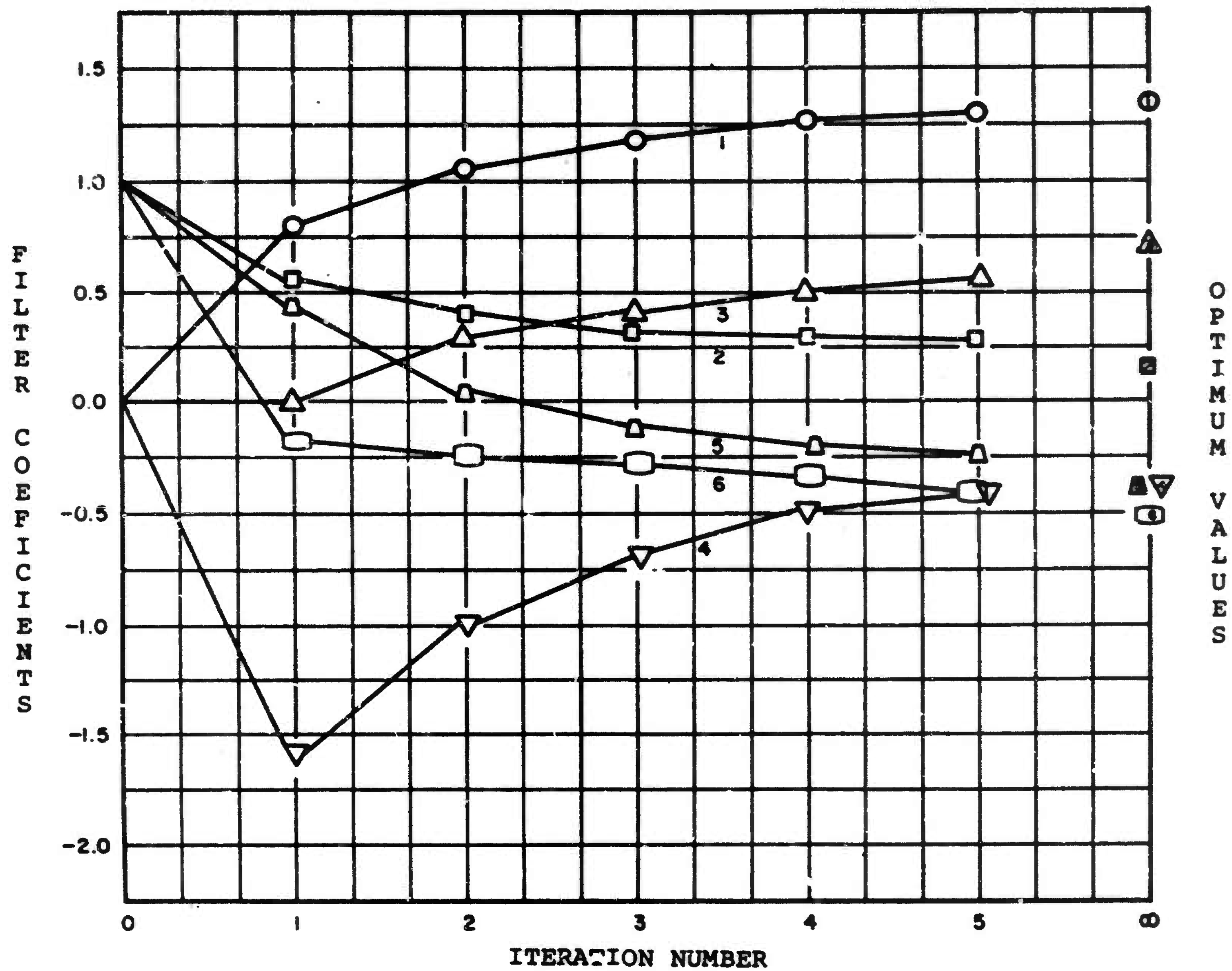
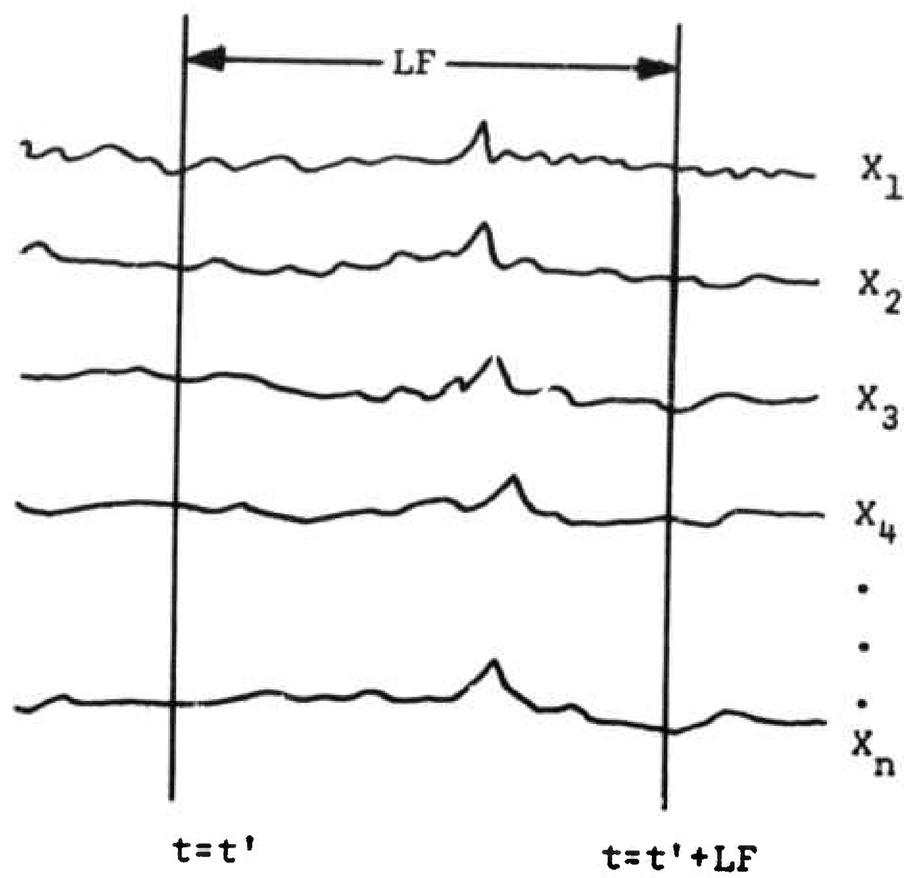


Figure 3. Illustrating the convergence of the iterative Wiener filter to the optimum filter as a function of the iteration number



Consider the possibility of taking all possible correlations of all the data points in the time window of length LF (LF = "length of fitting interval").

$$r_{ij}^{t'+LF}(k) = \sum_{s=t'}^{t'+LF} X_i(s)X_j(s+k) \qquad 2.45$$

Consider what happens to the correlation matrix when the time window is advanced one point.

$$r_{ij}^{t'+LF+1}(k) = \sum_{s=t'+1}^{t'+LF+1} X_i(s)X_j(s+k) \qquad 2.46$$

$$r_{ij}^{t'+LF+1}(k) = \sum_{s=t'}^{t'+LF} X_i(s)X_j(s+k) - X_i(t')X_j(t'+k) + X_i(t'+LF+1)X_j(t'+LF+1+k) \qquad 2.47$$

$$r_{ij}^{t'+LF+1}(k) = r_{ij}^{t'+LF}(k) - X_i(t')X_j(t'+k) + X_i(t'+LF+1-k)\cdot X_j(t=+LF+1) \qquad 2.48$$

Equation (2.48) states that the ij'th element of r(k) after a new data point arrives is equal to the previously calculated element minus one old term plus one new term.

Note that in the last term of the preceding expression, the lag parameter k was shifted to X_i from X_j and the sign of k was changed. This was done since data points ahead of t = t'+LF+1 are as yet unknown.

In the solution of the Wiener-Hopf equation RF = G, R is the known correlation matrix, F is the unknown filter vector and G is the crosscorrelation vector of the desired output with the input channels. We can update an element of G as easily as an element of R by simply replacing X_i in the Equations (1.1) - (1.4) by d_i and replacing r by g:

Hence

$$g_{ij}^{t'+LF+1}(k) = g_{ij}^{t'+LF}(k) - d_i(t')\, X_j(t'+k) + d_i(t'+LF+1-k)\cdot X_j(t'+LF+1) \qquad 2.49$$

In summary, there appear to be two basic methods of linear multichannel time-varying filtering.

Method 1 is to:

a) solve the Wiener-Hopf equation <u>once</u> to get the exact solution for F (or alternatively start with an arbitrary guess);

b) update R and G by the method in Equations (2.48) and (2.49);

c) periodically perform one of the iterative techniques to update the filter; or else use some parameter such as the variation of the mean square error to determine the necessity to update the filter.

Method 2 depends upon the estimation of the gradient of the mean square error:

a) start with an arbitrary set of filter coefficients

b) estimate RF - G, or equivalently, estimate $E\{\underset{\sim}{X}^t \xi\}$

c) then $F^{m+1} = F^m - k\,(RF^m - G)$ or $F^{m+1} = F^m - k\,E\{\underset{\sim}{X}^t \xi\}$

The technique used by Widrow (1966) and Burg <u>et</u> <u>al</u> (1967) is to estimate the expected value $E\{\underset{\sim}{X}^t \xi\}$ by $X\xi$. The technique used in this report estimates $E\{\underset{\sim}{X}\xi\}$ by use of a running cross-correlation vector between the input and the error trace, updated with the arrival of each new data point by the method of Equation (2.49).

SECTION III. EXPERIMENTAL RESULTS

In order to use the steepest descent algorithm Equation (2.19), it is first necessary to get some desired output trace from the set of input data channels. An extremely useful multichannel filter is the maximum-likelihood filter (Kelly and Levin, 1964) which attempts to pass signals which are the same on every channel and to reject signals (or noise) which are not the same on every channel (For a derivation of the maximum likelihood filter see McCowan, 1968). The maximum-likelihood filter may be cast into terms of a prediction-error filter (Burg _et al_, 1967). For the sake of clarity and coherence, the algorithm designed by Burg _et al_ (1967) is presented here.

The maximum likelihood filter is a multichannel filter that is designed subject to the constraint:

$$\sum_{i=1}^{N} f_{ij} = \delta_{js} \tag{3.1}$$

or

$$f_{1j} = \delta_{js} - \sum_{i=2}^{N} f_{ij} \tag{3.2}$$

Multichannel convolution of the filter with the inputs gives the output:

$$Y_{j-s} = \sum_{i=1}^{N} \sum_{k=1}^{L} f_{ik} X_{ij-k} \tag{3.3}$$

or

$$Y_{j-s} = \sum_{k}^{L} f_{1k} X_{1j-k} + \sum_{k=1}^{L} \sum_{i=2}^{N} f_{ik} X_{ij-k} \tag{3.4}$$

Substituting (3.2) into (3.4) gives

$$Y_{j-s} = \sum_{k=1}^{L} \left\{ \delta_{js} - \sum_{i=2}^{N} f_{ik} \right\} X_{1j-k} + \sum_{k=1}^{L} \sum_{i=2}^{N} f_{ik} X_{ij-k} \tag{3.5}$$

$$Y_{j-s} = X_{1j-s} - \sum_{k=1}^{L} \sum_{i=2}^{N} f_{ik} (X_{1j-k} - X_{ij-k}) \quad 3.6$$

Equation 3.6 is seen to be the error remaining in the attempt to predict the (j-s)'th value of X_1 from the differences between X_1 and each of the rest of the channels. The filters designed in this section have made use of this algorithm. The cross-correlation between the input and the error trace was calculated and updated with the arrival of each new data point by the method of Equation (2.49). The new filter for each new data point was calculated from Equation (2.19). The original guess for the optimum filter was, in both cases, 1/N, 0 , 0, 0... for each of the N channels used to predict the (N + 1)'st channel. The data was detrended, demagnified, time-aligned and bandpassed filtered by the "SDL Bandpass Filter" (Figure 4). The length of the adaptive filter was 30 points (1.5 sec). The best results were obtained by shifting the trace to be predicted 20 points (1 second ahead in time). This time shift corresponds to designing a filter to predict 20 points ahead. The filters were one-sided filters which operate only on past and present data. The length of the fitting interval was 320 points (16 seconds). The auto-adaptive filters in this report were capable of operating in one of three modes, namely:

Mode 1: Operate (filter the data), learn (update the filter with each new data point), and update (update the gradient of the mean square error with each new data point).

Mode 2: Operate and update only.

Mode 3: Operate only.

The data were stored in core in a block-multiplexed form (512 points from each channel). Thus there were 192 points outside the fitting interval which could be operated on without reading new data points from the disc. The filters were designed

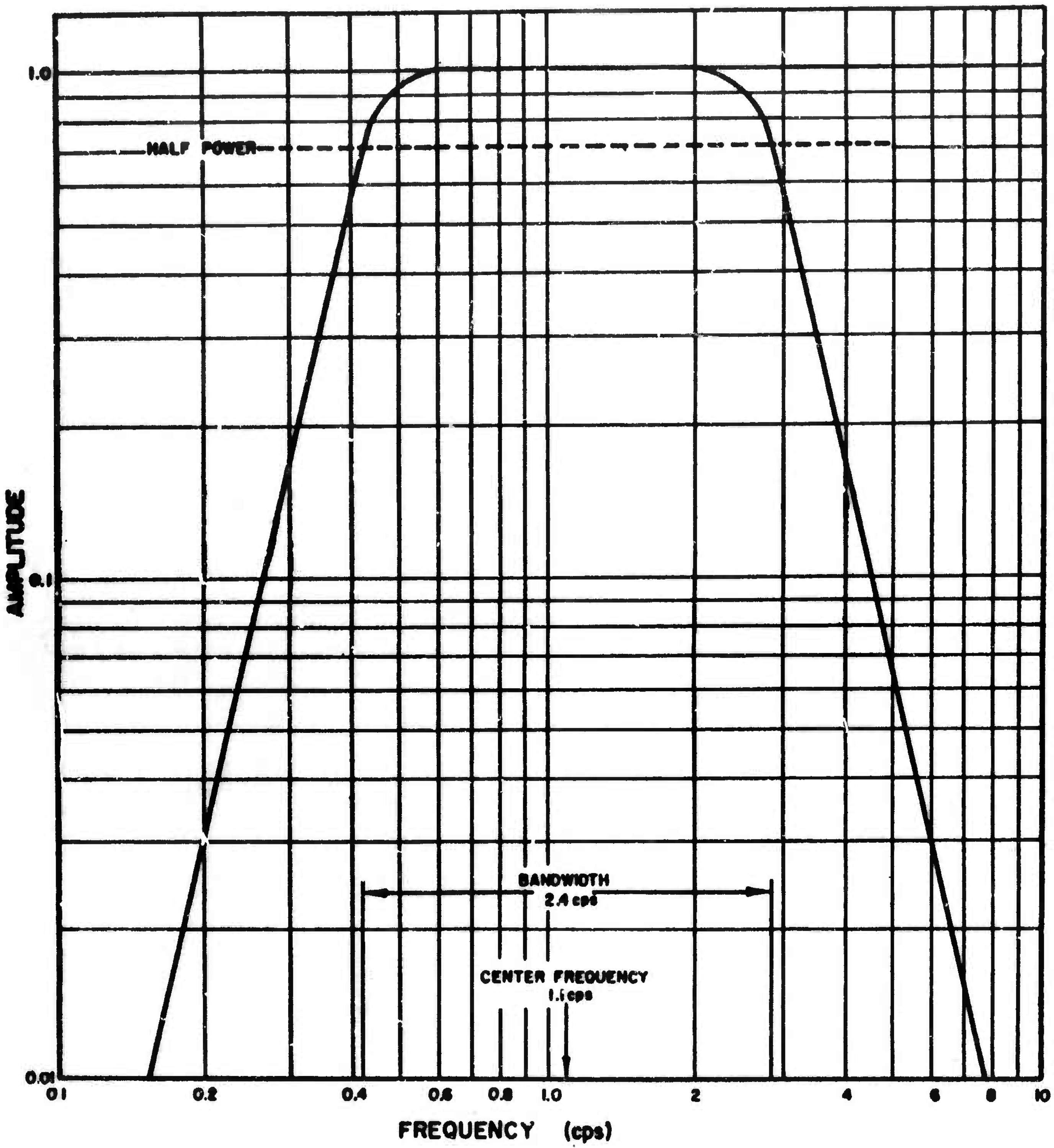


Figure 4. The "SDL" Bandpass Filter

to operate in Mode 1 for 100 points and in Mode 2 for 92 points, alternatively. After the signal arrived the filters were constrained to operate in Mode 2 for all 192 points. In an online situation it would be necessary to have in conjunction with the adaptive filter a signal detector system which, after a signal was detected, would cause the adaptive filter to operate only in Mode 2 until a preset time after the signal arrival. However, even if a signal is not detected, the gradient of the mean square error will not change drastically until the signal is well within the fitting interval, since the scale factor k in Equation (2.20) was made to depend upon the zero lag of the autocorrelation of each trace. The zero lag of the autocorrelation of each trace was updated with the arrival of each new data point in a manner similar to that of the gradient:

$$r_{ii}^{t'+LF+1}(0) = r_{ii}^{t'+LF}(0) - X_i^2\ (t') + X_i^2\ (t'+LF+1) \qquad 3.7$$

and thus the scale factor k in Equation (2.19) for the i'th channel was:

$$k(i) = k/r_{ii}^{t'+LF+1}(0) \qquad 3.8$$

In other words, the magnitude of the step taken opposite the direction of the gradient of the mean square error was inversly proportional to the magnitude of the zero lag of the autocorrelation of each channel. This procedure was equivalent to normalizing the variance of each channel to unity over the moving time window.

Figure 5 is a map of the ten sensors taken from the LASA D1 subarray. Figure 6 is the array response of the partial D1 subarray. The data and the array used are the same used in a previous report (McCowan, 1968), in order to compare directly the results of adaptive filtering with time stationary filtering. Table III-2 is a list of the sensor positions. Table III-3 is the pertainent data for the Aleutian event. Table III-4 gives the contour increment for the f-k response function.

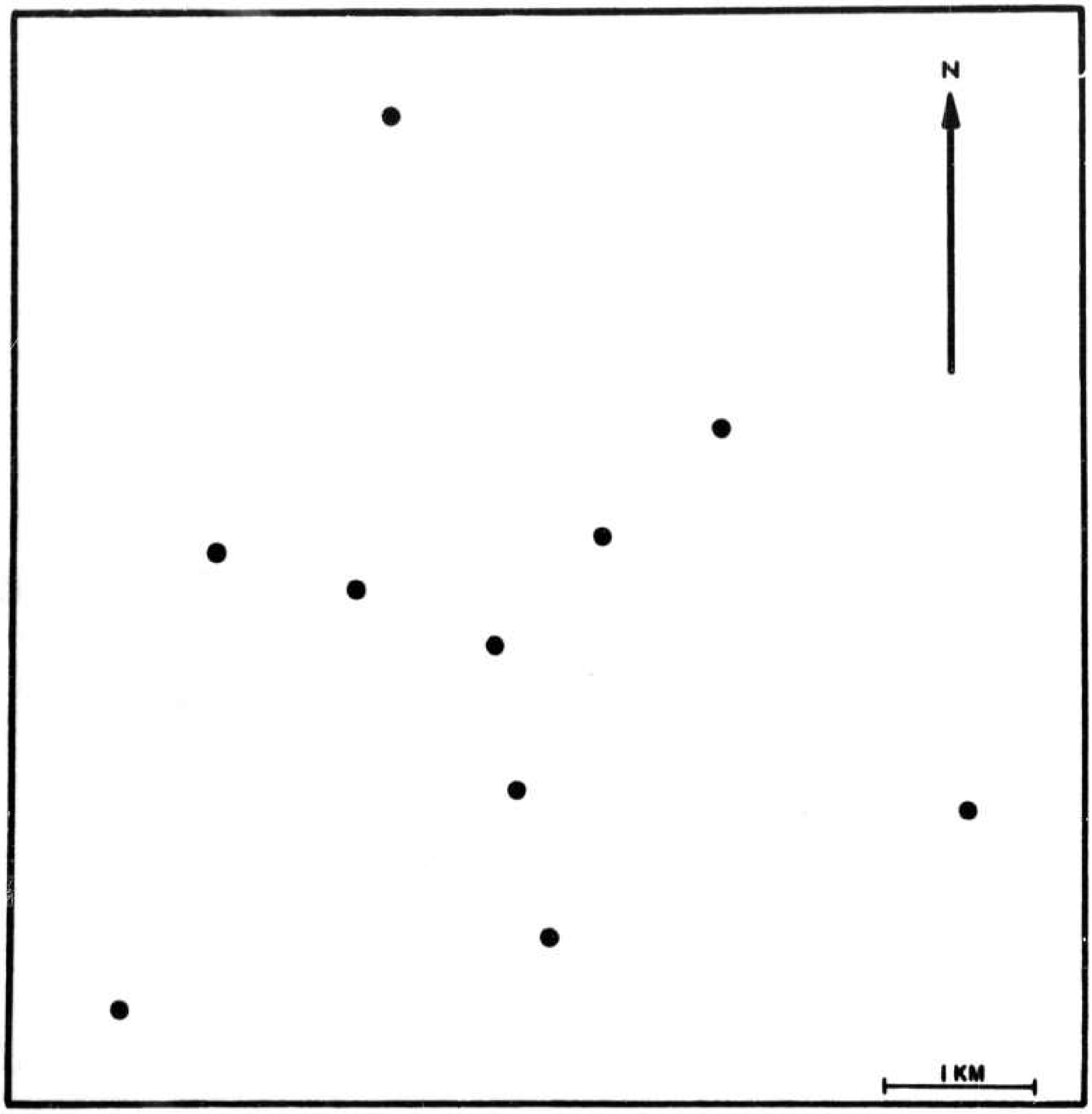


Figure 5. Array map of partial LASA D1 subarray

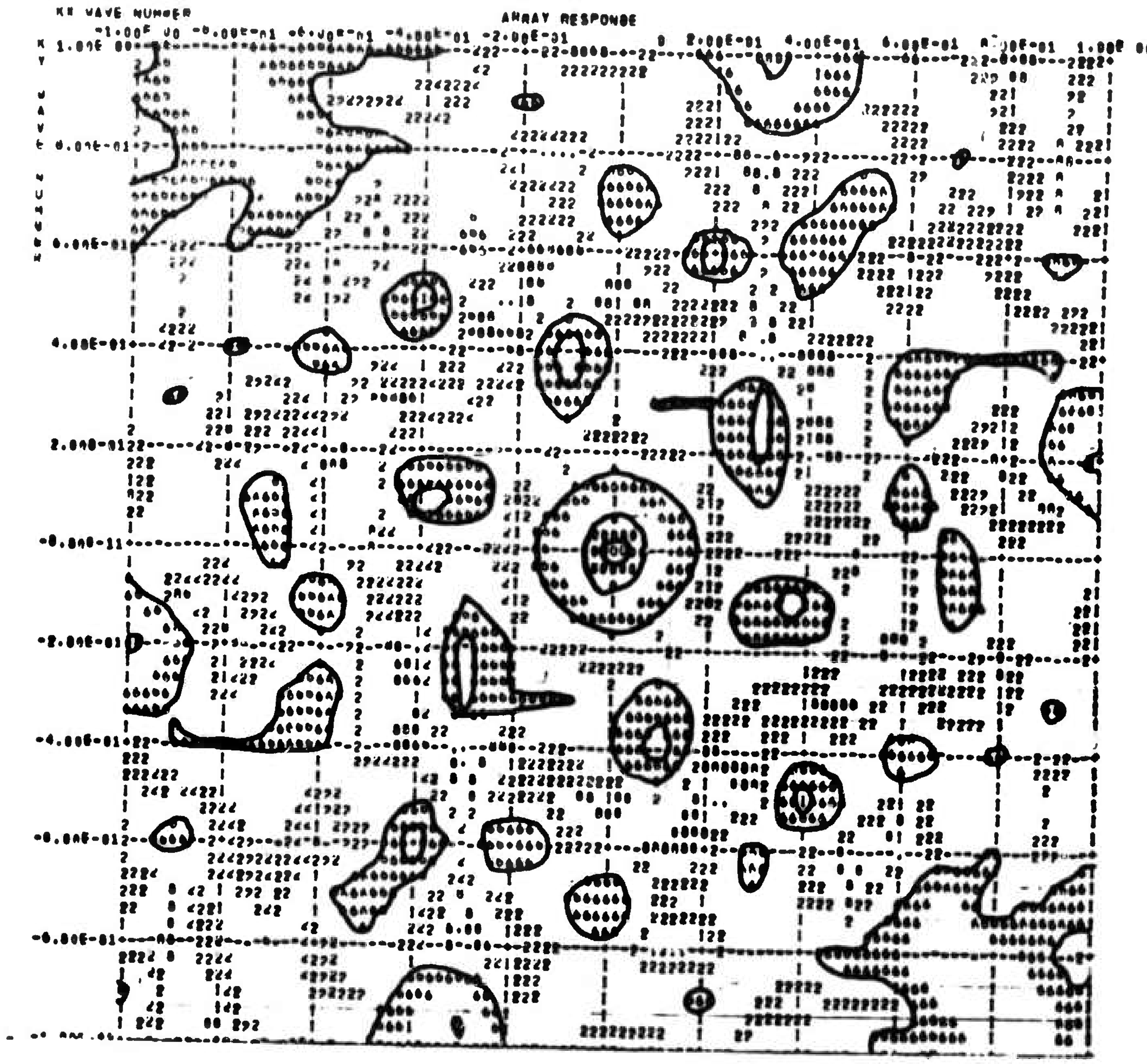


Figure 6. Array impulse response function

TABLE III-1 Description of MAP II Systems

Trace	System	System Type	Input Channels	Signal Model	Noise Model
2	MCF11	Multichannel Filter	Shallow-Buried Array (SZ1 - SZ10)	Velocity-Infinite	Measured Noise
3	MCF12	Multichannel Filter	Shallow-Buried and Vertical Array (SZ1 - 10, DH1 - 6)	Velocity-Infinite	Theoretical isotropic surface - mode noise
4	MCF13	Multichannel Filter	Vertical Array (DH1 - DH6)	Velocity-Infinite	Theoretical isotropic surface - mode noise
5	MCF14	Deghost filter	DH1, DH3, DH5 Up-going	Up-going Velocity-Infinite	Down-going infinite velocity, signal and theoretical surface - mode noise
6	MCF15	Deghost filter	DH1, DH3, DH5 Down-going	Down-going Velocity-Infinite	Up-going, infinite velocity, signal and theoretical surface - mode noise
7	MCF16	Deghost filter	DH2, DH4, DH6 Up-going	Up-going Velocity-Infinite	Down-going, infinite velocity, signal and theoretical surface - mode noise
8	MCF17	Deghost filter	DH2, DH4, DH6 Down-going	Down-going Velocity-Infinite	Up-going, infinite velocity, signal and theoretical surface - mode noise

Table III-1 Description of MAP II Systems, Continued

Trace	System	System Type	Input Channels	Signal Model	Noise Model
9	BSSV1	Beam Steered Summation	Vertical Array (DH1 - DH6)	Velocity-Infinite Up-going P-wave	Not Applicable
10	BSSV2	Beam Steered Summation	Vertical Array (DH1 - DH6)	8 km/sec velocity Up-going P-wave	Not Applicable
11	BSSV3	Beam Steered Summation	Vertical Array (DH1 - DH6)	8 km/sec velocity Up-going S-wave	Not Applicable
12	BSSV4	Beam Steered Summation	Vertical Array (DH1 - DH6)	Velocity-Infinite Down-going P-wave	Not Applicable
13	BSSV5	Beam Steered Summation	Vertical Array (DH1 - DH6)	8 km/sec velocity Down-going P-wave	Not Applicable
14	BSSV6	Beam Steered Summation	Vertical Array (DH1 - DH6)	8 km/sec velocity Down-going S-wave	Not Applicable
15	∑DVS	Simple Summation	Shallow-Buried and Vertical Array (SZ1-10, DH1-6)	---	Not Applicable

TABLE III-2

MULTICHANNEL FILTER REPORT

Array configuration

Channel No.	Channel ID	X Coordinate (km)	Y Coordinate (km)	Demagnification (counts/10-4 microns)
1	10	0.000	0.000	2.55
2	81	-0.716	3.426	3.09
3	32	0.745	0.667	2.78
4	52	1.610	1.440	3.18
5	83	3.325	-1.093	3.07
6	34	0.204	-0.979	2.48
7	54	0.409	-1.958	3.00
8	85	-2.609	-2.333	3.05
9	36	-0.950	0.312	2.69
10	56	-1.900	0.625	2.75

TABLE III-3

Location of Event

Date:	10 November 1965
Region:	Aleutian Islands
Magnitude*:	4.1
Origin Time*:	03:48:78.0 GMT
Latitude*:	51.2° N
Longitude*:	178.8° W
Focal Depth*:	"33 km"
Epicentral Distance to LASA A010:	46.1° (5126 km)
Azimuth of Epicenter From LASA A010:	304°
Surface Velocity at LASA A010:	14 km/sec

*Data from USC&GS Preliminary Determination of Epicenter.

TABLE III-4

DB	SYMBOL
0 - 1	0
1.001 - 3	0
6 - 9	6
12 - 15	2
18 - 21	8
24 - 27	.

Figure 7 is the result of applying the auto-adaptive filter to the input data. The first trace is a timing trace with 1-second pips, every fifth one of which is accentuated. Traces 2 through 11 are the observed data, trace 11 being the one to be predicted; this trace is shifted 1 second to the left. Trace 12 is the output of the auto-adaptive filter. Trace 13 (the error trace) is the difference between trace 11 and trace 12. Trace 14 is the sum of the nine input traces used to predict trace 11. The output of the adaptive filter is seen to be identical to the sum at the start of the record (as it should be, since the initial guess for the filters was 1/N, 0, 0, 0...).

Figure 8 is a continuation in time of the same data. We can see that the filter is indeed learning from the input data, since the noise power is getting progressively less in trace 12 as compared to trace 14. In general the same peaks seem to be coming through, but the power is quite a bit less in the output of the adaptive filter (trace 12) as compared to the beamed sum trace (trace 14) as time goes on. Figure 9 is a continuation of the same data. The first motion of the signal seems to be enhanced as does a secondary phase, probably pP. The signal-to-noise ratio improvement is 2 db over the bandpass phased sum, due to some signal rejection by the adaptive filter. The second phase occurs roughly nine seconds after the primary phase, the time delay being about right for pP.

It should be noted that the progressive decrease in the noise power with time is due to two factors: one is the convergence of the adaptive filter from the initial guess toward the optimum time-stationary filter; the other factor is the improvement achieved by reducing time-varying coherent noise. The small improvement (2 db) above the band limited phased sum seems to indicate that the amount of time-varying coherent noise power is low. Of course, another factor to consider is the poor array response (Figure 6) which causes coherent noise to alias back into the main lobe.

Table III-5 gives the pertainent data for the W.Andreanof Islands event.

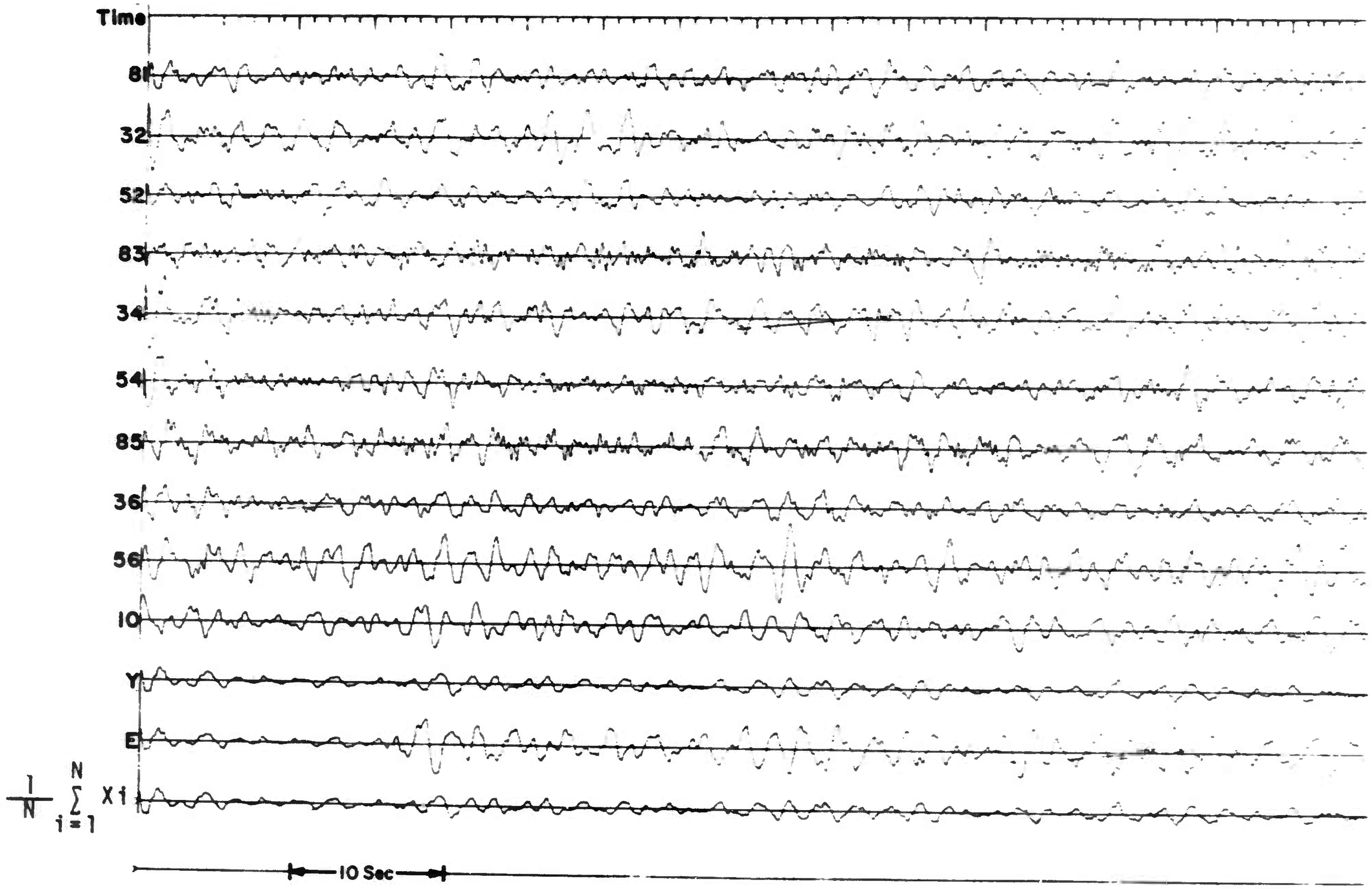


Figure 7. Noise and processed traces before the Aleutian event

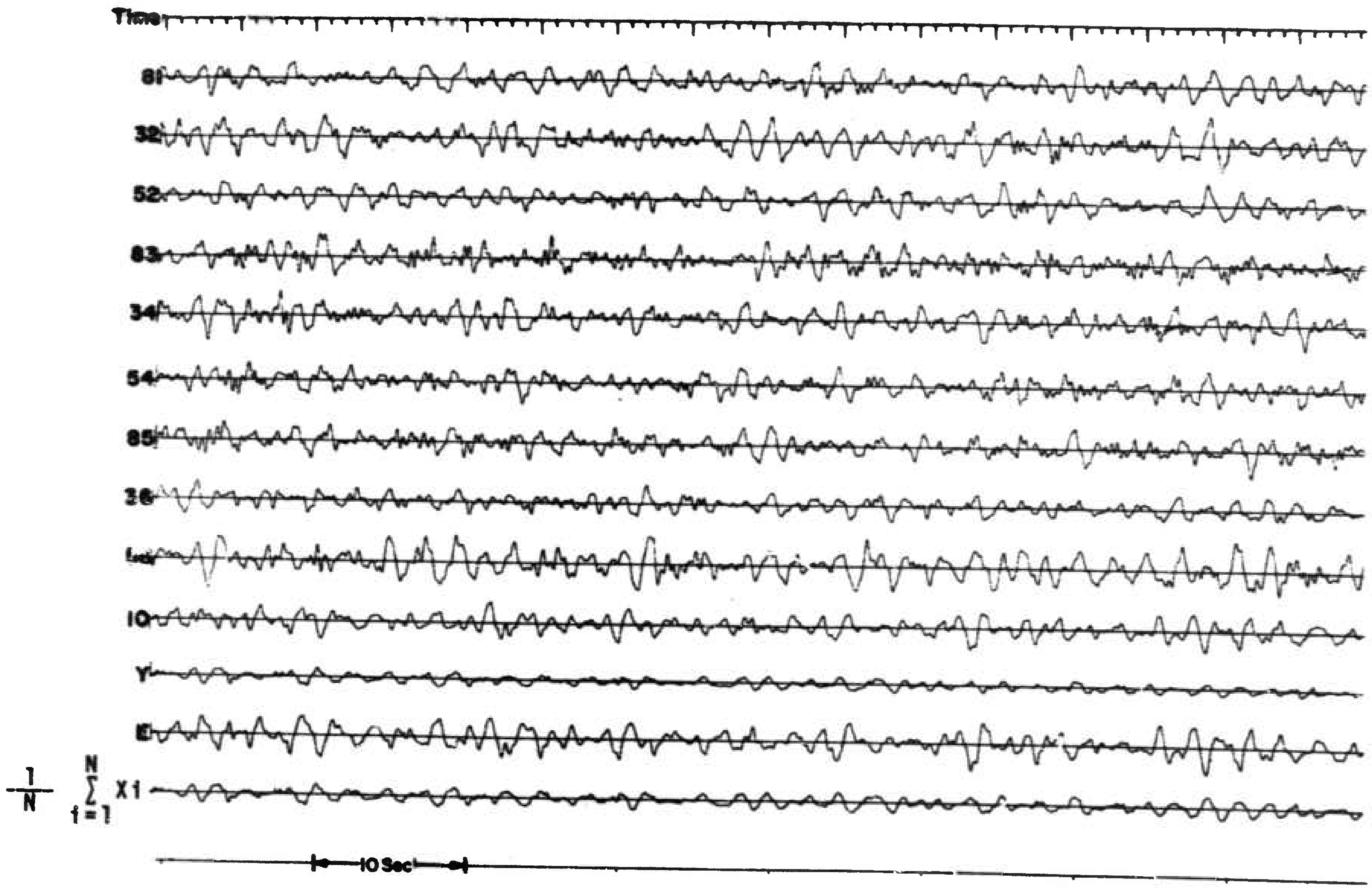


Figure 8. More Noise and processed traces before the Aleutian event

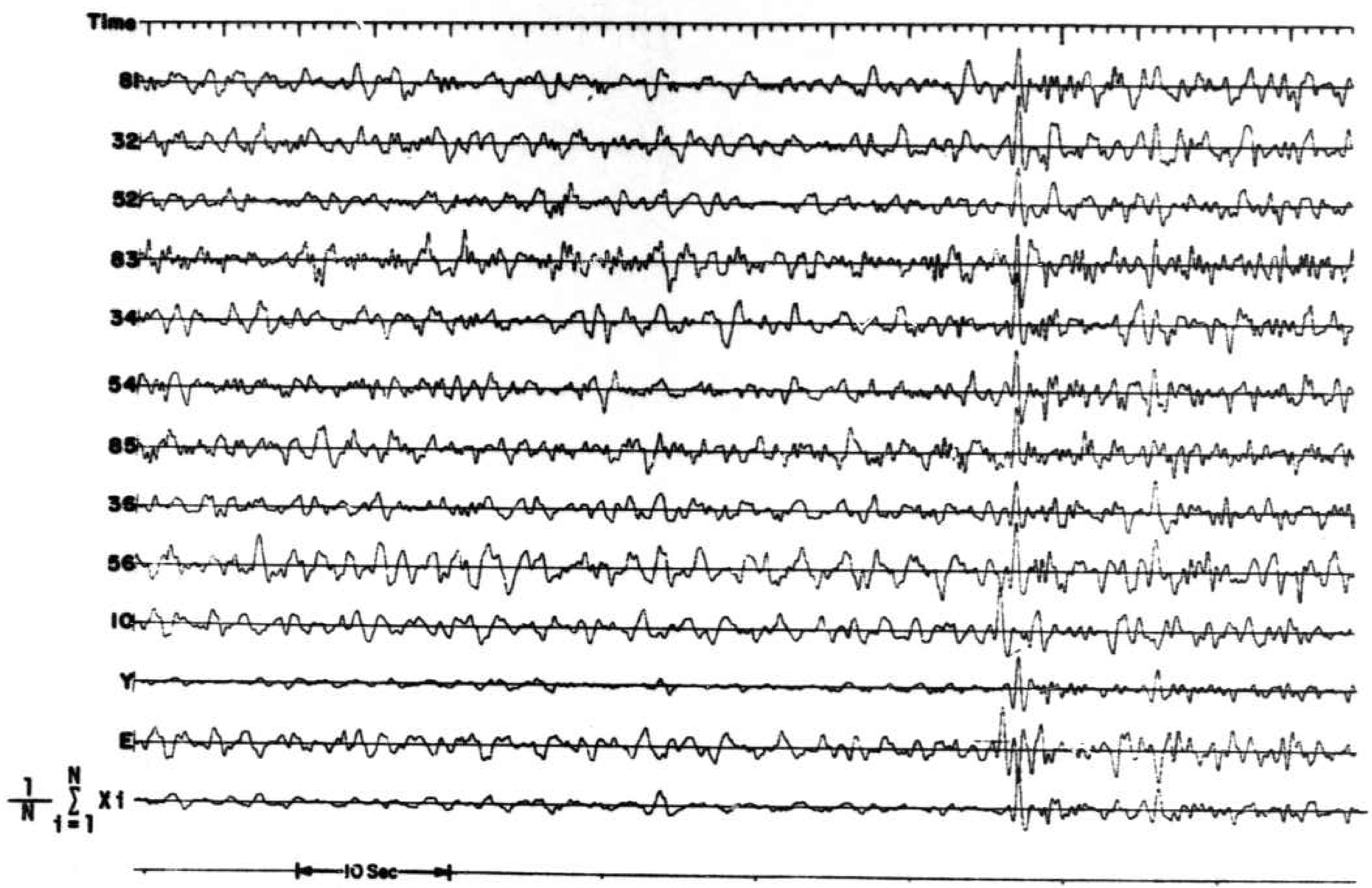


Figure 9. The Aleutian Event

TABLE III-5

Location of Event

Date:	22 February 1968
Region*:	West Andreanof Islands
Magnitude*:	3.8
Origin Time*:	14:43:46
Latitude*:	51.6N
Longitude*:	176.2W
Focal Depth*:	65 km
Epicentral Distance to UBO:	46.137° (5130 km)
Azimuth of Epicenter to UBO:	76.773°
Surface Velocity at UBO:	14 km/sec

* Data from USC&GS Preliminary Determination of Epicenter.

Figure 10 is a map of UBO nine-channel surface array. Figure 11 is the array response function. Figure 12 is the input data after it has been demagnified, bandpass filtered, and beamed; the various outputs are also shown. Trace 1 is the timing trace (one second pips, with every fifth pip accentuated). Traces 2 through 11 are the input data. Trace 11 is the predicted trace (shifted 1 second ahead in time). Trace 12 is the output of the adaptive filter. Trace 13 is the beam-formed sum and trace 14 (the error trace) is the difference between trace 12 and trace 13.

Once again, we note that at the beginning of the record, trace 12 and trace 13 are identical due to our original estimate of the optimum filter (the original estimate being 1/N,0,0,0... for each channel).

Figure 13 shows a continuation in time of the same data. Note that the auto-adaptive filter is performing better than the phased sum as the filter learns more about the noise characteristics. Figure 14 shows a continuation of the same data. The filter was constrained to operate in Mode 2 shortly before the expected arrival time of the signal (14:52:04). Although it was not practical to measure signal-to-noise ratios with such a small signal, it is obvious to the eye that the auto-adaptive filter is doing a very good job. In order to compare the results of the auto-adaptive filter with the on-line analog MCF's we have reproduced a section of film from Develocorder Number 5 at UBO (Figure 15). A description of the design of various filters may be found in a report by Edwards (1965). Table III-1 is a description from a report (Leichliter, 1967) about the on-line-filters. The auto-adaptive filter (Figure 14) seems (by eye) to be doing a better job than MCF 11 (trace 2) which is an infinite velocity measured-noise filter. Of course, it is not quite fair to compare an analog filter operating on analog data with a digital filter operating on digital data, since

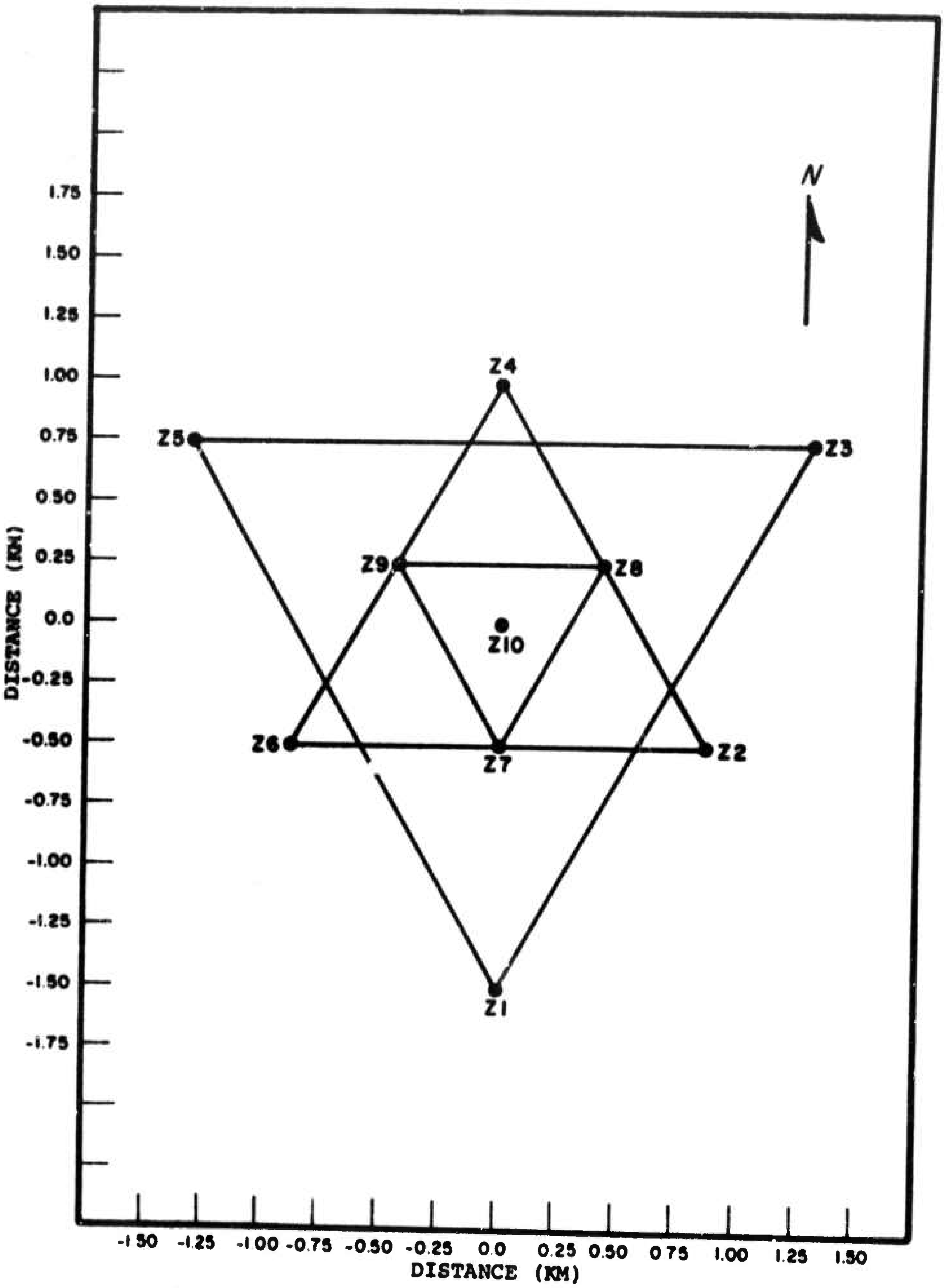


Figure 10. Array map of Uinta Basin Seismological Observatory

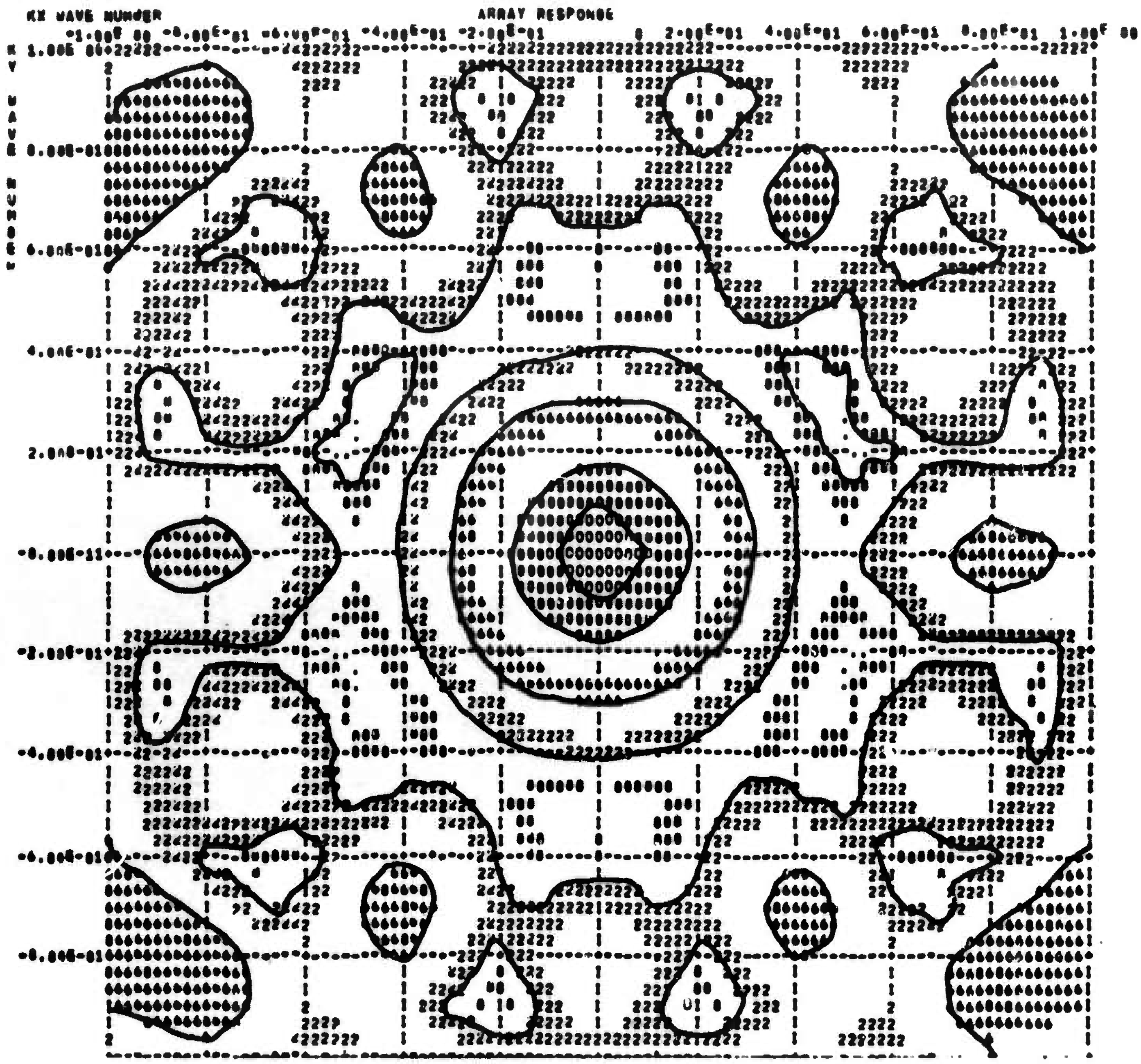


Figure 11. Array impulse response function

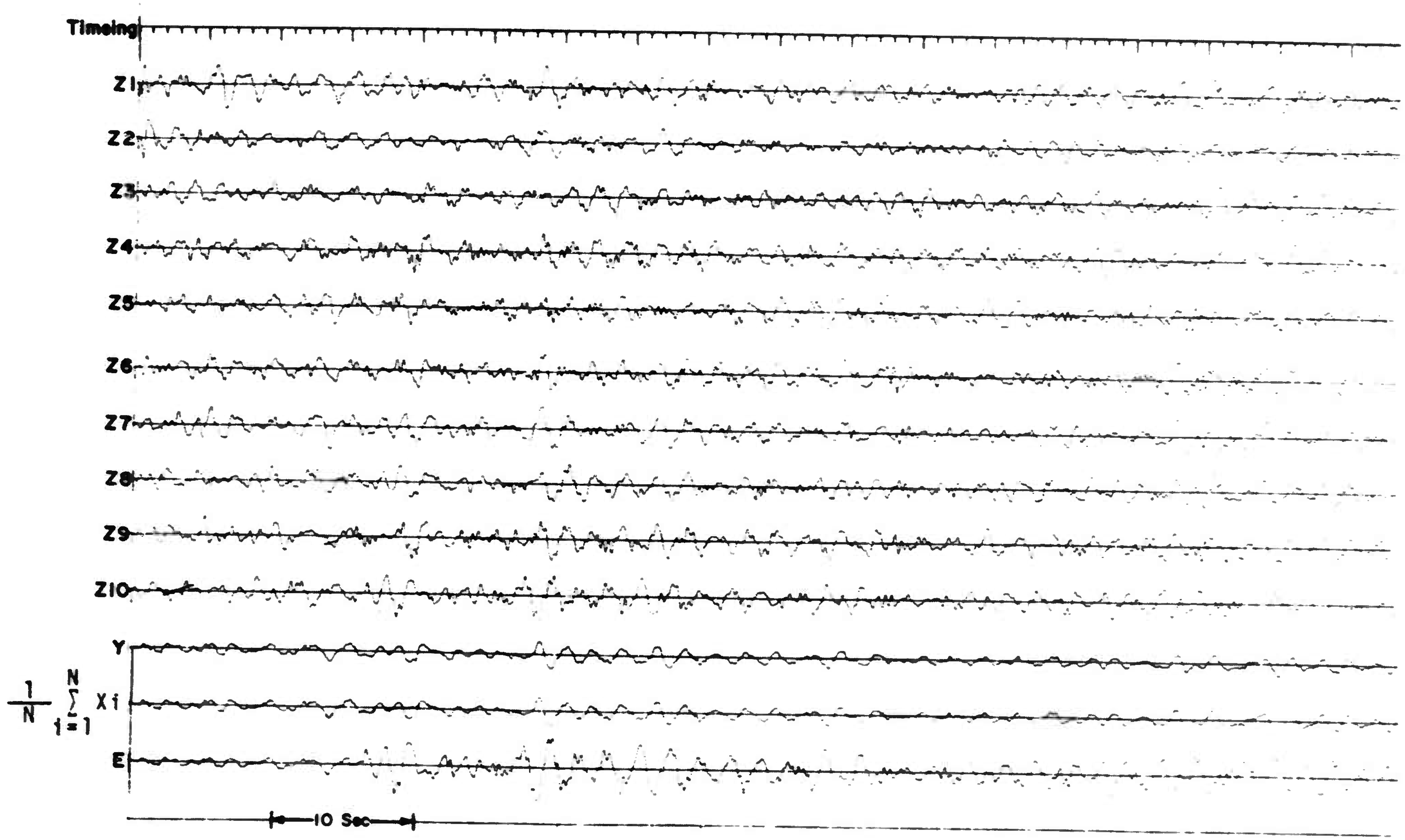


Figure 12. Noise and processed traces before the West Andreanof Islands Event

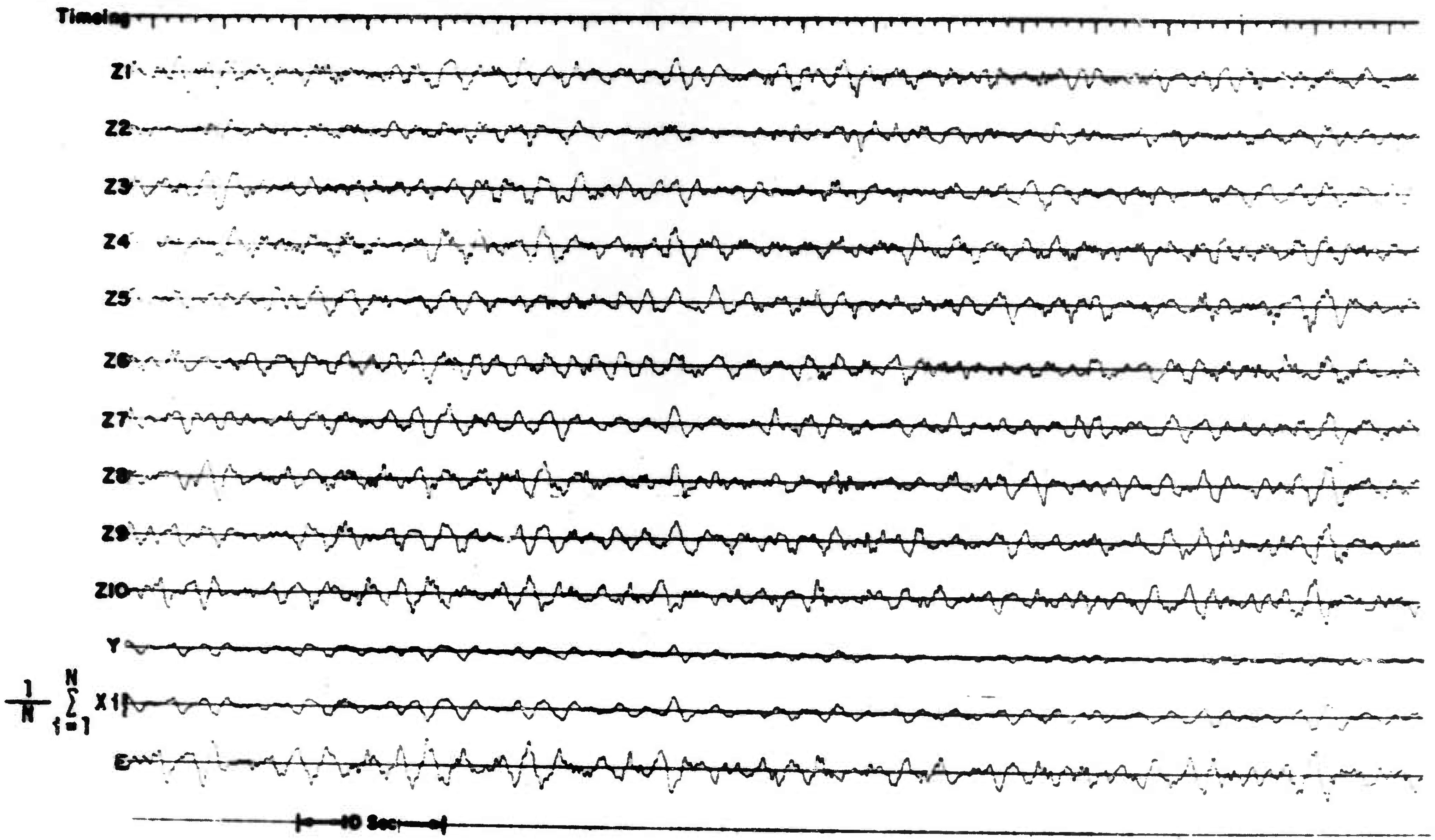


Figure 13. More noise and processed traces before the West Andreanof Islands Event

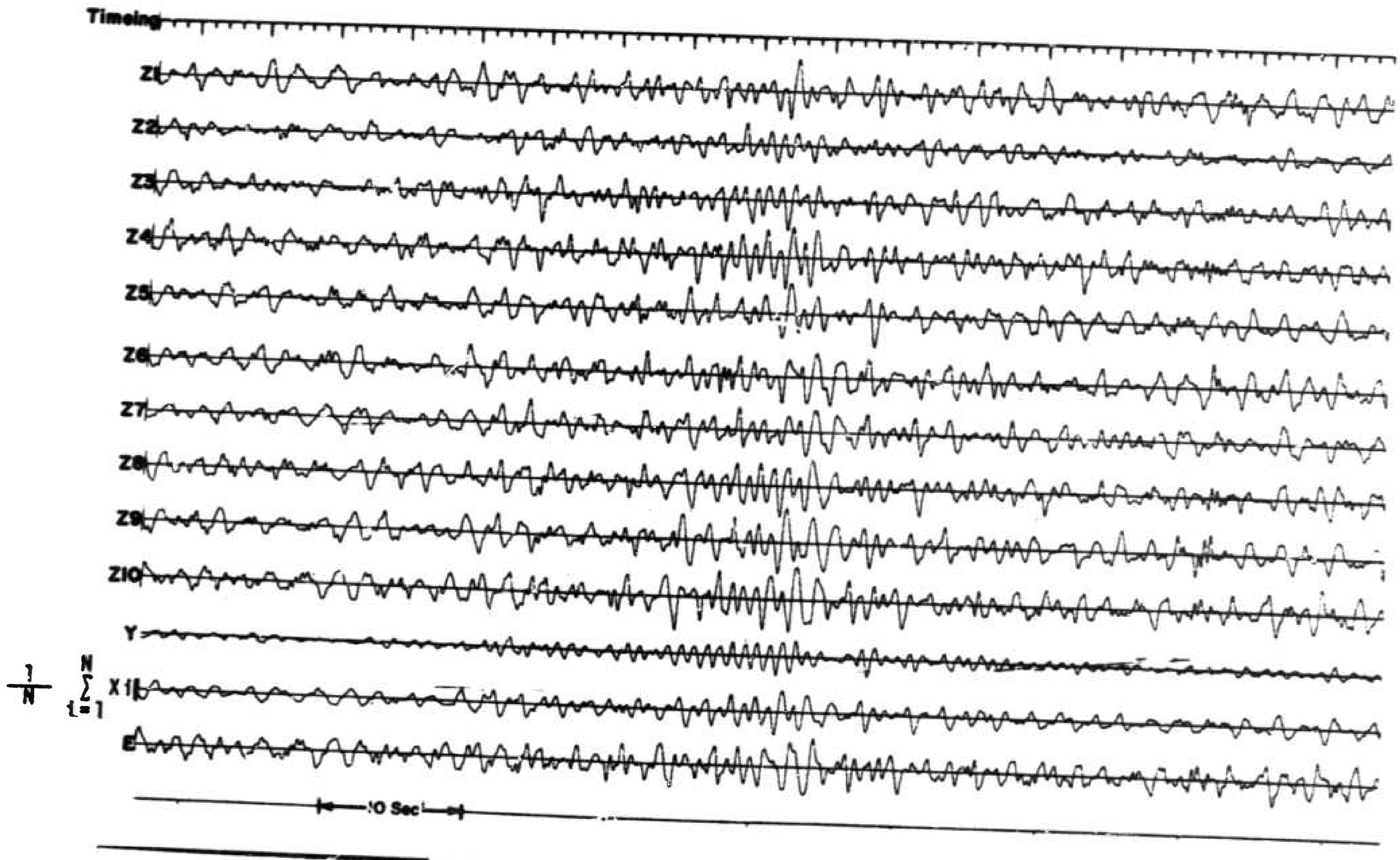


Figure 14. The West Andreanof Islands Event

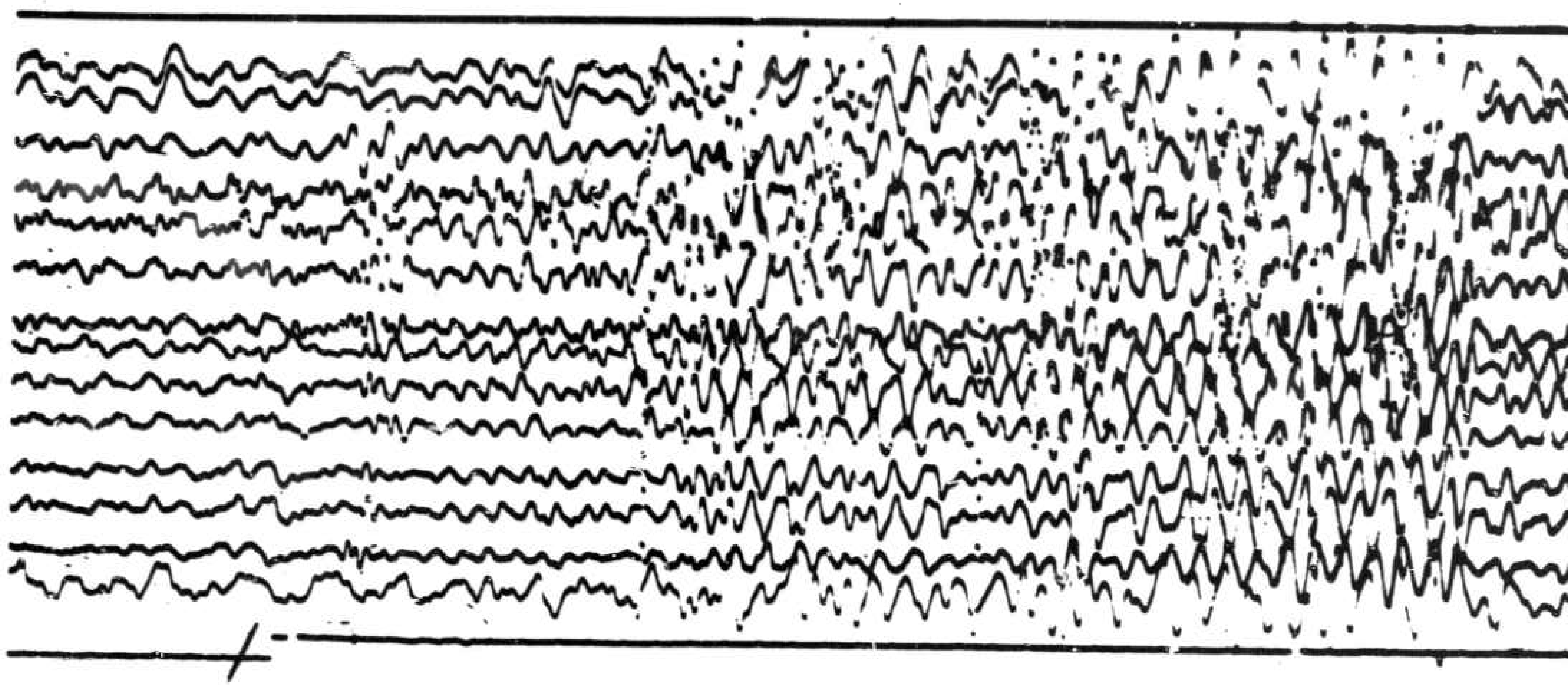


Figure 15. Output of the on-line Map II Processer at UBSO

all things being equal, the digital filter will always out-perform the analog filter. The interesting thing about Figure 14 and 15 is that the analog filters operating on the vertical array seem to be out-performing (in terms of first motion enhancement) even the digital auto-adaptive filter. The reason for the better performance of the vertical array stems from the fact that the noise level is lower on the vertical array, and the fact that the surface array is small.

Figure 16 illustrates the application of a time-stationary maximum likelihood filter to the input data. Trace 1 is the sum of the ten input traces (beamed and bandpassed). Trace 2 is the output of the maximum likelihood filter. This maximum likelihood filter was designed on 3800 points of noise before the signal. The filter was a one sided 21 point filter which operated on every other data point, hence the filter was effectively a 2.1 second filter.

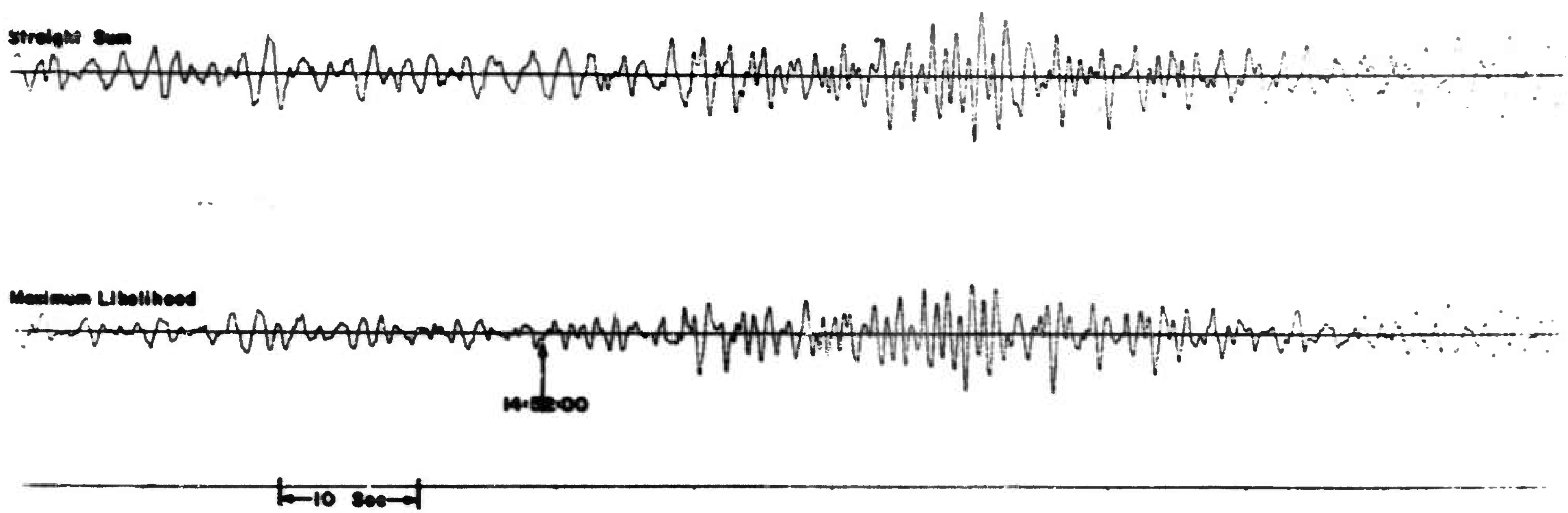


Figure 16. The direct sum and the time stationary maximum likelihood filtered output of the West Andreanof Islands Event

SECTION IV. CONCLUSIONS AND RECOMMENDATIONS

From this report it can be concluded that:

. The method of updating the correlation functions used in this report is a valid method for calculating relatively stable and slowly time varying correlation functions.

. The auto-adaptive filters designed from the time-varying noise correlation functions do converge, and they yield better results (2 db better on the LASA event) than simple bandpass filtering and beamforming. The concept of auto-adaptive filtering could be applied to single channel (prediction, prediction-error, and deghost) filtering as well as to multichannel filtering.

From this study, it can be recommended that:

. Additional iterative techniques such as the Gauss-Seidel method be investigated.

. A single-channel deghosting auto-adaptive filter program be written for vertical array data.

. A single channel prediction - error auto-adaptive filter program be written to process the output of isotropic multichannel time stationary filters or to process beamed sum traces.

ACKNOWLEDGEMENT

The author is in debt to Dr. E.A. Flinn for his many helpful discussions, in particular his suggestion to update the correlation functions with time, and for his editing of the manuscript. The author is also in debt to Dr. Hans Jarosch, on sabbatical from the Weizmann Institute, Rehovoth, Israel, for his discussions with the author concerning iterative solutions of the multichannel normal equations.

REFERENCES

Burg, J.P., Booker, A.H., Holyer, R.J., 1967, Adaptive Filtering of Seismic Array Data, Advanced Array Research Special Report No. 1: Dallas, Texas Instruments Inc., Contract F33657-67-C-0708-P001.

Edwards, J.P., 1965, Multiple Array Processor Final Report: Dallas, Texas Instruments, Inc., Contract AF 33(657)-13904.

Kelley, E.J., and Levin, M.J., 1964, Signal Parameter Estimation for Seismometer Arrays; Technical Report 339: MIT Lincoln Laboratory, Lexington, Massachusetts.

Leichliter, B.B., 1967, Evaluation of Multiple Array Processors at the Uinta Basin Seismological Observatory, Technical Report No. 67-75: Garland Texas, Teledyne Industries, Geotech Division, Contract No. AF 33(657)-16563.

Levinson, N., 1949, The Wiener RMS Error Criterion in Filter Design and Prediction: Appendix B in: Extrapolation, Interpolation, and Smoothing of Stationary Time Series, by Norbert Wiener, MIT Press, Cambridge, Massachusetts.

McCowan, D.W., 1968, Design and Evaluation of Certain Multi-Channel Filters, Seismic Data Laboratory Report No. 209: Alexandria, Earth Sciences, A Teledyne Company, Contract No. F33657-67-C-1313.

Meserve, W.J., 1968, Oral Communication to Author.

Papoulis, A., 1965, Probability, Random Variables, and Stochastic Processes: New York, McGraw-Hill Book Company

Robinson, E.A., 1967, Multichannel Time Series Analysis with Digital Computer Programs: San Francisco, Holden-Day, Inc.

Varga, R.S., 1965, Matrix Iterative Analysis: Englewood Cliffs, New Jersey, Prentic Hall, Inc.

Widrow B., 1966, Adaptive Filters 1: Fundamentals, Report No. SU-SEL-66-126: Stanford, California, Stanford Electronic Laboratories.

APPENDIX: PROGRAM LISTINGS

This section of the report presents the program listing for Program Adamax3 and Subroutine Upgrad. Program Adamax3 performs the operations for adaptive filtering in Mode 1 and Mode 2. Subroutine Upgrad performs the actual updating of the gradient of the mean square error and the actual filtering. Both routines are written for clarity rather than speed and economy. The program Adamax3 was written for demultiplexed data. The operation could be speeded up by operating on multiplexed data.

The statement

Call Disc63(NB,IRW,VAR,NPRW)

gives a read (write) from (onto) the disc if IRW is 0(1). NB is the block number (32 words per block) to read (write) from (onto) the disc. VAR is the variable location in core, and NPRW is the number of data points to read (write) from (onto) the disc.

It should be noted that subroutine Upgrad expects the filter vector F(J,I) to be stored backwards in core (i.e., F(L,I), F(L-1,I)....F(1,I).

```
      PROGRAM ADAMAX3
C
C     THIS IS THE MAIN PART OF THE PROGRAM ********************************
C
      IM=512-LF
      KK=IM/32
      K1=(NPTS-1)/IM+1
C     N2=N-1
C     N1=(NPTS-1)/32+1
C     JM=LF+1
C     LF IS A MULTIPLE OF 32
      DO 101 K=1,K1
      DO 9 I=1,N2
    9 CALL DISC63((I-1)*N1+(K-1)*KK,0,X((I-1)*512+1),512)
      CALL DISC63(N2*N1+(K-1)*KK,0,D,512)
      CALL DISC63((N+1)*N1+(K-1)*KK,0,Y,512)
      DO 13 I=1,N2
      DO 13 J=1,512
      LJ=(I-1)*512+J
   13 X(LJ)=D(J)-X(LJ)
      IF(K.EQ.K1)104,105
  104 IM=NPTS-(K1-1)*IM
  105 CONTINUE
      DO 100 M=1,IM
      ISW2=M-100
      IF(K.GE.K1-KCO)150,151
C
C     KCO IS THE CUTOFF VALUE FOR OPERATING IN MODE 1
C     LAST POINT FOR MODE 1 OPERATION=(K1-KCO)*(512-LF)
  150 ISW2=1
  151 CONTINUE
      CALL UPGRAD(X(M),GRAD,F,N2,JM,Y1,SK,E(M),L,Y2,R0,ISW2)
      MLF=M+LF
      Y(MLF)=Y1
      Y(MLF+1)=Y2
      E(MLF)=D(MLF)-Y1
  100 E(MLF+1)=D(MLF+1)-Y2
      CALL DISC63(N*N1+(K-1)*KK+LL,1,E(JM),IM)
  101 CALL DISC63((N+1)*N1+(K-1)*KK+LL,1,Y(JM),IM)
C   ************************************************************************
```

```
      SUBROUTINE UPGRAD(Y,G,F,N,LF,Y,SK,E,L,Z,R0,ISW2)
C     THIS SUBROUTINE UPDATES THE GRADIENT(OF THE ERROR SQUARED)WITH THE
C     ARRIVAL OF EACH NEW DATA POINT,AND CALCULATES THE NEW MCF FROM THE
C     OLD MCF.THE SUBROUTINE THEN OUTPUTS THE RESULT OF CONVOLVING THE
C     FILTER WITH THE INPUT DATA
      DIMENSION X(512,N),G(L,N),F(L,N),E(LF),R0(N)
C     LF=LENGTH OF FITTING INTERVAL + 1
C      L=LENGTH OF FILTER
      Y=0.0
      Z=0.0
      DO 11 J=1,L
      DO 11 I=1,N
   11 G(J,I)=G(J,I)-X(1,I)*E(J)+X(LF-J+1,I)*E(LF)
      DO 10 I=1,N
      R0(I)=R0(I)-X(1,I)**2+X(LF,I)**2
      SK1=SK/R0(I)
      IF(ISW2.LE.0)13,15
   13 DO 14 J=1,L
   14 F(J,I)=F(J,I)+SK1*G(L-J+1,I)
   15 DO 10 J=1,L
      Y=Y+F(J,I)*X(J+LF-L,I)
   10 Z=Z+F(J,I)*X(J+LF-L+1,I)
      RETURN
      END
```

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1 ORIGINATING ACTIVITY (Corporate author)

TELEDYNE, INC.
ALEXANDRIA, VIRGINIA

2a REPORT SECURITY CLASSIFICATION

Unclassified

2b GROUP

3 REPORT TITLE

PRINCIPLES OF WIENER AUTO-ADAPTIVE FILTERING

4 DESCRIPTIVE NOTES (Type of report and inclusive dates)

Scientific

5 AUTHOR(S) (Last name, first name, initial)

Lintz, P.R.

6 REPORT DATE

27 September 1968

7a TOTAL NO OF PAGES

48

7b NO OF REFS

11

8a CONTRACT OR GRANT NO.

F 33657-68-C-0945

b PROJECT NO

VELA T/6702

c ARPA Order No. 624

d ARPA Program Code No. 8F10

9a ORIGINATOR'S REPORT NUMBER(S)

224

9b OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10 AVAILABILITY/LIMITATION NOTICES

This document is subject to special export controls and each transmittal to foreign governments or foreign nationals may be made only with prior approval of Chief, AFTAC.

11 SUPPLEMENTARY NOTES

12 SPONSORING MILITARY ACTIVITY

ADVANCED RESEARCH PROJECTS AGENCY
NUCLEAR TEST DETECTION OFFICE
WASHINGTON, D. C.

13 ABSTRACT

This report discusses various techniques of designing auto-adaptive (self-changing) filters, and the application of one particular technique to two teleseisms. This report also discusses the calculation of running correlation functions using a moving time window, and iterative solutions to the multi-channel normal equations. A program listing is included (in the Appendix) which will calculate an auto-adaptive multi-channel filter by the use of running correlation functions and the method of steepest descent.

14 KEY WORDS

Auto-adaptive filtering
Running correlation functions
Moving time window
Adaptive multichannel filtering
Time-varying linear filtering

Unclassified

Security Classification