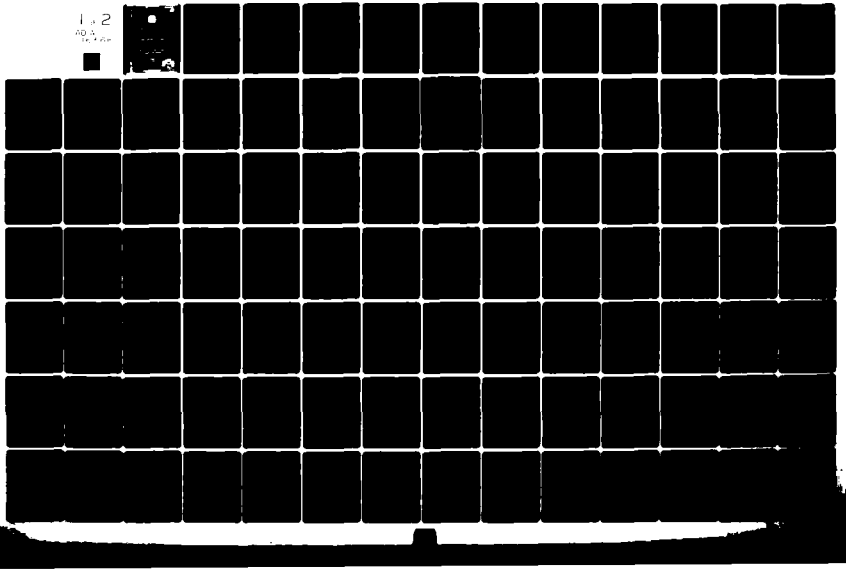
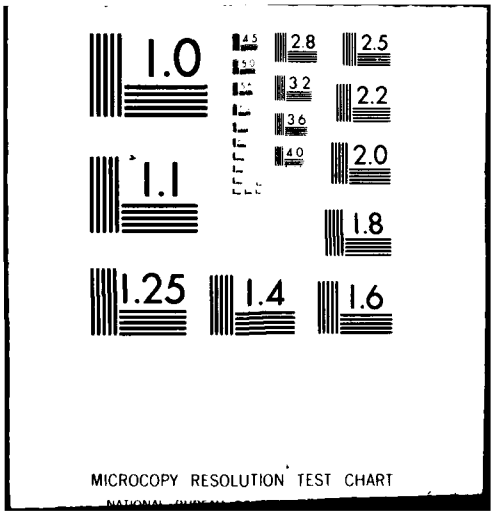
AD-A116 366 ATOMIC WEAPONS RESEARCH ESTABLISHMENT ALDERMASTON (EN--ETC F/G 8/11
THE ESTIMATION OF SEISMIC BODY WAVE SIGNALS IN THE PRESENCE OF --ETC(U)
DEC 81 A DOUGLAS, J B YOUNG
UNCLASSIFIED AWRE-O-14/81 DRIC-BR-81928 NL




1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
MICROCOPY RESOLUTION TEST CHART

Procurement Executive - Ministry of Defence

AWRE, Aldermaston

AWRE REPORT NO. O14/81

The Estimation of Seismic Body Wave Signals in the Presence of Oceanic Microseisms

A Douglas
J B Young

ISBN 0 8551845 1

CONTENTS

SUMMARY

Most studies of the processing of recordings from seismometer arrays to extract seismic body waves from noise have used data recorded by conventional narrow band systems. The aim in this type of processing has always been to achieve maximum signal-to-noise ratio. This report describes studies of the processing of broad band recordings (from a system with displacement response flat from around 0.1 to 10 Hz) for the estimation of signal shape rather than the maximisation of signal-to-noise ratio; processing of both single seismograph and array (multichannel) recordings is discussed. The data used come from a 4 element array of broad band seismometers in southern England. The predominant noise on the broad band recordings is oceanic microseisms with periods of around 6 s and the main purpose of any array processing is usually to suppress this type of noise.

By definition Wiener filtering gives the best estimate of signal shape in the sense that filters are designed to minimise the mean square of the difference between the true signal (desired output) and the estimated signal (actual output) consequently most of this report describes studies of the application of this type of filter. In the general (multichannel) case Wiener filters apply both spatial and frequency filtering to extract the signal. However, if the required noise reduction can be obtained by spatial filtering alone, then no frequency filtering is applied and so the signal is passed undistorted. From the data studied in this report it is possible to get noise reductions due to spatial filtering of up to 6 with the 4 element array.

Studies are also described of the use of filters designed to minimise the noise power at the output of the filters subject to the constraint that the desired signal is passed undistorted. It is shown that this method of processing (usually referred to as the maximum likelihood method) can be considered as a special case of Wiener filtering.

1. INTRODUCTION

Because of the large peak in the seismic noise spectrum of ground displacement at around 6 to 8 s period - the oceanic microseism peak - the body wave signals from all but large magnitude sources can only be seen above the background noise if the oceanic microseisms are attenuated by filtering; the bulk of this filtering is usually applied as frequency filtering by the seismograph. With this recording method signals with amplitudes that are above the noise over a wider band of frequencies than the limited pass band of the seismograph are filtered unnecessarily. Yet Berckhemer (1) has pointed to the need to record signals over as wide a band as possible, particularly for source studies, and Marshall, Burch and Douglas (2) illustrate the value of broad band seismograms for such studies by using large magnitude sources for which the signal is larger than the noise.

Recording with narrow band seismographs is only necessary when visual seismograms alone are recorded. Given magnetic tape recording, a better way of displaying seismograms for analysis would seem to be to use a recording system from which a wide range of frequencies can be recovered and to apply to these wide band seismograms just sufficient frequency filtering to extract the best estimate of signal shape. When array records are available, then differences in the spatial properties of the signal and noise can be exploited to reduce the noise and pass the signal unchanged - which is the object of using an array for noise suppression. It would seem then that any noise reduction that can be obtained from array processing should be applied before (or simultaneously with) the frequency filtering. In this way frequency filtering will not be applied where the required noise reduction can be obtained by spatial filtering.

Most studies of the use of seismometer arrays to extract seismic body waves from noise have used data recorded on conventional narrow band systems. For such arrays at sites where the signal does not vary greatly over the aperture of the array it has been found that satisfactory improvements in signal-to-noise ratio can be obtained by using simple delay and sum (DS) processing; the signals recorded at each seismometer are time shifted so that their onsets coincide, the channels are summed and this sum divided by the number of seismometers. The signal at the output of DS processing is thus the average over all channels. Often

the predominant frequency of the signal is obviously different from that of the noise and further improvements in signal-to-noise ratio can then be obtained by band pass filtering of the DS signal using a filter that passes the signal frequencies but attenuates the predominant noise frequencies.

If the noise has the same variance at each seismometer and is uncorrelated between pairs of seismometers, then DS processing of data from an n seismometer array gives on average $n^{\frac{1}{2}}$ improvement in signal-to-noise ratio, which is the greatest improvement that can be obtained (ignoring frequency filtering) for such noise. If the noise consists of propagating wave trains, sometimes described as spatially organised noise so that the outputs of pairs of seismometers are not all uncorrelated, DS processing does not in general give the best possible signal-to-noise improvement and other processing methods can be used which give improvements of greater than $n^{\frac{1}{2}}$.

However, no method of array processing that is capable of suppressing organised noise appears to have been widely used. The main reason for this seems to be that the value of any processing is usually assessed on the signal-to-noise improvement and it is commonly found that the same signal-to-noise improvements can be obtained using DS processing with additional band pass filtering, as with processing methods that attempt to exploit the spatial organisation of the noise. As DS processing is much simpler and quicker to carry out than other methods there is little incentive to use anything else. If signal-to-noise improvement is the only criterion used to measure the effectiveness of a processing method, then DS processing of narrow band recordings, with added band pass filtering if necessary, will probably always give the best results. The main result of this type of array processing is that the detection threshold is lowered below that of a single seismograph but only over the narrow band of frequencies where the noise is low anyway. There will usually only be an advantage in carrying out such processing for signals with amplitudes at or near the detection threshold of a single channel. Processing of narrowband signals that have amplitudes well above that of the noise simply to improve signal-to-noise ratio will usually be pointless.

In this report we investigate the application of processing methods for the estimation of signal shape rather than the maximisation of signal-to-noise ratio; processing of both array and single seismograph recordings are considered.

Now Wiener filtering by definition gives the best estimate of signal shape in the sense that the filters are designed to minimise the mean square of the difference between the true signal (desired output) and the estimated signal (actual output), consequently most of this report is concerned with the application of this type of filter for the extraction of signals from noise. The application of Wiener processing to narrow band recordings of array data has been investigated to Burg (3), Backus, Burg, Baldwin and Bryan (4) and Backus (5). We also investigate the processing method designed to minimise the noise power at the output subject to the constraint that the desired signal is passed undistorted. This method of processing has been studied in detail in several papers (see, for example, Capon, Greenfield and Kolker (6)) and is usually described as the maximum likelihood method but in this report we refer to it as the minimum power (MP) method. We consider the relationship between MP and Wiener filtering.

The simplest method of extracting high speed body wave signals from low speed oceanic microseisms is to install two seismometers half the microseismic wavelength apart and sum the outputs; the microseisms should then tend to cancel out and the signals to sum. This method of suppressing oceanic microseisms is suggested by Baker (7) and has been studied by Henger (8) using a 3 element triangular array in Germany (9,10). The spacing in the array was varied and the noise reduction as a function of seismometer spacing measured. Maximum noise reduction was obtained at spacings of ~ 12.5 km but, in general, this noise reduction was little better than $3^{\frac{1}{2}}$. The relationship of this simple approach to the suppression of microseisms and more complex processing schemes is considered later.

We are not concerned here with the detection of small signals from sources with unknown epicentre and origin times. We assume that any signal to be processed has been detected by narrow band systems and that the apparent velocity and rough onset time of the signal is known. At array sites the signal will differ from channel to channel (ideally these differences will be small) and we take the best representation of the signal to be the average of the outputs of all the channels.

2. THE ARRAY AND RECORDING SYSTEM

The data used in this study comes from an array of four seismometers situated in southern England near our laboratory at Blacknest about 20 km west of Reading, Berkshire (see inset figure 1); some initial results from the study of data from this Blacknest array (BNA) are given by Burton (11). The choice of suitable sites for seismometers is limited because the area is well populated so the four sites used (figure 1) were chosen mainly because they are the most convenient available.

The seismometers used in the array are Geotech S11 instruments with a natural frequency of 0.05 Hz (20 s period). Originally the output of these instruments was shaped electronically to produce the required flat displacement response between 0.1 and 10 Hz (the displacement broadband response: DBB) to simulate the Kirnos SKM system widely used in the USSR (2,11). Recording directly on to magnetic tape with such a response however does not make the best use of the available dynamic range of the tape; the DBB system was therefore altered so that the response as written on to tape is flat to ground velocity between 0.1 to 10 Hz (the velocity broad band response: VBB). The flat response to ground displacement is derived from the VBB recording on playback by integrating the output from the tape before writing the seismogram. Initially the magnetic tape recording was analogue only; now both analogue and digital recordings are made. As the digital recording is made at a sampling rate of 12.5 samples/s (for each seismometer channel) the analogue signal has to be low pass filtered to cut out power at frequencies above 6.25 Hz (the Nyquist frequency) to avoid aliasing on conversion from analogue to digital form. All broad band recordings shown in this report are either as recorded on the DBB response or are VBB recordings integrated (IVBB). When played back from tape using a recorder with a sensitivity of 1 V/cm the magnification at 1 Hz of the DBB system is 6600 and for the IVBB output is 5300. Figure 2 shows the response of the broad band systems.

Although the BNA is designed primarily to record broad band data it is possible to extract short period information by simply multiplying the spectrum of the BB recordings by $a_1(\omega)/a_2(\omega)$ and transforming back into time; $a_1(\omega)$ is the response of an SP seismograph at frequency ω and $a_2(\omega)$ that of the BB seismograph. In this report all SP seismograms displayed are as they would have been recorded by a WWSS SP seismograph with response as shown in figure 2.

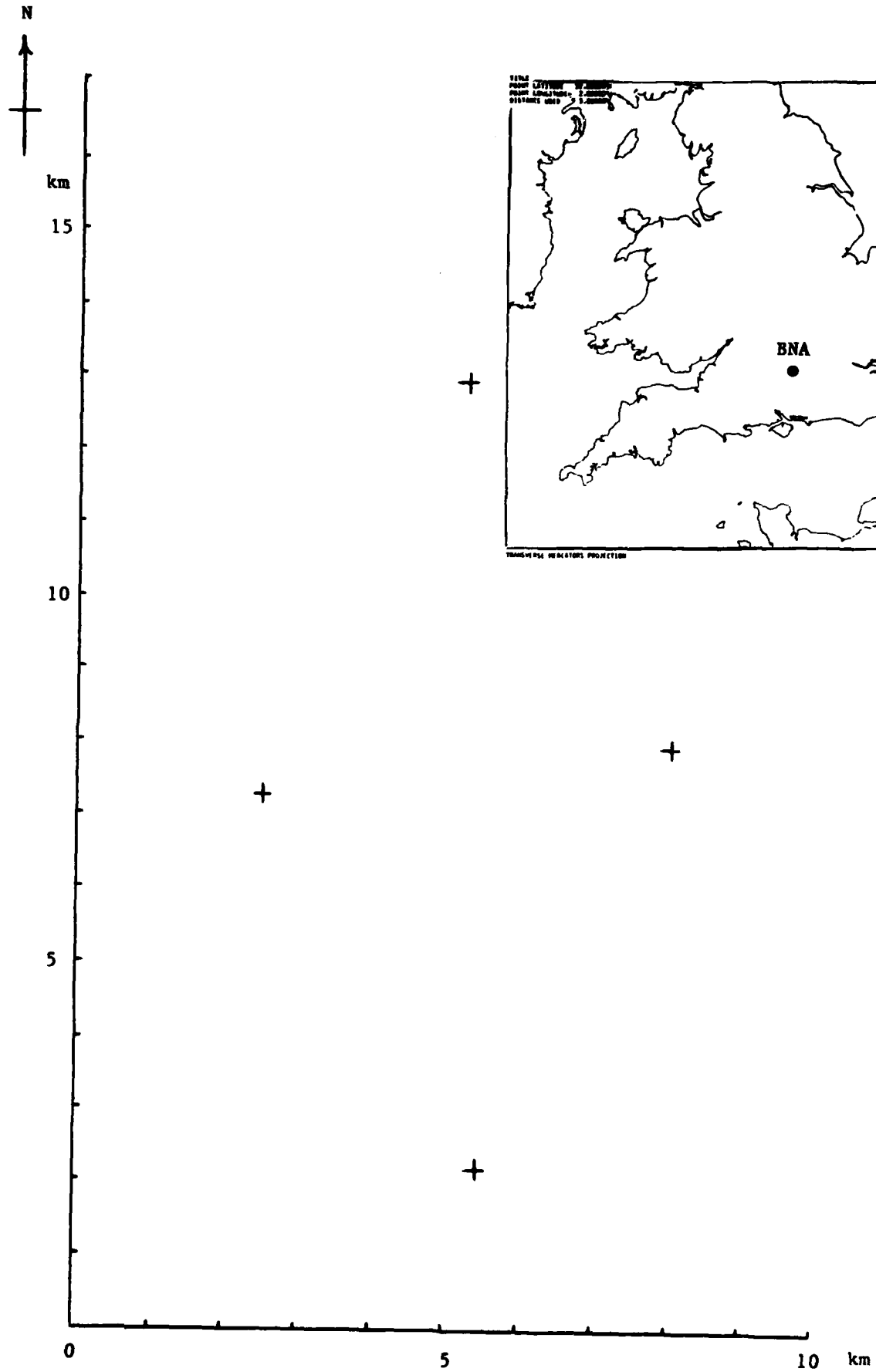


FIGURE 1. LAYOUT OF SEISMOMETERS IN THE BLACKNEST ARRAY (BNA) IN SOUTHERN ENGLAND. INSET SHOWS GEOGRAPHIC POSITION OF THE ARRAY. + SEISMOMETERS OF THE ARRAY. THE ORIGIN OF THE COORDINATE SYSTEM IS ARBITRARY.

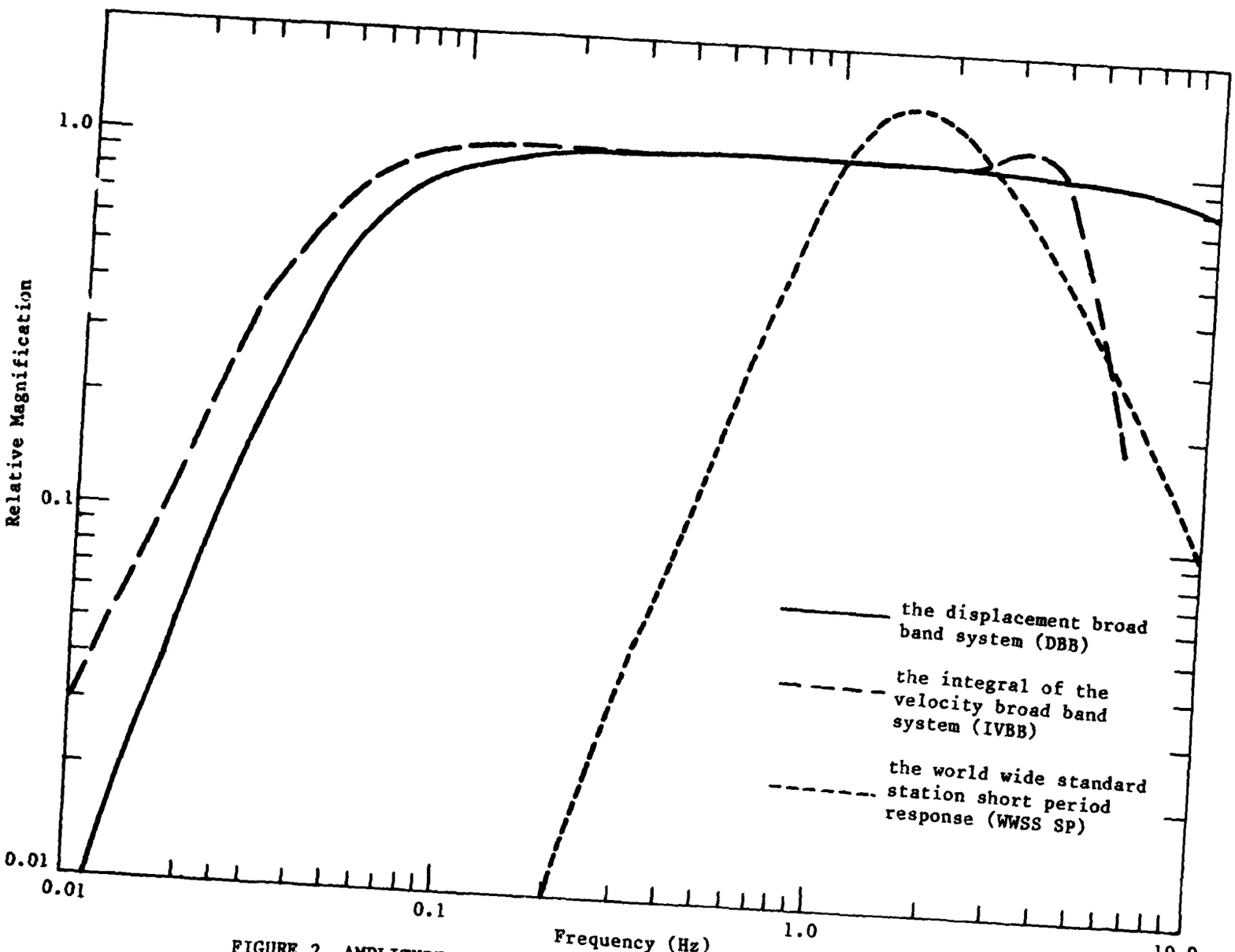


FIGURE 2. AMPLITUDE AND PHASE RESPONSE TO DISPLACEMENT AS A FUNCTION OF FREQUENCY FOR THE RECORDING-PLAYBACK SYSTEMS USED

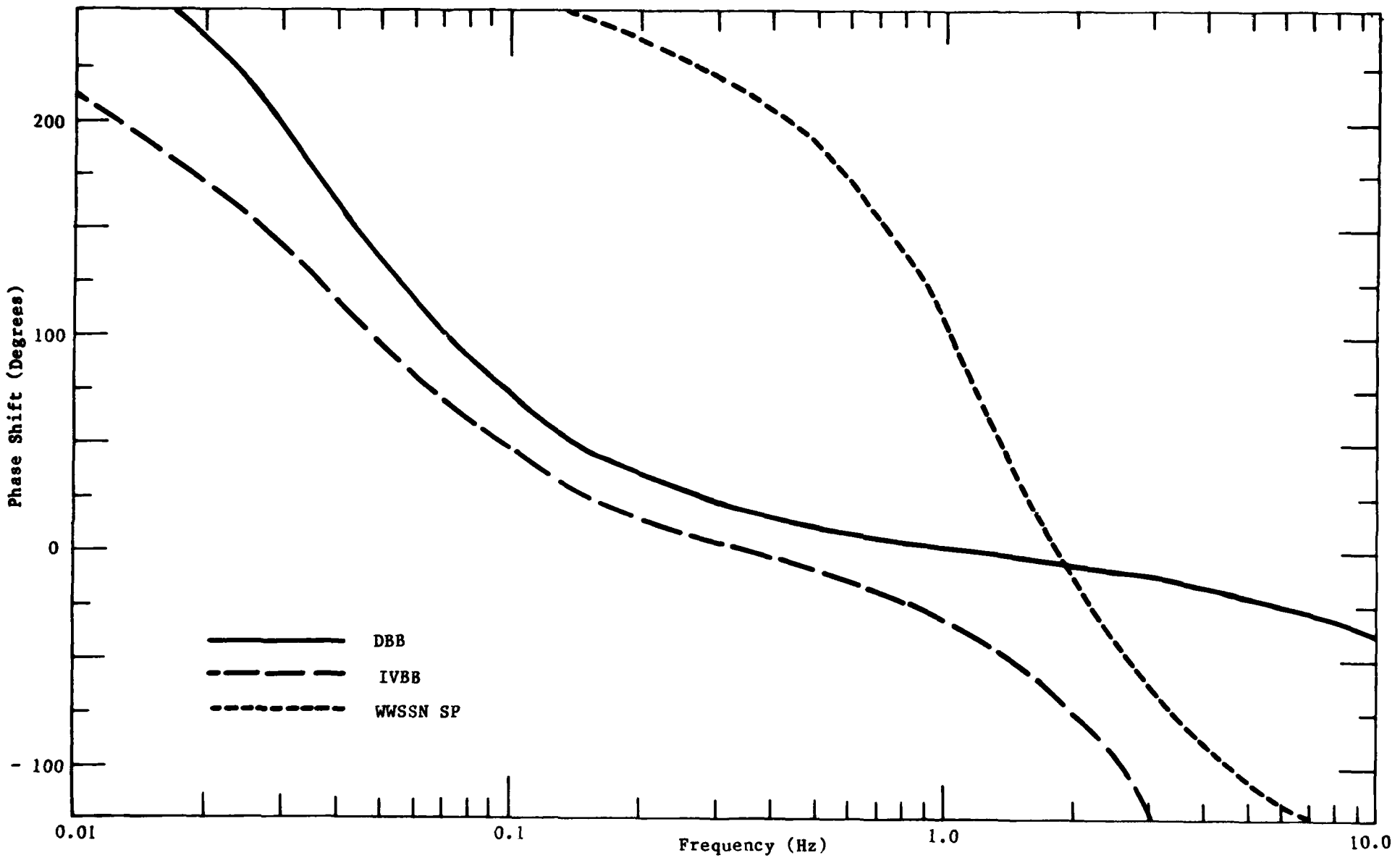


FIGURE 2 (CONTINUED)

The BNA is sited in a region where about 1 km of Mesozoic and Tertiary sediments lie unconformably on a Palaeozoic basement which is part of the London Platform. Superficial deposits of sands and gravels cover much of the region. The basement is cut by a fault zone striking roughly east-west across the array in the vicinity of the most southerly seismometer and the sediments increase rapidly in thickness from north to south in this region. The most southerly seismometer is emplaced in the Chalk (Cretaceous) whereas the other three seismometers are emplaced in the superficial sands and gravels. A seismic reflection survey has been carried out in the area in an attempt to obtain more detail about the geology beneath the array and a report on this is being prepared (12).

3. SEISMIC SIGNALS AND NOISE AT THE BLACKNEST ARRAY

In this section we demonstrate using a few samples some of the properties of seismic signals and noise as seen at the BNA. In particular we look at noise levels, at the coherence of signals and noise and at the value of SP as compared to BB signals.

The oceanic microseisms recorded at the BNA can be very large when there are storms in the North Atlantic. Such storms are commonest during the winter and so the noise level recorded during the winter is usually much higher than during the summer. Table 1 lists the average rms amplitude of the BB noise for a sample of noise taken on one day of each of the first six months of 1976; the average rms amplitude is

$$\left(\sum_{i=1}^{n} \sigma_i^2/n\right)^{\frac{1}{2}},$$

where σ_i^2 is the mean square amplitude of the noise on channel i and n (= 4) is the number of seismometers. Note that the BB noise level on 20 January is about 10 times that on 20 May. Table 1 gives the noise reduction obtained for each BB noise sample by DS processing; the quantity listed in table 1 is Φ_{DS} which is given by

$$\Phi_{DS} = \{(\sum_{i=1}^{n} \sigma_i^2)/(n\sigma_{DS}^2)\}^{\frac{1}{2}}, \qquad \ldots.(1)$$

TABLE 1

Blacknest Array Noise Amplitudes and the Noise Reductions Obtained by Delay and Sum Processing

	Broad Band		Short Period	
Noise Sample	Average RMS Amplitude, nm	Noise Reduction on Summing ($\bar{a}_{DS}$)	Average RMS Amplitude, nm	Noise Reduction on Summing ($\bar{a}_{DS}$)
20 January 1976	2069	1.56	43.1	1.95
21 February 1976	1447	1.72	24.1	2.03
20 March 1976	589	2.04	36.5	1.93
20 April 1976	453	2.31	21.8	1.91
20 May 1976	209	2.17	29.7	1.98
20 June 1976	260	1.77	27.6	1.88

where σ^2_{DS} is the mean square amplitude on the DS output and n and σ^2_i are as defined above. For uncorrelated noise the expected value of Φ_{DS} is 2. Table 1 lists Φ_{DS} for DS processing of the six BB noise samples for zero delays, that is for straight summing.

Measurements of the properties of the SP noise have been made from SP seismograms derived from the BB; table 1 shows the average rms amplitude of the SP noise (for magnification of unity at 1 Hz) computed for the same time window as for the BB noise. From the values given in table 1 it is clear that the BNA is on a site with high SP noise. The SP noise is apparently uncorrelated between channels and summing the 4 channels of the BNA without delays gives (for SP noise) noise reductions (table 1) of close to 2.

Figure 3(a) shows the power spectra of the 6 samples of BB noise (the spectra are averages over the power spectra of the 4 channels of the array). From these spectra it can be seen that, when the noise amplitude is large (20 January and 21 February), the spectra are sharply peaked at 6 to 8 s period, whereas when the noise is of lower amplitude (20 May and 20 June), there is only a weak maximum in the power spectrum and this lies at periods shorter than 6 s. Note also that, whereas the 6 to 8 s noise power varies by a factor of more than a 1000 over the 6 noise samples, at around 2 s period the power varies by only about a factor of 3.

In order to measure the coherency of the 6 to 8 s noise the normalised cross-correlation function for pairs of channels has been computed for the 20 January noise sample. The peak value of the cross-correlation function is about 0.7 for the two most widely spaced seismometers and is greater than 0.75 for the other pairs of seismometers in the array. These results show that the 6 to 8 s seismic microseisms were highly coherent for this noise sample. That the oceanic microseisms in the noise sample were highly correlated between seismometers can be seen from the seismograms and by measuring arrival times of peaks and troughs in the wave train the phase velocity of the large amplitude microseisms is estimated to be about 3 km/s in a direction N139^{o}E so on that day the source of the microseisms lay to the north-west of the array. Figure 3(b) shows a sample of the BB noise recorded on 20 January by the 4 elements of the BNA; the individual channels have been time shifted using the estimated velocity to bring the peaks and troughs into phase across the array. Also displayed in figure 3(b) is the DS output - the sum of the time shifted channels.

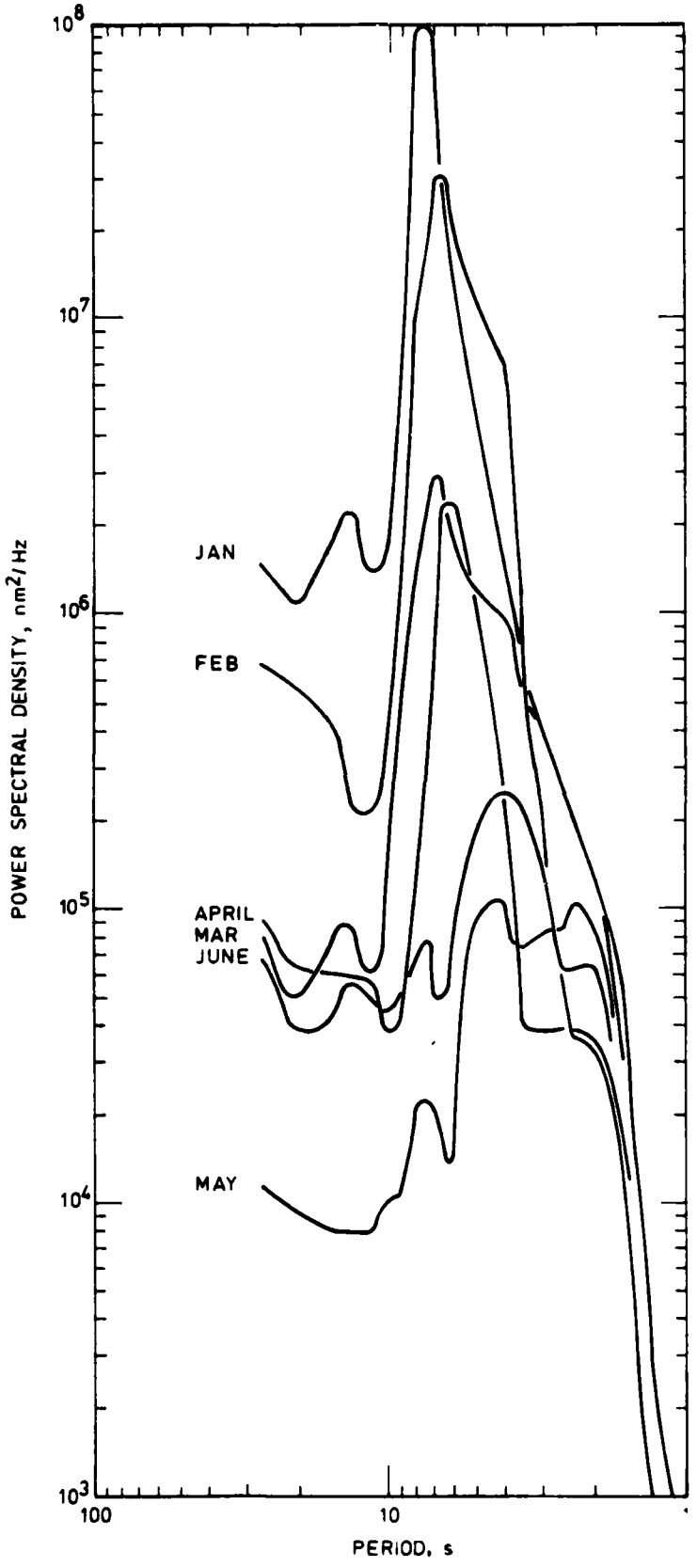


FIGURE 3(a) EXAMPLES OF BB(IVBB) NOISE RECORDED AT THE BNA. POWER SPECTRA OF THE 6 SAMPLES OF BB NOISE

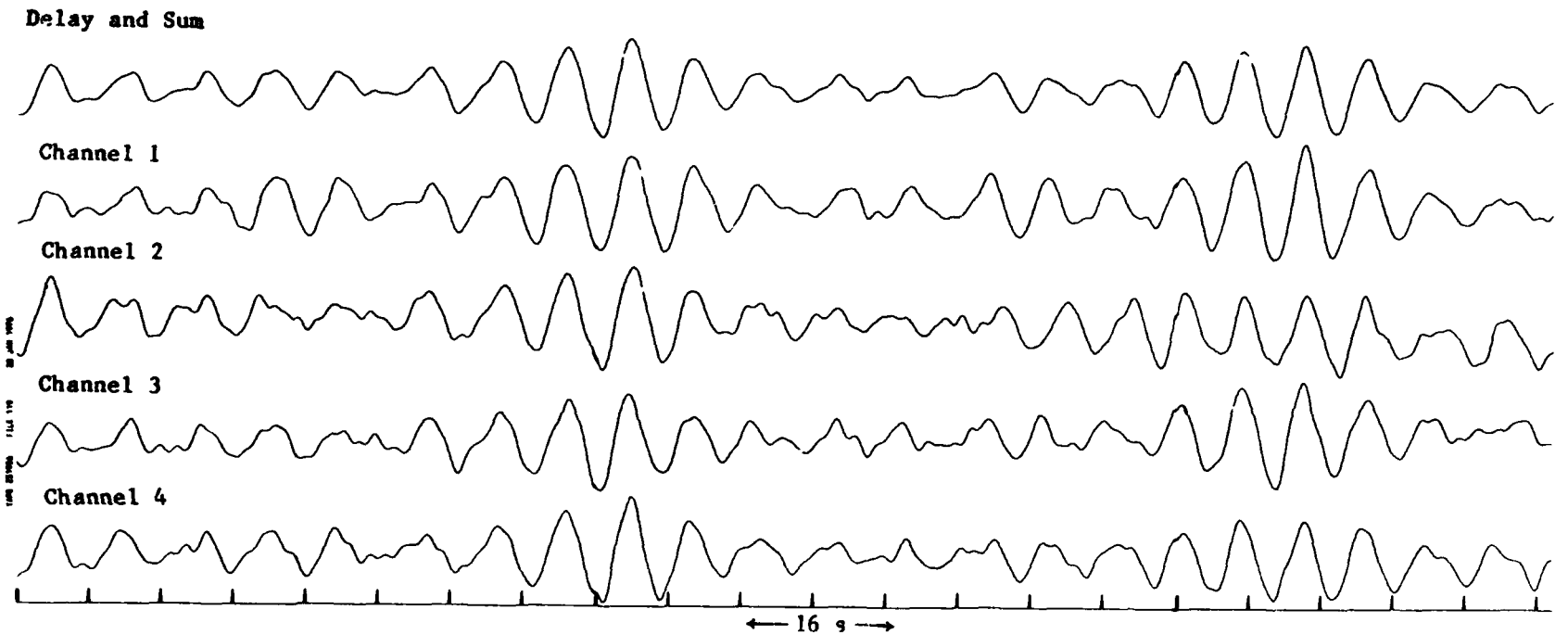


FIGURE 3(b) EXAMPLES OF BB(IVBB) NOISE RECORDED AT THE BNA. SAMPLE OF BB NOISE RECORDED ON 20 JANUARY 1976 FROM THE 4 SEISMOMETERS OF THE BNA; THE INDIVIDUAL CHANNELS HAVE BEEN TIME SHIFTED USING THE ESTIMATED VELOCITY OF THE OCEANIC MICROSEISMS TO BRING THE PEAKS AND TROUGHS INTO PHASE ACROSS THE ARRAY. ALSO DISPLAYED IS THE DS OUTPUT

When the amplitude of the 6 to 8 s period oceanic microseisms is low there is usually no significant correlation between the noise on pairs of channels. This may indicate that this low amplitude noise is not spatially organised or, what is more likely, that the noise field is more isotropic. At periods away from the 6 to 8 s period the coherence between pairs of channels is also low but again this may be because the noise field at these periods is isotropic rather than that the noise is not spatially organised. A detailed study of the noise is required to decide which is the most likely interpretation of these coherence measurements.

In order to measure the coherence of signals we use in this report Φ_{DS} as defined in equation (1) where now σ_i^2 and σ_{DS}^2 are the mean square amplitudes of the signal on channel i and on the DS output respectively. For a signal that is identical in all channels (in-phase and of equal amplitude), Φ_{DS} should be 1.0; if the signal is uncorrelated between channels, the expectation of Φ_{DS} is 2. Treating the section of BB noise shown in figure 3(b) as a signal and forming the DS output (as shown in figure 3(b)) using the estimated velocity of the noise, then $\Phi = 1.06$; this value of Φ_{DS} so close to 1.0 confirms yet again that these microseisms are highly correlated.

We now look at body wave signals recorded at the BNA to see how well they are correlated across the array. Figure 4 shows BB P signals from two earthquakes and an explosion (see table 2 for details) as recorded by each channel of the array. Consider first the Kodiak Island earthquake (figure 4(a)); clearly the main arrivals in the signal have similar shapes on all channels. The DS output for the Kodiak Island signals shown in figure 4(a) has been formed using the value of the apparent surface velocity computed using the estimated hypocentre. The value of Φ_{DS} computed using the mean square of the DS output is 1.04. This shows that any signal loss due to DS processing is negligible so that the effectiveness of DS processing for such highly coherent signals can be measured simply by computing the noise reduction.

The signal shown in figure 4(b) has too small a signal-to-noise ratio at the onset to give any indication of coherence. The signal does, however, show large amplitude low frequency arrivals 40 s after onset and from these it is possible to get an idea of the coherence of body wave signals at these low frequencies (~ 0.125 Hz); Φ_{DS} for these low frequency arrivals is 1.02, showing again that any signal loss on DS processing can be neglected.

TABLE 2

Earthquakes and Explosion Used in Coherence Studies

Date	Region	Origin Time	Depth, km	Distance, degrees	Obs SP Amp, nm	SP Ground Motion, nm	SP Period T, s	m_b^A	m_b (NEIS)	Max BB Amp	BB Period T, s
22 August 1973	Kodiak Island	18:14:37.2	38	69.3	305	305	1.0	6.4	5.9	1420	2.7
20 December 1976	Vancouver Island	20:33:7.8	10	70.5	105	538.3	1.7	6.4	5.9	3150	8.0
21 October 1975	Novaya Zemlya	11:59:57.3	0	32.3	1115	1115	1.0	6.7	6.5	2106	1.6

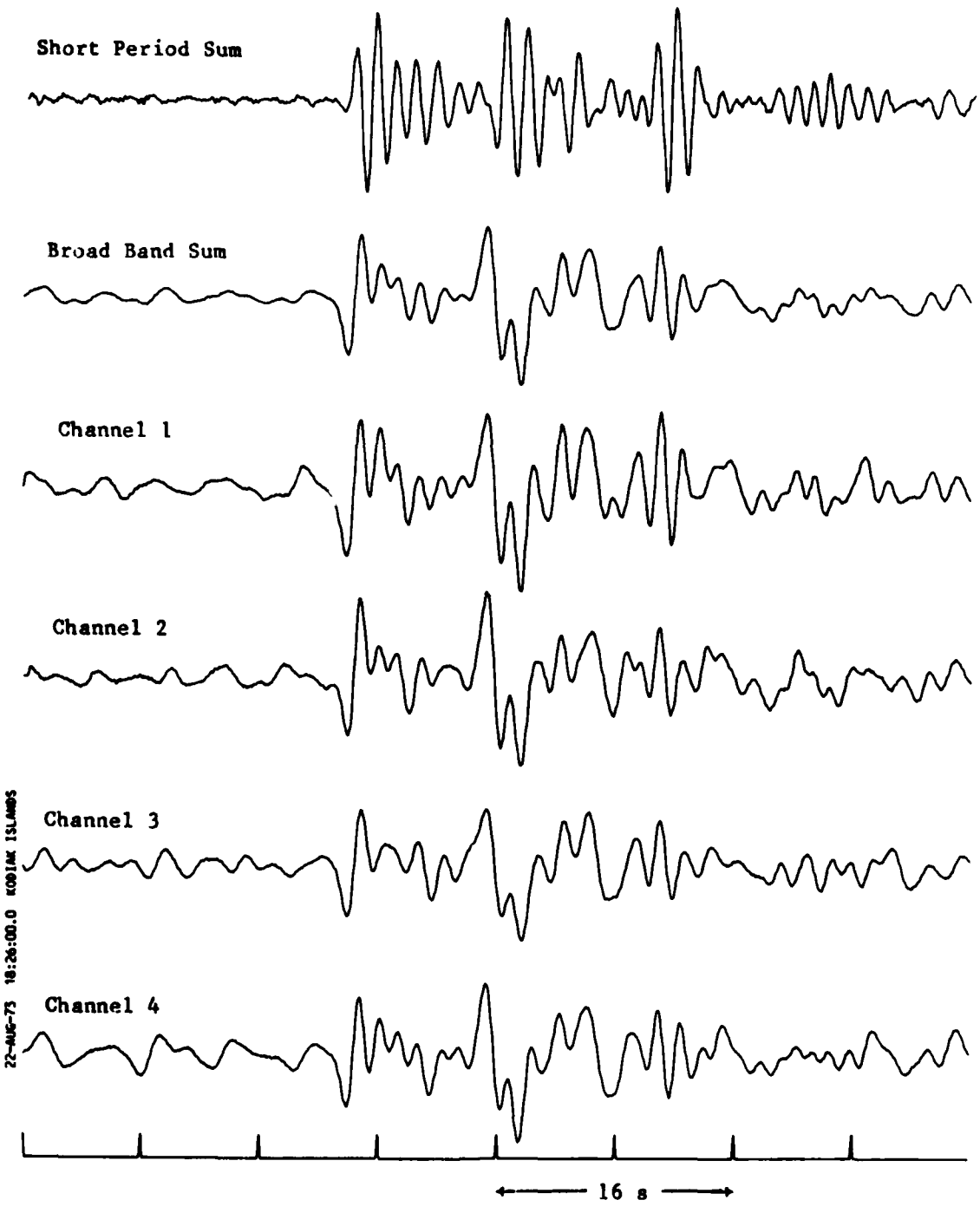


FIGURE 4(a) P-WAVE SEISMOGRAMS RECORDED AT THE BNA FOR THE KODIAK ISLAND EARTHQUAKE OF 22 AUGUST 1973. THE BB (DBB) SINGLE CHANNEL AND DELAY AND SUM OUTPUTS ARE SHOWN TOGETHER WITH THE SP DELAY AND SUM OUTPUT

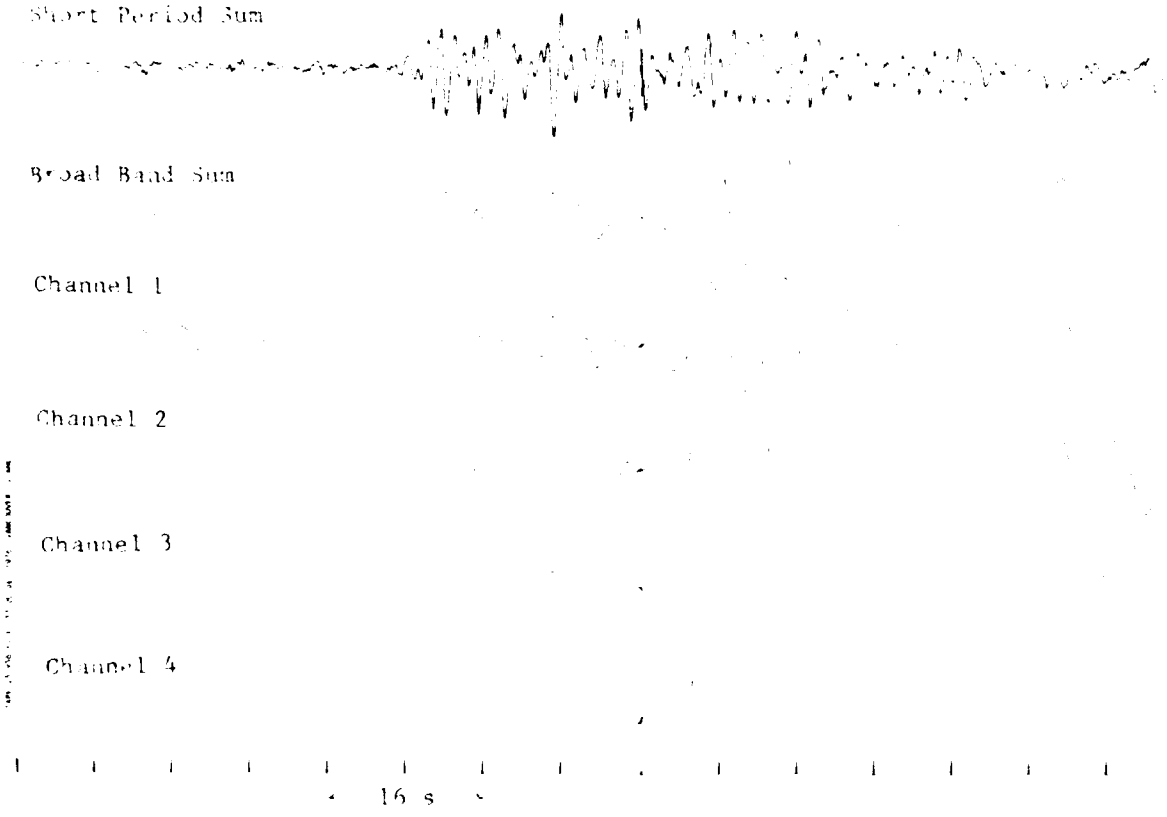


FIGURE [illegible] P-WAVE SEISMOGRAMS RECORDED AT THE BNA #3R FOR THE VANCOUVER ISLAND EARTHQUAKE OF 20 DECEMBER 1996, [illegible] (UTC). SINGLE CHANNEL AND BEAM AND SUM OUTPUTS ARE SHOWN TOGETHER WITH THE SP DELAY AND SUM OUTPUT

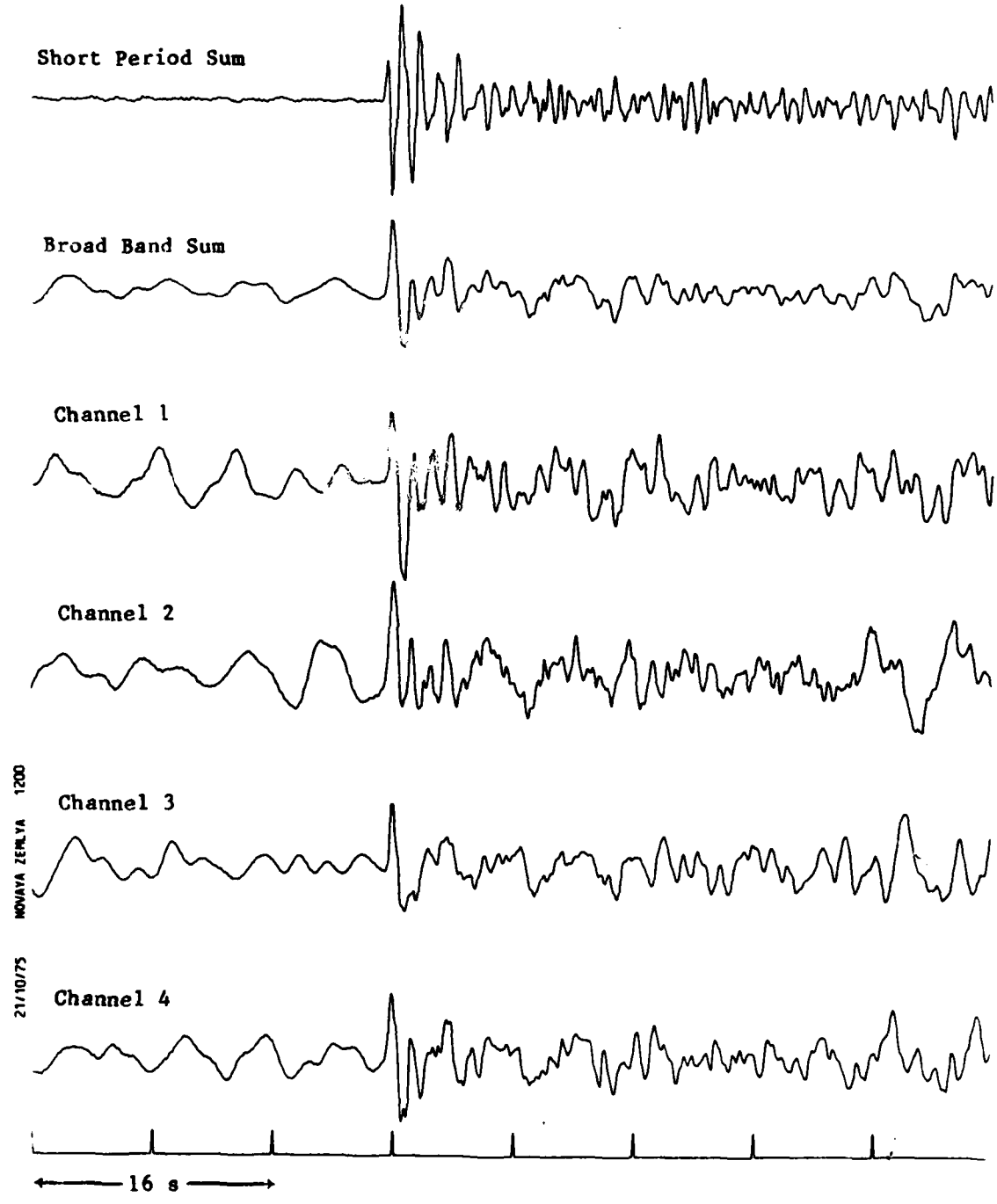


FIGURE 4(c) P-WAVE SEISMOGRAMS RECORDED AT THE BNA FOR THE NOVAYA ZEMLYA EXPLOSION OF 21 OCTOBER 1975. THE BB (DBB) SINGLE CHANNEL AND DELAY AND SUM OUTPUTS ARE SHOWN TOGETHER WITH THE SP DELAY AND SUM OUTPUT

The high coherence shown by the two earthquake signals (figures 4(a) and 4(b)) seems to be typical of signals recorded at the BNA. However, one group of P signals, those from explosions in Novaya Zemlya (NZ), seem to be less coherent than the earthquake signals.

An example of such a P signal is given in figure 4(c) and other details of the explosion in table 2. Inspection shows that the signal differs in shape across the array even in the first 2 or 3 s after onset; for the first 20 s of the signal $\Phi_{DS} = 1.27$ and for the 20 s of the signal starting 10 s after onset $\Phi_{DS} = 1.53$. These values of Φ_{DS} show that the explosion signal is not as coherent as the two earthquake signals and we show later that the departure from perfect coherence shown by the NZ signal can lead to difficulties when using Wiener filtering to suppress the noise.

In order to measure the coherence of P signals at frequencies around 1 Hz Φ_{DS} has been computed for the SP signals for the two earthquakes and the explosion discussed above. These values are listed in table 3, together with those computed for the BB signals. Note that for the earthquake and the first 20 s of the explosion signal the values of Φ_{DS} computed for the SP records are little different from those computed for the BB records. For the 20 s of the explosion signal starting 10 s after onset the value of Φ_{DS} is much larger on the SP than on the BB seismogram and is close to the value of 2 expected for uncorrelated noise.

Apart from the NZ explosion, however, the signals recorded at the BNA seem to be sufficiently coherent so that little signal loss is to be expected on DS processing and, as we see later, there is no difficulty in applying Wiener or MP filtering.

The reason for the low coherence of the NZ explosion is not clear. Explosion signals recorded from NZ are always more complex at the BNA than is usual for explosion signals. Some of this loss of coherence can presumably be attributed to scattering somewhere close to the array; the existence of such arrivals at other arrays has been demonstrated by Key (13). There is also evidence of structural complexity on the NZ-BNA path (14) and this may also contribute to the lack of coherence of the NZ explosion signals.

TABLE 3

Coherence Measurements for P Signals

Signal Source	Duration of Signal, s	Start Time of Signal	Coherence Measure (Φ_{DS}) Broad Band	Short Period
Kodiak Island Earthquake	20	Signal Onset	1.04	1.08
Vancouver Island Earthquake	20	Signal Onset	-	1.09
	20	40 s after onset	1.02	-
Novaya Zemlya Explosion	20	Signal Onset	1.27	1.34
	20	10 s after onset	1.53	2.12

The seismograms shown in figure 4 illustrate some of the advantages of broad band seismograms compared to SP. For BB explosion signals the first motion may be the largest amplitude on the record whereas on the SP the first motion is usually only a third or less of the maximum amplitude, as is illustrated by figure 4(c). For the Kodiak Island earthquake (figure 4(a)) the broad band seismogram shows two clear arrivals of opposite polarity; the first (P) has negative polarity whereas the second (presumably pP) has positive polarity. Both these pulses appear to have leading edges that show no discontinuity in gradient at the onset of the pulse. This absence of any discontinuity in the gradient shows up on the SP seismogram as a very small first motion (for both P and pP) so that, although two pulses are visible on the SP, their polarities are difficult to distinguish unambiguously. The BB shows something of the shape of the P and pP pulses; the leading edge of the pulses shows a smooth increase of amplitude with time whereas the trailing edge is marked by an abrupt drop in amplitude.

Turning now to the Vancouver Island earthquake, this shows that, whereas the SP seismogram displays only a small range of frequencies as is expected for such a narrow band system, on the BB seismogram the signal shows arrivals with frequencies ranging from around 0.125 to 1 Hz.

Table 2 lists the observed amplitude of the BB and SP signals; this amplitude is half the maximum peak-to-peak amplitude assuming each seismogram was recorded on a system with magnification unity at 1 Hz; this seems to be the most satisfactory way of specifying the signal amplitude to allow comparison of BB and SP amplitudes. In order to get the ground motion the observed amplitude must be divided by the relative magnification at the period of the signal. The observed SP and BB amplitudes given in table 2 show what is almost always found that the BB amplitude is greater than or equal to the SP amplitude (see also table 5, section 6).

Two magnitudes are given in table 2 for each earthquake and for the explosion; one is the body wave magnitude (m_b) taken from the National Earthquake Information Service (NEIS) bulletin, the other is the body wave magnitude (m_b^A) computed from the short period DS records shown in figure 4. If the data shown in table 2 are combined with that given in table 5 (section 5), it is clear that usually m_b^A is significantly greater than m_b, that is the BNA amplitudes are on average larger than would be expected given the NEIS m_b.

The few examples given of signal and noise at the BNA are typical. Thus, although the pass band of the BB seismograph contains the large noise peak at 6 to 8 s period, when the noise in this period band has large amplitude it is usually highly coherent and of low speed so that it should be possible to extract high speed body wave signals from such noise using an array. Further, the BB signals are often of larger amplitude than the SP so that, although noise may be large on BB seismograms, the signal may also be large so that the problem of extracting BB signals from noise is made easier. If BB signals can be extracted from the noise, then this will usually be worthwhile, because, as demonstrated here, these signals may show important features that cannot be seen on the SP.

4. PROCESSING METHODS

In this section we outline the theory of DS, Wiener and MP processing methods and compare their properties. The theory of these processing methods has been discussed in many papers (for DS processing see, for example, reference (15); for Wiener filtering see, for example, references (3) and (16); for MP filtering see, for example, references (6) and (17); all these papers give comprehensive lists of references). Here we do not attempt to give a full description of the theory; most of the discussion is based on the processing of a two channel array; the purpose of this is to illustrate some of the properties of the processing methods, particularly their similarities and differences.

All linear array processing methods for the extraction of signals from noise are examples of the general process of multichannel filtering; each data channel is passed through a filter (in the general case each channel has a different filter) and the filter outputs are summed. Multichannel filters can be constructed either in the time domain or in the frequency domain. Burg (3) states that, in practice, time domain methods of estimating Wiener filters give better results than those derived in the frequency domain. Capon et al. (6) find that there are some advantages in using the frequency domain rather than the time domain method for the estimation of MP filters, the main advantage being that estimating the filters in the frequency domain takes less computer time than the time domain method. In this report all filters are estimated in the time domain.

For simplicity we discuss multichannel processing as applied to sampled data although the process can be applied to continuous data (see, for example, reference (3)). We assume that the sampling interval is unity which simplifies the discussion without any loss of generality. Let the output of channel j be

$$\ldots;\ s(t-1) + x_j(t-1);\ s(t) + x_j(t);\ s(t+1) + x_j(t+1);$$
$$s(t+2) + x_j(t+2);\ \ldots;\ s(t+m-1) + x_j(t+m-1)\ldots,$$

where $x_j(k)$ is the noise on channel j and s(k) the signal (assumed to be the same on all channels) at time k; the signal is assumed to have been aligned on all channels before processing and this assumption is made throughout the report. Consider a two seismometer array and assume that the filter response for any channel can be represented by three points; thus for channel 1 the impulse response is $w_1(-1)$, $w_1(0)$, $w_1(1)$ and is zero elsewhere and for channel 2 the response is $w_2(-1)$, $w_2(0)$ and $w_2(1)$ (extension to the general case of n seismometers and p filter points per channel is not difficult but it is cumbersome).

The output of the two channel filter process z(t) can be written as a convolution:-

$$\begin{aligned} z(t) = {} & x_1(t+1)w_1(-1) + x_1(t)w_1(0) + x_1(t-1)w_1(1) \\ & + x_2(t+1)w_2(-1) + x_2(t)w_2(0) + x_2(t-1)w_2(1) \\ & + s(t+1)\{w_1(-1) + w_2(-1)\} + s(t)\{w_1(0) + w_2(0)\} \\ & + s(t-\ )\{w_1(1) + w_2(1)\}. \end{aligned} \qquad \ldots.(2)$$

Note that here the output at time t makes use, not only of data from time t and earlier, but also from later time, that is time t + 1. Clearly this is impossible if the multichannel filtering is operating in real time, for then the filters must be causal, that is the impulse response of the filters must be zero for time $t < 0$. In practice, however, multichannel filtering is usually carried out on recorded data so there is no difficulty in using non-causal filters. In this simple example the length of the non-causal filters before and after time zero are equal but this

need not be so. However, there are usually advantages, as illustrated later, in using such equal-sided filters and, when non-causal filters are used in this report, they are always of this type; we refer to these filters as two-sided filters and to causal filters as one-sided filters.

Signal-to-noise improvement by multichannel filtering is only possible if there are differences in the properties of the signal and noise; these differences in the signal and noise may be either in their frequency spectra or in their spatial properties. If the signal and noise at an array are made up of plane waves, then multichannel filtering can be viewed as multi-dimensional filtering (3). Plane waves at an array can be represented in three-dimensional frequency-wave number space; axes κ_x, κ_y and ω where κ_x and κ_y are two wave number axes at right angles and ω is the frequency axis; if the signals are aligned on all channels, then for the signal effectively $|\underline{\kappa}| = 0$. The signal can be enhanced relative to any noise at zero wave number by applying a frequency filter that passes those frequencies where the signal amplitude is large relative to noise and attenuating frequencies where the noise is relatively large. If at any frequency the noise and signals have different wave numbers, then a wave number filter can be used to attenuate the noise and pass the signal. As the vector wave number $\underline{\kappa} = \omega/\underline{c}$, where $\underline{c}$ is the apparent surface velocity, wave number filtering at frequency ω is essentially separating signal and noise on differences in velocity.

Wave number filtering is a form of spatial filtering. This is not the only form of spatial filtering that is possible for, as shown by Capon et al (17), signal-to-noise improvements can be obtained without using frequency filtering even when the signal and noise have the same wave number. For example, consider a two channel array where the noise is in-phase and perfectly correlated between the two channels (with correlation coefficient unity) but differs slightly in amplitude between the two channels. Thus, the signals and noise both have zero wave numbers ($|\underline{c}|$ is infinite) yet, as shown by Capon et al.(17), it is possible for such noise to find multichannel filters that reduce the noise to zero and pass the signal undistorted (the filters are in fact those given in equation (18) below).

In general, multichannel filtering applies both frequency and spatial filtering where spatial filtering is taken here to include both true wave number filtering and noise reduction that arises because of differences between channels in the amplitude of the noise.

4.1 Delay and sum processing

Putting $w_1(0) = w_2(0) = 1/2$ and $w_1(k) = w_2(k) = 0$ for $k \neq 0$ in equation (2) gives

$$z(t) = s(t) + \{x_1(t) + x_2(t)\}/2; \qquad \ldots.(3)$$

thus, the output z(t) is the mean of the channels at time t and is equivalent to DS processing for zero delay. The process of inserting delays can also be looked on as filtering so DS processing is a particular case of multichannel filtering.

For noise of equal variance on all channels and uncorrelated between channels (usually referred to simply as random noise), then DS processing gives a noise reduction of $n^{\frac{1}{2}}$ and this is the best that can be done. If the noise is uncorrelated but has variance σ_i^2 on channel i, then the best signal-to-noise improvement is obtained by weighting channel i by a factor proportional to $(\sigma_i^2)^{-1}$ and summing (15); for the two channel array this is equivalent to applying multichannel filters with:-

$$w_1(0) = (\sigma_1^2)^{-1}\ \{\sigma_1^2\sigma_2^2/(\sigma_1^2 + \sigma_2^2)\},$$
$$w_2(0) = (\sigma_2^2)^{-1}\ \{\sigma_1^2\sigma_2^2/(\sigma_1^2 + \sigma_2^2)\} \qquad \ldots.(4)$$

and $\quad w_1(k) = w_2(k) = 0$ for $k \neq 0$;

this process is usually referred to as weighted DS.

The effect in frequency-wave number space of DS processing is to apply a wave number filter with response that is the same at all frequencies. The response passes the signal (at $|\underline{\kappa}| = 0$) unattenuated and suppresses noise (which includes unwanted signals) at wave numbers away from zero. The wave number response for the BNA for DS processing is shown in figure 15. For random noise the noise at any frequency is uniformly distributed with wave number and the effect of DS processing can be thought of as applying to the noise at each frequency, wave number filters of the type shown in figure 15.

If the noise is concentrated at a wave number ($|\underline{\kappa}| \neq 0$) where the wave number response for the array is zero, then the noise will be reduced by much more than a factor of $n^{\frac{1}{2}}$; a two seismometer array with seismometers separated by half the wavelength of the predominant noise is an example of an array with such a response. In general, however, DS processing because it applies a fixed wave number filter will not give optimum noise reduction.

4.2 Wiener filtering

Wiener filters for application to sampled data can be derived as follows. Consider a section of array observations extending from time t to time t + m - 1, that is m observations of signal plus noise from each channel; the noise and signal are assumed to have stationary properties. For the two channel case we put

$$\underline{X}_1 = \begin{bmatrix} x(t+1), & x_1(t) & 0 \\ x_1(t+2), & x_1(t+1), & x_1(t) \\ \vdots & \vdots & \vdots \\ x_1(t+m-1), & x_1(t+m-2), & x_1(t+m-3) \\ 0 & x_1(t+m-1), & x_1(t+m-2) \end{bmatrix},$$

$$\underline{W}_1 = \mathrm{col}\{w_1(-1), w_1(0), w_1(1)\}$$

and define $\underline{X}_2$ and $\underline{W}_2$ in a similar way for channel 2, then if $\underline{Z} = \mathrm{col}\{z(t), z(t+1), \ldots z(t+m-1)\}$ is the output after multichannel filtering and

$$\underline{S} = \begin{bmatrix} s(t+1), & s(t) & 0 \\ s(t+2), & s(t+1), & s(t) \\ \vdots & \vdots & \vdots \\ s(t+m-1), & s(t+m-2), & s(t+m-3) \\ 0 & , s(t+m-1), & s(t+m-2) \end{bmatrix},$$

multichannel filtering can be written for the two channel case

$$(\underline{X}_1 + \underline{S})\underline{W}_1 + (\underline{X}_2 + \underline{S})W_2 = \underline{Z}. \qquad \ldots.(5)$$

Suppose now we require z(k) to be as close to s(k) as possible; the filters required are such as to convert signal plus noise into the best estimate of the signal. If the difference between z(k) and s(k) is

$$\varepsilon(k) = s(k) - z(k),$$

putting

$$\underline{\varepsilon} = \text{col}\,[\varepsilon(t), \varepsilon(t+1) \ldots., \varepsilon(t+m-1)]$$

z(k) will be the best least square estimate of s(k) if $\underline{\varepsilon}^T\underline{\varepsilon}$ is a minimum provided the expectation of $\underline{\varepsilon}$ is zero, but if $\underline{\varepsilon}^T\underline{\varepsilon}$ is a minimum, then $\underline{\varepsilon}^T\underline{\varepsilon}/m$, the mean square error is also a minimum so that the best least squares estimate of the signal can also be obtained by minimising this mean square error. The advantage of using the mean square error is that as t tends to $-\infty$ and m increases without limit $\underline{\varepsilon}^T\underline{\varepsilon}/m$ tends to a finite limit and other quantities converge to finite values as shown below.

Writing $\underline{W} = \text{col}\,[w_1(-1), w_1(0), w_1(1), w_2(-1), w_2(0), w_2(1)]$ the normal equations (assuming the expectation of the noise is zero and the noise and signal are uncorrelated) are

$$(\underline{R} + \underline{R}^S)\underline{W} = \begin{bmatrix} \underline{r}^S \\ \underline{r} \end{bmatrix}, \qquad \ldots.(6)$$

where

$$\underline{R} \approx \begin{bmatrix} r_{11}(0) & r_{11}(-1) & r_{11}(-2) & r_{12}(0) & r_{12}(-1) & r_{12}(-2) \\ r_{11}(1) & r_{11}(0) & r_{11}(-1) & r_{12}(1) & r_{12}(0) & r_{12}(-1) \\ r_{11}(2) & r_{11}(1) & r_{11}(0) & r_{12}(2) & r_{12}(1) & r_{12}(0) \\ r_{21}(0) & r_{21}(-1) & r_{21}(-2) & r_{22}(0) & r_{22}(-1) & r_{22}(-2) \\ r_{21}(1) & r_{21}(0) & r_{21}(-1) & r_{22}(1) & r_{22}(0) & r_{22}(-1) \\ r_{21}(2) & r_{21}(1) & r_{21}(0) & r_{22}(2) & r_{22}(1) & r_{22}(0) \end{bmatrix}$$

and $r_{ij}(1)$ is the cross-correlation of the noise on channels i and j at lag 1. As m increases the elements of $\underline{R}$ approximate more and more closely to $r_{ij}(1)$.

$$\underline{R}^s = \begin{bmatrix} \underline{A}, & \underline{C} \\ \underline{C}, & \underline{A} \end{bmatrix},$$

where $\underline{A} = \underline{S}^T\underline{S}$

$$S^TS = \begin{bmatrix} r^s(0), & r^s(-1), & r^s(-2) \\ r^s(1), & r^s(0), & r^s(-1) \\ r^s(2), & r^s(1), & r^s(0) \end{bmatrix}$$

and $r^s(1)$ is the autocorrelation of the signal at lag 1. The sub-matrix $\underline{C}$ is the cross-correlation of the signals between two channels, but as we have assumed that the signal has been time shifted so that the onset time is the same on all elements $\underline{C} = \underline{A}$. The vector $\underline{r}^s = \text{col}\,[r^s(1), r^s(0), r^s(-1)]$. Wiener filters can then be found by solving equation (6) provided that $\underline{R}$ and $\underline{R}^s$ are known.

We now re-write $\underline{W}$ as $\underline{W}(2|3)$ where

$$\underline{W}(2|3) = \text{col}\,[w_1(2|-1), w_1(2|0), w_1(2|1), w_2(2|-1), w_2(2|0), w_2(2|1)].$$

Defined in this way $\underline{W}(2|3)$ is to be understood as specifying a multichannel filter with 3 points per channel and with the filter coefficient for time t = 0 lying at position 2 in each filter. Then we can define two other 3 point filters:-

$$\underline{W}(1|3) = \text{col}\,[w_1(1|0), w_1(1|1), w_1(1|2), w_2(1|0), w_2(1|1), w_2(1|2)]$$

and

$$\underline{W}(3|3) = \text{col}\,[w_1(3|2), w_1(3|1), w_1(3|0), w_2(3|2), w_2(3|1), w_2(3|0)].$$

For $\underline{W}(1|3)$ the filter coefficients for time t = 0 lies at position 1 in each filter and the filters for each channel are causal filters, ie, the filter coefficients are zero for $t < 0$; similarly $\underline{W}(3|3)$ defines a filter for each channel that is zero for all $t > 0$. We can now re-write equation (6) to include all three filters, $\underline{W}(1|3)$, $\underline{W}(2|3)$ and $\underline{W}(3|3)$, as follows:-

$$(\underline{R} + \underline{R}^s)\,[\underline{W}(1|3), \underline{W}(2|3), \underline{W}(3|3)] = \begin{bmatrix} \underline{A} \\ \underline{A} \end{bmatrix}. \qquad \ldots.(6a)$$

In this simple case where there are 3 filter points per channel there are three possible sets of multichannel Wiener filters. In the general case of p filter points per channel there will be p possible sets of filters.

For an array of infinite extent and noise and signals that are plane waves the effect of Wiener filtering in frequency-wave number space as the spacing between seismometers tends to zero can be written (3)

$$H(\omega,\underline{\kappa}) = P^{S}(\omega,\underline{\kappa})/\{P(\omega,\underline{\kappa}) + P^{S}(\omega,\underline{\kappa})\}, \qquad \ldots.(7)$$

where $P^{S}(\omega, \underline{\kappa})$ is the signal power (assuming the signal is continuous) and $P(\omega, \underline{\kappa})$ the noise power at frequency ω and wave number $\underline{\kappa}$. As the required signal has a power spectrum that is strongly peaked at $|\underline{\kappa}| = 0$ and ideally zero elsewhere, then the effect of the filter is to suppress noise in the whole of frequency-wave number space, except at zero wave number. For $|\underline{\kappa}| = 0$ where the signal is large compared to noise the response tends to unity; where the signal is small, the response tends to $P^{S}(\omega, \underline{0})/P(\omega, \underline{0})$.

If the noise is all at wave numbers well away from zero, the noise can be suppressed completely and the signal left untouched. Usually, however, some of the noise will have zero wave number and then frequency filtering has to be applied to extract signals from noise. For practical arrays where the recordings are made at only a number of discrete points the effect of Wiener filtering is to apply filters that are smoothed versions of equation (7).

As the signal-to-noise ratio decreases, then at frequencies where the noise is large the response for $|\underline{\kappa}| = 0$ tends to $P^{S}(\omega, \underline{0})/P(\omega, \underline{0})$ and, if $P^{S}(\omega, \underline{0})$ is constant with ω, then the filter response is proportional to $P(\omega, \underline{0})^{-1}$. As short period narrow band seismographs have a response that is roughly the inverse of the noise spectrum, then in the frequency range where the noise is large, the amplitude response of the Wiener filters tends to that of conventional short period seismographs (provided that the signal amplitude is assumed to be constant with frequency).

4.3 Minimum power filtering

The method used in this section to derive MP filters is the direct method. An alternative method of derivation is to regard MP filters as a type of prediction error filter. Instead of computing MP filters which attempt to suppress the noise at the output, multichannel filters are computed which, acting on the noise in some time interval, predicts the noise on the DS output at some time t within the interval, with the constraint that any signals are suppressed. Signals should then be enhanced relative to noise on the prediction error output derived by subtracting the predicted noise from the DS output. The equivalence of this type of prediction error filtering and MP filtering is demonstrated in appendix A.

The expression for the minimum power filters is usually found by the method of maximum likelihood but it can be shown that the method of least squares yields the same result (6) and this latter method is used here. The basic equations of condition for the two channel case are

$$\underline{X}_1\underline{U}_1 + \underline{X}_2\underline{U}_2 = \underline{0}, \qquad \ldots.(8)$$

where

$$\underline{U}_1 = \text{col}\,(u_1(-1),\ u_1(0),\ u_1(1))$$

and

$$\underline{U}_2 = \text{col}\,(u_2(-1),\ u_2(0),\ u_2(1))$$

are the MP filters; that is filters are required that reduce the noise at the output to zero plus an "error". In order to ensure any signal present is passed unattenuated the filters are estimated by solving equation (8) with the constraints that $u_1(0) + u_2(0) = 1$ and $u_1(k) + u_2(k) = 0$ for $k \neq 0$. In order to apply these constraints to the two channel case with 3 point two-sided filters the following equations must also be satisfied:-

$$\underline{Q}^T\underline{U} = \underline{V}_2, \qquad \ldots.(9)$$

where $\underline{Q}^T$ is the transpose of $\underline{Q}$ and is given by

$$\underline{Q}^T = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}, \quad \underline{V}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \text{ and } \underline{U} = \begin{bmatrix} \underline{U}_1 \\ \underline{U}_2 \end{bmatrix}.$$

Using the method of Lagrangian multipliers it can be shown that the required filters are given by the solution of the normal equations

$$\begin{bmatrix} \underline{R} & \underline{Q} \\ \underline{Q}^T & \underline{0} \end{bmatrix} \begin{bmatrix} \underline{U} \\ \underline{\lambda} \end{bmatrix} = \begin{bmatrix} \underline{o} \\ \underline{V}_2 \end{bmatrix}, \qquad \ldots.(10a)$$

where $\underline{\lambda}$ is a vector of Lagrangian multipliers, $\underline{0}$ is a 3 × 3 matrix of zeros and $\underline{o}$ is a 6 element vector of zeros. Provided that $\underline{R}$ has an inverse then $\underline{U}$ can be found by eliminating $\underline{\lambda}$ from equation (10a) which gives

$$\underline{U} = \underline{R}^{-1}\underline{Q}(\underline{Q}^T\underline{R}^{-1}\underline{Q})^{-1}\underline{V}_2. \qquad \ldots.(10b)$$

The MP method can be thought of as multichannel filtering with a restricted set of filters, all filters derived by the method being purely spatial filters. To estimate the filters requires that $\underline{R}$ be specified.

Now let $\underline{V}_1 = \text{col}\,(1,0,0)$ and $\underline{V}_3 = \text{col}\,(0,0,1)$ and define three sets of MP filters $\underline{U}(1\,|\,3)$, $\underline{U}(2\,|\,3)$ and $\underline{U}(3|\,3)$ in the same way as the three sets of Wiener filters $\underline{W}(1\,|\,3)$, $\underline{W}(2\,|\,3)$ and $\underline{W}(3\,|\,3)$ are defined. Then we can re-write equation (10b) to include the three sets of MP filters:-

$$\begin{aligned} (\underline{U}(1\,|\,3),\underline{U}(2|\,3),\underline{U}(3|\,3)) &= \underline{R}^{-1}\underline{Q}(\underline{Q}^T\underline{R}^{-1}\underline{Q})^{-1}(\underline{V}_1,\underline{V}_2,\underline{V}_3), \\ &= \underline{R}^{-1}\underline{Q}(\underline{Q}^T\underline{R}^{-1}\underline{Q})^{-1}, \end{aligned} \qquad \ldots.(10c)$$

as $(\underline{V}_1,\underline{V}_2,\underline{V}_3)$ is the identity matrix.

If equation (10c) is used to estimate the MP filters, then the Lagrangian multipliers need not be computed. However, as shown in appendix B, it may be useful to obtain at least some of these multipliers as they provide a convenient way of estimating the mean square noise at the output after MP filtering.

4.4 Comparison of delay and sum, Wiener and minimum power processing

Consider first the case where the matrix $\underline{R}$ (equation (10a)) has a zero eigenvalue (the inverse of $\underline{R}$ cannot then be defined) and so there is an eigenvector $\underline{\mu}$ say such that $\underline{R}\underline{\mu} = 0$. Then if $\underline{\mu}$ also satisfies the equation

$$\underline{Q}^T\underline{\mu} = \alpha\underline{V}, \qquad \ldots.(11)$$

where α is a scalar constant and $\underline{V}$ is either $\underline{V}_1$, $\underline{V}_2$ or $\underline{V}_3$, then $\alpha^{-1}\mu$ defines a set of MP filters which reduce the noise to zero but pass the signal undistorted. Substituting for $\alpha^{-1}\mu$ in equation (6), however, shows that the Wiener filters will also satisfy this equation and so when $\underline{R}$ has a zero eigenvalue and the equivalent eigenvector satisfies equation (11), then MP and Wiener filters are identical; for this case Wiener filters are thus pure spatial filters. Kelly (18) has compared Wiener and MP filters for the more general case where all the eigenvalues of $\underline{R}$ are non-zero (they are then all positive) and thus $\underline{R}$ has an inverse and it can be shown that the solution of equation (6) can be written

$$\underline{W} = \underline{R}^{-1}\underline{Q}(\underline{Q}^T\underline{R}^{-1}\underline{Q})^{-1}\underline{F}^{-1}\underline{r}_s,$$

where

$$\underline{F} = \{(\underline{Q}^T\underline{R}^{-1}\underline{Q})^{-1} + \underline{A}\}.$$

Thus, using equation (10c),

$$\underline{W} = (\underline{U}(1|3),\underline{U}(2|3),\underline{U}(3|3))\underline{F}^{-1}\underline{r}_s \qquad \ldots.(12a)$$

and from equation (6a)

$$(\underline{W}(1|3),W(2|3),\underline{W}(3|3)) = (\underline{U}(1|3),\underline{U}(2|3),\underline{U}(3|3))F^{-1}\underline{A}. \qquad \ldots.(12b)$$

For simplicity we now confine the discussion to the relationship between the Wiener filter $\underline{W}(2|3) = \underline{W}$ and the MP filters. First, we consider how $\underline{W}$ varies as the amplitude of the signal increases assuming that $\underline{R}$ and the signal shapes are kept fixed. Writing the signal as $\beta s(t)$ where β is a simple multiplier, then $\underline{F}^{-1}\underline{r}_s$ is replaced in equation 12(a) by

$$\{(\underline{Q}^T\underline{R}\underline{Q})^{-1} + \beta^2\underline{A}\}^{-1}\beta^2\underline{r}_s$$

which can be written

$$\{\beta^{-2}\underline{A}^{-1}(\underline{Q}^T\underline{R}^{-1}\underline{Q})^{-1} + \underline{I}\}\begin{bmatrix}0\\1\\0\end{bmatrix}.$$

Thus, as β increases, that is as the signal-to-noise ratio increases, the term $\underline{F}^{-1}\underline{r}_s$ tends to col $(0,1,0)$ and so $W(2|3) \to \underline{U}(2|3)$, that is the Wiener filters tend to MP filters.

Now consider the case where the noise is not small relative to the signal. Then it turns out that $\underline{F}^{-1}\underline{r}_s$ can be regarded as the frequency filter applied to the signal (and to any component of noise that lies at $|\underline{\kappa}| = 0$ and has equal amplitude on all channels). For the effect of the Wiener filters on the signal is given by

$$(\underline{S}, \underline{S})\underline{W} = \begin{bmatrix} \cdot & \cdot & \cdot \\ s(t+1) & s(t) & s(t-1) \\ s(t+2) & s(t+1) & s(t) \\ \cdot & \cdot & \cdot \end{bmatrix}\underline{F}^{-1}\underline{r}_s, \qquad \ldots.(13)$$

which is simply the convolution of a frequency filter $\underline{F}^{-1}\underline{r}_s$ and the signal. $\underline{F}^{-1}\underline{r}_s$ can be thought of as the response of the multichannel filters to an impulse applied to all channels at time $t = 0$. Note that as $(\underline{S},\underline{S})\,\underline{W}$ can be written $\underline{S}\underline{W}^1$, where $\underline{W}^1 = W_1 + W_2$, then from equation (13) $\underline{F}^{-1}\underline{r}_s = \underline{W}^1$; that is summing the Wiener filters across channels gives $\underline{W}^1$ the impulse response of the frequency filter: which is a convenient way of obtaining this impulse response. We refer to $\underline{W}^1$ as the frequency component of the multichannel Wiener filters.

The effect of the Wiener filters on the noise can be written

$$\begin{bmatrix} \cdot & \cdot & \cdot \\ x^b(1|3,t+1), & x^b(2|3,t), & x^b(3|3,t-1) \\ x^b(1|3,t+2), & x^b(2|3,t+1), & x^b(3|3,t) \\ \cdot & \cdot & \cdot \end{bmatrix}\underline{F}^{-1}\underline{r}_s,$$

where $x^b(i|3,t)$ is the noise at the output after applying the MP filters $\underline{U}(i|3)$. In general, $\underline{F}^{-1}\underline{r}_s$ cannot be thought of simply as applying a frequency filter to the noise at the output of an MP filter unless, for all t,

$$x^b(1|3,t) = x^b(2|3,t) = x^b(3|3,t). \qquad \ldots.(14)$$

The conditions given by expression (14) will be fulfilled exactly for some special cases. For example, if the noise is uncorrelated between channels and $r_{11}(0) = \sigma_1^2$ and $r_{22}(0) = \sigma_2^2$, then the MP filters are given by

$$\left(\underline{U}(1|3),\ \underline{U}(2|3),\ \underline{U}(3|3)\right) = \begin{bmatrix} (1+\sigma_1^2/\sigma_2^2)^{-1} & 0 & 0 \\ 0 & (1+\sigma_1^2/\sigma_2^2)^{-1} & 0 \\ 0 & 0 & (1+\sigma_1^2/\sigma_2^2)^{-1} \\ (1+\sigma_1^2/\sigma_2^2)^{-1} & 0 & 0 \\ 0 & (1+\sigma_1^2/\sigma_2^2)^{-1} & 0 \\ 0 & 0 & (1+\sigma_1^2/\sigma_2^2)^{-1} \end{bmatrix},$$

that is the MP filters reduced to weighted DS or when $\sigma_1^2 = \sigma_2^2$ to DS filters and the Wiener filters $\underline{W}$ (= $\underline{W}(2|3)$) are given by

$$\underline{W} = \begin{bmatrix} \underline{U}_1^T(2|3) \\ \underline{U}_2^T(2|3) \end{bmatrix} \underline{F}^{-1} \underline{r}_s, \qquad \ldots.(15)$$

that is the Wiener filters are equivalent to frequency filtering of the MP (or weighted DS) output. In addition, for all cases where the only non-zero filter coefficients in the MP filters are at t = 0 the relationship between Wiener and MP filters will have the form shown in equation (15).

Note that $\underline{F}^{-1}$ always turns out to be a matrix with elements $f_{ij}^{-1} = f_{i+1,\,j+1}^{-1}$ and $f_{ij}^{-1} = f_{ji}^{-1}$, that is $\underline{F}^{-1}$ is a symmetric Toeplitz matrix. Thus, as r_s is symmetric about its mid-point $\underline{W}^1$ is symmetric and the frequency component of the Wiener filter $\underline{W}(2|3)$ is phaseless. In general, if the Wiener filters used are non-causal, and have an impulse response that is of equal length before and after time zero, then the frequency components of the filters are phaseless.

Capon et al. (17) have used the simple case of two channel 1 point filters to illustrate some of the properties of MP filters; here we follow their example and use the same simple model as a further illustration of the similarities and differences of Wiener and MP filters. For this model the $\underline{R}$ matrix reduces to

$$\begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix},$$

where σ_1^2 and σ_2^2 are the mean square values of the noise on channels 1 and 2 respectively and ρ is the correlation coefficient for the noise on the two channels.

For Wiener filters the normal equations for the two channel one point case can be written

$$\begin{bmatrix} \sigma_1^2 + s^2 & \rho\sigma_1\sigma_2 + s^2 \\ \rho\sigma_1\sigma_2 + s^2 & \sigma_2^2 + s^2 \end{bmatrix} \begin{bmatrix} w_1(0) \\ w_2(0) \end{bmatrix} = \begin{bmatrix} s^2 \\ s^2 \end{bmatrix}, \qquad \ldots.(16)$$

where s^2 is the mean square value of the signal and $w_1(0)$ and $w_2(0)$ are the filter coefficients. Note that as only one point filters are used no frequency filtering can be applied other than simple attenuation of all frequencies by a constant factor.

For MP filtering the normal equations are, from reference (17)

$$\begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 & 1 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} u_1(0) \\ u_2(0) \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \qquad \ldots.(17)$$

λ is a Lagrangian multiplier which has to be introduced when the constraint $u_1(0) + u_2(0) = 1$ is applied.

If $\rho = 1$, that is the noise is perfectly correlated, or $\rho = -1$, that is the noise is exactly out of phase, then $\underline{R}$ has a zero eigenvalue and for both cases $\rho = 1$ and $\rho = -1$ the corresponding eigenvector satisfies the equation $u_1(0) + u_2(0) = \alpha$ where α is a constant. Thus, from the above discussion the Wiener and MP filters should be identical and this is in fact so; the solutions of (16) and (17) when $\rho = 1$ are

$$w_1(0) = u_1(0) = -\sigma_2/(\sigma_1 - \sigma_2); \; w_2(0) = u_2(0) = \sigma_1/(\sigma_1 - \sigma_2); \quad \ldots.(18)$$

and when $\rho = -1$ the solutions are

$$w_1(0) = u_1(0) = \sigma_2/(\sigma_1 + \sigma_2); \; w_2(0) = u_2(0) = \sigma_1/(\sigma_1 + \sigma_2). \quad \ldots.(19)$$

When $\sigma_1^2 = \sigma_2^2$, then for $\rho = 1$ there is no solution; when $\rho = -1$, $u_1(0) = u_2(0) = w_1(0) = w_2(0) = 0.5$ and the solution reduces to DS processing.

If $\rho = 0$, that is the noise is uncorrelated, the solutions for the two types of filter differ thus; from equation (17) the MP filters are given by

$$\begin{aligned} u_1(0) &= (1 + \sigma_1^2/\sigma_2^2)^{-1}, \\ u_2(0) &= (1 + \sigma_2^2/\sigma_1^2)^{-1} \end{aligned} \quad \ldots.(20)$$

and from equation (16) the Wiener filters by

$$\begin{aligned} w_1(0) &= u_1(0)\underline{F}^{-1}\underline{r}_s, \\ w_2(0) &= u_2(0)\underline{F}^{-1}\underline{r}_s, \end{aligned} \quad \ldots.(21)$$

where $\underline{F}^{-1}\underline{r}_s$ is now

$$\{\sigma_1^2\sigma_2^2/(\sigma_1^2 + \sigma_2^2) + s^2\}^{-1}s^2.$$

MP filtering for $\rho = 0$ is thus equivalent to weighted DS processing (equation (4)) and for $\sigma_1^2 = \sigma_2^2$ MP filtering reduces to DS processing in agreement with results derived earlier. Note that if $s^2 \gg \sigma_1^2$ and σ_2^2, then $\underline{F}^{-1}\underline{r}_s \to 1$ and again as expected from earlier discussion $w_1(0) \to u_1(0)$ and $w_2(0) \to u_2(0)$.

The above discussion shows that whatever the optimum multichannel filters required to extract the signal from noise either spatial filters (ie, MP filters which include filters equivalent to DS and weighted DS processing) or combined spatial and frequency filters it is only necessary to estimate the Wiener filters because the MP filters can be thought of as a special case of the Wiener filters which is chosen when spatial filters are adequate to extract the signal from noise. Also if for a given signal and noise spatial filters rather than combined spatial and frequency filters are desired, then it is only necessary to impose the condition that the signal is much larger than the noise to ensure the

estimated Wiener filters tend to pure spatial (MP) filters. Assuming that for any given signal there are some frequencies where the signal-to-noise ratio is greater than unity, Wiener filters should always give an output which shows the signal above the noise and so the detection threshold on broad band recordings should be about the same as for narrow band recording; for pure spatial filtering the threshold will usually be above this.

As, in general, Wiener filters apply some frequency filtering to the desired signal it will usually be useful to have some way of measuring the extent to which frequency filtering contributes to the noise reduction. Now Kelly (18) states that the expected mean square signal error σ_E^2, say, at the output of a multi-channel Wiener filter is given by

$$\sigma_E^2 = r_s(0) - \underline{W}^1 \underline{r}_s .$$

When $\underline{W}^1$ is an all pass filter, that is no frequency filtering is applied to the signal, $\sigma_E^2 = 0$. As the effect of frequency filtering increases $\underline{W}^1 \underline{r}_s$ tends to zero and thus $\sigma_E^2 \rightarrow r_s(0)$. Thus, as the frequency filter varies from an all pass filter to a filter with all coefficients zero, then

$$\gamma = 1 - \sigma_E^2 / r_s(0) \qquad \ldots.(22)$$

ranges from one to zero; this suggests that γ might be used as a measure of the performance of the Wiener filters; if $\gamma \approx 1$, then little frequency filtering is applied, ie, the performance is high; if γ is significantly less than unity, this indicates that the desired signal can only be extracted from the noise by substantial frequency filtering, ie, the performance of the filters is low.

As the absolute detection threshold of Wiener filters as applied to broad band recordings should be about the same as for narrow band recordings some way is required of describing the ability of Wiener filters to extract signals from noise, which depends not only on whether the signal is seen above the noise or not, but also measures how much frequency filtering has to be applied to extract the signal; the variation in the performance figure γ with magnitude is a way of doing this.

4.5 Signal coherence

So far in discussing multichannel filters we have assumed that the signal is identical on all channels. In general, this is not so; the effects of the topography and other lateral variations in structure at an array result in variations in the signal between channels. The result of this may be that, although the noise is reduced on applying multichannel filtering, the signal may also be reduced or distorted or both. This can be illustrated using the simple case of two channel one point filters discussed above where the noise is in-phase and perfectly correlated ($\rho = 1$) and the variance of the noise on channel 1 is σ_1^2 and on channel 2 is σ_2^2; assuming $\sigma_1^2 \neq \sigma_2^2$, then for this model the multichannel filters (given in equation (18)) reduce the noise to zero but pass the signal unchanged provided that the signal is identical on each channel. We now assume that the signals are not the same on the two channels but that on channel 1 the signal is $s_1(t) = s(t) + e_1(t)$ and on channel 2 the signal is $s_2(t) = s(t) + e_2(t)$ where $s(t)$ is the true signal and $e_1(t)$ and $e_2(t)$ are deviations from the true signals due to the effects of the array site. If $e_1(t)$ and $e_2(t)$ are uncorrelated, the estimated signal when the noise is reduced to zero is

$$s(t) + \{\sigma_1 e_2(t) - \sigma_2 e_1(t)\}/(\sigma_1 - \sigma_2).$$

Thus, as σ_1 approaches σ_2, the error term will tend to be large and so swamp the signal. Thus, although the noise is reduced to zero the signal is highly distorted. In the absence of noise the best estimate of the signal given $s_1(t)$ and $s_2(t)$ is given by DS processing. For DS processing of an n element array the deviations from the true signal will tend to be reduced to $n^{\frac{1}{2}}$, assuming they are uncorrelated between channels, just as random noise is reduced. Key (13) has shown that at least for some SP arrays the coda of some seismograms are reduced in this way.

The above discussion shows that, as well as considering the ability of spatial filters to suppress noise, it is necessary to consider how they distort signals that are not perfectly correlated across the array. We return to this problem in section 5.

5. PROCESSING IN PRACTICE

In order to carry out Wiener filtering for the general case requires the auto-correlation function of the signal and the auto- and cross-correlation functions of the noise. The $\underline{R}$ and $\underline{A}$ matrices (equation (6)) can then be constructed where these matrices are now re-defined (by analogy with the two-channel three point filter cases) to cover the general case of n channels and p filter points per channel. When $n = 1$ the Wiener filter becomes a simple frequency filter of the form $(\underline{R} + \underline{A})^{-1}\underline{r}^{S}$ where $\underline{R}$ is constructed from the auto-correlation of the noise in the same way as $\underline{A}$ is constructed from the auto-correlation of the signal.

For true Wiener filtering $\underline{R}$ and $\underline{A}$ should by definition be derived from stochastic models of the signal and noise processes (16); such models have been used by Burg (3) and Backus et al. (4) for most of their work on the Wiener processing of SP data. An alternative way of constructing $\underline{R}$ is to estimate the noise properties from a section of observed noise; this can be done in seismology where the signals are transient and so the noise can be observed free from signals; in this report we refer to filters derived using observed data as data-dependent Wiener (DW) filters.

To estimate MP filters requires only that $\underline{R}$ be specified; as with Wiener filters $\underline{R}$ could be derived from a stochastic model of the noise process but in all the published work on the application of the minimum power method, $\underline{R}$ appears to have been estimated from sections of the observed noise. We refer to filters derived in this way using observed noise as data-dependent minimum power (DMP) filters.

When stochastic models of noise and signal are used the filters are not designed to fit any specific section (realization) of the noise or signal so that one set of filters should be sufficient for all time. Thus, for example, one of the multichannel filters investigated by Backus et al. (4) specifies the signal as all plane waves with apparent speeds greater than 8.1 km/s independent of azimuth, and the noise as 80% plane waves with apparent speeds of 2.5 to 3.5 km/s independent of azimuth plus 20% of uncorrelated noise. The noise and signal power spectra are assumed to be constant with frequency. Filters derived from these general models should thus always be able to extract high speed signals from low speed noise and, as the filters are not designed on a particular section of noise, should not on average give better results for any one section of noise compared to another.

Use of stochastic models allows an analysis of the stability of the computed filters to be made and where the normal equations for estimating the filters turn out to be ill-conditioned the origin of this ill-conditioning can be identified and steps taken to control it. The most obvious way ill-conditioning can arise is when the noise on any two channels is assumed to be identical (in-phase, perfectly correlated, equal in amplitude), then the matrix of coefficients is singular and there are an infinite number of possible solutions (assuming, as we do here, that the signal is also identical on all channels). If it is assumed that two or more channels are identical, all except one of these channels can be dropped and filters estimated for the remainder. Alternatively, the constraint can be applied that, say, all filters for channels with identical noise are equal and again a unique solution can be obtained. Usually observed noise will not be identical on any two (or more) channels but the noise could be similar and then the stability of the filters could be low. When the $\underline{R}$ matrix is constructed from a stochastic model of the noise, then instabilities can be identified and controlled because the noise properties are specified exactly; when the $\underline{R}$ matrix is constructed from observed noise some way is needed either to measure the stability of the estimates or to ensure that any ill-conditioning is avoided.

Usually it is impossible by definition to construct $\underline{A}$ from the observed signal for, if $\underline{A}$ is known, then there is little point in estimating it; a signal model has thus to be used. For the noise, however, despite the advantages of using stochastic models, it would seem intuitively that better signal-to-noise improvements could be obtained by constructing $\underline{R}$ from the observed noise. For example, the noise at any time might be highly directional but the direction might vary from day-to-day. So filters designed to suppress noise in all azimuths might not achieve the noise suppression that would be obtained by using filters designed to suppress noise from a specific azimuth. As the object of the present study is to extract chosen signals from noise, rather than to process all data to extract all possible signals, it seems sensible to use the noise just ahead of the signal in constructing $\underline{R}$ and this we always do.

In constructing $\underline{R}$ from observed data it is necessary to decide on the length of data to be used to estimate the auto- and cross-correlation functions of the noise; this length is referred to by Capon et al. (6,17) as the fitting interval. The longer the fitting interval, the greater the amount of computation, and if the

noise properties are changing with time, the more the measured properties may differ from the noise properties just before the onset of the signal. If the fitting interval is short, on the other hand, spectacular noise reductions may be obtained in the fitting interval but little reduction (and possibly amplification) outside this interval; this effect is referred to by Capon et al. (6) as "supergain". The reason for "supergain" is as follows. Suppose the true filters would reduce the variance of the noise to σ_o^2, the filters computed using a sample of the observed noise in some fitting interval are only estimates of the true filters so when applied to noise data outside the fitting interval they are unlikely to achieve a noise reduction at the output to σ_o^2. Inside the fitting interval the best estimate $\hat{\sigma}^2$ of σ_o^2 given m values of the output $\varepsilon(k)$ is

$$\sum_{k=1}^{m} \varepsilon^2(k)/q,$$

where q is the degrees of freedom; for DMP filtering and DW filtering at large signal-to-noise $q = m - (n - 1)p$ where p is the number of points in each single channel filter and n the number of channels. The apparent variance σ_A^2 of the residual noise in the fitting interval, however, is

$$\sum_{k=1}^{m} \varepsilon^2(k)/m$$

and this is less than $\hat{\sigma}^2$. Thus, in general, data-dependent filters give an apparently better noise reduction within the fitting interval than outside it. In order to keep the effects of differences between the apparent variance within the fitting interval and outside it to a minimum, m must be much larger than np.

In most of the examples of multichannel filtering in this report we use 2048 points from each channel which at 12.5 samples/s is a fitting interval of about 164 s. With this length of fitting interval the estimated filters appear to be stable (the maximum filter length used for array processing is 39 points) and the noise reduction in the fitting interval is not noticeably different from the noise reduction outside the interval. Usually it is computationally convenient to estimate the mean square noise on the filter output in the fitting interval; when this is done $\hat{\sigma}^2$ is used and not σ_A^2. For the array studies described in this report instabilities in the estimates due to similarities in the noise on two channels do not appear to arise but for arrays with closer spaced seismometers than used here this problem could become more serious.

We now consider the choice of signal model for the construction of the $\underline{A}$ matrix. For the application of Wiener filtering to the extraction of signals that (at least ideally) arrive at the same time (after time shifting) and are of equal amplitude on all channels, the auto- and cross-correlation functions are identical. Where a unique solution exists that reduces the noise exactly to zero using only spatial filtering, then the solution is independent of the shape assumed for the signal (and hence on the assumed signal power spectrum). Also it can be seen from equation (7) that, as the noise power becomes small compared to the signal power, the filter response tends to unity whatever the assumed signal power spectrum.

When the noise cannot be adequately suppressed by spatial filtering the assumed signal auto-correlation is more important. However, some of the general properties of the body wave spectra to be expected at teleseismic distances are known. For example, P signals contain little energy above frequencies of 3 Hz. Also the spectrum of source pulses radiated by earthquakes is roughly flat from zero frequency to some high frequency limit (corner frequency) above which the amplitude spectrum falls off rapidly. For processing broad band signals that are recorded on a seismograph with flat response from 0.1 to 6.25 Hz a signal spectrum that is flat from 0.1 to, say, 3 Hz could be used in the absence of more detailed knowledge of the signal.

For particular seismic sources it may be that it is possible to guess at the rough form of the spectrum. Thus, if a signal is from an explosion, then the spectrum of a model explosion signal can be used. If the signal is an earthquake that looks like an explosion so that m_b measured on SP seismograms is much greater than M_s, which is measured on LP seismograms, then this suggests that the body waves contain a large proportion of their energy at frequencies around 1 Hz. If $M_s \gg m_b$, then the high frequency energy in the body wave pulses is likely to be small. The ratio of m_b to M_s could thus be used as a guide to the signal model to use. Figure 5 shows the signals used in this report in constructing $\underline{A}$. Signal A is a theoretical earthquake signal and signal B a theoretical explosion signal. These two signals were computed using the method of Hudson (19,20) and Douglas et al. (21); the details of the computation are not important. Signal C is the impulse response of the earth, assuming that the only effect of the earth is anelastic attenuation, convolved with the impulse response of the BB seismograph. The anelastic attenuation is allowed for using the method of Carpenter

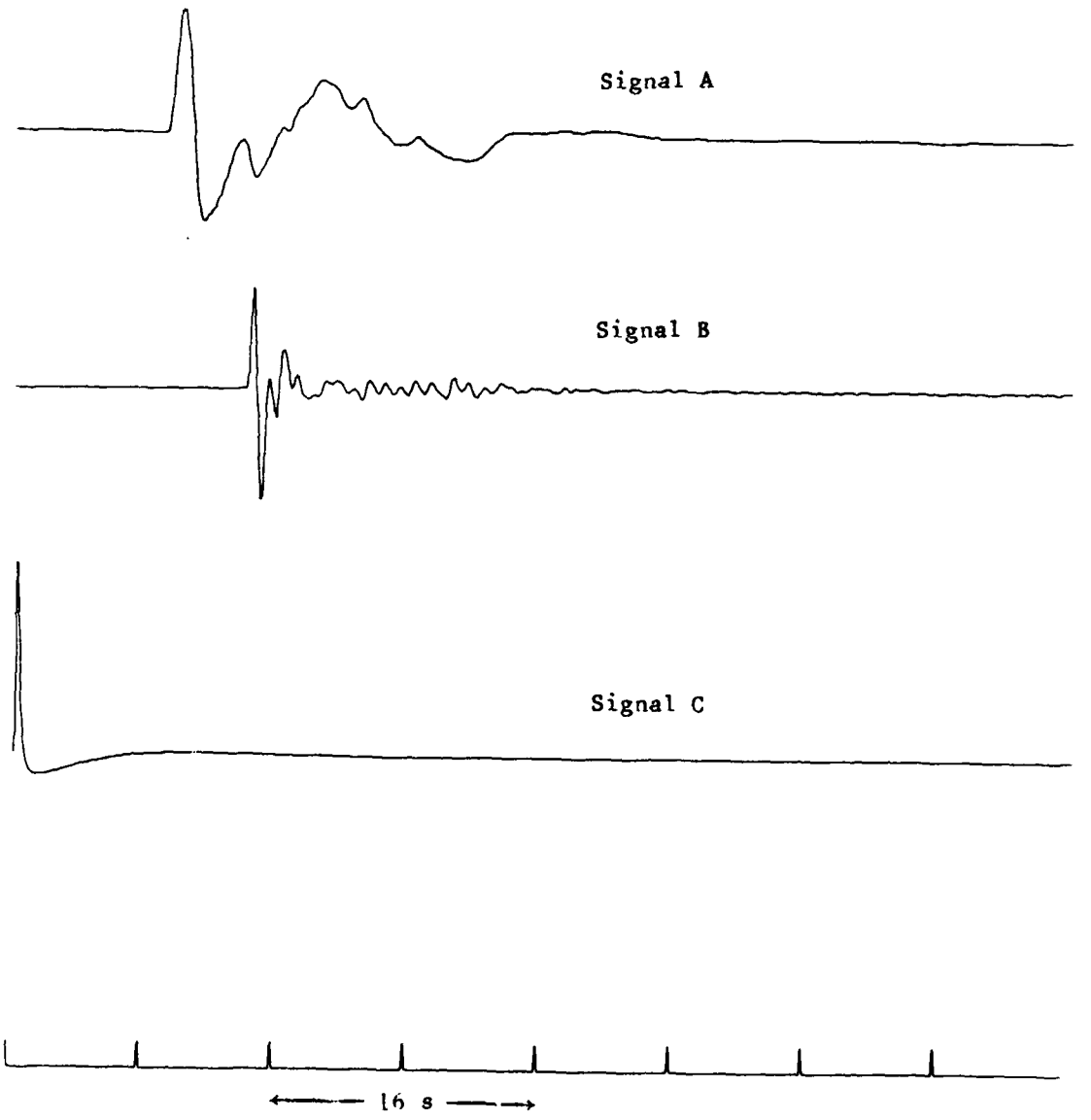


FIGURE 5. MODEL SIGNALS USED IN DESIGNING DATA-DEPENDENT WIENER FILTERS: (A) EARTHQUAKE (B) EXPLOSION (C) ATTENUATION - SEISMOGRAPH IMPULSE RESPONSE

(22); the amplitude spectrum of the pulse is assumed to have the form $\exp(-|\omega|t^*/2)$ where ω is angular frequency, $t^* = T/Q_{av}$, T is the total travel time and Q_{av} the average Q on the path; $t^* = 0.4$ s has been used to compute signal C. We refer to signal C as the attenuation - seismograph impulse response.

The theoretical development of Wiener filtering given earlier assumes that the signal is a stationary process extending from $-\infty$ to $+\infty$. Observed seismic signals are transients so some way is required of modelling a transient by a continuous signal. One way of doing this is to choose a model transient to represent the observed signal and then to assume that this model is replaced by a continuous signal that has an auto-correlation function.

$$r^s(k) = \sigma_c^2 r_o^s(k)/r_o^s(0),$$

where $r_o^s(k)$ is the auto-correlation of the model transient signal with arbitrary amplitude and σ_c^2 is the mean square value of the continuous signal. A suitable value can be assigned to σ_c^2 by setting $\sigma_c = b/a$ where b is an estimate of the maximum amplitude of the signal to be extracted from the noise and a is some chosen constant, for example, a might be set equal to 3; this is equivalent to replacing the model transient by a continuous stationary signal that has an auto-correlation function that is the same shape as that of the transient and has an rms amplitude that is a third of the maximum amplitude of the transient. If the signal is visible above the noise, then b can be set to the observed signal amplitude. If the signal is not visible on the broad band records, then an estimate of b can be made from a knowledge of the magnitude of the earthquake (or explosion). In practice it seems that the best value for b can most easily be found by trial and error; two runs of the program are usually sufficient to determine a suitable value for b. It is easy to check if too small a value has been chosen for b for then the frequency component of the DW filters attenuates all frequencies; this happens because the assumed signal-to-noise ratio is not greater than unity at any frequency so the only way the noise can be reduced is by lowering the magnification of the filter. When this happens b must be increased so that for at least one frequency the amplitude response of the filter is unity. In order to apply DW filters that are purely spatial filters (equivalent to MP filters) the true signal amplitude is simply ignored and b set to some large value, say, ten times the maximum noise amplitude.

In order to assess the effectiveness of DW filtering three measures of noise reduction are used: (1) Φ_{DW}, the total noise reduction obtained by DW filtering, (2) Φ_{DS} (defined in equation (1)), the noise reduction obtained by DS processing, and (3) Φ_S, the maximum noise reduction that can be obtained by spatial filtering; that is, the noise reduction that would be obtained using MP filters. We define Φ_{DW} as $(\sum\sigma_i^2/n\hat{\sigma}^2)^{\frac{1}{2}}$ where $\hat{\sigma}^2$ is the mean square amplitude of the noise after DW filtering and Φ_S as $(\sum\sigma_i^2/n\hat{\sigma}_s^2)^{\frac{1}{2}}$ where $\hat{\sigma}_s^2$ is the mean square amplitude of the noise after spatial filtering only; a convenient way of obtaining $\hat{\sigma}_s^2$ is given in appendix B. If Φ_{DW} and Φ_S are roughly equal, then this indicates that the noise reduction due to Wiener filtering is essentially spatial filtering; if Φ_{DW} is much greater than Φ_S, then this indicates that the noise reduction due to Wiener filtering is mainly frequency filtering. Similarly, comparing Φ_{DS} and Φ_S shows the additional noise reduction that can be obtained by spatial filtering compared to simple DS processing. Table 4 gives the noise reduction factors, Φ_{DW} and Φ_{DS} for each of the six BB noise samples described in section 3; the DW filtering was carried with the assumption that the S/N ratio was large (= 64) so that the filters obtained are purely spatial filters and, thus, $\Phi_S = \Phi_{DW}$ and the performance factor $\gamma = 1$. Note that in general the larger the noise amplitude, the greater Φ_{DW}.

We now demonstrate some of the properties of DW (and DMP) filtering using a known signal in observed noise; this known signal is an artificial earthquake seismogram signal A (figure 5) and is chosen because a range of frequencies are visible so that the effects of any frequency filtering should be obvious to the eye. Signal A has been scaled and added to four channels of noise to simulate a set of array recordings with signal-to-noise ratio varying from 4:1 to 0.0625:1 in binary steps (here a signal-to-noise ratio of h:1 means that the peak amplitude of the signal is h times larger than the peak amplitude of the noise). The noise used is that shown in figure 3(b) and is highly coherent. The result of DS processing for this artificial earthquake seismogram in noise is shown in figure 6; the noise reduction obtained is 1.5.

TABLE 4

Comparison of Noise Reduction Obtained by Delay and Sum Processing and Data-Dependent Wiener (Spatial) Filtering for Six Noise Samples

Noise Sample	BB RMS Amplitude, nm		Φ_{DS}	Φ_{DW}
	Average over Single Channels	After DW (Spatial) Filtering		
20 January 1976	2069	387	1.56	5.4
21 February 1976	1447	346	1.72	4.1
20 March 1976	589	192	2.04	3.0
20 April 1976	453	114	2.31	4.0
20 May 1976	209	92	2.17	2.3
20 June 1976	260	118	1.77	2.2

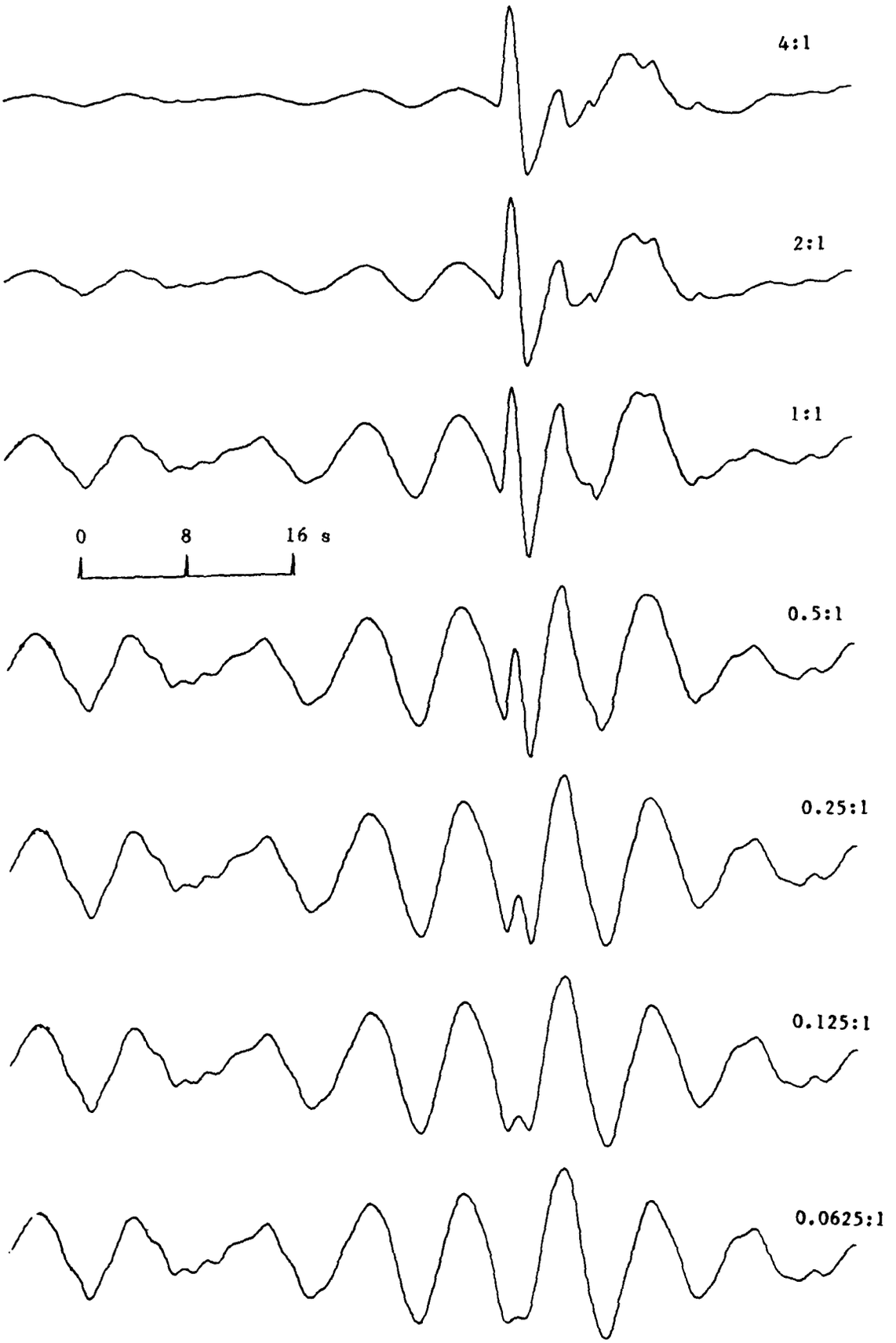


FIGURE 6. DELAY AND SUM PROCESSING OF AN ARTIFICIAL EARTHQUAKE SIGNAL (SIGNAL A, FIGURE 5) FOR SIGNAL-TO-NOISE RATIOS RANGING FROM 4:1 TO 0.0625:1. NOISE AMPLITUDES ARE REDUCED BY 1.5 RELATIVE TO THE AVERAGE NOISE AMPLITUDE ON A SINGLE CHANNEL

Figure 7 shows the results of DW filtering using a 164 s fitting interval (2048 points/channel), 39 point filters and the artificial signal itself (signal A) as the model signal used to construct the filters. Now $\Phi_S = 6.1$ for this sample of noise so it can be seen from the values of Φ_{DW} shown on figure 7 that, for signal-to-noise ratios of 0.5:1 or greater, there is little frequency filtering ($\Phi_{DW} \approx \Phi_S$). At lower signal-to-noise ratios the effect of frequency filtering is more important.

In practice, the true signal will not be available to use as the signal model. Figure 8 is an example of how the estimated signal is distorted if an incorrect model is used; these results were obtained using signal B (figure 5), the high frequency explosion signal, as the signal model in designing the DW filters. As the model signal has little energy at the periods of the microseisms the computed filters attenuate the microseisms by frequency filtering. Only at large signal-to-noise ratios is the effect of DW filtering to apply mainly spatial filtering. At low signal-to-noise ratios most of the gain comes from frequency filtering. Figure 9 shows the results of DW filtering with a more sensible choice of signal model: signal C, the attenuation-seismograph impulse response. The results for signal C are similar to those with signal A, except that at low signal-to-noise ratios the proportion of frequency filtering is larger than when using signal A. However, it appears that signal C gives satisfactory results and we have used this for all the examples shown in this report. It is obvious from inspection of figures 7 to 9 that the effect of spatial filtering is to reduce the amplitude of the coherent oceanic microseisms.

Figure 10 shows the amplitude response of the frequency filter applied in DW filtering of the artificial signal in noise when the filters are designed using signal C (figure 5) as the signal model; the amplitude response to ground displacement of the WWSS SP seismograph is also shown for comparison. Note that at large signal-to-noise ratios the effect of the frequency component of the DW filters is negligible but as the signal-to-noise ratio decreases the noise at periods below 1 Hz is progressively attenuated and the frequency response tends to the SP response. This is to be expected as the largest noise amplitudes are at periods of greater than 1 s and the DW filters should tend to the inverse of the power spectrum of the noise at low signal-to-noise ratio (from equation (7)); the SP response is also designed to suppress the long period noise. Hence, the tendency of the amplitude response of the Wiener filters to coincide with the SP response.

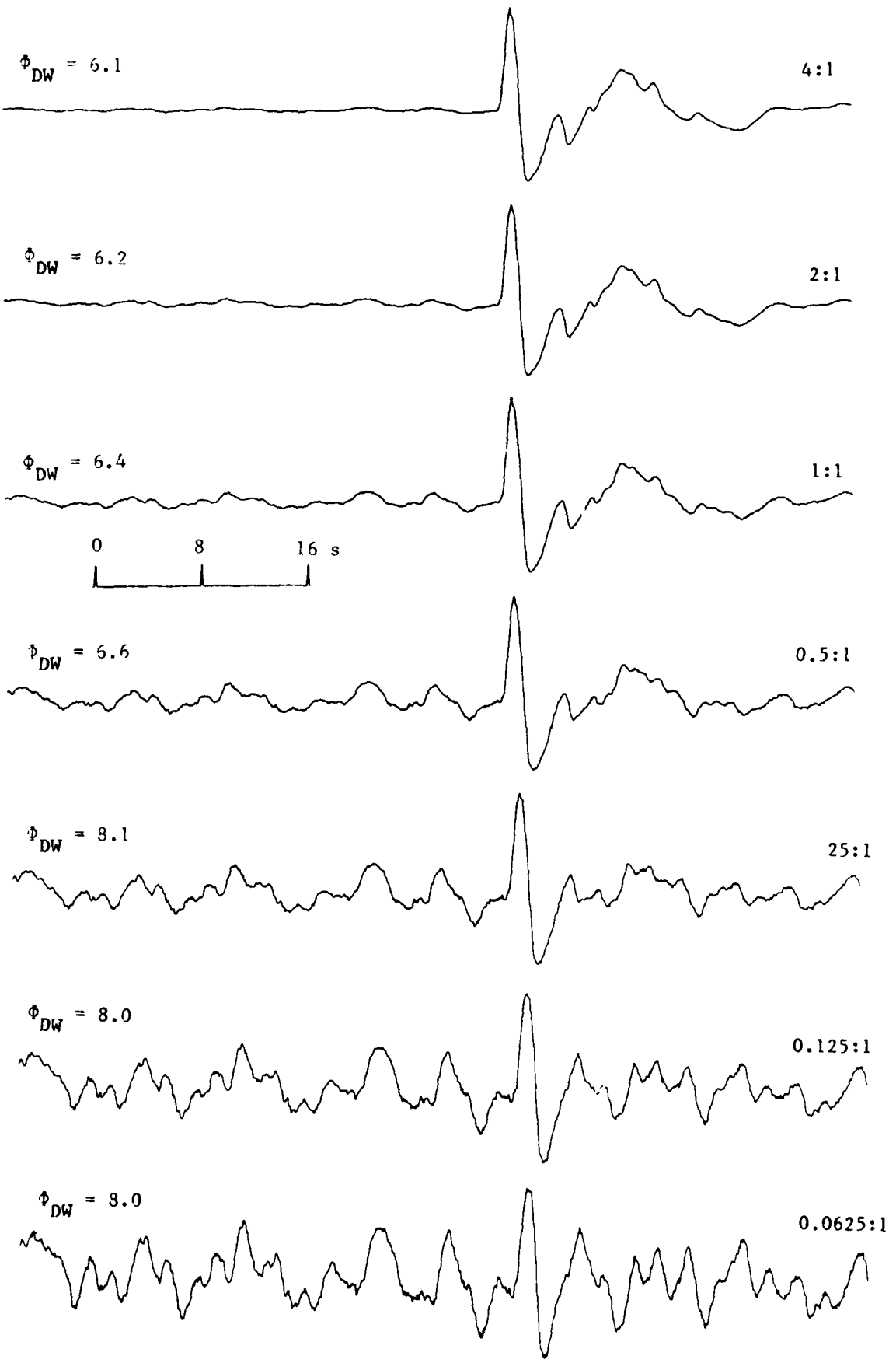


FIGURE 7. DATA-DEPENDENT WIENER PROCESSING OF AN ARTIFICIAL EARTHQUAKE SIGNAL (SIGNAL A, FIGURE 5) FOR SIGNAL-TO-NOISE RATIOS RANGING FROM 4:1 TO 0.0625:1. THE SIGNAL USED TO DESIGN THE FILTERS IS THE ARTIFICIAL SIGNAL ITSELF. AGAINST EACH PROCESSED RECORD IS SHOWN Φ_{DW}, THE FACTORS BY WHICH DW FILTERING HAS REDUCED THE NOISE, FOR THIS NOISE SAMPLE.

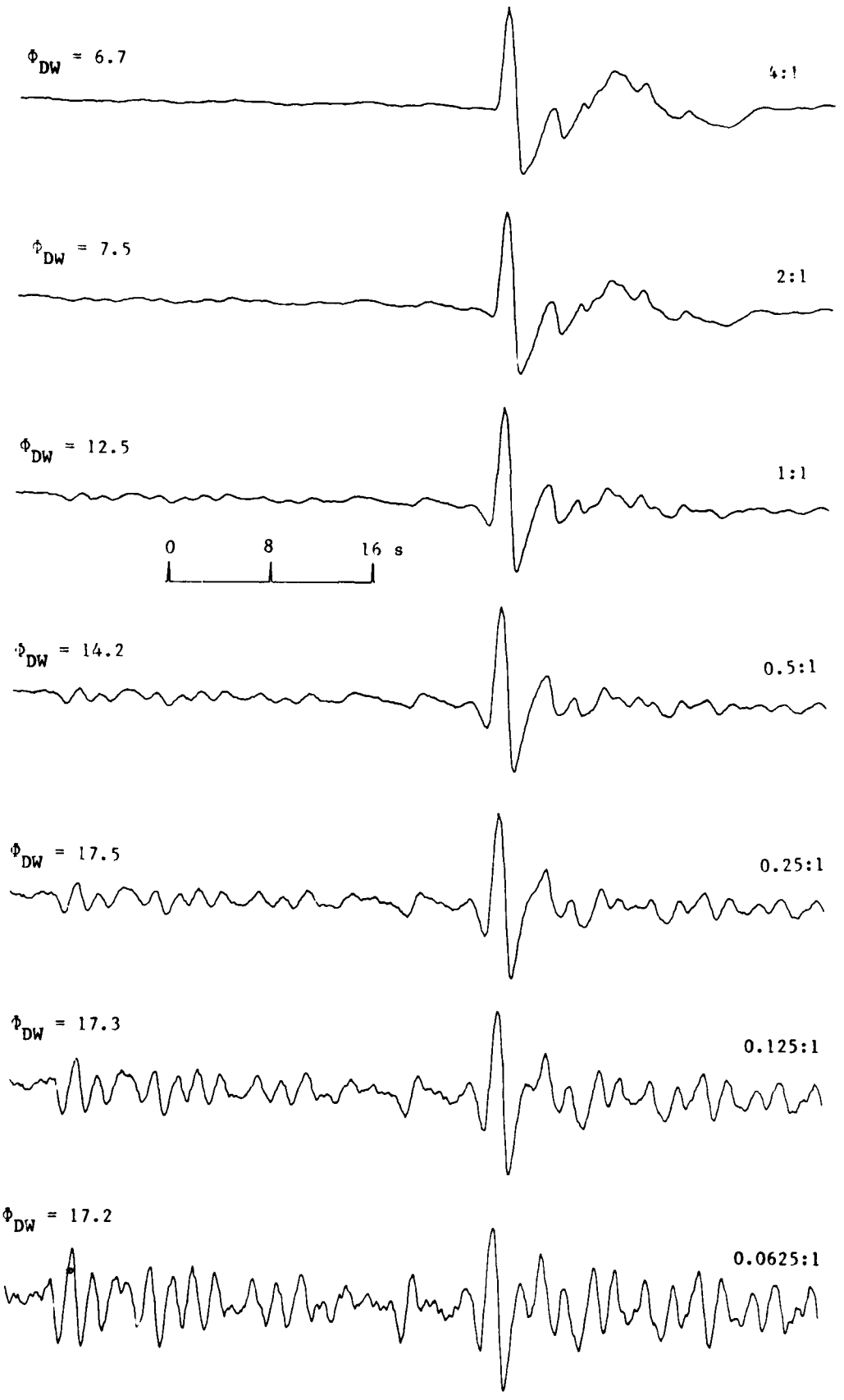


FIGURE 8. DATA-DEPENDENT WIENER PROCESSING OF AN ARTIFICIAL EARTHQUAKE SIGNAL (SIGNAL A, FIGURE 5) FOR SIGNAL-TO-NOISE RATIOS RANGING FROM 4:1 TO 0.0625:1. THE SIGNAL USED TO DESIGN THE FILTERS IS THE EXPLOSION SIGNAL (SIGNAL B, FIGURE 5). AGAINST EACH PROCESSED RECORD IS SHOWN THE FACTOR BY WHICH DW FILTERING HAS REDUCED THE NOISE. FOR THIS NOISE SAMPLE Φ_S = 6.1

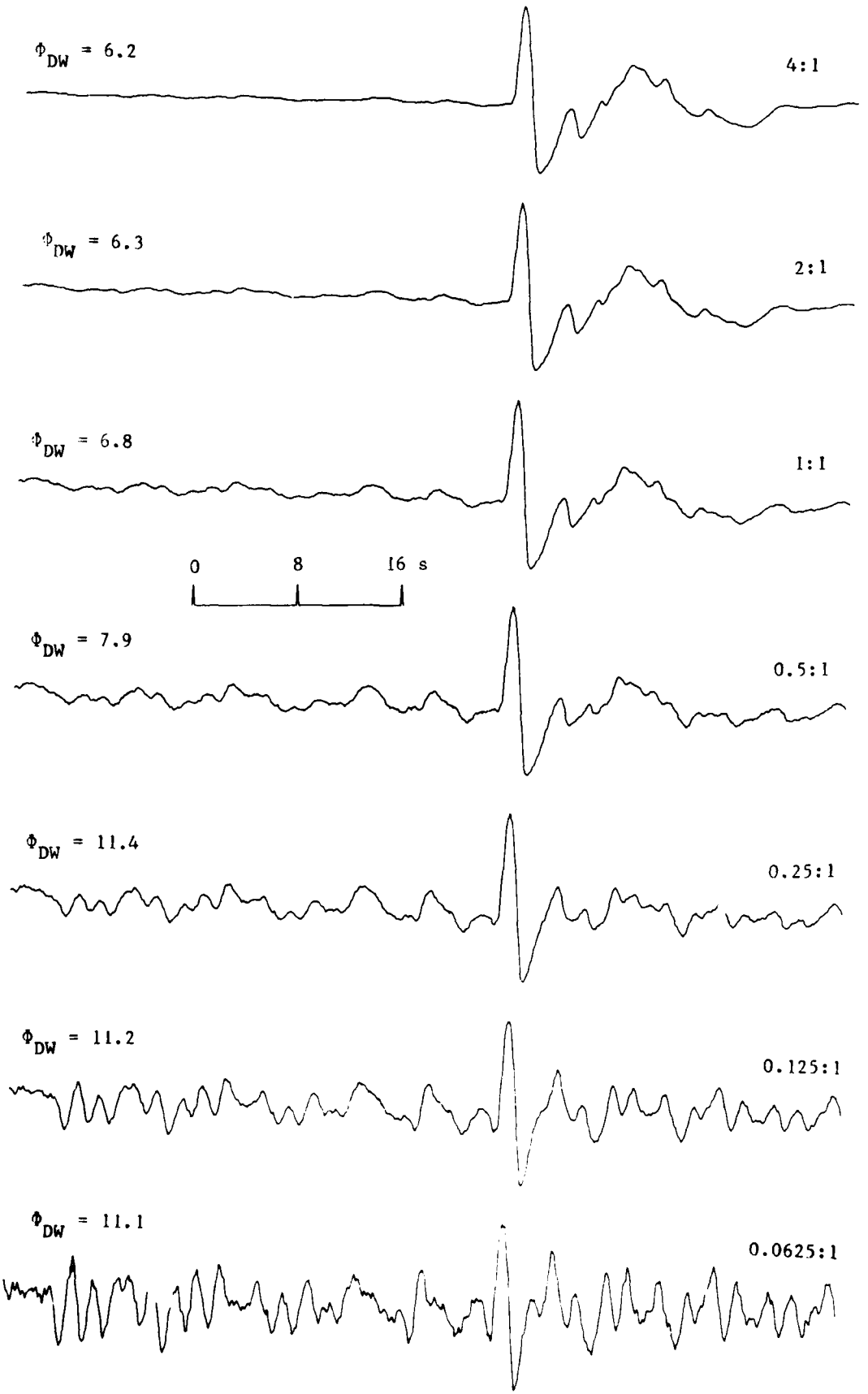


FIGURE 9. DATA-DEPENDENT WIENER PROCESSING OF AN ARTIFICIAL SIGNAL (SIGNAL A, FIGURE 5) FOR SIGNAL-TO-NOISE RATIOS RANGING FROM 4:1 TO 0.0625:1. SIGNAL USED TO DESIGN THE WIENER FILTERS IS THE ATTENUATION - SEISMOGRAPH IMPULSE RESPONSE (SIGNAL C, FIGURE 5). AGAINST EACH PROCESSED RECORD IS SHOWN Φ_{DW}, THE FACTOR BY WHICH DW FILTERING HAS REDUCED THE NOISE. FOR THIS NOISE SAMPLE Φ_S = 6.1

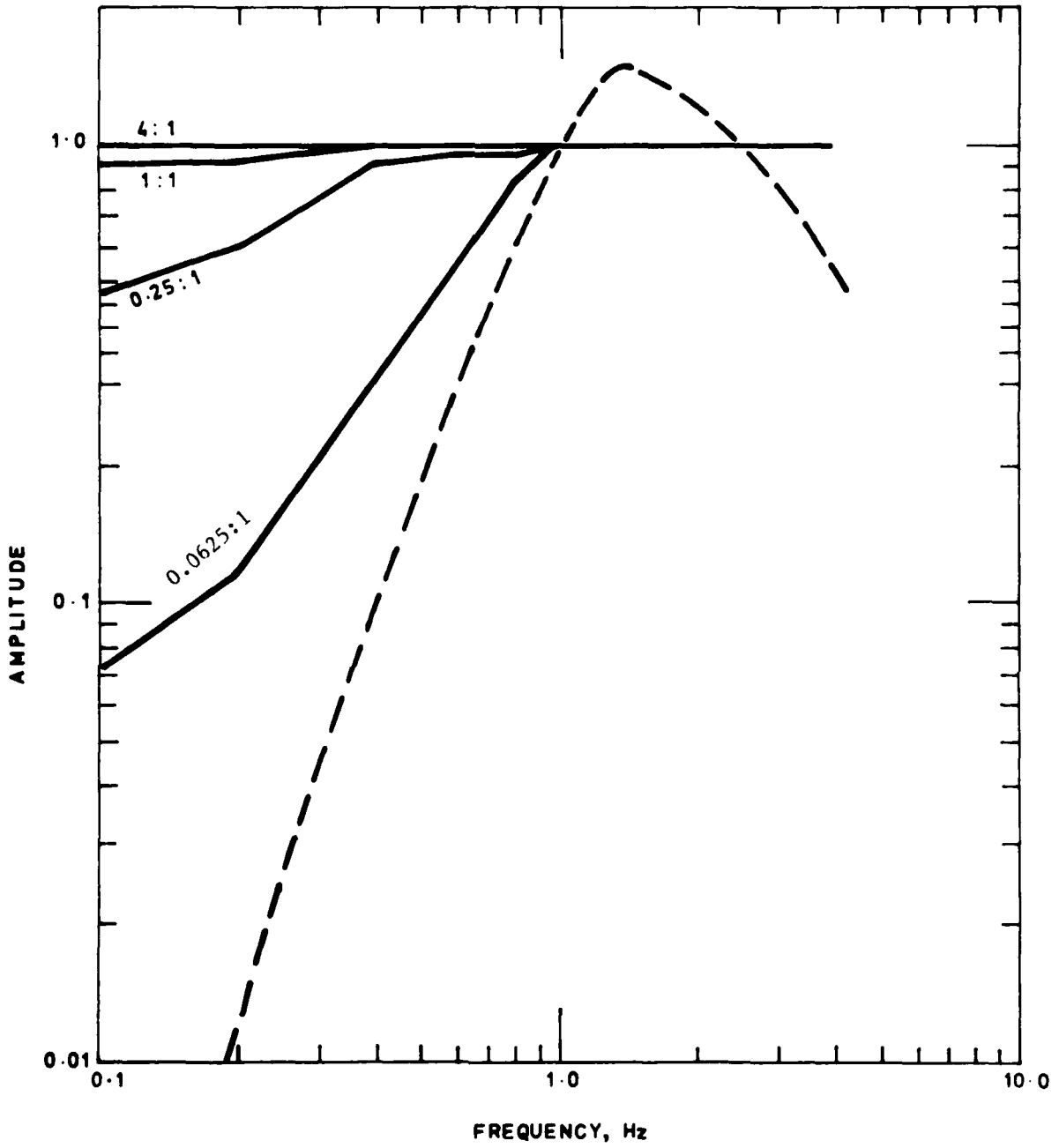


FIGURE 10. AMPLITUDE RESPONSE OF FREQUENCY FILTERING COMPONENT OF THE DATA-DEPENDENT WIENER FILTERS USED TO EXTRACT THE ARTIFICIAL EARTHQUAKE SIGNAL FROM NOISE USING SIGNAL C (FIGURE 5) AS THE SIGNAL MODEL FOR THE FILTER DESIGN. RESPONSES SHOWN FOR SIGNAL-TO-NOISE RATIOS OF 4:1, 1:1, 0.25:1 AND 0.0625:1. THE AMPLITUDE RESPONSE OF THE WWSS SP SEISMOGRAPH IS SHOWN FOR COMPARISON

Figure 11 shows two examples of the standard output that we use for routine DW filtering. Again the data are the artificial earthquake signal in noise; the DW filters were designed using signal C (figure 5) as the signal model. For each example figure 11 shows the WWSS SP seismogram, the DW filtered output, the DS output after filtering with $\underline{W}^1$, the frequency component only of the DW filters (FDS), the DS output, and the output of one seismometer of the array. This display provides a guide to the effect of each component of the processing of the BB records and allows the BB output to be compared with the SP. Thus, from figure 11(a) it can be seen that the DS and FDS outputs are about the same so that the frequency filtering applied by the DW filters is small; from figure 11(b), on the other hand, it is obvious that the frequency filtering is significant. Note that the signal-to-noise ratio on the BB after DW filtering is about the same or better than on the SP seismogram. The WWSS SP seismogram was derived from the broad band DS output in the way described in section 2 and is the SP DS output.

Figure 12 shows the standard output after DW filtering applied to the artificial earthquake signal in noise (signal-to-noise ratio 0.5:1) using a fitting interval of only 160 points. The effect of "supergain" is clearly seen; the apparent noise reduction in the fitting interval (see figure 12) is 6.3. However, using $\hat{\sigma}$, the best estimate of the rms noise in the fitting interval, the noise reduction is insignificant which is obviously a more sensible figure.

Figure 13 demonstrates that DW for large signal-to-noise ratios (where the DW filtering is mainly spatial filtering) and DMP filtering give nearly identical results; to obtain the DW output the actual signal-to-noise ratio was ignored and a ratio of 64:1 was used. Figure 13 shows the outputs after DW and DMP filtering for the artificial earthquake signal in noise, together with the outputs of each channel after DW and DMP filtering. Note the striking similarity between the outputs from the two methods, particularly on the outputs of individual channels. This shows that the DW and DMP filters are very similar, as is expected from the discussion given in section 4.

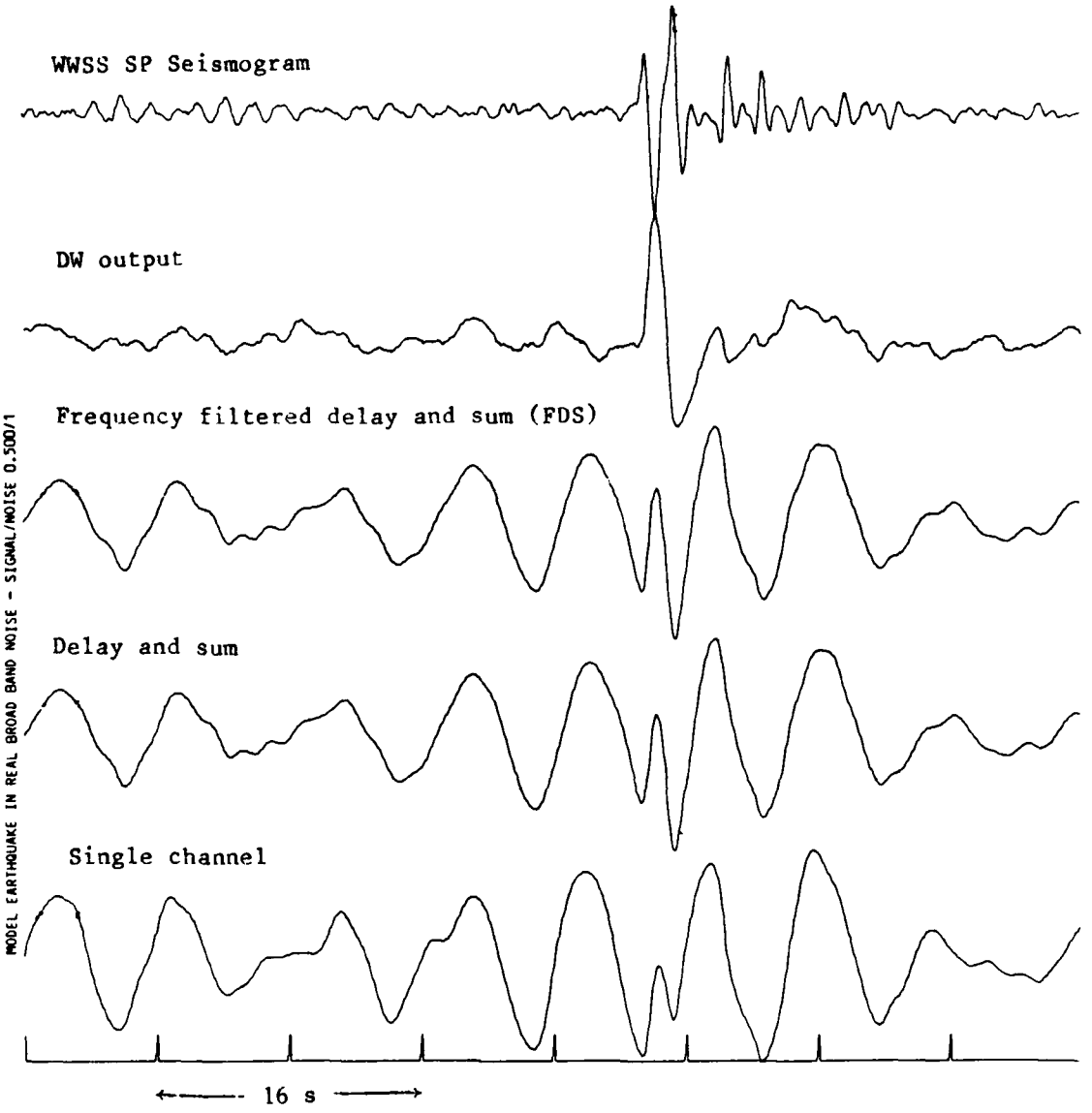


FIGURE 11(a) EXAMPLE OF THE STANDARD OUTPUT OF THE DATA-DEPENDENT WIENER (DW) FILTERING PROGRAM. DATA IS THE ARTIFICIAL EARTHQUAKE SIGNAL IN NOISE FOR SIGNAL-TO-NOISE RATIO 0.5:1. THE FILTERS WERE DESIGNED USING SIGNAL C (FIGURE 5) AS THE SIGNAL MODEL. THE FOLLOWING SEISMOGRAMS ARE SHOWN: WWSS SP SEISMOGRAM; DW OUTPUT; DELAY AND SUM OUTPUT AFTER FILTERING WITH THE FREQUENCY COMPONENT OF THE DW FILTERS; THE DELAY AND SUM OUTPUT; AND THE OUTPUT FROM A TYPICAL SEISMOMETER OF THE ARRAY

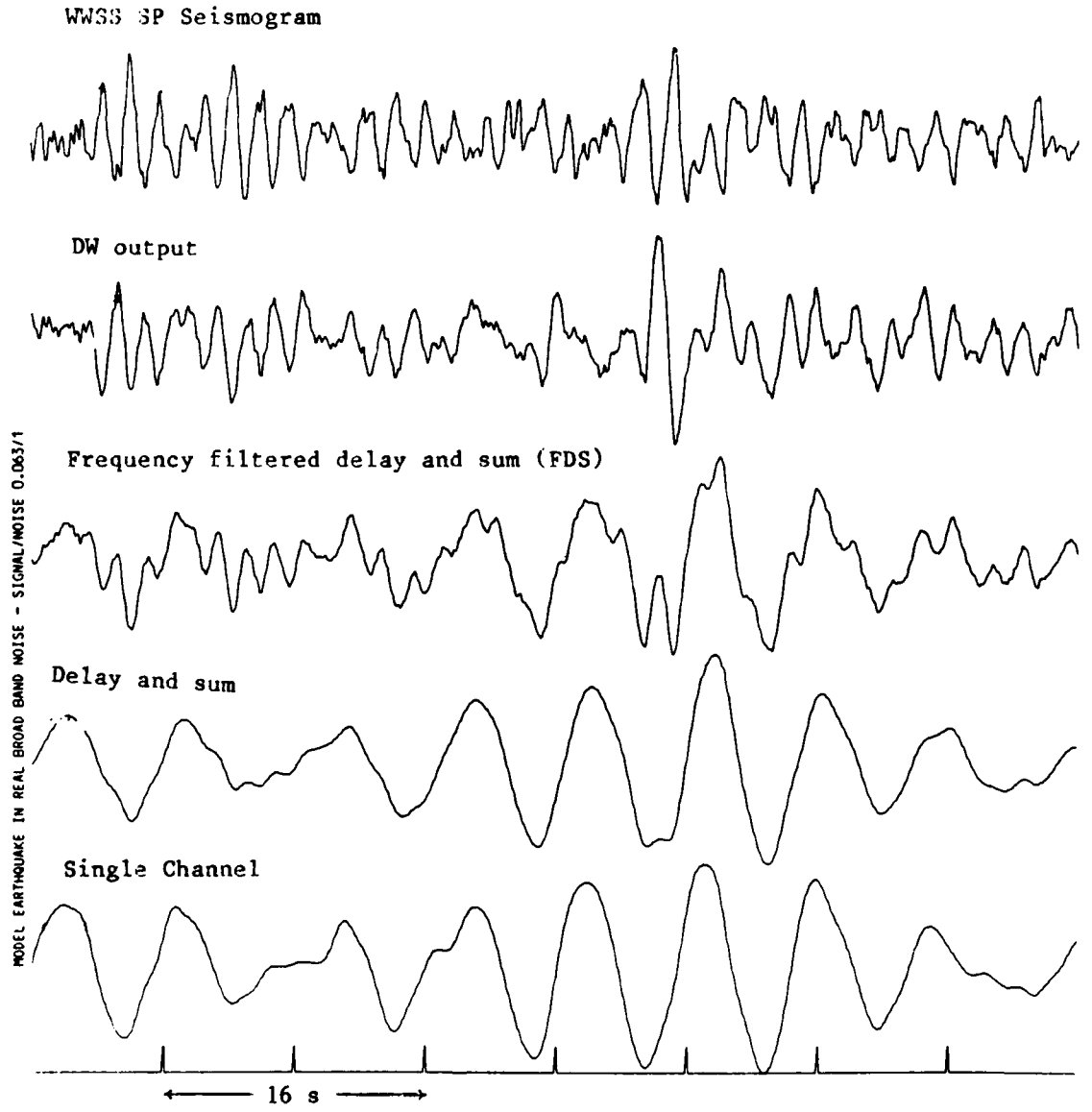


FIGURE 11(b) EXAMPLE OF THE STANDARD OUTPUT OF THE DATA-DEPENDENT WIENER (DW) FILTERING PROGRAM. DATA IS THE ARTIFICIAL EARTHQUAKE SIGNAL IN NOISE FOR SIGNAL-TO-NOISE RATIO 0.0625:1. THE FILTERS WERE DESIGNED USING SIGNAL C (FIGURE 5) AS THE SIGNAL MODEL. THE FOLLOWING SEISMOGRAMS ARE SHOWN: WWSS SP SEISMOGRAM; DW OUTPUT; DELAY AND SUM OUTPUT AFTER FILTERING WITH THE FREQUENCY COMPONENT OF THE DW FILTERS; THE DELAY AND SUM OUTPUT; AND THE OUTPUT FROM A TYPICAL SEISMOMETER OF THE ARRAY

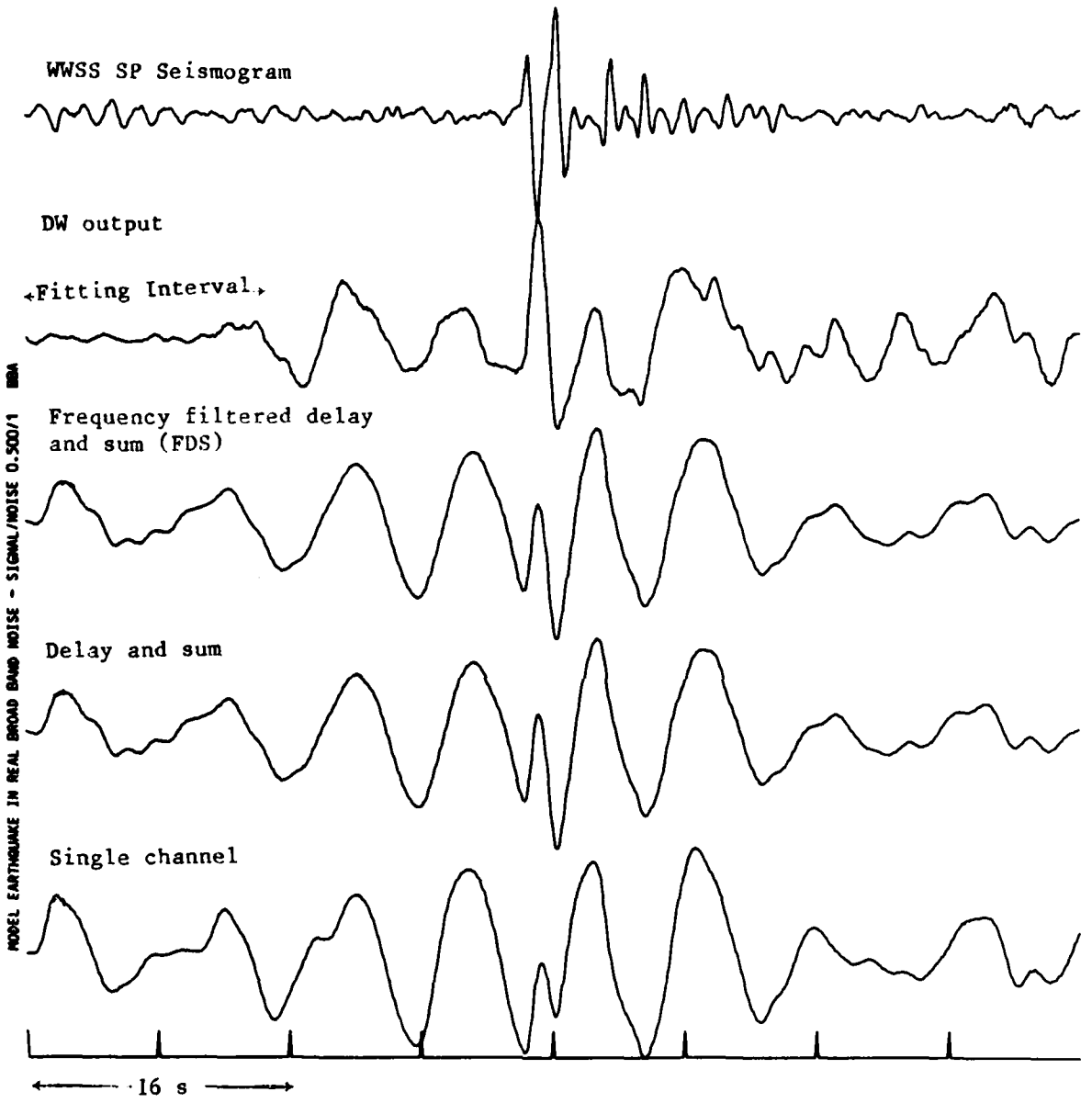


FIGURE 12. STANDARD OUTPUT OF DW FILTERING PROGRAM TO ILLUSTRATE "SUPERGAIN". DATA IS THE ARTIFICIAL EARTHQUAKE SIGNAL IN NOISE FOR SIGNAL-TO-NOISE RATIO 0.5:1. THE FILTERS WERE DESIGNED USING SIGNAL C (FIGURE 5) AS THE SIGNAL MODEL. NOTE THAT IN THE FITTING INTERVAL WHICH HERE IS ONLY 160 POINTS THE NOISE REDUCTION IS APPARENTLY VERY GREAT; OUTSIDE THE INTERVAL HOWEVER VERY LITTLE NOISE REDUCTION IS OBTAINED

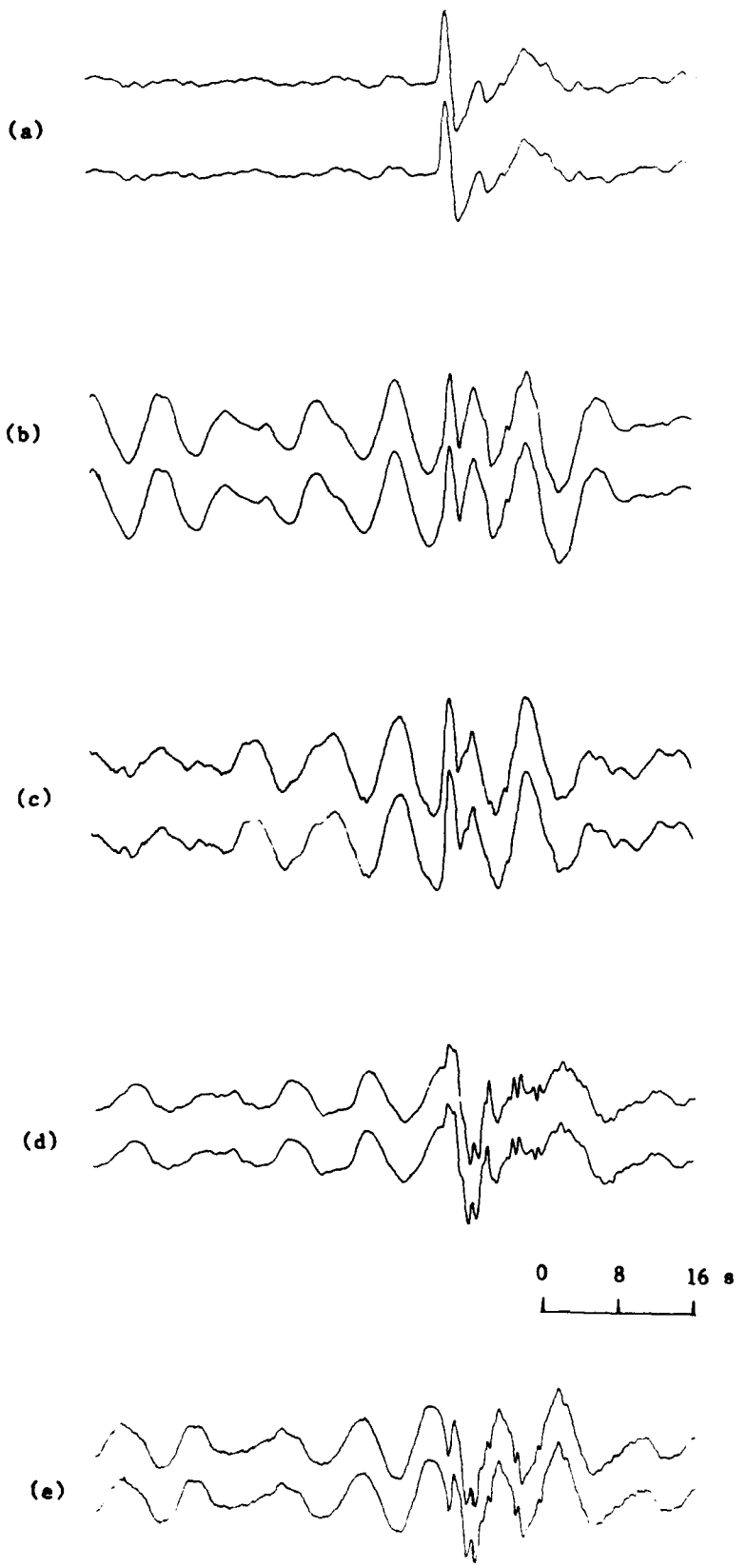


FIGURE 13. COMPARISON OF DATA-DEPENDENT WIENER (DW) FILTERING AND DATA-DEPENDENT MINIMUM POWER (DMP) FILTERING. THE OUTPUTS OF THE FILTERS FOR EACH INDIVIDUAL CHANNEL (b, c, d and e) ARE SHOWN TOGETHER WITH THE SUMMED OUTPUT (a). THE TOP TRACE OF EACH PAIR IS THE RESULT OF DMP FILTERING, THE LOWER THE RESULT OF DW FILTERING. DATA IS THE ARTIFICIAL EARTHQUAKE SIGNAL IN NOISE FOR SIGNAL-TO-NOISE RATIO 0.5:1 BUT A SIGNAL-TO-NOISE RATIO OF 64:1 HAS BEEN ASSUMED IN DESIGNING THE DW FILTERS

All the examples of processing shown above were made with filters of 39 points per channel. This length was chosen because it is the longest filter that can conveniently be fitted into the computer we use. Figure 14 shows how the noise reduction varies with filter length for DW filtering using purely spatial filters. Figure 14 shows that, although the noise is reduced to lower and lower amplitudes as the filter length increases, there is little to be gained from using much longer filters. Note that for the sample of noise used to derive this graph the noise reduction for one point filters is about the same as for DS processing, showing that the noise reduction does not come from weighted DS rather than wave-number filtering.

Figure 15(a) shows the relative power response of the BNA as a function of wave number for straight summing of the four channels of the array; the wave number of the coherent component in the noise sample of 20 January 1976 is shown. This response predicts a noise power reduction of about 2.0 will be obtained by straight summing of the array channels. This noise reduction is about 1.4 in amplitude which is close to 1.5, the value actually obtained. Figure 15(b) shows the BNA power response as a function of wave number for 0.13 Hz (7.6 s period) after applying the DW filters estimated for the noise sample of 20 January 1976. Note that now the response shows a null at the wave number of the coherent noise, yet the response to signals at zero wave number is unity.

The examples of DW (and DMP) processing shown above are idealised because the artificial signal is identical on all channels whereas observed signals will always show some differences between channels and we point out in section 4 that it is important to consider the effects of departures of the signal from perfect coherence.

Consider the Novaya Zemlya explosion signal discussed in section 3 and shown to depart significantly from perfect coherence. Figure 16 shows the result of applying DW filtering and, although significant noise reduction (due to spatial filtering) is obtained, the signal is distorted; the DW filtered signal (figure 16(b)) shows precursors to the signal onset and large amplitude high frequency arrivals in the coda which are not shown by the DS output (figure 16(d)); the DW filtered output does not look like the DS with the noise removed. The distortion

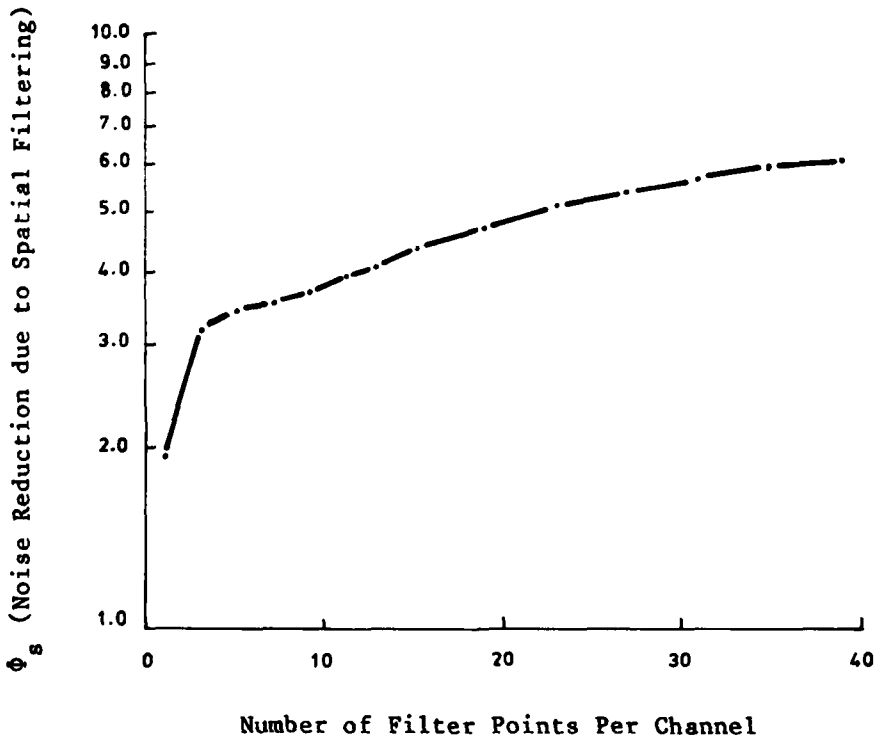


FIGURE 14. VARIATION OF NOISE REDUCTION WITH FILTER LENGTH FOR DATA-DEPENDENT WIENER FILTERS APPLIED TO NOISE SAMPLE RECORDED ON 20 JANUARY 1976. NOISE REDUCTION IS ALL DUE TO SPATIAL FILTERING

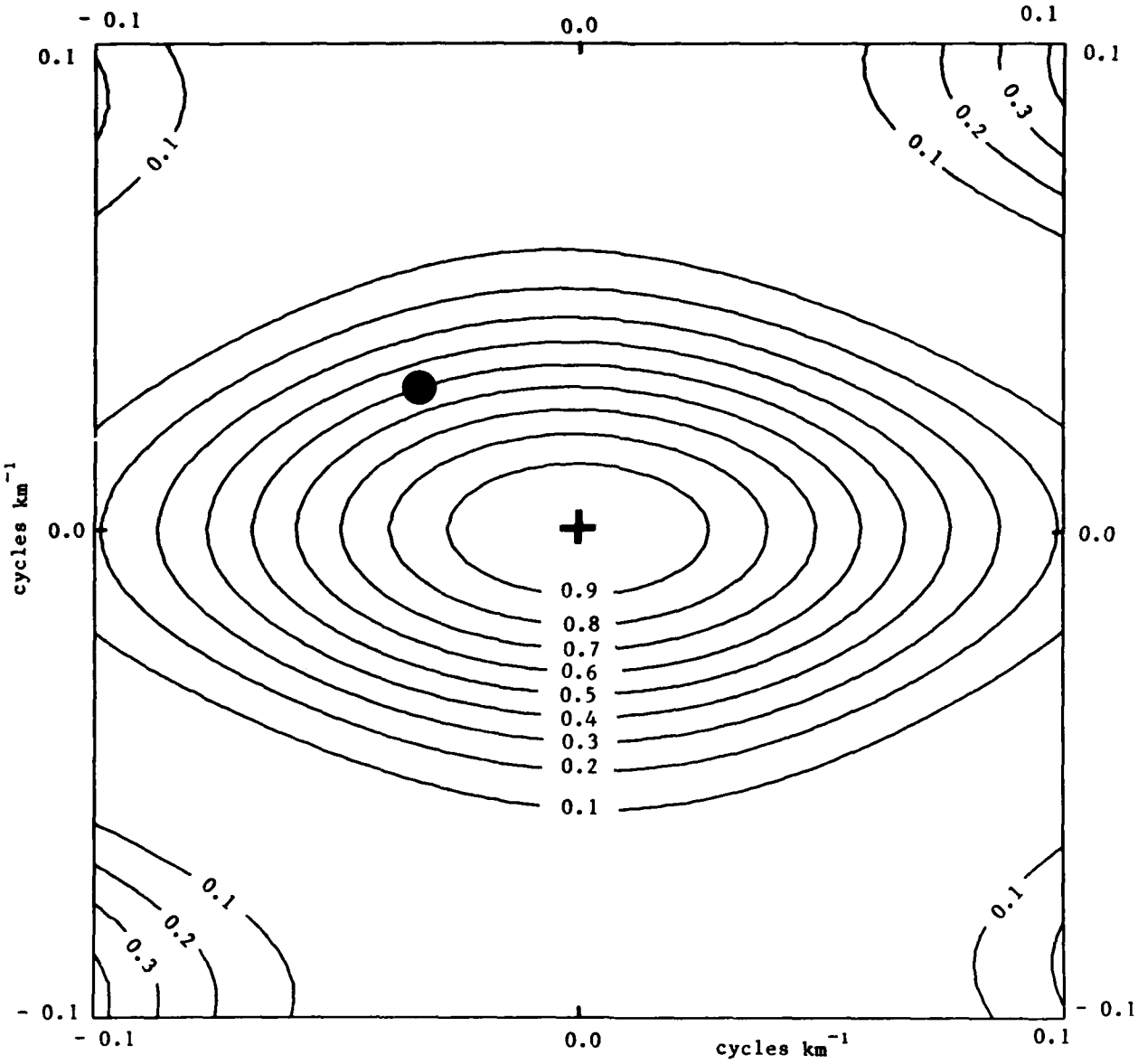


FIGURE 15(a) POWER RESPONSE OF THE BLACKNEST ARRAY AS A FUNCTION OF WAVE NUMBER FOR STRAIGHT SUMMING OF THE FOUR CHANNELS OF THE ARRAY; THE WAVE NUMBER OF THE COHERENT COMPONENT IN THE NOISE SAMPLE OF 20 JANUARY 1976 IS MARKED

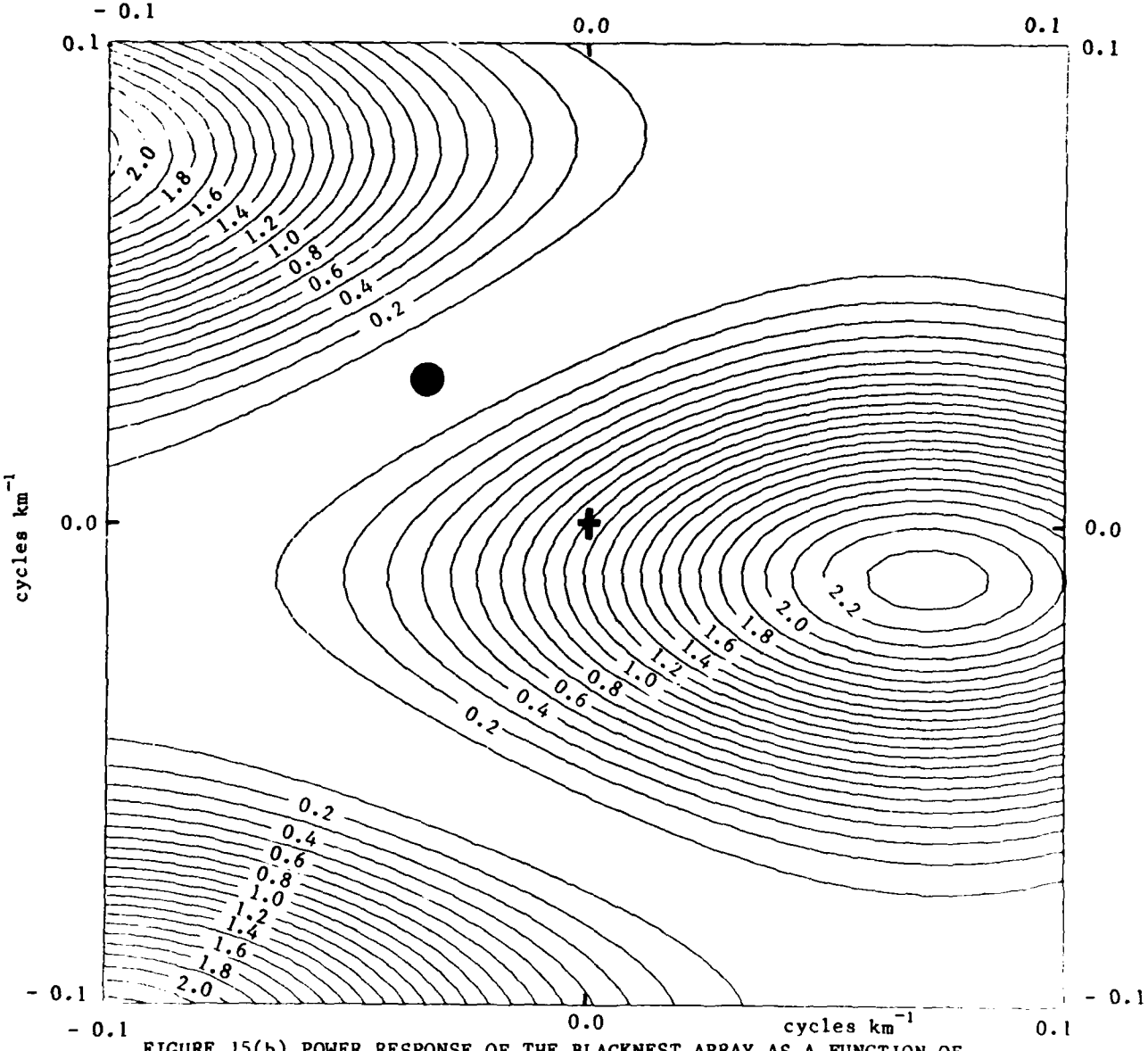


FIGURE 15(b) POWER RESPONSE OF THE BLACKNEST ARRAY AS A FUNCTION OF WAVE NUMBER AT 0.13 Hz (7.6 PERIOD) AFTER APPLYING DW FILTERS ESTIMATED FOR THE NOISE SAMPLE OF 20 JANUARY 1976. THESE FILTERS ARE DESIGNED TO SUPPRESS THE MAIN SOURCE AND PASS SIGNALS AT ZERO WAVE NUMBER

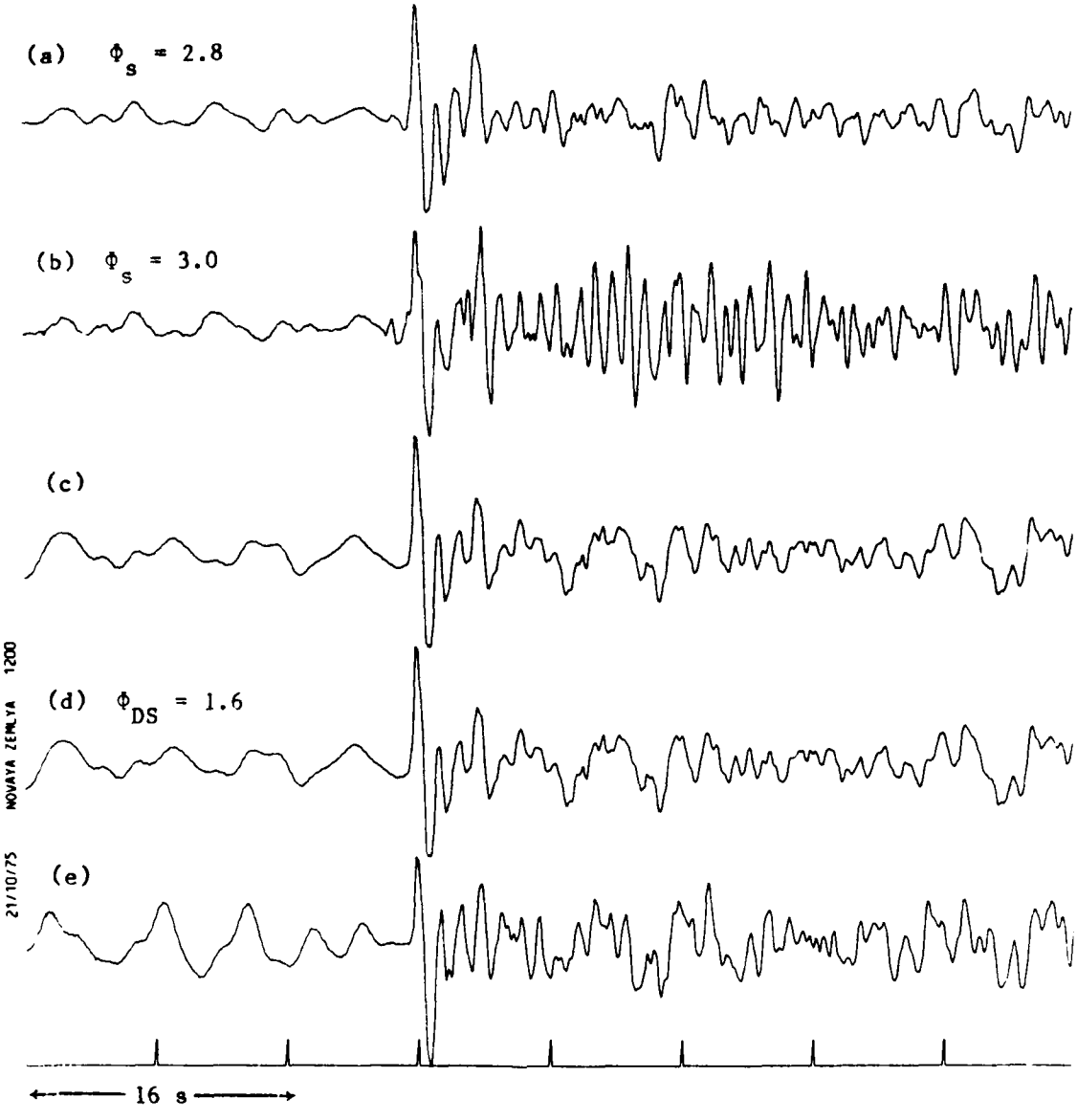


FIGURE 16. DATA-DEPENDENT WIENER (DW) FILTERING OF SIGNALS FROM THE NOVAYA ZEMLYA EXPLOSION ON 21 OCTOBER 1975 TO SHOW SIGNAL DISTORTION DUE TO SIGNAL INCOHERENCE AND SUPPRESSION OF DISTORTION BY ADDITION OF WHITE NOISE
(a) DW OUTPUT WITH ADDITION OF WHITE NOISE
(b) DW OUTPUT WITHOUT WHITE NOISE
(c) DELAY AND SUM OUTPUT FILTERED WITH FREQUENCY COMPONENT OF DW FILTERS USED IN (b)
(d) DELAY AND SUM OUTPUT
(e) OUTPUT FROM A SINGLE CHANNEL OF THE BNA

of the signal appears to arise from the spatial component of the DW filters because, when the DS is filtered with the frequency component of the DW filters, no distortion is seen (figure 16(c)). Capon et al. (6) found similar effects when applying DMP filtering to the Large Aperture Seismic Array (LASA) data and demonstrated that the precursors arise because of differences in the amplitude of the signal between channels. The large amplitude high frequency arrivals on the DW output appear to arise because effectively the noise properties are different before and after the signal onset; before onset the noise is predominantly low frequency oceanic microseisms; after onset the noise is a mixture of this low frequency noise plus high frequency noise generated by the scattering of the signal and which is shown in section 3 to have low coherence. Consequently, when filters designed on one type of noise, that before the onset, are applied to noise with different properties after onset it is perhaps not surprising that the additional noise that is present after the onset of the signal is amplified rather than reduced.

As the noise generated by scattering appears to be incoherent between channels one possible way to suppress it when using DW filtering is to add a component of incoherent noise to the observed noise; filters are then designed to suppress not only the observed noise but also any incoherent components. To include the effects of a component of incoherent white noise with mean square amplitude σ^2 in the filter estimation process it is only necessary to add σ^2 to the first element of the auto-correlation function of the noise on each channel. Figure 16(a) shows the results of DW filtering using filters estimated with a component of incoherent white noise with mean square amplitude 0.1 times the mean square amplitude of the noise on channel 1. Now the estimated signal is effectively the DS minus the low frequency noise.

The way the addition of the small component of uncorrelated white noise works to suppress the precursors to the signal and the high frequencies in the coda and yet still allows reduction of the oceanic microseisms appears to be as follows. The amplitude of the low frequency components of the noise (0.125 to 0.15 Hz) are so large relative to the high frequencies that the addition of a small proportion of white noise has little influence on the low frequency response of the computed filters. The high frequency components of the observed noise which have very low amplitude relative to the low frequency components are coherent

between channels at some frequencies, probably because some of this noise is instrumental. If a component of white noise uncorrelated between channels is not added, then the designed filter attempts to reduce the coherent components of the instrumental noise by spatial filtering. When the computed filters are now applied to a signal that contains a large proportion of scattered high frequency seismic energy with properties very different from the instrumental noise, the effect of the filters is to amplify the scattered signal. The addition of the uncorrelated white noise swamps the coherent components in the observed noise and weights the high frequency components of the filters to be DS filters. Note that the signal shown in figure 16 was recorded on the DBB system (see section 2) for which the system noise at frequencies above about 2 Hz tends to be larger than the seismic noise. Signal distortion on DW filtering appears to be less serious using IVBB recordings probably because for such recordings the system noise at high frequencies lies well below the seismic noise.

The procedure of adding a component of uncorrelated white noise may also have other advantages to that described above. For if the noise on two channels is identical, so that normally there would be no unique solution, then adding white noise stabilises the solution. For a two channel array the solution would be DS filters. The addition of uncorrelated noise might also result in estimated filters that give a much smaller reduction in the noise than obtained by filters estimated from the observed noise only. Thus, for the two channel example discussed in section 4, where the noise is perfectly correlated but differs slightly in amplitude, it is possible to find filters (specified in equation (18)) that reduce the noise exactly to zero. Addition of a component of white noise would attempt to weight the solution to DS filters for which noise reduction could be almost nil. This would appear to be a disadvantage of adding white noise but, as pointed out in section 4, filters that make use of small amplitude differences between two channels of correlated noise will usually be undesirable because, unless the signals are identical on both channels, such filters may distort the signal and the expected signal-to-noise improvement will not be obtained. The effect of adding a component of uncorrelated white noise to weight the solution towards DS filters will in such circumstances be an advantage.

The distortion of the signal in the absence of white noise can be looked upon as an example of the effect of ill-conditioning of the normal equations for simply changing the diagonal elements of $\underline{R}$ by small amounts to simulate the presence of a small amount of uncorrelated white noise results in a large change in the filters. This is illustrated in figure 17(a) which shows filters estimated with no uncorrelated white noise added and those with noise with mean square amplitude 0.01 of the mean square amplitude of the observed noise. Note the large change in the filter coefficients that results from this small change in the $\underline{R}$ matrix.

For single channel Wiener there also appears to be an advantage in adding a small proportion of white noise because, at frequencies above, say, 4 Hz, the signal and noise amplitudes are very small relative to the amplitudes at low frequencies (say, around the frequencies of the oceanic microseisms) so that the response of the Wiener filter will only be well defined at the low frequencies; at high frequencies, provided that the signal and noise is not greatly amplified, the effect on the mean square amplitude at the output would be negligible. As shown in figure 17(b) the addition of a small component of white noise ensures that at high frequencies the response is well controlled and falls off systematically towards the high frequencies.

From the experiments described above with adding white noise it would appear that the best way of constructing $\underline{R}$ is from a mixture of observed data and a stochastic (white) noise model. In all the examples that follow this way of setting up $\underline{R}$ is used. Experiments have been carried out using different proportions of white noise but it appears that, apart from rather special cases such as that shown in figure 16, a mean square amplitude of the white noise of about 0.01 of the mean square of the observed noise is adequate to stabilise the filters. All the examples shown here of the effects of the addition of white noise are for DW filtering but similar effects are obtained for DMP filtering.

The DW and DMP filtering described above was done using two-sided filters although almost the same noise reduction is obtained using one-sided filters for DMP filtering or for DW filtering with predominantly spatial filters (as is expected, see appendix B). For DW filtering where there is a significant component of frequency filtering there are advantages in using two-sided rather than one-sided filters. The advantage is illustrated in figure 18 which shows the

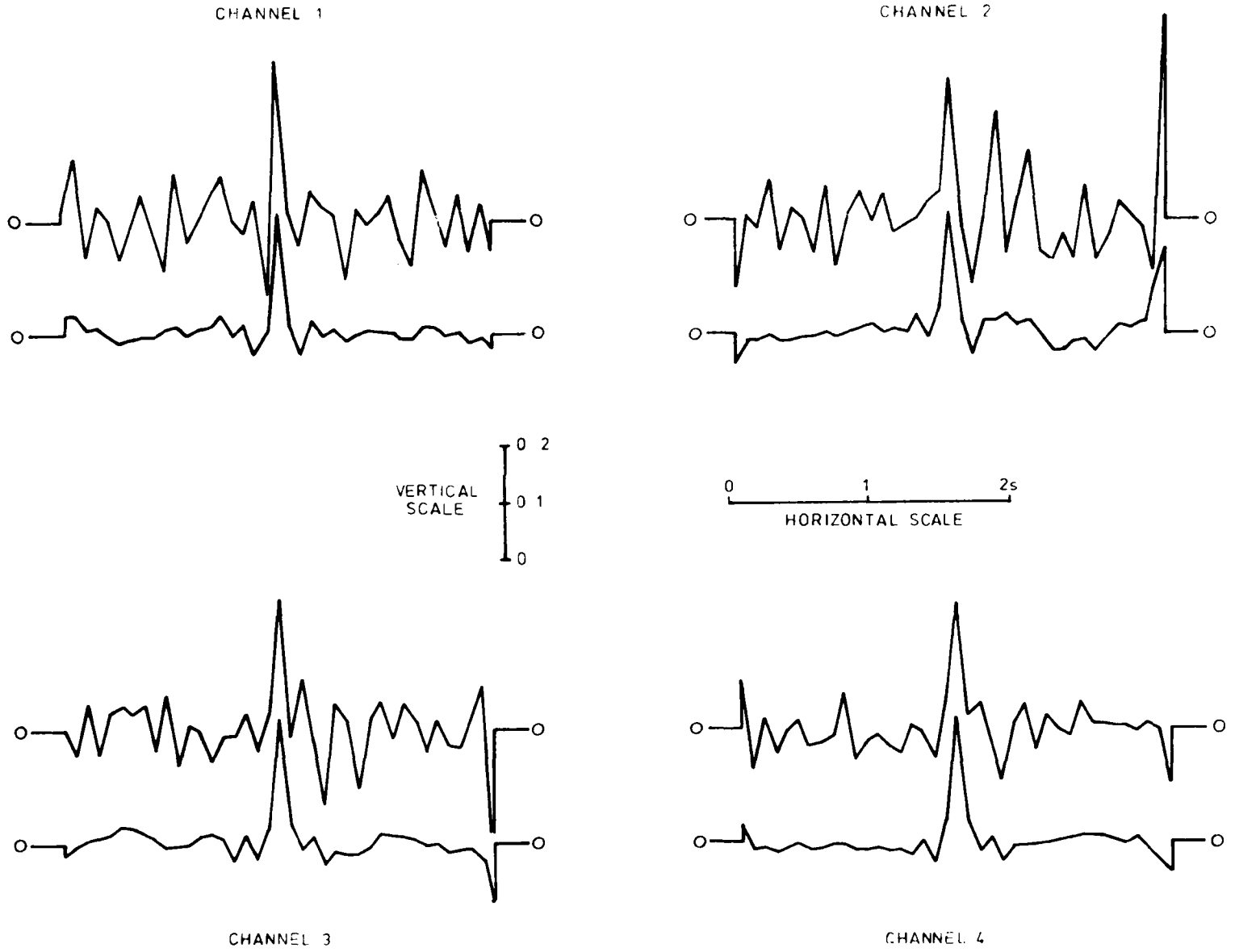


FIGURE 17(a) DATA-DEPENDENT WIENER FILTERS FOR THE 4 CHANNELS OF THE BLACKNEST ARRAY ESTIMATED FOR THE SEA OF OKHOTSK EARTHQUAKE (FIGURE 20) SHOWING THE EFFECT OF ADDING A COMPONENT OF WHITE NOISE TO THE OBSERVED NOISE

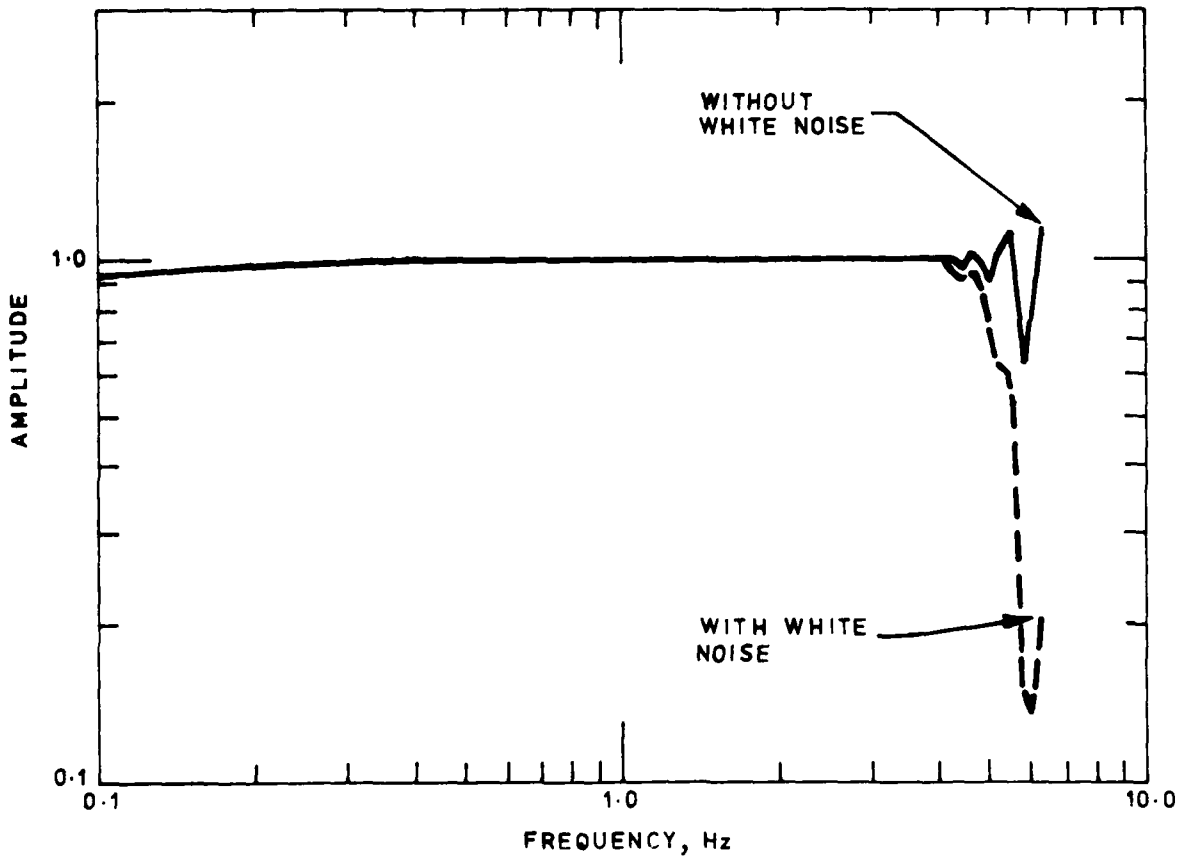


FIGURE 17(b) AMPLITUDE RESPONSES OF SINGLE CHANNEL DATA-DEPENDENT WIENER FILTERS ESTIMATED FOR THE DELAY AND SUM OUTPUT OF THE EAST KAZAKHSTAN EXPLOSION (FIGURE 25) SHOWING THE EFFECT OF ADDING A COMPONENT OF WHITE NOISE TO THE OBSERVED NOISE

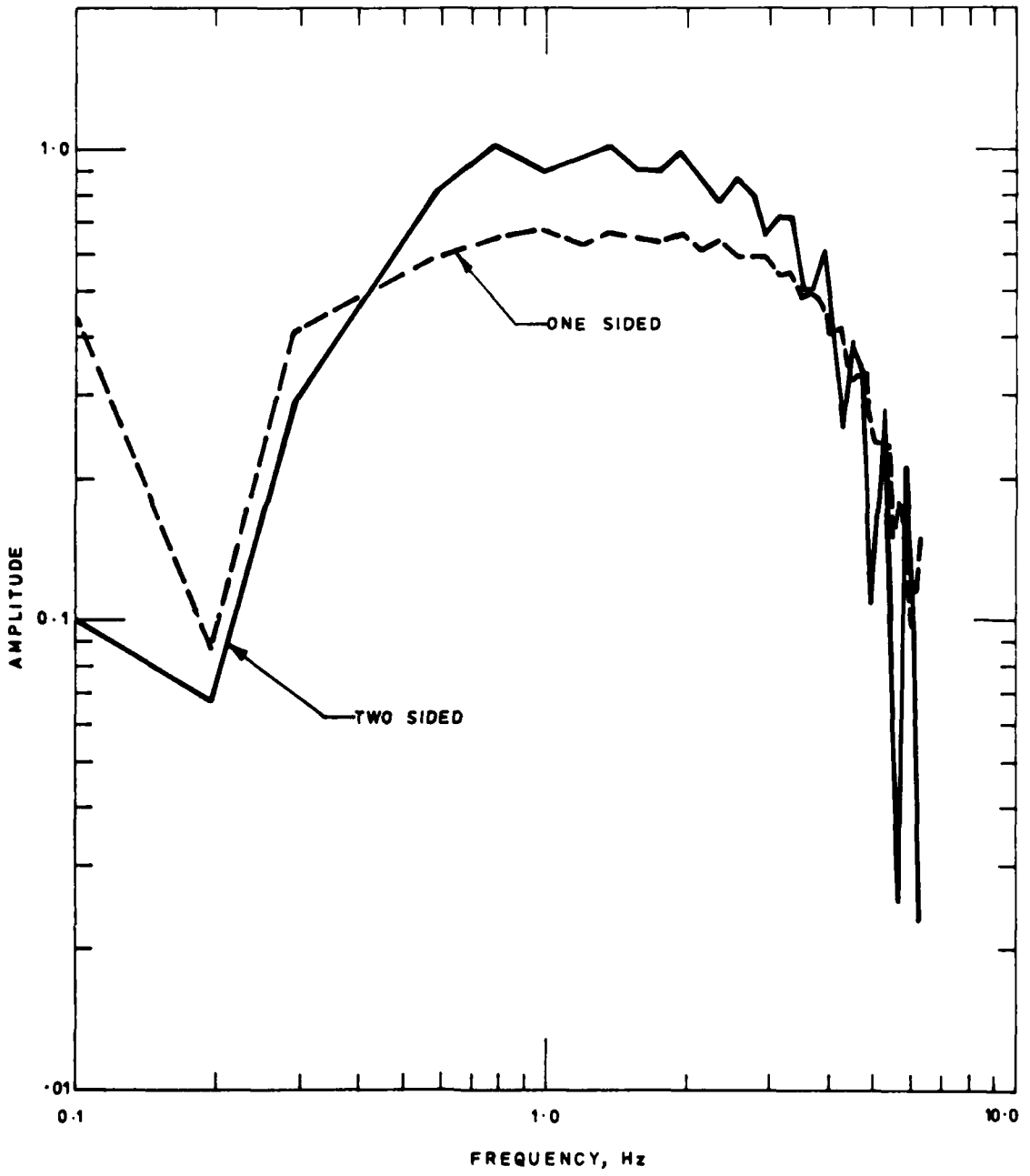


FIGURE 18. COMPARISON OF AMPLITUDE RESPONSE OF ONE-SIDED AND TWO-SIDED DATA-DEPENDENT WIENER FILTERS. FILTERS ESTIMATED WITH IDENTICAL NOISE AND SIGNAL DATA

amplitude response at zero wave number of two-sided and one-sided filters derived from the same data; figure 19 shows the DW output after applying these filters. Note that one-sided filters give a noise reduction over all frequencies (amplitude response everywhere less than unity) whereas the two-sided filter does approach unity at around 0.8 Hz. Clearly the one-sided filter is unsatisfactory because an apparent noise reduction is obtained by lowering the magnification of the filters. The reason why this occurs for the one-sided filter appears to be as follows. The signal-to-noise ratio requires a certain noise reduction which cannot be obtained by spatial filtering. To obtain the noise reduction by frequency filtering without lowering the magnification requires a frequency filter with an amplitude spectrum similar to that of the impulse response of the two-sided filter which has steeper gradients than that of the one-sided filters. To avoid signal distortion due to phase shifts when using a one-sided filter requires that the phase shifts be small. However, the phase and amplitude response of one-sided filters cannot be set independently; given the amplitude response the minimum phase is set. Thus, it would appear that there is no one-sided filter with the required amplitude and phase spectrum - the only remaining option for reducing the noise is to reduce the magnification of the filters below unity. Two-sided filters should thus be used wherever possible and this is done in what follows.

Two-sided filters do have a disadvantage if the frequency filtering component is large because they then usually generate precursors (see figure 8 for examples) which may make measurements of the onset time or first motion difficult. In practice, however, we find that making allowance for the precursors is not difficult and, as the narrow band SP seismogram is also available, this can be used to assist in picking the arrival time.

Figures 20 to 25 show examples of DW filtering applied to the BB signals from the 5 earthquakes and one explosion listed in table 5; the array magnitude m_b^A and the observed SP and BB amplitudes listed were measured as described in section 3. All the examples shown use a 2048 point fitting interval, except the earthquake of 7 January 1976 (figure 24) where 1024 points had to be used because the presence of tape faults made use of a longer fitting interval impossible. Most of the signals were recorded during the winter when the oceanic microseisms were of large amplitude and so it was expected significant noise reduction due to spatial filtering might be possible. The model signal used in the filter design is the attenuation - seismograph impulse response (signal C, figure 5). We now consider each of the examples briefly, noting some of the most important features.

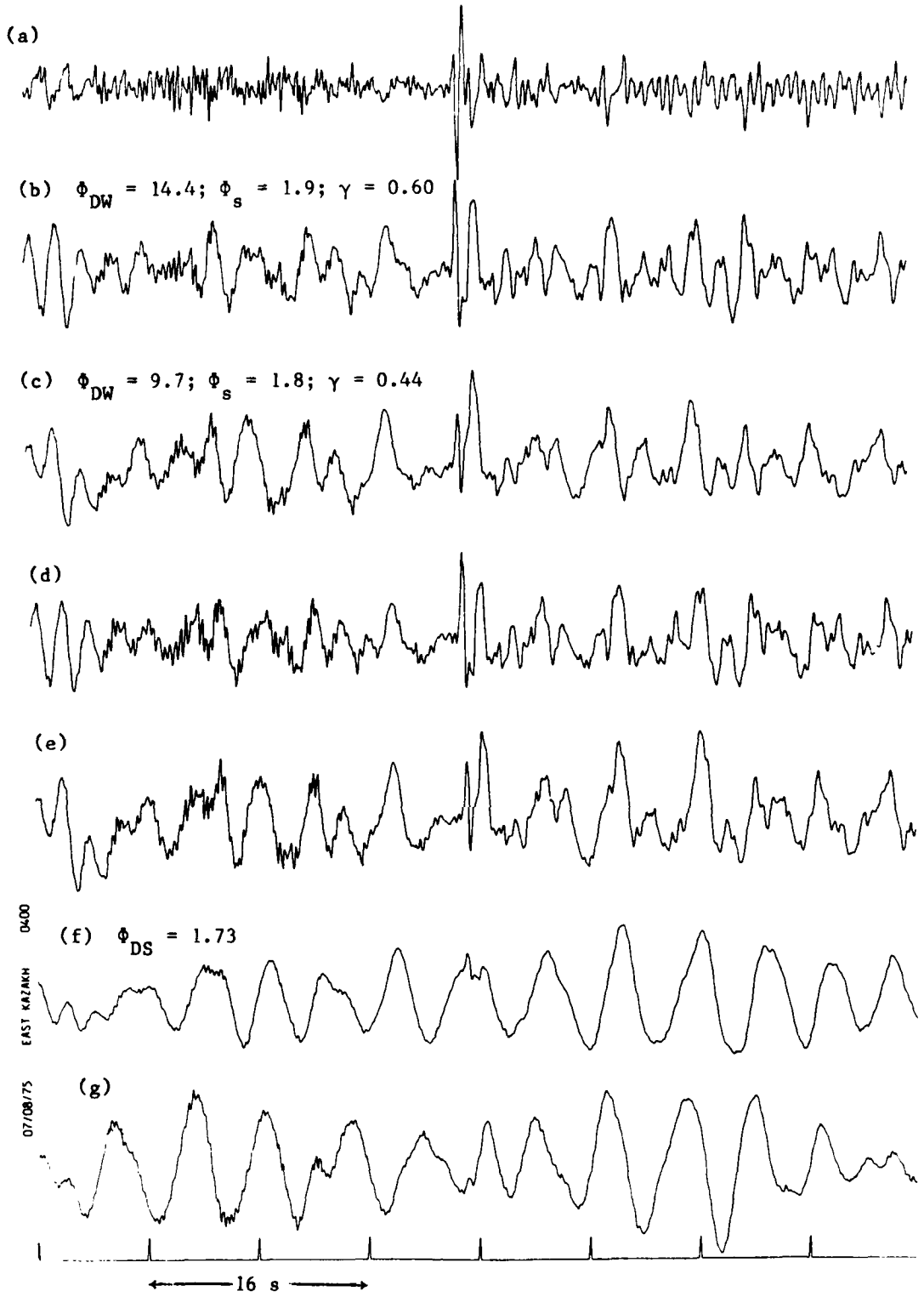


FIGURE 19. COMPARISON OF DATA-DEPENDENT WIENER (DW) FILTERING USING ONE-SIDED AND TWO-SIDED FILTERS. OBSERVED SIGNAL IS FROM AN EXPLOSION IN EAST KAZAKHSTAN ON 7 AUGUST 1975

(a) WWSS SP SEISMOGRAM
(b) DW OUTPUT: TWO-SIDED FILTER
(c) DW OUTPUT: ONE-SIDED FILTER
(d) DELAY AND SUM OUTPUT FILTERED WITH FREQUENCY COMPONENT OF DW FILTERS: TWO-SIDED FILTERS
(e) DELAY AND SUM OUTPUT FILTERED WITH FREQUENCY COMPONENT OF DW FILTERS: ONE-SIDED FILTERS
(f) DELAY AND SUM OUTPUT
(g) SINGLE CHANNEL FROM THE BNA ARRAY

TABLE 5

Earthquakes and Explosions Used for Studies of Data-Dependent Wiener Filtering

Date	Region	Origin Time	Depth, km	Distance, degrees	Obs SP Amp, nm	SP Ground Motion, nm	T, s	m_b^A	m_b	Maximum BB Amp, nm
21 December 1975	Sea of Okhotsk	10:54:17.7	554.0	69.3	1170	1170	1	7.1	6.0	7684
9 January 1976	New Hebrides	23:54:35.6	168.0	143.3	100	1064	2.3	6.7	6.1	2000
20 January 1976	Tonga	19:16:12.0	33	150.0	129	129	1	5.6	5.4	850
6 January 1976	Off East Coast of Kamchatka	22:17:47.9	33	75.6	382	272	0.75	6.5	5.6	595
7 January 1976	Off East Coast of Kamchatka	04:32:46.8	33	75.5	56	40	0.75	5.6	5.0	53
7 August 1975	East Kazakhstan	03:56:57.5	0	48.0	52.5	42.3	0.5	5.8	5.2	48

21 December 1975 - Earthquake: Sea of Okhotsk

This earthquake has a large amplitude on both the SP (1170 nm) and BB seismogram (7684 nm) and can clearly be seen above the noise on the single channel of the BNA (figure 20(e)). The DW filter output (figure 20(b)) gives a noise reduction due to DS processing of 2.0; the total noise reduction due to spatial filtering is 5.9. The effects of frequency filtering are negligible as can be seen by comparing the DS and FDS outputs (figure 20(c) and (d)). Note that the BB seismograms show two distinct pulses (A_1 and A_2) separated by about 4 s suggesting a double earthquake. On the SP seismogram, on the other hand, the signal is complex and cannot be interpreted. The onset is as clear on the BB as on the SP and the first motion on the BB seismogram is almost as large as the maximum amplitude on the record, whereas on the SP the first motion is only about 0.125 of the maximum amplitude.

9 January 1976 - Earthquake: New Hebrides

The P arrival shown in figure 21 (which is PKP) contains little high frequency energy so that on the SP seismogram the predominant period of the signal is not around 1 s but is about 2.3 s. Consequently the amplitude as seen on the SP is much smaller (100 nm) than on the BB seismograms (2000 nm). After DW filtering, which applies negligible frequency filtering, the signal-to-noise ratio on the BB seismograms is almost three times that on the SP. Note that the onset is more difficult to pick on the SP than on the BB; the first motion on the BB is clearly downwards, whereas on the SP it is not clear which is first motion. The DW filtered channel appears to have revealed a low frequency arrival (A_3) about a minute after onset which is not shown up clearly by any of the other channels.

20 January 1976 - Earthquake: Tonga

The PKP signal shown is virtually invisible on the single channel (figure 22(e)) although it can be picked out on the DS output (figure 22(d)) by comparison with channels (a), (b) and (c). It is clear from a comparison of the DS channel (figure 22(d)) with the FDS channel (figure 22(c)) that a significant component of frequency filtering is applied by the DW filters. Nevertheless the

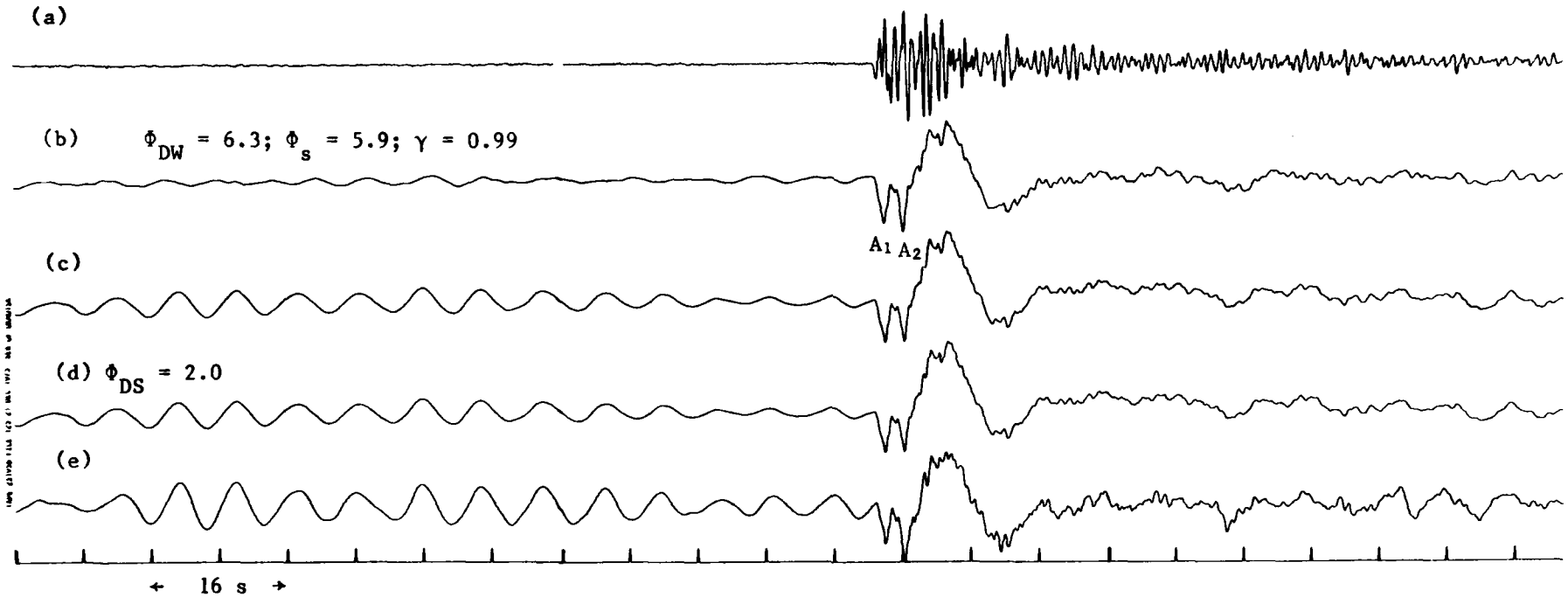


FIGURE 20. DATA-DEPENDENT WIENER (DW) FILTERING OF SEA OF OKHOTSK EARTHQUAKE RECORDED AT THE BNA: 21 DECEMBER 1975
(a) WWSS SP SEISMOGRAM
(b) DW FILTERED OUTPUT
(c) DELAY AND SUM OUTPUT FILTERED WITH FREQUENCY COMPONENT OF DW FILTERS
(d) DELAY AND SUM OUTPUT
(e) OUTPUT FROM SINGLE CHANNEL OF THE ARRAY

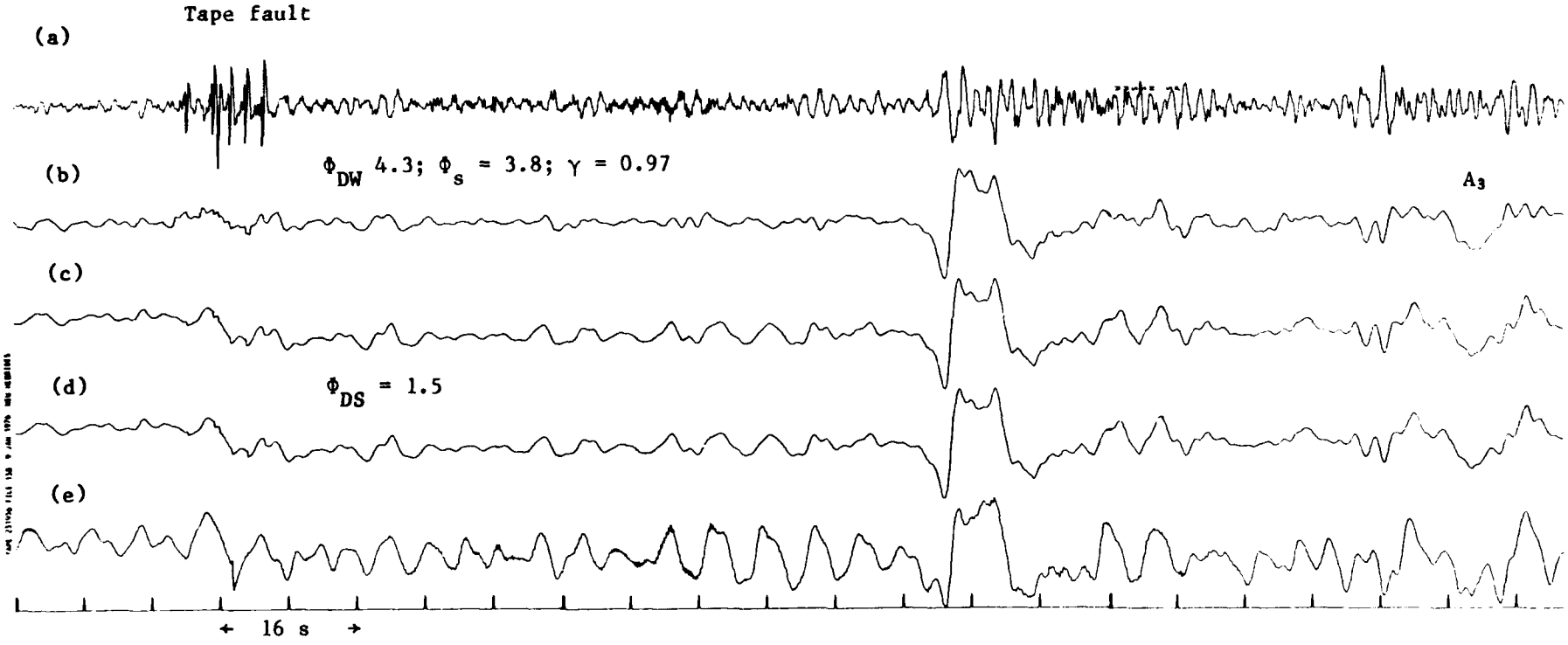


FIGURE 21. DATA-DEPENDENT WIENER (DW) FILTERING OF NEW HEBRIDES EARTHQUAKE RECORDED AT THE BNA: 9 JANUARY 1976
(a) WWSS SP SEISMOGRAM
(b) DW FILTERED OUTPUT
(c) DELAY AND SUM OUTPUT FILTERED WITH FREQUENCY COMPONENT OF DW FILTERS
(d) DELAY AND SUM OUTPUT
(e) OUTPUT FROM SINGLE CHANNEL OF THE ARRAY

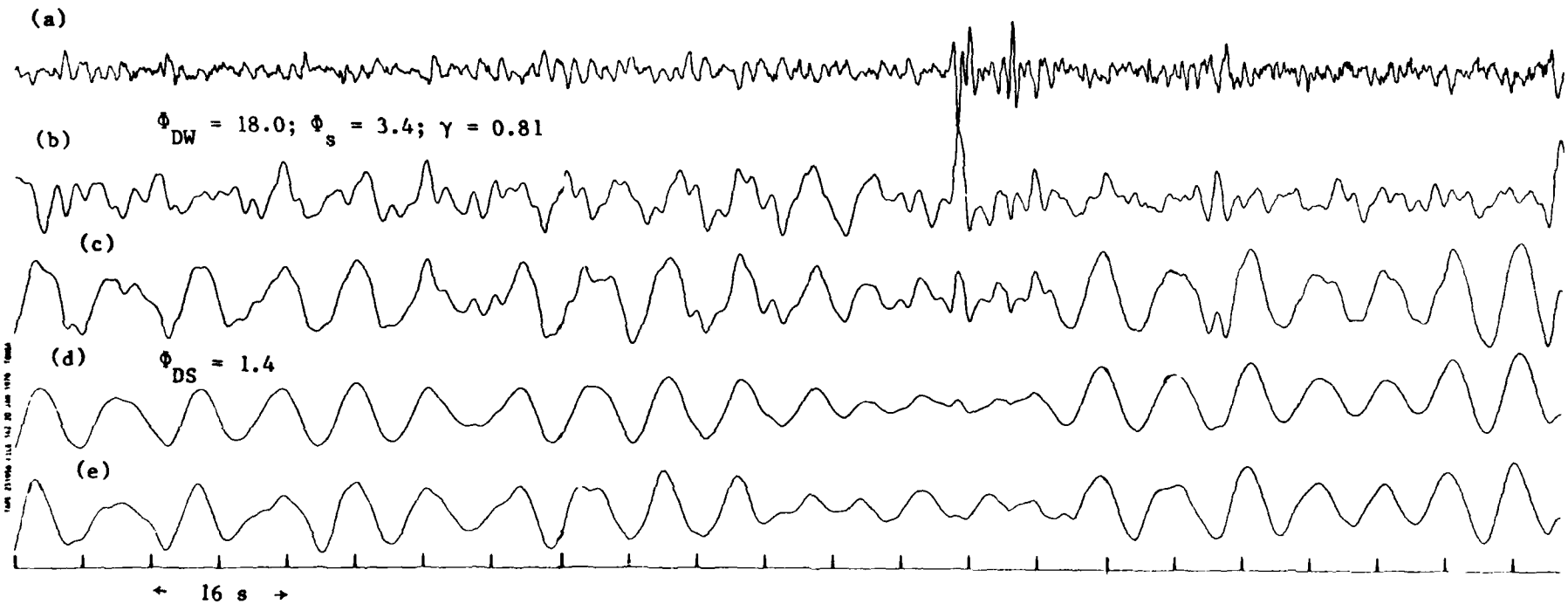


FIGURE 22. DATA-DEPENDENT WIENER (DW) FILTERING OF TONGA EARTHQUAKE RECORDED AT THE BNA: 20 JANUARY 1976
(a) WWSS SP SEISMOGRAM
(b) DW FILTERED OUTPUT
(c) DELAY AND SUM OUTPUT FILTERED WITH FREQUENCY COMPONENT OF DW FILTERS
(d) DELAY AND SUM OUTPUT
(e) OUTPUT FROM A SINGLE CHANNEL OF THE ARRAY

spatial component of the DW filters also contributes significantly to the noise reduction; this is particularly clear after signal onset where the low frequency noise seen on channel (c) is much reduced in amplitude on channel (d). the signal-to-noise ratio is slightly poorer on the DW output than on the SP but first motion on the DW output is the largest amplitude on the record. Note that the second arrival on the SP about 8 s after onset has a similar amplitude to the first SP arrival whereas on the DW output the second arrival is much smaller than the first showing that the second arrival has relatively more high frequency energy than the first arrival.

6 January 1976 - Earthquake: Off East Coast of Kamchatka

The P signal shown is visible on the single channel (figure 23(e)) even though its amplitude is less than the maximum amplitude of the noise because the signal has a higher predominant frequency than that of the noise. Considerable signal-to-noise improvement with little distortion of the signal would be possible by frequency filtering only. Some noise reduction can be obtained by spatial filtering however and the DW filtering takes advantage of this. Note that as the signal is predominantly high frequency the amplitude seen on the SP and the DW filtered BB are similar.

7 January 1976 - Earthquake: Off East Coast of Kamchatka

This earthquake is included here because it has been given a rather low magnitude (m_b 5.0) by the NEIS; the amplitude of the SP signal shown (figure 24(a)) gives m_b^A 5.6. Whatever the true magnitude it is clear that the signal is close to the detection threshold of the BNA. Despite the poor signal-to-noise ratio seen on the DW output there do seem to be advantages in having this BB seismogram in addition to the SP, for the BB seismogram indicates that the pulse A_4 (figure 24(b)) has a smooth leading edge similar to that shown by the Kodiak Island earthquake (figure 4(a)) which results in a low amplitude first motion on the SP seismogram. It is also possible that the large amplitude low frequency arrival A_5 on the DW output is part of the P signal and not low frequency noise but without further data it is impossible to check this.

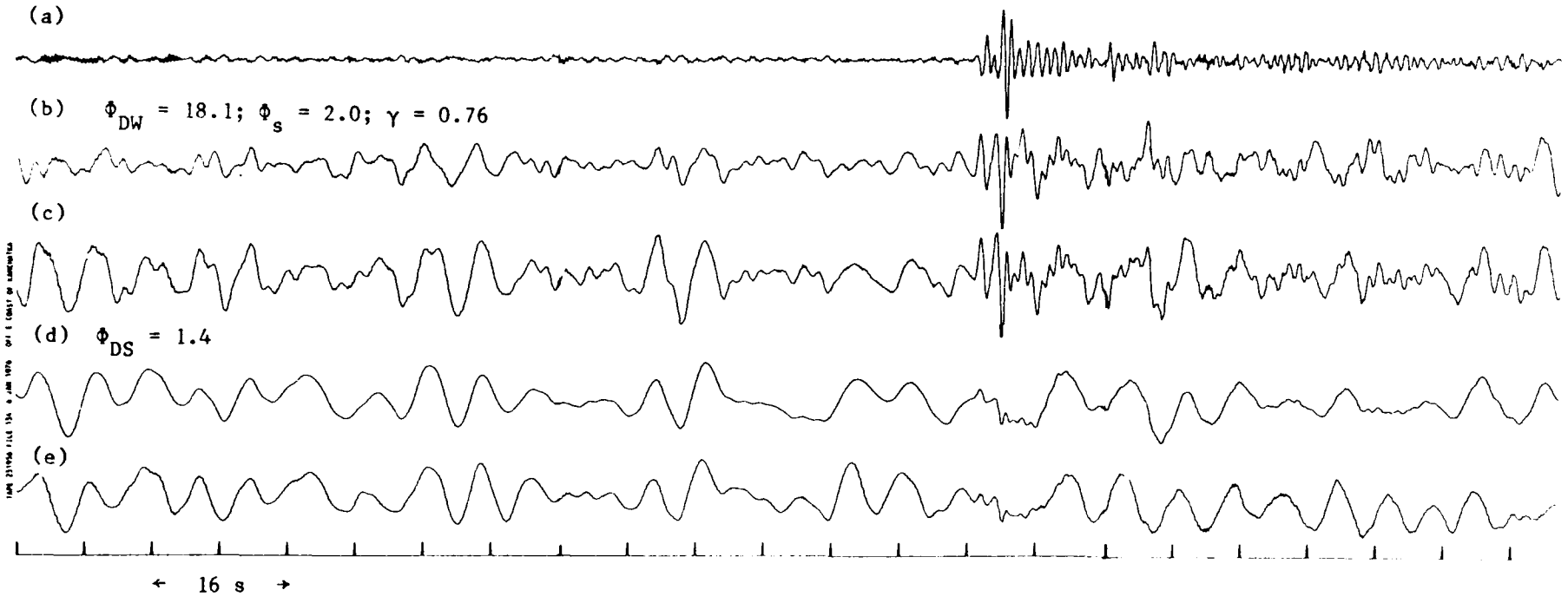


FIGURE 23. DATA-DEPENDENT WIENER (DW) FILTERING OF AN EARTHQUAKE OFF THE EAST COAST OF KAMCHATKA RECORDED AT THE BNA: 6 JANUARY 1976
(a) WWSS SP SEISMOGRAM
(b) DW FILTERED OUTPUT
(c) DELAY AND SUM OUTPUT FILTERED WITH THE FREQUENCY COMPONENT OF THE DW FILTERS
(d) DELAY AND SUM OUTPUT
(e) OUTPUT FROM A SINGLE CHANNEL OF THE ARRAY

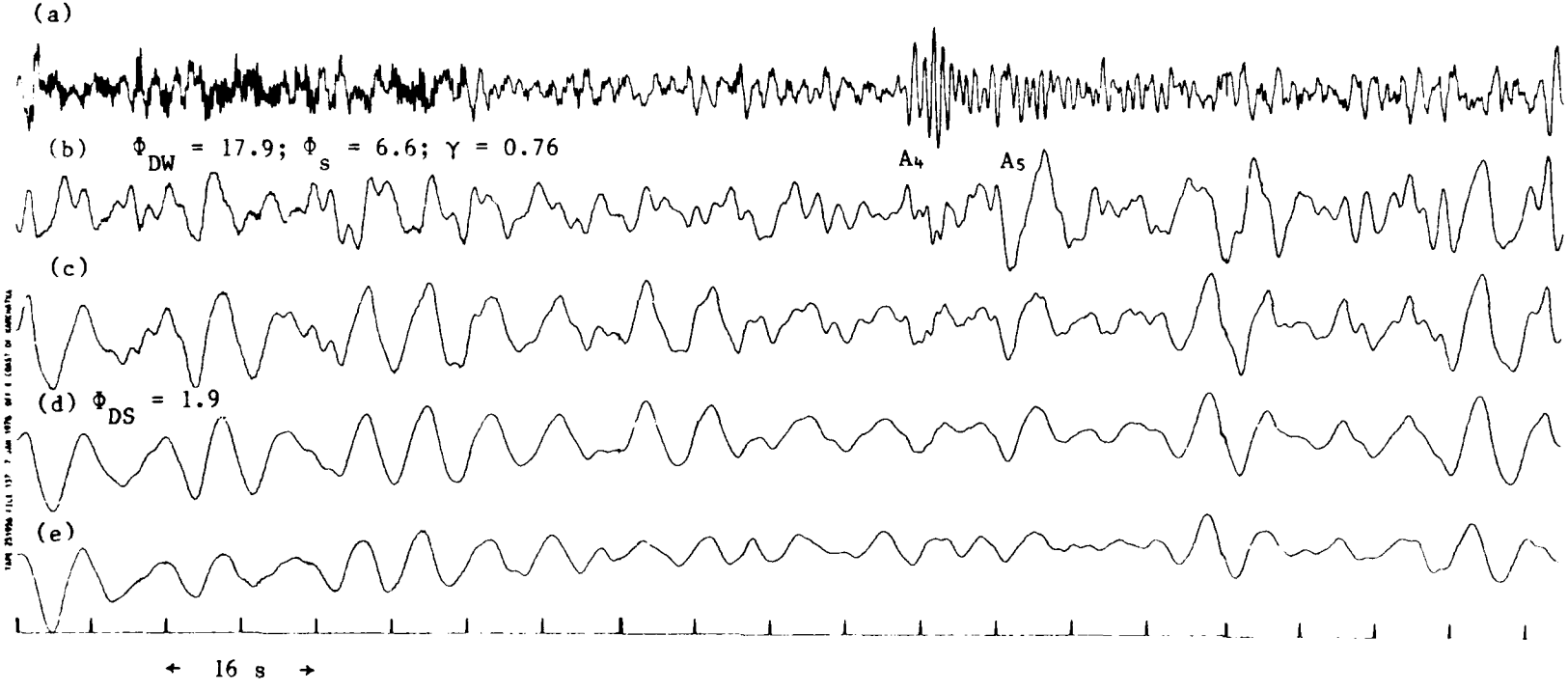


FIGURE 24. DATA-DEPENDENT WIENER (DW) FILTERING OF AN EARTHQUAKE OFF THE *EAST COAST OF KAMCHATKA* RECORDED AT THE BNA: 7 JANUARY 1976
(a) WWSS SP SEISMOGRAM
(b) DW FILTERED OUTPUT
(c) DELAY AND SUM OUTPUT FILTERED WITH THE FREQUENCY COMPONENT OF THE DW FILTERS
(d) DELAY AND SUM OUTPUT
(e) OUTPUT FROM A SINGLE CHANNEL OF THE ARRAY

7 August 1975 - Explosion: East Kazakhstan

The results of DW filtering of the BNA data for this explosion is shown in figure 19 which shows that little spatial filtering is obtained, the main noise reduction coming from frequency filtering. Thus, similar results might be obtained simply by applying a single channel DW filter to the DS output, that is, by applying a single channel DW filter that is purely a frequency filter. The results of applying such a DW filter are shown in figure 25; comparison with the results obtained using multichannel filters (figure 19(b)) shows that the results are very similar.

6. DISCUSSION

In the foregoing section we show that Wiener filtering can be applied satisfactorily to both array and single seismograph recordings to estimate broad band signals in noise and that the estimated signals have significant features that are not shown by the SP signals. Further, attempting to extract BB signals has no disadvantage in that SP seismograms can also be obtained from the same basic recordings.To have both BB and SP seismograms can indeed be instructive as comparison of the two types of seismogram may, for example, show up variations with time in the frequency content of the signal (see reference (23) for further examples of the advantage of having both SP and BB seismograms).

For signals such as those shown in figures 20, 21 and 24, which contain significant energy at frequencies less than say 0.5 Hz, the advantage of achieving noise reduction by spatial filtering of the oceanic microseisms is obvious. For high frequency signals (frequencies around 1 Hz) such as those shown in figure 23 and 25 the signal can be seen on the output from a single seismometer, riding on the 6 to 8 s period microseisms and, as the differences in the predominant frequency of the signal and noise is so great, little would seem to be lost in using only frequency filtering to suppress the noise; a single channel DW filter designed using a high frequency model signal extracts such signals satisfactorily from the DS output. Such processing, however, presupposes that any low frequency energy in the signal is insignificant and if, for example, the signals were known to be from an explosion such an assumption would be justified. However, for signals which are not known to be definitely explosions and for all earthquakes it will usually be safer to assume that there is low frequency energy present even when the visible parts of the signal appear to be predominantly high frequency and to achieve noise reduction over as wide a band of frequencies as possible by spatial filtering.

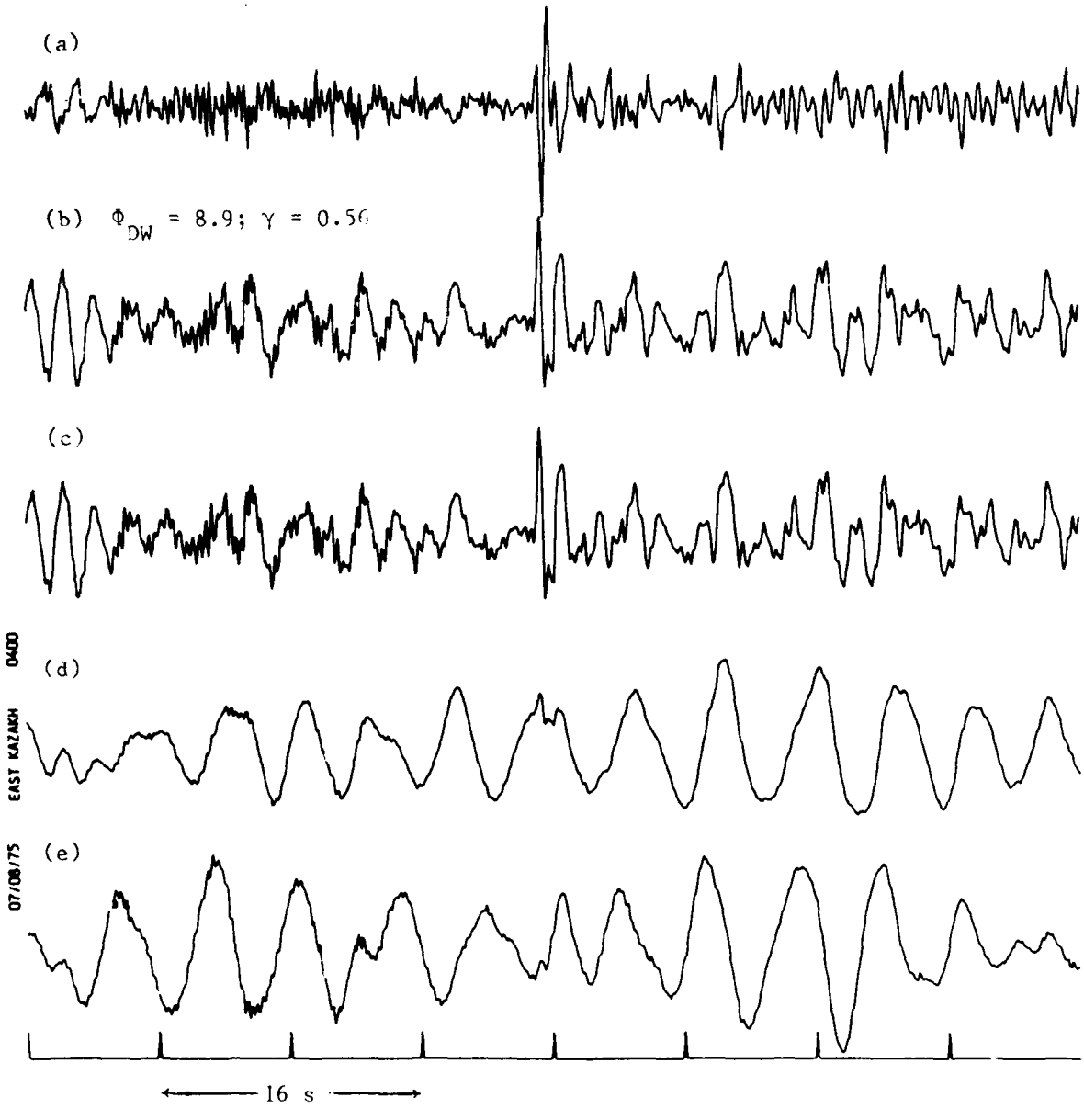


FIGURE 25. DATA-DEPENDENT WIENER (DW) FILTERING OF AN EXPLOSION IN EAST KAZAKHSTAN: 7 AUGUST 1975
(a) WWSS SP SEISMOGRAM
(b) DW FILTERED OUTPUT
(c) DELAY AND SUM OUTPUT FILTERED WITH THE FREQUENCY COMPONENT OF THE DW FILTERS
(d) DELAY AND SUM OUTPUT
(e) OUTPUT FROM A SINGLE CHANNEL OF THE ARRAY

For the four element BNA, noise reductions can be obtained by spatial filtering that are significantly greater than 2, the value expected from random noise; most of the noise reduction arises from wave number filtering of the oceanic microseisms of 6 to 8 s period. It is possible that similar noise reductions could have been obtained by simple DS processing with an array specifically designed to have nulls in the array response at the wave number of the oceanic microseisms. This could have been done using the method of Henger (8) of varying the design of the array until the noise on the DS output is a minimum. For such an array design, Wiener filtering coincides with DS processing. This method of suppressing oceanic microseisms has two main advantages: the method is easy to apply and the signal is the average of all the channels and so should be a reliable estimate of the signal. The disadvantage of the method is that it requires that the noise properties be stable over time. If the noise properties vary as they always do, then an array design that gives an optimum noise reduction at one time will not usually be optimum at other times. The optimum array for suppressing oceanic microseisms would then seem to be one with an aperture at least equal to the wavelength of the microseisms so that there are nulls in the vicinity of the wave number of the predominant noise. Wiener filtering would then be applied to the recording from such an array; this process can be thought of as trimming the array response to take account of the particular noise properties and such solutions will, in general, be close to the DS solutions and the noise reductions will arise from wave number filtering.

For the BNA it is obvious that the ability of the array to suppress oceanic microseisms independently of azimuth would be improved if the aperture of the array on an east-west axis were increased to about 20 km from the present aperture of 9 km. Such a modified array would be able to suppress 6 to 8 s period microseisms from most azimuths. As these oceanic microseisms seem to be highly coherent across the array, then further improvements in the ability of the BNA to suppress such microseisms could be obtained by adding further seismometers with the objective of improving the ability to reject 6 to 8 s period oceanic microseisms with speeds of around 3 km/s. In addition to considering the 6 to 8 s period microseisms, it is also necessary to find ways of suppressing noise at around 2 s; such noise, as can be seen from figure 3(a), is always present and during the summer such noise may predominate. If the 2 s noise is surface waves

propagating at speeds of around 3 km/s, then it has wavelengths of around 6 km and it is clear that the spacing of most of the seismometers in the BNA is too large to resolve such wavelengths. There is no evidence from the studies we have made so far that the 2 s noise is coherent across the array but this apparent lack of coherence may arise simply because the 2 s noise is arriving from a wide range of azimuths and the seismometers are spaced at a wavelength or more. Before any conclusions can be drawn about the properties of the 2 s noise it is thus necessary to add further seismometers to the array to reduce the spacing between seismometers to about 3 km.

To extrapolate from the results presented here for a four element array to what an array at the same site but with more elements would achieve cannot be done with any certainty. The best results for spatial filtering obtained here with data from the BNA is when the noise is large amplitude storm microseisms; this is because the noise is effectively from a single source and DW filters can be found that can reduce the array response at the wave number of these microseisms to reject noise from this one source and still pass the signal at zero wave number undistorted. The reason that the BNA is not as effective in reducing noise in the oceanic microseism band during quiet times is probably because, although the noise has low surface speed, it is arriving at the array from many azimuths and with only 4 seismometers it is impossible to design filters that will reduce noise over a wide range of azimuths simultaneously. Increasing the number of seismometers in the array might therefore improve the ability of the array to reduce low amplitude oceanic microseisms and give noise reductions of better than $n^{\frac{1}{2}}$ on quiet days.

Suppose the number of elements were increased to 16 and the signal coherence over these 16 elements was not significantly less than over the array of 4 elements, then it should be possible to achieve at the very least a further noise reduction of 2 by spatial filtering only, so giving an rms amplitude at the output of the DW filters of 200 nm or less (from table 4) at all times of the year; thus, the DW output would have an rms amplitude for broad band noise that is never greater than the rms amplitude on a single seismometer of the array during the quietest times (which usually occur in summer).

To relate this noise level to a magnitude threshold is impossible without further study because BB signal amplitudes tend to be larger than SP amplitudes and body wave magnitudes (m_b) computed from BNA SP amplitudes tend to be greater than the average m_b published by NEIS anyway. However, it would be surprising, judging from the processing so far carried out, if all signals from sources with NEIS m_b of, say, 5.5 or greater (and from some with lower m_b) could not be extracted from BB recordings by spatial filtering alone (performance parameter $\gamma \approx 1$). For many of the other earthquakes that might be detected on the SP seismograms, but cannot be satisfactorily extracted from BB seismograms by spatial filtering alone, it will be necessary to apply frequency filtering. It is possible that most of these low magnitude signals have source dimensions that are so small that the pulses radiated by such sources are only around a second or less in duration. If this is correct, then most of the energy in these pulses will be at frequencies higher than those of the oceanic microseisms and so the fact that frequency filtering has to be applied to extract the signal from the noise at these low magnitudes may not then be important.

The BNA lies close to an ocean and is consequently very noisy and so it is probable that other more ideal sites can be found to establish a BB array. Choosing a site for an array is influenced by many factors some of which are considered below.

The results obtained from the application of Wiener filtering to data from small aperture (3 to 4 km) SP arrays (4,5) shows that SP noise in general consists of two components: a high speed component (called teleseismic noise or mantle P wave noise) arriving as body waves from distant sources. Superimposed on this is low speed noise with sources close to the array. At sites where the organised low speed noise is large, then Wiener filtering gives signal-to-noise improvements better than $n^{\frac{1}{2}}$ over some frequency intervals. At sites where the noise is high speed teleseismic noise, Wiener filtering gives less than $n^{\frac{1}{2}}$ improvement in signal-to-noise ratio with these small aperture arrays.

The model of high and low speed noise appears to apply in the oceanic microseism band; both high and low speed components have been identified (see, for example, references (24) and (25)) and at low noise mid-continental sites the bulk of the noise seems to be high speed P waves (26). An array with an aperture similar to the BNA can only significantly reduce the low speed noise so that the best that can be hoped for with such an array is to reduce the oceanic microseisms to the level of the amplitude of the mantle P wave noise.

Backus (5) considers the design of an SP array to suppress both high and low speed noise and suggests using small aperture arrays (arrays of 3 to 4 km aperture; roughly 1 wavelength of low speed 1 Hz noise) as sub-arrays of a large aperture array, the spacing of the sub-arrays to be about half the wavelength of the high speed 1 Hz noise. Processing each sub-array should reduce the organised low speed noise and summing the processed outputs of the sub-array the teleseismic noise. Two arrays with this general design have been built and operated: the LASA (Montana) and the NORSAR (Norway). The results of spatial filtering (DMP rather than DW filtering was used) mainly carried out on LASA data have been disappointing for the maximum noise reduction obtained for these SP data was little better than $n^{\frac{1}{2}}$ compared to a single channel. In addition, there was some loss of signal amplitude on processing because many SP signals were not coherent across the array. However, this work focussed on extracting from the noise the very weak SP signals from low magnitude sources and it may be that, had the array been used to extract broad band signals from noise for larger magnitude sources, more satisfactory results would have been obtained.

At sites where the noise in the oceanic microseism band consists of a mixture of high and low speed noise, then to reduce both components of noise would seem to require an array design similar to that suggested by Backus (5) of sub-arrays within a larger array but now the sub-arrays should have apertures of around 20 km (~ a wavelength of low speed 6 s microseisms) and be spaced at ~ 50 km intervals.

Most sites for SP arrays have been selected because the noise level in the SP band is low. However, Phinney (27) has pointed out that the quietest sites are not necessarily the best sites for installing arrays. Possible disadvantages of very quiet sites for arrays are:-

(a) The quietest sites for SP noise tend to be in orogenic areas where the signals recorded from simple explosion sources may be complex and where from the LASA experience the signals are incoherent.

(b) At many sites where the SP noise is below average the signal amplitude at 1 Hz is also below average.

(c) At low noise sites the noise consists of high speed noise so that a large array is required to separate noise and signal but as the signal may not be coherent at such sites expected S/N improvements will not usually be obtained.

In selecting a site for a broad band array the level of the SP noise is not that important because the predominant noise will always be the oceanic microseisms with periods of 6 to 8 s. It is essential, however, to choose a site where the signal is coherent over the aperture of the array at all frequencies of interest and sites on shields would seem to be the ideal sites. The work of Kulhanek (28) suggests that there are sites at least on the Baltic Shield where the signal is coherent over distances of 100 km. Note, however, that broad band signals will usually have a predominant frequency less than 1 Hz and these lower frequencies are likely to be more coherent than those recorded on standard short period seismographs. Shield sites also have the advantage that the signal amplitude at least around 1 Hz tends to be greater than in orogenic regions. It is possible to find sites on shields where at least for part of the year the SP noise is very low; at Yellowknife, Canada (YKA) for instance the SP noise amplitude is around 1 nm during the winter, although during summer the noise amplitude may be ten times this. The summer noise is uncorrelated, however, and is reduced by $n^{\frac{1}{2}}$ by DS processing.

At sites in the middle of shields the main noise in the oceanic microseism band is probably high speed noise and so an array at such a site would have to have a large aperture to obtain significant reductions in the amplitude of this noise. If sites in the centre of shields are available, they will usually be preferable to coastal sites even if a large array is not installed because the amplitude of the low speed component of the oceanic microseisms will be of

much lower amplitude than at coastal stations. If the only sites available are coastal sites (which is true of all sites in the UK), then any convenient site where the signal is coherent would probably suffice. In addition, it is preferable to choose a site where the SP signal amplitude is above rather than below average if such sites can be found. For such sites, however, significant noise reduction should be obtained by spatial filtering using arrays of the dimensions of the BNA (10 to 20 km) particularly during periods of large microseisms. The value of such arrays is that they allow use to be made of signals from large magnitude earthquakes and explosions which would otherwise be completely obscured by noise. During periods when the amplitude of the low speed noise is low such arrays can be expected to produce at best $n^{\frac{1}{2}}$ improvement in S/N. On quiet sites it is possible that low speed noise is absent and if this is so, then arrays of 10 to 20 km aperture will usually have little value for noise reduction in the oceanic microseism band.

Given a recording system that allows broad band recordings to be recovered on playback from tape, we propose the following general scheme for obtaining the best estimates of signal shape. For sites where low speed surface waves are the predominant form of noise in the oceanic microseism band, then arrays with apertures of 20 km or so can be used to suppress this noise by spatial filtering. At low noise sites, where the noise in the oceanic microseism band is likely to be mantle P wave noise, then the only way to suppress such noise by spatial filtering is to have a large array (aperture 100 km at least) and such an array will only be worth installing if sites can be found where the signal is coherent over such a large aperture. Smaller arrays on low noise sites are unlikely to have the resolution to allow high speed oceanic microseisms to be suppressed by spatial filtering and the most efficient method of processing would then seem to be simply DW filtering of the DS output. The DS processing will reduce any random noise and DW processing will apply the required frequency filtering to extract the best estimate of signal shape. This scheme is similar to that currently applied to much SP array data where band pass filtering of the DS output is used to remove frequencies which by eye are seen to be different from the predominant frequencies in the signal. Such filtering, however, is only applied to signals with low signal-to-noise ratio on the SP system. The advantage of starting with broad band records and applying DW filtering is that just sufficient filtering is applied to produce the minimum distortion of the signal. At high signal-to-noise ratio then ideally the whole signal spectrum within the wide

recording band of the instruments will be passed by the DW filter; at low signal-to-noise ratio the output will tend to that seen on a conventional narrow band seismograph. The detection threshold should then, theoretically at least, be the same for both SP and Wiener filtered broad band recordings. Note, however, that the broad band seismograph has a phase response which in the pass band gives a much smaller phase shift than standard SP systems (figure 2). If a two-sided DW filter is used to extract signals from broad band recordings, the estimated signal will be as recorded by a near phaseless seismograph; the same recording on an SP seismograph would be distorted by the large phase shifts introduced by the seismograph.

To record broad band ground displacement directly is inefficient because most of the dynamic range of the system is taken up in recording the oceanic microseisms (1). The ideal recording system for an array station would appear to be one in which the response is the inverse of the spectrum of the incoherent components of seismic noise; the incoherent seismic noise as recorded by such a system would then be white. (Berckhemer (1) discusses the design of a seismograph which has a response which is the inverse of the noise spectrum.) The amplitude of the system noise must then be significantly smaller than the amplitude of the smallest signal that can be extracted from the incoherent noise by DS processing. At stations that are not arrays the ideal recording system should have a response that is the inverse of the noise spectrum for the quietest time and the system noise should be such that the amplitude of the seismic noise as recorded is just larger than the system noise. With such a recording system all signals that can possibly be extracted by Wiener processing from the noise will be recorded with sufficient signal to system noise ratio to allow processing to be carried out satisfactorily.

Now the response of an SP seismograph from around 1 Hz down to the frequency of the oceanic microseism peak is roughly the inverse of the seismic noise spectrum (which is not surprising as the attraction of such a response for visual recording is that it flattens the noise spectrum). So it should be possible to pass narrow band SP signals through a filter to compensate for the effects of the recording system and obtain broad band signals down to the frequency of the microseism peak. There is a limit to the band width that can be recovered in this way because at some frequency the signal level on the SP system falls below the level of the instrumental noise. However, Douglas et al. (29) show that

seismograms that display ground displacement at constant magnification in the range 0.3 to 10 Hz can be derived from SP seismograms using spike filtering. Good estimates of broad band signals can also be derived from SP seismograms by just reversing the process described in section 2 for obtaining SP seismograms for BB; the spectrum of the SP seismogram is simply multiplied by $a_2(\omega)/a_1(\omega)$ and transformed back into time. Examples of BB seismograms derived from SP are shown by Douglas et al. (23).

The SP data used by Douglas et al. (23) come from a system which was not specifically designed to be used for deriving broad band signals and the recordings were made on analogue tape recorders for which the dynamic range is less and the system noise greater than modern digital recorders. Key (30) has shown that with modern digital recording systems it is possible to recover broad band seismograms from the short period with little interference from system noise, at least out to the period of the oceanic microseisms, so it appears that in future at stations where digital systems are installed there will be no need to make special provision for broad band recording, simply recording narrow band SP signals should be sufficient.

Douglas et al. (23) derive the broad band seismogram from the narrow band in two steps: the first step is to convert from the response as recorded to the desired broad band displacement response, the second step is to estimate and apply Wiener filters to extract the signals from noise. An alternative method of processing is to combine the two steps so that filters are computed that give the best estimate of the broad band displacement signal given the data as recorded, the response of the recording system and the system noise level; Franklin (31) has extended the Wiener filter theory to cover this case but this theory does not appear to have been applied yet to the extraction of seismic signals from noise.

8. CONCLUSIONS

The main conclusions of this study are as follows:-

(a) The most flexible processing method appears to be Wiener filtering. In the general (multichannel) case the estimated filters apply both spatial and frequency filtering to extract the signal, but if

the required noise reduction can be obtained by spatial filtering alone, then frequency filtering is not applied. This is a desirable property of multichannel filters because ideally spatial filtering passes the signal undistorted. From the data studied in this report it is possible to get noise reductions due to spatial filtering of up to 6 with a 4 element array.

(b) There seems to be no advantage in using minimum power (MP) filtering as opposed to Wiener filtering. In the MP method multichannel filters are designed to minimise the noise power at the output subject to the constraint that the desired signal is passed undistorted; the noise reduction can be thought of as arising purely from spatial filtering. If spatial filtering alone is sufficient to allow the signal to be extracted from the noise, then the Wiener and MP methods give the same results (which is to be expected, as is shown by theoretical considerations). On the other hand, if the signal can only be extracted from the noise by frequency filtering, the MP method fails, whereas the Wiener method (ideally) always shows signal above noise provided the signal amplitude is greater than the noise amplitude in some frequency band. The detection threshold for Wiener filtering (in the ideal case) should never be worse (and could be better) than for narrow band SP recordings.

(c) To construct Wiener filters the auto- and cross-correlation functions of the signal and noise are required. The most satisfactory way of constructing the noise correlations appears to be to use a section of observed noise ahead of the signal to which it is assumed a small proportion of white noise (uncorrelated between channels) has been added. If the white noise is not added, the effect of the filters is sometimes to distort the estimated signal shape as compared to the shape as seen on the delay and sum (or single channel). This distortion arises because the signal is not perfectly coherent across the array.

To construct the signal correlations the power spectrum of the signal must be roughly known. Power spectra based on simple model signals seem to be adequate for this purpose. Wher the noise reduction arises from spatial filtering the assumed form of the power spectrum has no effect on the estimated signal.

(d) The most widely used method of processing SP array data is delay and sum combined at low magnitude with band pass filtering to cut out those frequencies where the noise amplitudes are large relative to the signal amplitude. This process can be thought of as applying a crude one channel Wiener filter to the delay and sum output, the filter being designed on a general knowledge of signal and noise properties. Usually a Wiener filter estimated from the observed noise properties and a signal model will give a better estimate of signal shape than routine application of a fixed band pass filter.

(e) If oceanic microseisms have well defined wave numbers, then it is possible to suppress them by applying delay and sum processing to data from an array which has a null in its wave number response at the wave number of the noise. In this case Wiener filtering, MP filtering and delay and sum processing (and common sense) all coincide. Such arrays require apertures of the order of a wavelength of the noise or greater. It may be possible to suppress noise using smaller arrays by exploiting small differences in amplitude between the noise on different channels but filters estimated from such arrays are likely to be unstable and, unless the signal is highly coherent, could result in distortion and suppression of the signal.

It will usually not be possible to design an array so that the array response always has a null at the required point to suppress the noise. An array should thus be designed to have nulls in the vicinity of the wave number of the principal noise sources. Wiener filtering can then be used to trim the array response and obtain the optimum noise suppression for the particular noise sample. Such Wiener filters will, it is hoped, usually not depart markedly from the delay and sum solution and so the estimated signal should be close to the average of the signal over all channels.

(f) There is some evidence that, as with SP noise, oceanic microseisms consist of two components: a low speed component (~ 3 km/s) propagating horizontally as surface waves and a high speed component (> 8.0 km/s) which is propagating as body waves and is travelling steeply upwards from the mantle. The very large amplitude oceanic microseisms seem to be mainly the low speed component; arrays of 10 to 20 km aperture are required to suppress them. The high speed component usually has low amplitude and is then only seen when the low speed component is small or absent; to suppress high speed oceanic microseisms requires an array of around 100 km aperture at least.

At sites where both components of the oceanic microseisms are to be suppressed an array design such as that used at the Large Aperture Seismometer Array in Montana with sub-arrays within a larger array could be used, the sub-arrays being 10 to 20 km aperture and the total array aperture 100 km or more.

REFERENCES

1. H Berckhemer: "The Concept of Wide Band Seismometry". European Seismological Commission, Luxemburg, 1-7 (1970)

2. P D Marshall, R F Burch and A Douglas: "How and Why to Record Broad Band Seismic Signals". Nature, 239, 154-155 (1972)

3. J P Burg: "Three-Dimensional Filtering with an Array of Seismometers". Geophysics, 29, 693-713 (1964)

4. M M Backus, J P Burg, D Baldwin and E Bryan: "Wide Band Extraction of Mantle P Waves from Ambient Noise". Geophysics, 29, 672-692 (1964)

5. M M Backus: "Teleseismic Signal Extraction". Proc Roy Soc A, 290, 343-367 (1966)

6. J Capon, R J Greenfield and R J Kolker: "Multidimensional Maximum-Likelihood Processing of a Large Aperture Seismic Array". Proc IEEE, 55, 192-211 (1967)

7. R G Baker: "Suggested Use of a Medium Period Seismograph Array". Bull Seis Soc Am, 60, 1735-1737 (1970)

8. M Henger: "Wave Number Filtering of Microseisms with Triangular Arrays", in Exploitation of Seismograph Networks. Ed K G Beauchamp. NATO Advanced Study Institute Series E: Applied Sciences - No. 11, Noordhoff-Leiden, 205-218 (1975)

9. H-P Harjes: "Wide-Band Seismic Recording at the Graefenberg (GRF) Station", in Exploitation of Seismograph Networks. Ed K G Beauchamp. NATO Advanced Study Institute Series E: Applied Sciences - No. 11, Noordhoff-Leiden, 197-204 (1975)

10. H-P Harjes and D Seidl: "Digital Recording and Analysis of Broad-Band Seismic Data at the Graefenberg (GRF) Array". J Geophys, 44, 511-523 (1978)

11. P W Burton: "An Array of Broad Band Seismographs: Some Initial Results." XIVth General Assembly of the ESC, National Komitee fur Geodasie and Geophysik bei der Akademie der Wissenschaften der Deutschen Democratishen Republik, Berlin, 145-155 (1975)

12. D T Hopkins: Private Communication

13. F A Key: "Some Observations and Analyses of Signal Generated Noise". Geophys J R astr Soc, 15, 377-392 (1968)

14. A Douglas, P D Marshall, P G Gibbs, J B Young and C Blamey: "P Signal Complexity Re-examined". Geophys J R astr Soc, 33, 195-221 (1972)

15. J W Birtill and F E Whiteway: "The Application of Phased Arrays to the Analysis of Seismic Body Waves". Phil Trans Roy Soc A, 421-493 (1965)

16. A J Berkhout and P R Zaanen: "A Comparison between Wiener Filtering and Kalman Filtering and Deterministic Least Squares Estimation". Geophys Prosp, 24, 141-197 (1976)

17. J Capon, R J Greenfield, R J Kolker and R T Lacoss: "Short-Period Signal Processing Results for the Large Aperture Siesmic Array". Geophysics, 33, 452-472 (1968)

18. E J Kelly: "A Comparison of Seismic Array Processing Schemes". Technical Note 1965-21, Lincoln Laboratory, Lexington, Mass. (1965)

19. J A Hudson: "A Quantitative Evaluation of Seismic Signals at Teleseismic Distances. I: Radiation from Point Sources". Geophys J R astr Soc, 18, 233-249 (1969)

20. J A Hudson: "A Quantitative Evaluation of Seismic Signals at Teleseismic Distances. II: Body Waves and Surface Waves from an Extended Source". Geophys J R astr Soc, 18, 353-370 (1969)

21. A Douglas, J A Hudson and C Blamey: "A Quantitative Evaluation of Seismic Signals at Teleseismic Distances. III: Computed P and Rayleigh Wave Seismograms". Geophys J R astr Soc, 28, 385-410 (1972)

22. E W Carpenter: "Absorption of Elastic Waves - an Operator for a Constant Q Mechanism". AWRE Report O43/66, HMSO, London (1966)

23. A Douglas, J A Hudson and B J Barley: "The Complexity of Short Period P Seismograms: What Does Scattering Contribute?" AWRE Report O3/81, HMSO, London (1981)

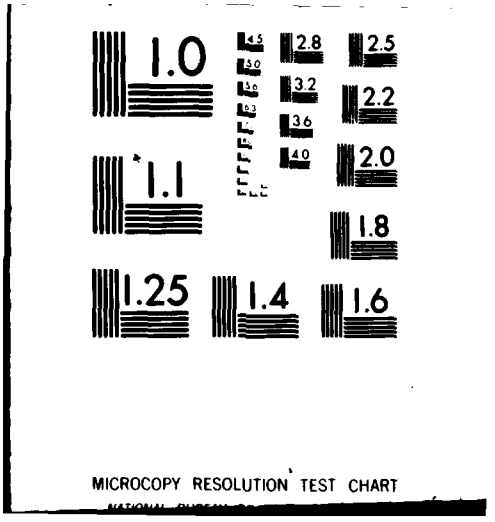
24. M N Toksoz and R T Lacoss: "Microseisms: Mode Structure and Sources". Science, 159, 872-873 (1968)

AD-A116 366 ATOMIC WEAPONS RESEARCH ESTABLISHMENT ALDERMASTON (EN--ETC F/G 8/11
THE ESTIMATION OF SEISMIC BODY WAVE SIGNALS IN THE PRESENCE OF --ETC(U)
DEC 81 A DOUGLAS, J B YOUNG

UNCLASSIFIED AWRE-O-14/81 DRIC-BR-81928 NL

2 of 2

END
DATE FILMED
8-82
DTIC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
MICROCOPY RESOLUTION TEST CHART

25. E Hjortenberg: "Direction of Approach of P- and Rayleigh Waves in the Greenland Microseisms". Symposium sur les Microseisms IUGG Monograph 31, Paris (1970)

26. L P Vinnik: "Sources of Microseismic P Waves". PAGEOPH, 103, 282-289 (1973)

27. R A Phinney: "Discussion". Proc Roy Soc A, 290, 467-471 (1966)

28. O Kulhanek: "Signal and Noise Coherence Determination for the Uppsala Seismograph Array Station". PAGEOPH, 109, 1653-1671 (1973)

29. A Douglas, D J Corbishley, C Blamey and P D Marshall: "Estimating the Firing Depth of Underground Explosions". Nature, 237, 26-28 (1972)

30. F A Key: Private Communication

31. J N Franklin: "Well-Posed Stochastic Extensions of Ill-Posed Linear Problems". Journal of Mathematics Analysis and Applications, 31, 682-716 (1970)

APPENDIX A

COMPARISON OF MINIMUM POWER AND UNBIASED PREDICTION ERROR FILTERING USING A TWO CHANNEL ARRAY

Let

$$\underline{U}_1^P = \text{col}\left[u_1^P(-1),\ u_1^P(0),\ u_1^P(1)\right]$$

and

$$\underline{U}_2^P = \text{col}\left[u_2^P(-1),\ u_2^P(0),\ u_2^P(1)\right]$$

be filters that predict the noise on the delay and sum output but do not pass the signal of interest. If the filters are to suppress the signals, we must have

$$u_1^P(k) + u_2^P(k) = 0 \text{ for all } k.$$

If $\underline{b}^P$ is the predicted output given by such filters, we can write

$$\underline{X}_1\underline{U}_1^P + \underline{X}_2\underline{U}_2^P = \underline{b}^P.$$

If $\underline{b}$ is the delay and sum output, then the best least-squares estimates of $\underline{U}_1^P$ and $\underline{U}_2^P$ can be found by minimising the sum of the squares of the differences between $\underline{b}$ and $\underline{b}^P$; these estimates are given by the solution of

$$\begin{bmatrix} \underline{X}_1^T\underline{X}_1, & \underline{X}_1^T\underline{X}_2, & \underline{Q}_1 \\ \underline{X}_2^T\underline{X}_2, & \underline{X}_2^T\underline{X}_2, & \underline{Q}_1 \\ \underline{Q}_1^T, & \underline{Q}_1^T, & \underline{0} \end{bmatrix} \begin{bmatrix} \underline{U}_1^P \\ \underline{U}_2^P \\ \underline{\lambda}^P \end{bmatrix} = \begin{bmatrix} \underline{X}_1^T\underline{b} \\ \underline{X}_2^T\underline{b} \\ \underline{0} \end{bmatrix}, \qquad \ldots.(A1)$$

where

$$\underline{Q}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } \underline{\lambda}^P$$

is a three element column vector of Lagrangian multipliers. If $\underline{d}$ = col (0, 0.5, 0), then we can write

$$X_1 d + X_2 d = \underline{b}.$$

The unbiased prediction error output is $\underline{b} - \underline{b}^P$, so

$$\underline{b} - \underline{b}^P = \underline{X}_1 (\underline{d} - \underline{U}_1^P) + \underline{X}_2 (\underline{d} - \underline{U}_2^P).$$

Suppose now we wish to choose $\underline{U}_1^P$ and $\underline{U}_2^P$ in such a way that minimum power and unbiased prediction error are equal, then we must have $\underline{U}_1 = \underline{d} - \underline{U}_1^P$ and $\underline{U}_2 = \underline{d} - \underline{U}_2^P$. Substituting in equation (10) for $\underline{U}_1$ and $\underline{U}_2$ we have:-

$$\begin{bmatrix} \underline{X}_1^T \underline{X}_1, & \underline{X}_1^T \underline{X}_2, & \underline{Q}_1 \\ \underline{X}_2^T \underline{X}_1, & \underline{X}_2^T \underline{X}_2, & \underline{Q}_1 \\ \underline{Q}_1^T, & \underline{Q}_1^T, & \underline{0} \end{bmatrix} \begin{bmatrix} \underline{d} - \underline{U}_1^P \\ \underline{d} - \underline{U}_2^P \\ \underline{\lambda} \end{bmatrix} = \begin{bmatrix} \underline{0} \\ \underline{0} \\ \underline{V}_2 \end{bmatrix},$$

or

$$\begin{bmatrix} (\underline{X}_1^T \underline{X}_1 + \underline{X}_1^T \underline{X}_2)\underline{d} \\ (\underline{X}_2^T \underline{X}_1 + \underline{X}_2^T \underline{X}_2)\underline{d} \\ (\underline{Q}_1^T + \underline{Q}_1^T)\underline{d} \end{bmatrix} - \begin{bmatrix} \underline{X}_1^T \underline{X}_1, & \underline{X}_1^T \underline{X}_2, & \underline{Q}_1 \\ \underline{X}_2^T \underline{X}_1, & \underline{X}_2^T \underline{X}_2, & \underline{Q}_1 \\ \underline{Q}_1^T, & \underline{Q}_1^T, & \underline{0} \end{bmatrix} \begin{bmatrix} \underline{U}_1^P \\ \underline{U}_2^P \\ -\underline{\lambda} \end{bmatrix} = \begin{bmatrix} \underline{0} \\ \underline{0} \\ \underline{V}_2 \end{bmatrix},$$

or

$$\begin{bmatrix} \underline{X}_1^T \underline{b} \\ \underline{X}_2^T \underline{b} \\ \underline{V}_2 \end{bmatrix} - \begin{bmatrix} \underline{X}_1^T \underline{X}_1, & \underline{X}_1^T \underline{X}_2, \underline{Q}_1 \\ \underline{X}_2^T \underline{X}_1, & \underline{X}_2^T \underline{X}_2, \underline{Q}_1 \\ \underline{Q}_1^T, & \underline{Q}_1^T, \underline{0} \end{bmatrix} \begin{bmatrix} \underline{U}_1^P \\ \underline{U}_2^P \\ -\underline{\lambda} \end{bmatrix} = \begin{bmatrix} \underline{0} \\ \underline{0} \\ \underline{V}_2 \end{bmatrix},$$

or

$$\begin{bmatrix} \underline{X}_1^T\underline{X}_1, & \underline{X}_1^T\underline{X}_2, & \underline{Q}_1 \\ \underline{X}_2^T\underline{X}_1, & \underline{X}_2^T\underline{X}_2, & \underline{Q}_1 \\ \underline{Q}_1^T, & \underline{Q}_1^T, & \underline{0} \end{bmatrix} \begin{bmatrix} \underline{U}_1^P \\ \underline{U}_2^P \\ -\underline{\lambda} \end{bmatrix} = \begin{bmatrix} \underline{X}_1^T\underline{b} \\ \underline{X}_2^T\underline{b} \\ \underline{0} \end{bmatrix} . \qquad \ldots.(A2)$$

However, equation (A2) is identical to equation (A1) putting $\underline{\lambda}^P = -\underline{\lambda}$. Thus, the filtered output obtained by applying minimum power filters estimated using equation (10) is identical to that obtained by applying the prediction filters estimated using equation (A1) and subtracting this predicted output from the delay and sum output.

APPENDIX B

THE SIGNIFICANCE OF THE LAGRANGIAN MULTIPLIERS IN MINIMUM POWER FILTER ESTIMATION

Writing the Lagrangian multipliers obtained estimating $\underline{U}(1|3)$, $\underline{U}(2|3)$ and $\underline{U}(3|3)$ (equation 10c) as:-

$$\underline{\lambda}(1|3) = \begin{bmatrix} \lambda(1|0) \\ \lambda(1|1) \\ \lambda(1|2) \end{bmatrix}, \quad \underline{\lambda}(2|3) = \begin{bmatrix} \lambda(2|-1) \\ \lambda(2|0) \\ \lambda(2|1) \end{bmatrix} \text{ and } \underline{\lambda}(3|3) = \begin{bmatrix} \lambda(3|-2) \\ \lambda(3|-1) \\ \lambda(3|0) \end{bmatrix},$$

respectively then it is easy to show that

$$\left(\underline{\lambda}(1|3),\ \underline{\lambda}(2|3),\ \underline{\lambda}(3|3)\right) = -\left(\underline{Q}^T\underline{R}^{-1}\underline{Q}\right)^{-1}.$$

Now the Lagrangian multiplier $\lambda(i|o)$ gives the expected mean square noise at the output after applying the MP filters $\underline{U}(i|3)$(18). However, $(\underline{Q}^T\underline{R}^{-1}\underline{Q})^{-1}$ is a symmetric Toeplitz matrix so that $\lambda(i|o)$ for i = 1, 2, 3 is constant so that the noise reduction due to MP filtering is the same for all three sets of filters $\underline{U}(i|3)$ i = 1, 2, 3.

Writing (from expression (14)) the noise at the output after applying the MP filters $\underline{U}(i|3)$, i = 1, 2, 3 as $x^b(i|3,t)$, i = 1, 2, 3 respectively, then it can be shown that the expected value of the covariance between $x^b(i|3,t)$ and $x^b(j|3,t)$ is $\lambda(k|i-j)$. This result can be generalised to p filter points and n channels.

If the expected mean square noise after applying MP filters is σ_s^2, then given m data points $\hat{\sigma}_s^2$ the best estimate of σ_s^2 is $\{\lambda(i|o)m\}/q$ where q is the number of degrees of freedom.

DOCUMENT CONTROL SHEET

Overall security classification of sheet ..... UNCLASSIFIED .....

(As far as possible this sheet should contain only unclassified information. If it is necessary to enter classified information, the box concerned must be marked to indicate the classification eg (R), (C) or (S)).

1. DRIC Reference (if known)	2. Originator's Reference	3. Agency Reference	4. Report Security Classification
-	AWRE ~~Report~~ No. 014/81	-	UNLIMITED

5. Originator's Code (if known)	6. Originator (Corporate Author) Name and Location
-	Atomic Weapons Research Establishment, Aldermaston, Berkshire

5a. Sponsoring Agency's Code (if known)	6a. Sponsoring Agency (Contract Authority) Name and Location
-	-

7. Title

The Estimation of Seismic Body Wave Signals in the Presence of Oceanic Microseisms

7a. Title in Foreign Language (in the case of Translation)

-

7b. Presented at (for Conference Papers). Title, Place and Date of Conference

-

8. Author 1.Surname, Initials	9a. Author 2	9b. Authors 3, 4	10. Date	pp	ref
Douglas A	Young J B	-	December 1981	101	30

11. Contract Number	12. Period	13. Project	14. Other References
-	-	-	-

15. Distribution Statement

No Restriction

16. Descriptors (or Keywords) (TEST)

Seismic array
Seismology
Seismographs
Wiener filtering
Oceanic microseisms
Broad band seismographs

Abstract The report describes studies of the processing of broad band seismograms (from a system with displacement response flat from around 0.1 to 10 Hz) for the estimation of signal shape rather than the maximisation of signal-to-noise ratio. By definition Wiener filtering gives the best estimate of signal shape in the sense that filters are designed to minimise the mean square of the difference between the true signal and the estimated signal, consequently most of the report describes studies of the application of this type of filter. For seismometer array recordings Wiener filters apply both spatial and frequency filtering to extract the signal. However, if the required noise reduction can be obtained by spatial filtering alone, then no frequency filtering is applied and so the signal is passed undistorted. From the data studied in this report it is possible to get noise reductions due to spatial filtering of up to 6 with an array of 4 seismometers.

Some Metric and SI Unit Conversion Factors

(Based on DEF STAN 00-11/2 "Metric Units for Use by the Ministry of Defence", DS Met 5501 "AWRE Metric Guide" and other British Standards)

Quantity	Unit	Symbol	Conversion
Basic Units			
Length	metre	m	1 m = 3.2808 ft 1 ft = 0.3048 m
Mass	kilogram	kg	1 kg = 2.2046 lb 1 lb = 0.45359237 kg 1 ton = 1016.05 kg
Derived Units			
Force	newton	N = kg m/s^2	1 N = 0.2248 lbf 1 lbf = 4.44822 N
Work, Energy, Quantity of Heat	joule	J = N m	1 J = 0.737562 ft lbf 1 J = 9.47817 × 10^{-4} Btu 1 J = 2.38846 × 10^{-4} kcal 1 ft lbf = 1.35582 J 1 Btu = 1055.06 J 1 kcal = 4186.8 J
Power	watt	W = J/s	1 W = 0.238846 cal/s 1 cal/s = 4.1868 W
Electric Charge	coulomb	C = A s	-
Electric Potential	volt	V = W/A = J/C	-
Electrical Capacitance	farad	F = A s/V = C/V	-
Electric Resistance	ohm	Ω = V/A	-
Conductance	siemen	S = 1 Ω^{-1}	-
Magnetic Flux	weber	Wb = V s	-
Magnetic Flux Density	tesla	T = Wb/m^2	-
Inductance	henry	H = V s/A = Wb/A	-
Complex Derived Units			
Angular Velocity	radian per second	rad/s	1 rad/s = 0.159155 rev/s 1 rev/s = 6.28319 rad/s
Acceleration	metre per square second	m/s^2	1 m/s^2 = 3.28084 ft/s^2 1 ft/s^2 = 0.3048 m/s^2
Angular Acceleration	radian per square second	rad/s^2	-
Pressure	newton per square metre	N/m^2 = Pa	1 N/m^2 = 145.038 × 10^{-6} lbf/in^2 1 lbf/in^2 = 6.89476 × 10^3 N/m^2
	bar	bar = 10^5 N/m^2	- 1 in. Hg = 3386.39 N/m^2
Torque	newton metre	N m	1 N m = 0.737562 lbf ft 1 lbf ft = 1.35582 N m
Surface Tension	newton per metre	N/m	1 N/m = 0.0685 lbf/ft 1 lbf/ft = 14.5939 N/m
Dynamic Viscosity	newton second per square metre	N s/m^2	1 N s/m^2 = 0.0208854 lbf s/ft^2 1 lbf s/ft^2 = 47.8803 N s/m^2
Kinematic Viscosity	square metre per second	m^2/s	1 m^2/s = 10.7639 ft^2/s 1 ft^2/s = 0.0929 m^2/s
Thermal Conductivity	watt per metre kelvin	W/m K	-
Odd Units*			
Radioactivity	becquerel	Bq	1 Bq = 2.7027 × 10^{-11} Ci 1 Ci = 3.700 × 10^{10} Bq
Absorbed Dose	gray	Gy	1 Gy = 100 rad 1 rad = 0.01 Gy
Dose Equivalent	sievert	Sv	1 Sv = 100 rem 1 rem = 0.01 Sv
Exposure	coulomb per kilogram	C/kg	1 C/kg = 3876 R 1 R = 2.58 × 10^{-4} C/kg
Rate of Leak (Vacuum Systems)	millibar litre per second	mb l/s	1 mb = 0.750062 torr 1 torr = 1.33322 mb

*These terms are recognised terms within the metric system.


DAT
FILM