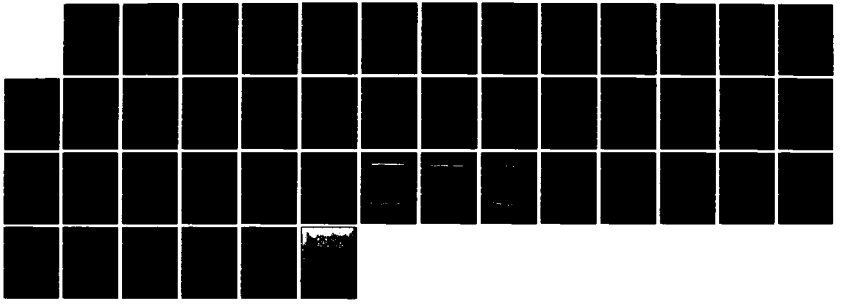
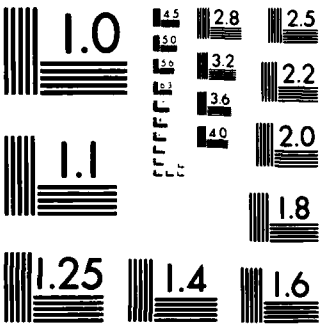
AD-A130 397 A CHARACTERIZATION OF ARCTIC UNDERSEA NOISE(U) 1/1
PRINCETON UNIV NJ DEPT OF STATISTICS J G VEITCH ET AL.
JUN 83 TR-12 N00014-81-K-0146

UNCLASSIFIED F/G 14/2 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

(12)

REPORT NUMBER 12

A CHARACTERIZATION OF ARCTIC UNDERSEA NOISE

J.G. VEITCH and A.R. WILKS

Department of Statistics
Princeton University
Princeton, New Jersey 08544

JUNE 1983

Prepared for

OFFICE OF NAVAL RESEARCH (Code 411SP)
Statistics and Probability Branch
Arlington, Virginia 22217
under Contract N00014-81-K-0146
SRO(103) Program in Non-Gaussian Signal Processing

S.C. Schwartz, Principal Investigator

Approved for public release; distribution unlimited

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 12	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): A CHARACTERIZATION OF ARCTIC UNDERSEA NOISE	5. TYPE OF REPORT & PERIOD COVERED: Technical Report Sept. 1982-May 1983
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): James G. Veitch and Allan R. Wilks	8. CONTRACT OR GRANT NUMBER(s): N00014-81-K-0146
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Department of Statistics Princeton University Princeton, New Jersey 08544	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR SRO-103
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research (Code 411 SP) Department of the Navy Arlington, Virginia 22217	12. REPORT DATE: June 1983 13. NUMBER OF PAGES: 42
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release; Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

arctic noise
noise analysis and models
arctic noise spectrum analysis

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Arctic undersea noise is investigated and partially characterized by a probability model suggested by exploration of a sample of such noise. The large size of the dataset makes this possible. The model is that of a mixed spectrum, with stationary Gaussian noise intermixed with random sinusoids, and occasional high intensity impulsive noise bursts, which are very short-lived. Due to the complex behavior of the random sinusoids, the noise cannot be considered as stationary in the time frame (10 minutes) investigated.

(con't.)

(Abstract con't.)

The validity of the model is checked by testing model predictions against the actual data. In particular, this model makes some predictions about the type of spectrum observed, and these predictions are tested by a non-linear filtering of the periodogram. A running median filter is used and is shown to asymptotically estimate the spectrum of the Gaussian noise. This filter seems to work fairly well on the data.

A Characterization of Arctic Undersea Noise

by

James G. Veitch and Allan R. Wilks

Department of Statistics

Princeton University

Princeton, New Jersey, 08544

Abstract

Arctic undersea noise is investigated and partially characterized by a probability model suggested by exploration of a sample of such noise. The large size of the dataset makes this possible. The model is that of a mixed spectrum, with stationary Gaussian noise intermixed with random sinusoids, and occasional high intensity impulsive noise bursts, which are very short-lived. Due to the complex behavior of the random sinusoids, the noise cannot be considered as stationary in the time frame (10 minutes) investigated. The validity of the model is checked by testing model predictions against the actual data. In particular, this model makes some predictions about the type of spectrum observed, and these predictions are tested by a non-linear filtering of the periodogram. A running median filter is used and is shown to asymptotically estimate the spectrum of the Gaussian noise. This filter seems to work fairly well on the data.

This research was supported by the Office of Naval Research under Contract N00014-81-K-0146.

1. Description of Data and Goals of Analysis

This report describes the analysis of a particular dataset. In this section we describe the dataset, and the goals of the analysis. In section 2 we look at some previous work done by others in analyzing this and other similar datasets. In the remaining sections we describe the steps in our data analysis, along with some theoretical justification of several novel techniques.

The dataset is a large univariate time series. It is a sample of underwater noise observed on 23-24 April 1980 from a pack ice camp in the Arctic Ocean, located at 86 degrees north latitude, 25 degrees west longtitude. A broadband, omnidirectional hydrophone was suspended to a depth of 91 meters in 4000 meter deep water, and the underwater acoustic noise was recorded on an analog device. Later, this noise was filtered by a lowpass filter with cutoff around 2500 Hz, and then digitized at 10 kHz, ie, the analog signal was sampled at the rate of 10,000 times per second, and each of these sampled values was rounded to the nearest of a finite set of numbers. Our dataset is this digitized record during the period approximately 11:30-11:40 pm on the 23rd. Ten minutes at 10 kHz yields six million numbers; we actually have 6,150,144 numbers which is slightly longer than ten minutes. This conveniently works out to 6006 groups of 1024 numbers each.

Such a large, single time series presents both problems and opportunities; this is perhaps a general comment relevant to the analysis of large datasets. The main problem is that the statistics we usually compute such as moments, autocovariances and Fourier transforms, are extremely inconvenient to compute for large amounts of data, and may even make little sense. Just handling the data presents logistic problems. On the other hand, we can do things which are impossible or impractical for smaller datasets. Many such techniques will be described in the following pages and perhaps what makes these techniques

attractive is the general idea that by blocking the data into contiguous chunks we can imagine that we have an ensemble of time series. This contrasts with the usual case in time series analysis in which we only have one time series from a *conceptual* ensemble of *possible* time series.

The particular dataset analyzed here appears to exhibit several different sorts of non-stationarity and non-Gaussianity. Our main goal in analyzing this so-called 'ice noise' is to try to characterize it. We wish to find a model which adequately describes the noise. This characterization could be used to enhance knowledge about underwater acoustics near ice, and might also allow more efficient detection of signals in the presence of such noise.

We have taken a data-oriented approach in our analysis. We begin by picturing the data in certain simple ways and then letting these pictures guide our further analysis. The end result is an approximate theoretical model, suggested to us purely by consideration of the data itself. It is to be expected that detailed knowledge of underwater acoustics under ice would lead to at least some enhancement of our model. This point will be taken up later in greater detail.

2. Previous Work

This dataset has been looked at several times before. In particular, Roger Dwyer (1981, 1982), has done considerable work in which he investigates running moments in the time domain and does analysis in the frequency domain. He concludes that the noise is both non-Gaussian and nonstationary. In particular he sees very large sample kurtoses, which are attributed to impulsive noise, and notes the presence of pronounced tonals in the spectrum estimates which are hypothesized to be squeaks due to ice dynamics. Similar results have been reported by Milne and Ganton (1964) in other data sets.

we confirm the presence of such impulsive noise and such tonals, but believe that the remainder of the under ice noise is essentially Gaussian in nature. This is explored in Section 4.

3. Data Blocking and Some Simple Traces

Many of the commonly used techniques in time series rest on the twin assumptions of stationarity and Gaussianity. The largeness of our dataset gives us the opportunity to examine these assumptions for validity. Since computing some of the traditional statistics such as moments, quantiles and Fourier transforms is impractical for six million numbers, we are naturally led to the idea of blocking the data. To do this we group consecutive observations into groups of size N, say, where, if convenient, N may be taken as a divisor of T, the total number of data points. By computing, for example, moments for each of these $R = T/N$ records, we may ascertain whether stationarity and Gaussianity are reasonable assumptions.

3.1. Choosing a Block Size

From physical considerations, there is no natural record size N, for our ten minutes of acoustic noise, and we are thus free to choose it as we please. In most of our analysis we have used $N = 1024$ for the following reasons:

(i) Dwyer has used this value in his analyses, so our work and his may be compared at appropriate points.

(ii) A power of two is convenient for the computation of discrete Fourier transforms via the FFT algorithm.

(iii) N should be greater than about 100 in order to achieve satisfactory standard errors for record statistics, and less than about 10,000 in order that the computations do not get too cumbersome. Our chosen N is roughly the geometric mean of these two extremes.

(iv) N should be sufficiently small that within a single record the data is roughly stationary. 1024 seems empirically to satisfy this condition, though more rigorous testing would be possible.

(v) N should be at least 512 since this blocksize corresponds to .05 seconds of data so that the frequency of blocks is 20 Hz, and signals of 10 Hz or less would appear as a trend from block to block. Such frequencies are unlikely to be detected with the recording equipment used, though with a smaller blocksize, there might start to be block to block trends.

Thus, our first step is to break the data into $R = 6006$ records each of size $N = 1024$ numbers.

3.2. Running Means and Variances

Our first plots replicate those in Dwyer (1981), page 8. Denote the jth record by $x_{j,0}, \ldots, x_{j,N-1}$, for $j = 1,\ldots,R$, with $N = 1024$ and $R = 6006$. Thus the entire time series is just the numbers

$$\{x_{1,0}, \ldots, x_{1,1023}, \ldots, x_{6006,0}, \ldots, x_{6006,1023}\}.$$

We compute for the jth record the mean

$$m_j = \frac{1}{1024}\sum_{i=0}^{1023} x_{j,i}$$

and the sample central moments

$$m_j^{(k)} = \frac{1}{1024}\sum_{i=0}^{1023} (x_{j,i} - m_j)^k, \quad k = 2,3,4.$$

Each of these sample moments is computed for records which do not consist of independent observations, but in fact are highly correlated. Thus the sample moments will be biased estimates of the true moments. Notice that we are tacitly assuming that the data is relatively stationary within records, so that it makes sense to speak of the theoretical moments of a record. As an example of this bias, consider the sample variance in the presence of positive serial

correlation. In this case sample variance would estimate the variance with a negative bias, i.e. it would underestimate the true variance. We choose to ignore the effects of bias in what follows because the sample size is relatively large (1024) and because the records tend to have sample autocorrelation functions which drop off fairly rapidly.

The actual center and variability of the raw data are artifacts of the digitization method used on the original analog signal. Thus, we denote by m the sample average of the entire time series, and by $m^{(2)}$ the sample variance of the time series, and plot in Exhibits 3.1 and 3.2 the values of $(m_j - m)/(m^{(2)})^{1/2}$ and $m_j^{(2)}/m^{(2)}$, as functions of j.

It is clear from Exhibit 3.1 that the data seem trendless, and that except for records 2220, 2248 and 3236, we are just seeing sample fluctuations in the sample average. This trendlessness is to be expected, since the hydrophone was measuring longtitudinal pressure waves which must clearly average out to zero in any length of time relatively long compared to the period of the longest wave present. (See reason (v), above, for the use of a blocksize of 1024.)

The sample variance trace, Exhibit 3.2 clearly demonstrates that over the time period of the data there is marked non-stationarity. In fact it would appear that even in the short run there is no stationarity, with the possible exception of a few sections of the data where the sample variance trace is relatively flat. Sample variance is a measure of volume or signal energy, and thus it appears that the volume of the noise is fluctuating. In later analysis we will show that much of this variability in the variance can be attributed to the momentary presence of strong sinusoidal components superimposed upon relatively stable noise, and that the variance fluctuation is, in fact, not being caused by changes in the volume of this 'background' noise.

3.3. Running Skews and Kurtoses

Exhibits 3.3 and 3.4 are plots of the traces of the sample skew

$$S_j = \frac{m_j^{(3)}}{(m_j^{(2)})^{3/2}},$$

and the sample kurtosis

$$K_j = \frac{m_j^{(4)}}{(m_j^{(2)})^2}$$

for each of the 6006 records.

The skew and kurtosis for any Gaussian distribution are 0 and 3. Exhibits 3.3 and 3.4 show skews and kurtoses which are usually reasonable for a Gaussian distribution, with occasional dramatic deviations from the baselines of 0 and 3. In looking at these two plots together, it is striking that the large deviations in skew and kurtosis appear to occur together much of the time.

3.4. Explanation for Large Skews and Kurtoses

To illustrate this point the absolute skews and the kurtoses were each sorted, and the 50 largest selected from each list. Those records which are common to the two lists are sorted by kurtosis in Table 1 below. Nearly half of the spikes in the skew and kurtosis traces occur for the same records among the largest 50 of each. In addition, the largest 12 absolute skews and the largest 11 kurtoses appear on this list. Thus large absolute skew and large kurtosis are highly correlated. Exhibit 3.5 is a time series plot of the data for record 1362, which has the second highest kurtosis and the largest absolute skew. The reason for the large moments is immediately clear, viz, the short burst of very large observations around observation number 695. The duration of the burst is about 10 observations which in real time is about 1 millisecond. This behavior is, in fact, characteristic of records with large skew and large kurtosis, and accounts for large kurtoses at least as small as 6, as illustrated by Exhibit 3.6, a

time series plot of record 2238. This record has the 50th largest kurtosis (5.73) and a relatively small skew (.39).

Table 1
Largest Skews and Kurtoses

Record	Kurtosis	Skew
2066	29.51	-1.57
1362	28.17	-3.12
4947	27.91	-2.55
1487	24.15	0.51
2220	22.97	-2.72
2248	18.49	-2.92
41	17.67	-0.68
3236	17.07	-2.76
5776	14.95	0.62
2236	14.54	-1.39
2177	13.41	-1.05
1281	12.42	-0.70
2041	11.25	-0.89
2246	10.24	1.02
4902	10.00	-1.04
5546	9.99	-0.67
5557	9.55	-0.84
5705	9.25	-0.64
4808	8.05	-0.52
2042	7.43	0.62
1938	7.10	-0.55
2724	6.59	-0.63
4836	6.12	-0.66

Kurtosis is classically used to measure "fatness" in the tails of a distribution and these bursts appears to be the major factor contributing to the high K_n. If, for sample 1362, we excise the points from 689 to 699 and recalculate kurtosis, K_{1362} drops from 28.17 to 3.24. Skew is also significantly affected by these bursts; for sample 1362, the same excision causes S_{1362} to go from -3.12 to 0.039.

As stated previously, kurtosis measures "fatness" in the tails of the distribution. However, a large kurtosis does not tell us how far from the center of the distribution we must go before we see "fat" tails. We wished to check that the observed high sample kurtoses were being caused by bursts (i.e. the tails are

"fat" only far from the center). For we expect that the body of the distribution of the numbers from a record will remain unchanged by a burst, but that the few largest and smallest values in the sample will be much larger (or smaller, respectively). We devised a check by computing the statistic:

$$F_n = Y_n / H_n$$

where Y_n is the difference between the second highest and second lowest observations and H_n is the interquartile range.

This statistic is an analog of the sample kurtosis, in that it measures how long the tails are (Y_n) relative to the spread of the distribution (H_n). Sample kurtosis achieves this effect by dividing the average fourth power of deviations from the mean by the average second power. The latter measures spread, while the former causes greater magnification of those observations which deviate further from the mean. The problem with using kurtosis as a measure of the length of the tails of a distribution is that kurtosis can also be large for a distribution which is peaked, i.e. a distribution whose tails are fat relative to the center, but whose extreme tails are not particularly large. F_n will only be large in the presence of large extreme tails in the distribution and can therefore be used to check our belief that the observed large sample kurtoses for the ice noise are due to very high amplitude bursts and not merely to a peaked distribution.

For a comparison of the two measures see Exhibit 3.7, which illustrates the trace of F_n, limited to between 2 and 7, and the trace of K_n limited to between 2 and 5. The values were so limited in order that the behaviour of the majority of the values was not swamped by the few large values in each trace. It is clear that the two traces are virtually the same in character. The correlation between the logs of the measures is .93. (The reason for taking logs is to reduce the effect the few large values have on the correlation coefficient.) The group of

records with the 127 highest kurtoses and the group of records with the 127 highest F_n's share 117 members in common! Kurtosis and F_n measure essentially the same thing, and therefore kurtosis is large precisely because of the stray values represented by bursts.

A fair question to ask is the following: If we wish to automatically detect bursts in this noise, how large should the kurtosis of a record be before we flag such a burst? We settled on 4, after considering a portion of the data analyzed in section 4.1.1. In this portion of the data all but three kurtoses are under 3.9; the remaining three are all much larger. Theoretical considerations also make this a generous upper bound -- see section 4.1.1 for a discussion of simulation results.

It would appear, then, that all the large kurtosis records have a common cause: a transient large amplitude signal which disappears very rapidly compared to the time scale of each record. The erratic behavior of this transient seems to account for the large skew values often observed along with large kurtoses. Because K_n demarcates these records, we have a simple method for identifying the records containing bursts.

3.5. Other Behavior in the Skews and Kurtoses

Since these bursts are so large, they obscure other behavior in the moment traces for skew and kurtosis in Exhibits 3.3 and 3.4. In order to see this other behavior, we replot in Exhibit 3.8 the skew and kurtosis traces in the regions of the 0 and 3 baselines. The variance trace is also replotted in this Exhibit. This Exhibit contain most of the skews and kurtoses; in fact 99.6% of the skews and 98.8% of the kurtoses. In the skew plot, there is a region of relatively large negative skews from about 300 to 1400 and two segments where the skews seem to be compressed toward the origin: 1600-1900 and 3300-4300. In the kurtosis plot there seem to be three regions where the kurtosis drops below 3: 1000-1400,

1600-1900 and 3300-4300. These regions also seem to form natural boundaries on the variance plot. The coincidences of these regions bears investigation; we will suggest a model in the next section which appears to account for these effects. To this end, we divide and name the data as follows:

Region	Records	Variance	Skew	Kurtosis
1	1-300	normal	normal	normal
2	301-1000	high	stretched	normal
3	1001-1400	high	stretched	low
4	1401-1600	normal	normal	normal
5	1601-1900	high	squeezed	low
6	1901-3300	normal	normal	normal
7	3301-4300	high	squeezed	low
8	4301-6006	normal	normal	normal

4. The Basic Model.

On inspection of the variance trace, it is clear that the volume of the noise appears relatively constant over fairly long stretches of time. One can see this on casual inspection of the variance trace in Exhibit 3.2, for record numbers 1900-3300 or 4300-6006 (regions 6 and 8). In the regions where the variance is fluctuating much more wildly, many of the log periodograms of single records have very marked peaks at a few frequencies. Further, for these log periodograms, on ignoring peaks "by eye", the underlying periodogram looks very much the same from record to record. This suggests that the ice noise might be well modeled by a continuous spectrum, stationary series with added discrete sinusoids of high power whose frequencies change from record to record. An added complexity would be to allow the continuous portion to vary slowly through time. We shall call such discrete sinusoids *tones*, or *tonals*. Mathematically, let $P_n(\omega)$ be the observed periodogram of record n; ω ranges through the Fourier frequencies. Then the former hypothesis is essentially one of a *mixed spectrum*:

$$P_n(\omega) = h(\omega) + \sum_{i=1}^{I} A_{in} F_i(\omega - \varphi_{in}) + \varepsilon_n(\omega) \qquad (1)$$

The first term on the right is the continuous portion represented by a spectral density function $h(\omega)$. The sum represents discrete tones at frequencies $\varphi_{1n},\ldots,\varphi_{In}$ for some small integer I, and presents the effect of computing a periodogram as a sum Fejer kernels $F_1,\ldots,F_I$ weighted by the amplitudes $A_{1n},\ldots,A_{In}$ of the tones. If we allow variation of the continuous part, the model becomes:

$$P_n(\omega) = h_n(\omega) + \sum_{i=1}^{I} A_{in} F_i(\omega-\varphi_{in}) + \varepsilon_n(\omega) \qquad (2)$$

where the h_n may be slowly varying with n in a complex fashion. The tones are allowed to be quite arbitrary from record to record; the only constraint is that there be no more than I of them at any one time. If we assume in both hypotheses that the part of the process with spectrum $h(\omega)$ (or $h_n(\omega)$, in Model (2)) is stationary Gaussian, then the $\varepsilon_n(\omega)$ are independent, and exponentially distributed with mean 0 and variance $h(\omega)$. (Note that the $\varepsilon_n(\omega)$ are exponential because P_n is the raw periodogram; it is the sum of the squares of two independent normals - the real and imaginary parts of the discrete Fourier transform). We shall call the component of the noise with spectral density h the *background noise*.

We shall further assume (as is suggested by the log periodograms that we looked at) that h (or h_n) is fairly smooth. In particular, we assume that any peaks in h are much wider than the width of the peaks of tones appearing in the periodogram.

Both of these models have some flaws that are apparent on casual inspection. First, the high frequency part of all the raw periodograms are very flat (e.g. Exhibit 4.1 pictures a periodogram from record 4601) - the error distribution $\varepsilon_n(\omega)$ is obviously only poorly modeled here. This high frequency part of the spectrum has such low power that modifying any single observation can change the high frequency portion of the periodogram by several orders of magnitude.

We should recall that the data was filtered by a low-pass filter with a cut-off at 2500 Hz, so it is not surprising the models fail in this region. We shall thus ignore the frequencies above 2500 Hz. Second, Model (1) appears flawed in that even in the absence of tones h seems to vary systematically over time -- look at the variance trace for records 4430-4450. That there are no discrete frequencies for these records will be further explored in the section on background noise, Section 4.1.1. It is clear that any such model as (1) can be at best only approximately true.

There is a third flaw with Model (1) In some of the records very large transient pressure changes are obvious, as discussed in Section 3.4. These transient bursts are on the order of 1/100 of a second or less. They seem to occur in only a small fraction of the records. We could modify Model (1) or (2) to accommodate these bursts by adding a "burst" to the process at random intervals, but as records with bursts can be identified as discussed in section 3.4 we shall assume that conclusions about Model (1) or (2) apply to these records (except for bursts), and we shall base our study of Model (1) on records without these bursts. We do not attempt to model bursts in this analysis.

Our aim is to check whether Model (1) is a reasonable description of the data, and this will be done in the following sections.

4.1. Confirmation of Model (1)

4.1.1. Gaussianity in the Background Noise.

We can reasonably assume there are no tones in areas where the variance trace is flat. We thus wish to check whether the noise in these regions is stationary Gaussian; this would be predicted by Model (1). In order to substantiate this assumption we considered the sequence of records 4601-4901. A log periodogram of one of these (4601) is given in Exhibit 4.1. We ignore three of these

records, viz 4692, 4710 and 4808, where large kurtoses of 6.38, 11.36 and 8.05 flagged bursts in the noise. Otherwise, all the moment traces are very flat, with the variance trace ranging between .378 and .764. The observed kurtoses remain close to three with an average of 3.02 and a standard deviation of .233. In a simulation of 10,000 sample kurtoses from records of 1024 simulated independent Normals we obtained an average kurtosis of 2.995, with a standard deviation of .152; this is not strictly comparable since we have dependence in the ice noise, but it suggests that the ice data kurtoses are in fact what one might expect under an assumption of Gaussianity. We get very similar results for skew; the ice data skews for these records average .017 with a standard deviation of .125, to be compared with an average of -.00007, and a standard deviation of .0785 for the simulated data. This is encouraging evidence that records 4601 to 4901 may be Gaussian noise.

We substantiate this further by considering the actual distribution of the data for part of this region. However, the difficulty of handling so many numbers at once means we shall do this for only a small proportion of these numbers. The technique we use is to get a large independent sample by taking every 64th number in this region. This is reasonable since autocorrelations tend to fluctuate around zero by lag 8 or so. This gives us 298x16, or 4768 numbers which should behave as an independent identically distributed sample of Gaussian random variables. These numbers had a kurtosis of 3.167 and a skew of .058 respectively. A test against the hypothesis of kurtosis exactly 3 would reject at the 2% level; however a similar test for zero skew is only significant at about the 10% level. Considering how sensitive sample kurtosis is to tails larger than expected from normality, this result is remarkably close to the normal theory prediction.

QQ plots provide a very good qualitative measure of Gaussianity- deviations

from the expected straight line are easy to see. This plot is shown in Exhibit 4.2. Qualitatively, the fit to a straight line seems excellent. Quantitatively, we can test this fit. This test is equivalent to a Kolmogorov-Smirnov test of whether the distribution is Gaussian and only rejects the hypothesis that the distribution is Gaussian at a significance level of greater than 20%. Thus the subsample of 4768 points passes this test also. A histogram of the subsample with a superimposed exact normal curve is shown in Exhibit 4.3. This picture also qualitatively confirms the hypothesis of Gaussianness.

We conclude that the assumption of a Gaussian distribution for this stretch of the data appears well substantiated, except for three records containing bursts. It thus seems a reasonable hypothesis that the background noise remains Gaussian, even in the high variance regions, especially if we can demonstrate that the background noise keeps much the same power. This is one of the conclusions in the next section.

4.1.2. Existence and Identification of Tones.

In the trace of record variances, there are several high variance regions: 300-1400, 1600-1900 and 3300-4300 (regions 2, 3, 5 and 7). Minor fluctuations are apparent in regions 2600-2800, and perhaps 5000-5400.

A reasonably typical log periodogram (smoothed) of a group of these records (3508-3512) is pictured in Exhibit 4.4. Notice there is an obvious double spike at about 1250 Hz and a possible harmonic of these spikes at 2500 Hz. So tones occur some of the time. Inspection of periodograms in the first group (300-1300) reveal a more complicated structure (e.g. see Exhibit 4.5, record 730), but we suspect that the high variance records observed in this group have the same cause. We shall treat this group separately, later in this section.

We first note one possible consequence of Model (1). The addition of a

discrete tone adds a sinusoidal component to the observed data. If the tone has large power compared to the background noise, then the distribution of observed series will be similar to the distribution obtained by randomly sampling X from a distribution with distribution function G, where

$$G(x) = \frac{1}{2} + \frac{1}{\pi} Arcsin\ x$$

The kurtosis of this distribution is 1.5. It is to be expected that the addition of a small number of tones with high power will have much the same effect, namely lowering the observed kurtosis to less than 3. This need not necessarily be true; if there are harmonics with power close to that of the fundamental tone, the kurtosis may vary over a range, including values larger than 3.

This suggests that a comparison of low kurtosis records with their corresponding variances might be revealing. These plots (Exhibit 3.8 and Exhibit 3.2) clearly pick out the regions mentioned above, except for 5300-5500 and perhaps 2600-2800. The regions 1600-1900 and 3300-4300 are particularly striking, with almost no kurtoses greater than three. The case of few tones, with few harmonics seems to be true here. However, the region 300-1300 retains many records with kurtoses greater than 3.

Note that this does not prove Model (1), even in the regions of low kurtosis. Strengthening the case for Model (1) requires a demonstration that the background noise is not varying, and that larger kurtoses in the 300-1300 record number region are actually due to tones with many harmonics, and not to other reasons.

It is certainly not unreasonable to expect the background noise to change at the same time as we observe mixtures of discrete frequencies (a failure of Model (1)). But, as discussed in the beginning of Section 4, the background noise appears very constant with virtually no change in power for long record stretches.

To check this we wished to find the power due to the tones and remove it, leaving only the power due to the background noise. We did this by smoothing the periodogram in such a way as to remove the influence of tones at or around discrete frequencies while preserving the overall shape of the periodogram. Since the total variance or power of a process is just the area under the periodogram, we could then compute the "residual variance" of the part of the noise modelled by h by taking the area under the smoothed periodogram. Such a smoothing procedure should have the property that it preserves the shape of periodograms from records with no apparent spikes due to tones. Classical smoothing methods are not applicable, as they merely average the spikes at discrete frequencies into the neighboring frequencies. In our circumstance, where the variance attributable to a tone is between two or three times the variance at all other frequencies, classical smoothing will clearly produce a highly biased estimate of the power due to h_n.

The usual method of removing tones is to use complex demodulation to identify the frequency at which the tone occurs, and remove that frequency with some sort of average of surrounding frequencies. The major problem with this approach is that it does not lend itself well to automation, for it requires automatic identification of the frequencies at which tones occur. Alternatively, one can attempt to use some sort of sequential testing procedure like the $P(\lambda)$ test Priestley (1981) advocates, and then remove the peaks. A double windowing technique like that suggested by Hannan (1961), is also a possibility. All these approaches require considerable work which is not easy to automate.

Some sort of non-linear smoother seems to answer these problems. A non-linear smoother will produce a biased estimate of h, but under the assumption of sufficiently smooth h, and Gaussianity in the background noise, this bias can be adjusted for reasonably well. The smoother we chose to use was that of a

$(2m+1)$ point running median. Denote this smoother by S_m. Then the value of the smoothed periodogram at Fourier frequency $\omega_i = 2\pi i / N$ is:

$$S_m P(\omega_i) = \underset{i-m \le j \le i+m}{\text{median}} \{P(\omega_j)\}$$

We need an endpoint rule: we chose to simply left extend P for values of $i \le m$; similarly for right extension. One of the virtues of this smoother is that it gives the same smooth regardless of whether one operates on the raw periodogram and takes logs, or operates with the log periodogram, since log is a monotone transformation.

Intuitively, this smoother will remove any narrow spikes in the periodogram. As an estimator of h it has several sources of bias. Firstly, it will smooth out peaks and troughs in h. That this bias is kept small will depend on h being sufficiently smooth in relation to the "peakiness" of the spikes in the periodogram due to tones. Secondly, even if h is quite smooth, $S_m P$ will have a consistent bias due to the fact that the median of the exponential distribution is log(2) rather than 1. This can easily be adjusted for. Thirdly, it will fail if there are too many spikes too close together in the periodogram. In this case it is unclear how to mechanize a procedure for tone removal -- this appears to be a problem in the 300-1300 group, as discussed a little later.

The major theoretical justification for using S_m can be found in Theorem 1 of Appendix A, which states that, under mild conditions on h and m, $S_m(P_N(\omega))/\log(2)$ converges to $h(\omega)$ uniformly in probability. Further, uniform convergence to h in probability implies that the residual variance given by

$$\frac{1}{N}\sum_{i=1}^{N-1} S_m(P(\omega)) \to \int h(\omega)d\omega = \text{var}(X) \quad \text{in prob.}$$

where X is the Gaussian part of the time series. (For an exact statement, see Corollary 1 in Appendix A).

We chose $m=25$. This removes the discrete spikes almost totally, and, since

h looks locally smooth, the residual variance should not have too much bias in it. Model (1) predicts that the residual variance trace should be constant. We calculated this residual variance for each record and the result of our procedure can be seen in Exhibit 4.6, where the original sample variance trace is plotted above the residual variance trace. (All records containing suspected bursts, i.e. those with sample kurtoses bigger than 4, are not plotted -- our smoothing procedure can only handle tones, not bursts.) The procedure seems to do a reasonable job of normalizing the variance trace everywhere except for the region of records 300-1000. That these records still fit into the framework of Model (1) remains to be shown. It is clear though, that Model (1) is only approximately true elsewhere.

The region 300-1000 is anomalous in other ways. The kurtoses are not consistently low. An inspection of the skew trace, with skew values limited to between +.6 and -.6 (Exhibit 3.8) clearly indicates that this region is qualitatively different from the region 3300-4300, where skew appears to decrease in absolute value. Record 730 is displayed in Exhibit 4.5, and clearly shows a strong non-sinusoidal but periodic component. The periodogram of this record shows the fundamental at about 240 Hz and six or seven harmonics, five of which have about the same power. The running median smoother cannot deal with this case, as the Fejer kernels about each peak leak substantial amounts of power which reinforce each other, at least for the length of running median we chose. It is obvious that for a sample length of 1024, no running median smoother will work well in this case. Other periodograms we looked at have a yet more complex structure with two or three fundamentals apparent, together with their harmonics.

We believe that Model (1) is still valid for the whole region, but this is on purely observational grounds - the periodograms we considered look like they

retain the same underlying structure. This remains an area where we need to do more rigorous research.

There remains open the question of how one might model the frequencies and amplitudes of these tones, as a random process. This is not an easy task, as it seems apparent that the tones at different frequencies are connected with each other -- to wit, it is clear that we sometimes observe a fundamental with its harmonics. Any reasonable model must take such structure into account. Consideration of the physics of underwater acoustics and of acoustic energy released in the fracture of solids might suggest some avenues of approach.

That these tones have a complex structure even without considering the dependence structure at one time is illustrated by the fact that even in periods of a second or so, the tone may wander in both frequency and in amplitude. A plot of the spikes in the estimated spectra of a series of records 1770-1830 (see Exhibit 4.7) illustrates this point. This plot was made by first computing a spectrum estimate for each of the records. The estimate used was a length five Daniell window applied to the periodogram. Then the difference between this estimate and a smoothed version (smoothed using running medians of length 41) was computed in order to leave only the spikes. The plot shows, for each record, only the upper 15% of these differences.

In the early records, 1770 to 1790 or so, there is quite a strong tone at around 1100 or 1200 Hz, which appears to split into two tones; one staying at the same frequency until late in the sequence, the other wandering up to about 1500 Hz. The two tones seem to be merging again by record 1830. There is also an apparent harmonic at 2200-2300 Hz, which increases in frequency just as its fundamental does.

5. Summary

This report has modeled the undersea noise with a model suggested by an exploration of the dataset itself. The model consists of three parts; the first part is a relatively stable Gaussian process (called the *background noise* in this report) which is assumed to have a smooth continuous spectrum. Superimposed on this Gaussian process are a small number of random tones. The frequencies amplitudes of these tones are changing over time and are sometimes absent entirely. The third component in the model is the appearance of sporadic, extremely transient, high intensity bursts of impulsive sound energy (bursts). These bursts are quite rare in the data we examined.

The model we suggest is incomplete in several respects. It is only approximate, but we believe this is not a problem. This belief could best be tested by trying out the model with another dataset. The model is unproven in some regions of the data, particularly where a large number of tones are difficult to distinguish from the background Gaussian noise.

There are several areas remaining open for investigation. One such area is in further modeling. We do not specifically estimate the parameters associated with the Gaussian process. In addition, we do not attempt to model any probabilistic process for the impulsive noise bursts, nor do we attempt to model the tones by a probability model. Trying to model the bursts, perhaps as Poisson events, may be fruitful. The problems associated with modeling the tones appear more difficult. For one thing, we cannot always dissect out the tones, as mentioned earlier, and for another, the tones probably have a complex structure dependent on the physics of acoustic energy released due cracking and strains in the ice. However, it may be that a simple model useful in signal detection can adequately describe these tones.

Another area of interest is in signal detection. Knowledge of the noise can

suggest useful methods of signal detection. There are a variety of approaches -- for example, one might derive theoretical results which can be tested by adding signal to the noise data. This might best be done in conjunction with more complete modeling Another interesting approach would be to compare the effectiveness of existing detection schemes in the presence of this noise.

Appendix A

The aim of this appendix is to give the technical results used in the paper. These are given in Theorem 2 and its Corollary.

Suppose that we observe

$$Z_i = X_i + Y_i, \quad i=1, 2,..., N, \tag{1}$$

where X_i is Gaussian with non-normalized spectral density $h(\omega)$, positive for all ω, and Y_i is a mixture of sinusoids at I discrete frequencies $\{\varphi_i\}$ with amplitudes $\{A_i\}$ and arbitrary phase shifts ($i = 1, 2,..., I$). Assume that X_i is a linear process:

$$X_i = \sum_{t=-\infty}^{\infty} \beta_t \varepsilon_{i-t} \tag{2}$$

where the ε_t are independent and Gaussian, $E(\varepsilon_t) = 0$, $E(\varepsilon_t^2) = 1$, and the β_t are constants satisfying the condition

$$\sum_{t=-\infty}^{\infty} t^c |\beta_t| < \infty \tag{3}$$

for some $c > 0$. Condition (3) implies that the spectral density h is continuous and bounded on $[0,\pi]$. Lastly, assume $h(\omega) > 0$, for all ω.

Form the standard periodogram of Z_i at Fourier frequencies $\omega_{jN} = 2\pi j / N$, $0 < j < N/2$ (we ignore the end points 0, $N/2$ so that Theorem 1 holds true). So

$$P_N(\omega_{jN}) = \frac{1}{2\pi}\frac{1}{N}\left|\sum_{j=1}^{N}(Z_j - \bar{Z})e^{-i\omega_{jN}}\right|^2 \tag{4}$$

We need two preliminary results- the first is a lemma due to Hoeffding (1963):

Lemma 1

Let W_i, $i=1, 2,..., m$ be independent random variables bounded between zero and one. Then, for each $t > 0$,

$$P\left[\sum_{i=1}^{m} W_i - \sum_{i=1}^{m} E(W_i) \geq mt\right] \leq e^{-2mt^2}$$

The second result is due to Walker (1965) and concerns the asymptotic independence of $P_N(\omega_{jN})$ for different Fourier frequencies:

Theorem 1

Let X_t be given by (2) and satisfy (3). Then:

$$P_N(\omega_{jN}) = V_{jN} + R_N(\omega_{jN}) \tag{5}$$

where the V_{jN} are independent, each with distribution $\frac{1}{2}h(\omega_{jN}).\chi^2_2$. (Notice the Fourier frequencies do not include 0, $N/2$).

Further:

$$\max_{0<j<N/2} |R_N(\omega_{jN})| = o_p(1) \tag{6}$$

Under these conditions we prove that the running median estimate of h converges uniformly in probability by proving the following theorem.

Theorem 2

Let Z_t be a process as given in (1), satisfying condition (3). Let m be an integer depending on N such that as $N \to \infty$, the following holds:

$$\frac{m}{N} \to 0 \text{ as } N \to \infty \tag{7}$$

and there exists a $\gamma > \frac{1}{2}$ so that:

$$\frac{m}{N^\gamma} \to \infty \text{ as } N \to \infty \tag{8}$$

and assume h bounded away from zero.

Given ω, let $\{\omega_j\}$ be the set of m ω_{jN}'s nearest to ω (mod π). Let

$$h_{mN}(\omega) = \text{median of } \{P_N(\omega_j)/\log(2)\} \tag{9}$$

Let $\varepsilon > 0$ be given. Then, as $N \to \infty$,

$$P[\sup_{\omega} |h_{mN}(\omega) - h(\omega)| > \varepsilon] \to 0 \tag{10}$$

Proof

First note that $P_N(\omega_{jN})$ are approximately distributed as N independent $\frac{1}{2}h(\omega_{jN}).\chi^2_2$, (i.e. exponentials with parameter $h(\omega_{jN})^{-1}$), but with a fixed

constant added, $g(\omega_{jN})$, where

$$g(\omega) = \frac{1}{2}\sum_{i=1}^{I} A_i^2\{F_N(\omega-\varphi_i)+F_N(\omega+\varphi_i)\} \tag{11}$$

where the A_i's and φ_i's come from the discrete frequencies assumed in the theorem and F_N is a Fejer kernel;

$$F_N(\vartheta) = \frac{1}{2\pi N}\frac{\sin^2(N\vartheta/2)}{\sin^2(\vartheta/2)} \tag{12}$$

Given the decomposition in Theorem 1, we know the distribution functions of V_{jN}, say they are given by G_{jN}. For simplicity in the following argument, let us further suppose $I=1$. The case of $I>1$ is proved along the same lines, except for increased notation. Fix an ω_0. Let the median of an exponential, parameter $h(\omega_0)^{-1}$, be given by:

$$\xi_0 = \log(2).h(\omega_0)$$

Let $\xi_{mN}'(\omega_0) = \log(2).h_{mN}(\omega_0)$. We shall abuse notation by dropping the argument to ξ_{mN}' when it is clear to which ω_0 we are referring. Then ξ_{mN}' is an estimate of ξ_0, the median of an exponential, parameter $h(\omega_0)$. We shall work with ξ_0 and the estimate ξ_{mN}', so without loss of generality, replace ε by $\varepsilon.\log(2)$ in the statement of this theorem.

Make the decomposition (5). We work first with the case of considering $\xi_{mN}(\omega_0)$ = the median of $\{V_{jN}\}$ and extend to ξ_{mN}' later. Again, we drop the argument to ξ_{mN} if it is clear which argument is implicitly referred to. So:

$$P[|\xi_{mN}-\xi_0| > \varepsilon] = P[\xi_{mN} > \xi_0 + \varepsilon] + P[\xi_{mN} < \xi_0 - \varepsilon] \tag{13}$$

Let

$$\alpha = \xi_0 + \varepsilon$$

Then

$$\begin{aligned} P[\xi_{mN} > \alpha] &= P[\tfrac{m}{2}\text{or more } P_N(\omega_j) > \alpha] \\ &= P[\sum_{i=1}^{m} I[P_N(\omega_j) > \alpha] \geq \tfrac{m}{2}] \\ &= P[\sum_{i=1}^{m}(I[P_N(\omega j) > \alpha]-G_{jN}(\alpha)) \geq \sum_{i=1}^{m}(\tfrac{1}{2}-G_{jN}(\alpha))] \end{aligned}$$

$$\leq \exp\left[-2m\left(\frac{1}{m}\sum\left(G_{jN}(\alpha)-\frac{1}{2}\right)\right)^2\right]$$

$$= \exp(-2m\,\delta_{mN}^2)$$

by Lemma 1, where

$$\delta_{mN} = \frac{1}{m}\sum\left(G_{jN}(\alpha)-\frac{1}{2}\right)$$

We now show for m, N large enough

$$\delta_{mN}^2 \geq \delta_0^2 > 0$$

Put $L = \max\{h(\omega)^{-1};\omega \text{ in } [0,2\pi)\}$. L exist by the assumption h is positive. As maximum of h is bounded, put $M = \min\{h(\omega)^{-1}\}$, and for N large enough,

$$A_1/N < \frac{L\varepsilon}{4M}, \text{ for all } \omega$$

So assume this holds. Also:

$$A_1 F_N(\omega_{jN}-\varphi_1) \geq h(\omega_{jN})^{-1} \text{ for at most } k_N \text{ of the } \omega_{jN}$$

We can choose k_N to satisfy $\frac{N}{k_N^2} \to 0$. where k_N comes from consideration of the form of the Fejer kernel, which is bounded by $O(N/k_N^2)$, for the $N-k_N$ Fourier points farthest from φ_1.

Simultaneously, we can choose k_N to satisfy $\frac{k_N}{m} \to 0$, by considering condition (8); e.g. choose $k_N = N^\zeta$, where $\frac{1}{2} < \zeta < \gamma$.

So, for at most k_N points, $\xi_0+\varepsilon-g(\omega_j) < 0$. So:

$$\delta_{mN} = \frac{1}{m}\sum_j G_{jN}(\alpha)-\frac{1}{2}$$

$$= \frac{1}{m}\sum_j\left[e^{-\xi_0 h(\omega_0)^{-1}}-e^{-h(\omega_j)^{-1}(\xi_0+\varepsilon-g(\omega_j))}\right]$$

except for at most k_N ω_j,

$$= \frac{1}{2m}\sum_j\left[1-\exp\{-h(\omega_j)^{-1}\varepsilon+[h(\omega_0)^{-1}-h(\omega_j)^{-1}]\xi_0+h(\omega_j)^{-1}g(\omega_j)\}\right] \qquad (15)$$

except for at most k_N ω_j. We use now uniform continuity of h^{-1} and the fact $m/N \to \infty$ to note that for m large enough,

$$|h(\omega')^{-1} - h(\omega)^{-1}| < \varepsilon' < \frac{L\varepsilon}{4\xi_0}$$

if $|\omega_j - \omega_0| < m/N$. So (15) $\geq$

$$\geq \frac{1}{2m}\sum_j \left[1 - \exp(-L\varepsilon + \varepsilon'\xi_0 + MA_i/N)\right] - \frac{k_N}{2m}$$

where we have dealt with the "at most k_N ω_j's" by bounding any exceptional term $G_{j'N}(\alpha) - \frac{1}{2}$ from below by $-1/2$.

$$\geq \frac{1}{2m}\sum_j [1 - e^{-\frac{L\varepsilon}{2}}] - \frac{k_N}{2m} > \delta_0 > 0$$

for m, N large enough, recalling $\frac{k_N}{m} \to 0$. It is clear that this bound does not depend on ω_0. So we have proved

$$P[\xi_{mN} > \xi_0 + \varepsilon] \leq e^{-2m\delta_0^2}$$

for m, N large enough. We prove an exactly similar bound on the second term in the right side of (13). Let δ_1 be the minimum of the two δ_0's. Hence

$$P[|\xi_{mN} - \xi_0| > \varepsilon] \leq 2e^{-2m\delta_1^2} \tag{16}$$

Now we bound the effect of including $R_N(\omega_{jN})$:

$$P[\sup_{0\leq\omega_0\leq\pi} |\xi_{mN}'(\omega_0) - \xi_0| > \varepsilon] = P[\sup_{0<j<N/2} |\xi_{mN}'(\omega_{jN}) - \xi_j| > \varepsilon]$$

(where $\xi_j = \log(2).h(\omega_{jN})$)

$$\leq P[\sup_{0<j<N/2} \{|\xi_{mN}(\omega_{jN}) - \xi_j| > \varepsilon - \sup_{\omega_{jN}} |R_N(\omega_{jN})|\}]$$

$$\leq P[\sup_{0<j<N/2} |\xi_{mN}(\omega_{jN}) - \xi_j| > \varepsilon/2]$$

$$+ P[\sup_{0<j<N/2} |R_N(\omega_{jN})| > \varepsilon/2] \tag{17}$$

The second term of the (17) is $o_p(1)$ by (6). We show that the first term is also $o_p(1)$. But by applying (16),

$$P[\sup_{0<j<N/2} |\xi_{mN}(\omega_{jN}) - \xi_j| > \varepsilon/2] \leq Ne^{-2m\delta^2}$$

(by applying Lemma 1 and sub-additivity of probability measures)

$$= \exp(-2m\delta^2 + \log N)$$

$$\to 0 \text{ if } \frac{m}{\log N} \to \infty$$

This is implied by condition (5), so we are done.

We have an immediate corollary to Theorem 1:

Corollary 1

Let

$$V_{mN} = \frac{1}{N}\sum_{i=1}^{N} h_{mN}(\omega_{jN})$$

and let

$$V = \int h(\omega)d\omega = \mathrm{var}(X)$$

Then, given $\varepsilon > 0$, and under the assumptions of Theorem A,

$$P[\,|V_{mN} - V| > \varepsilon\,] \to 0$$

as $N \to \infty$.

Proof

The proof is immediate on considering the result of the theorem and recalling that h is uniformly continuous.

1. Dwyer, R. (25 November 1981), "FRAM II Single Channel Ambient Noise Statistics", *NUSC Technical Document 6583*.

2. Dwyer, R. (May, 1982), "Arctic Ambient Noise and their Implications To Sonar Performance Improvements", *Proc. of "Undersea Ambient Noise", SACLANT ASW Research Centre, La Spezia, Italy*.

3. Hannan, E.J. (1961), "Testing for a jump in the spectral function", *J Roy. Statist. Soc. Ser. B*, **23**, 394-404.

4. Hoeffding, W. (1963), "Probability inequalities for sums of bounded random variables", *J. Amer. Statist. Assoc.*, **58**, 369-408.

5. Milne, A.R., and Ganton, J.H. (May 1964), "Ambient Noise Under Arctic-Sea Ice", *J. Acoust. So. Am.*, Vol. 36, No. 5.

6. Priestley, M.B. (1981), "Spectrum Analysis and Time Series", Academic Press.

7. Walker, A.M. (1965), "Some asymptotic results for the periodogram of a stationary time series", *J. Austral. Math. Soc.*, **5**, 107-128.

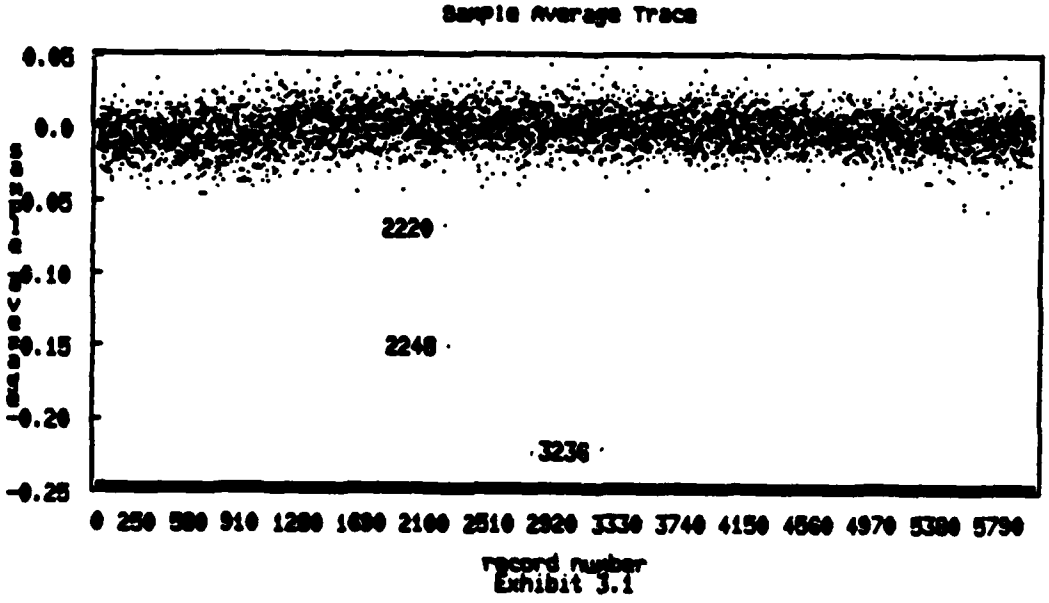


Exhibit 3.1

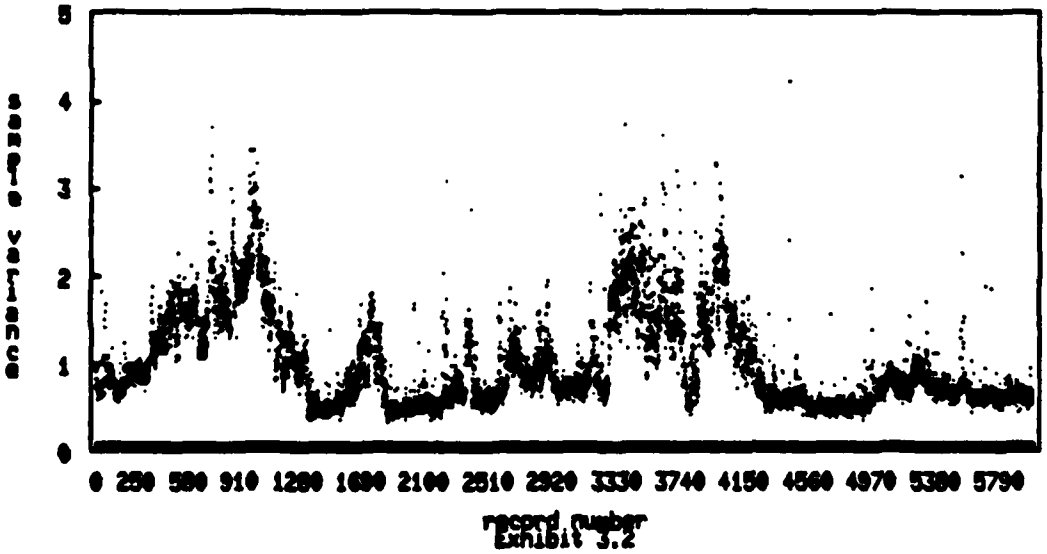


Exhibit 3.2

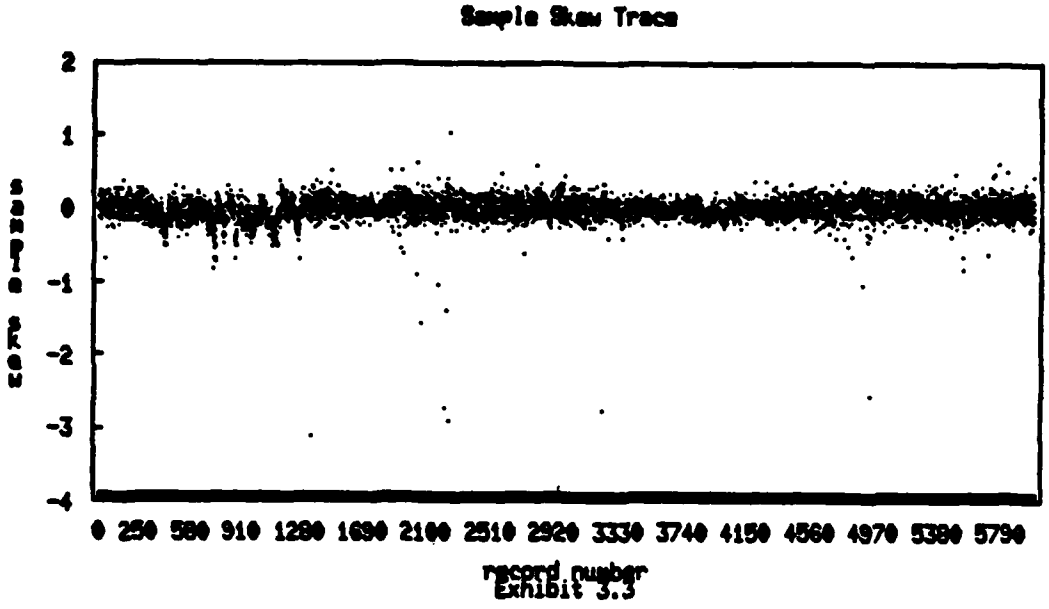


Exhibit 3.3

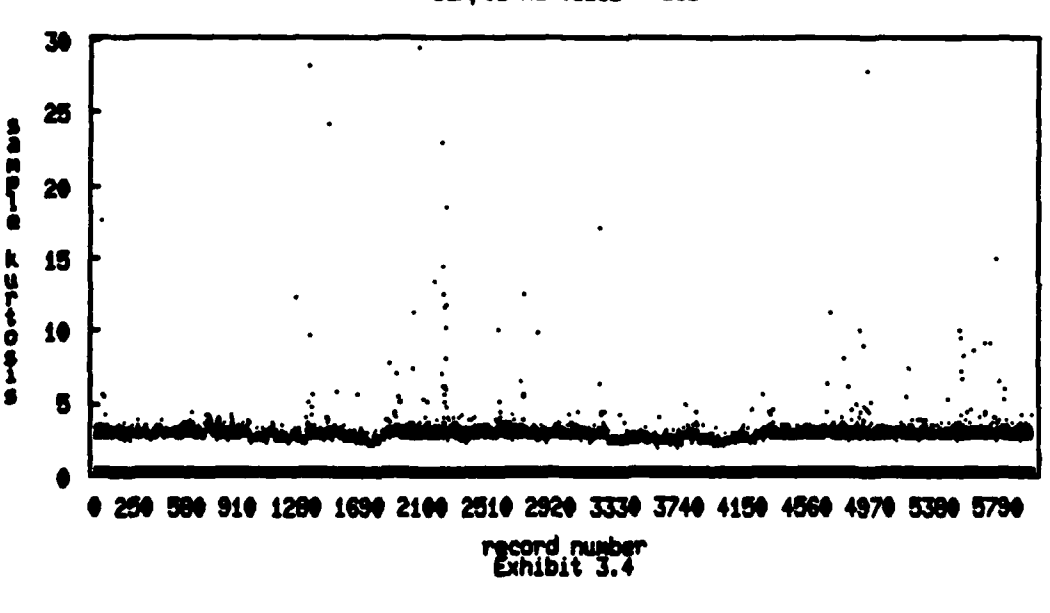


Exhibit 3.4

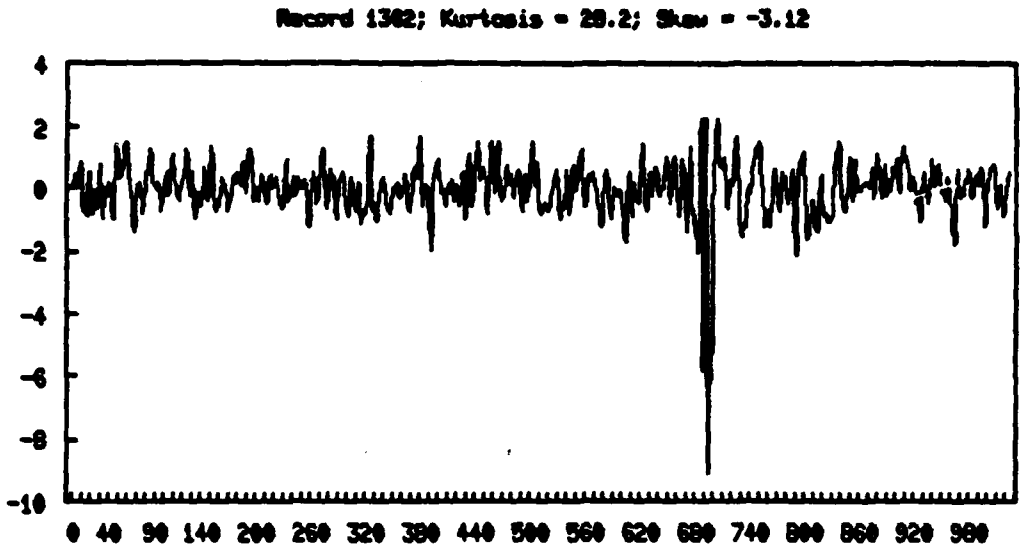


Exhibit 3.5

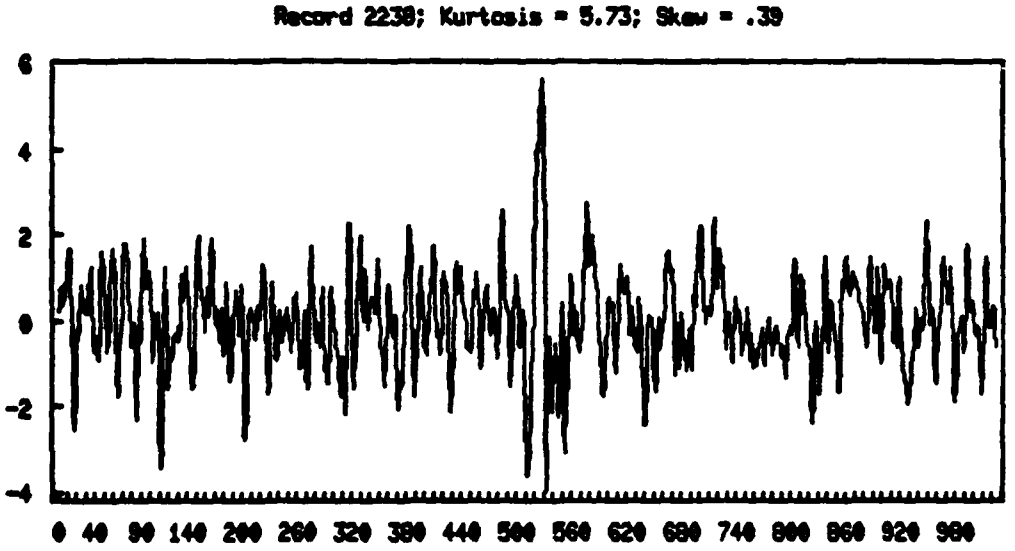


Exhibit 3.6

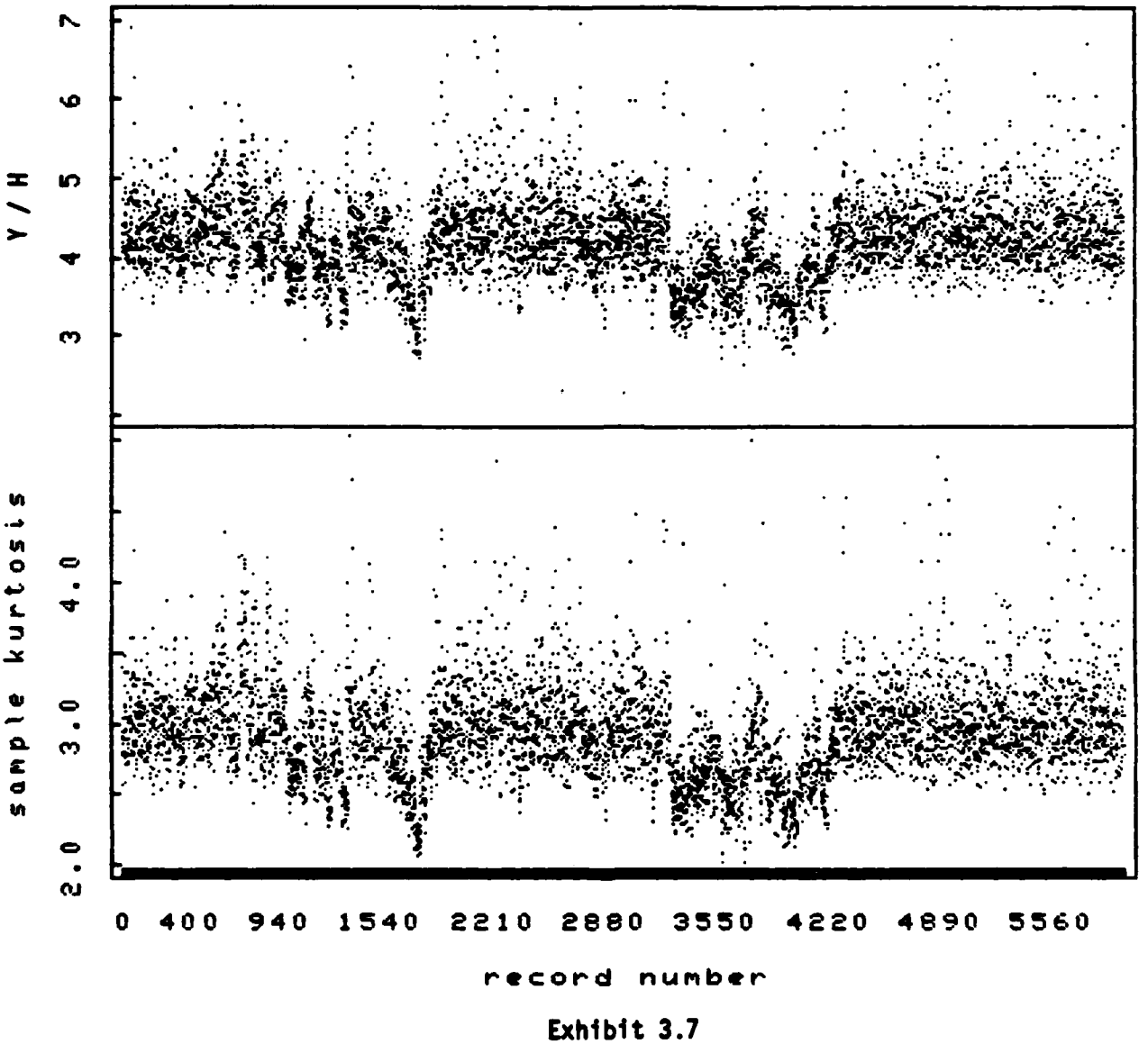

Trace Y-Spread / H-Spread and Kurtosis in [2,5]
Y / H
3
4
5
6
7
sample kurtosis
2.0
3.0
4.0
0
400
940
1540
2210
2880
3550
4220
4890
5560
record number

Exhibit 3.7

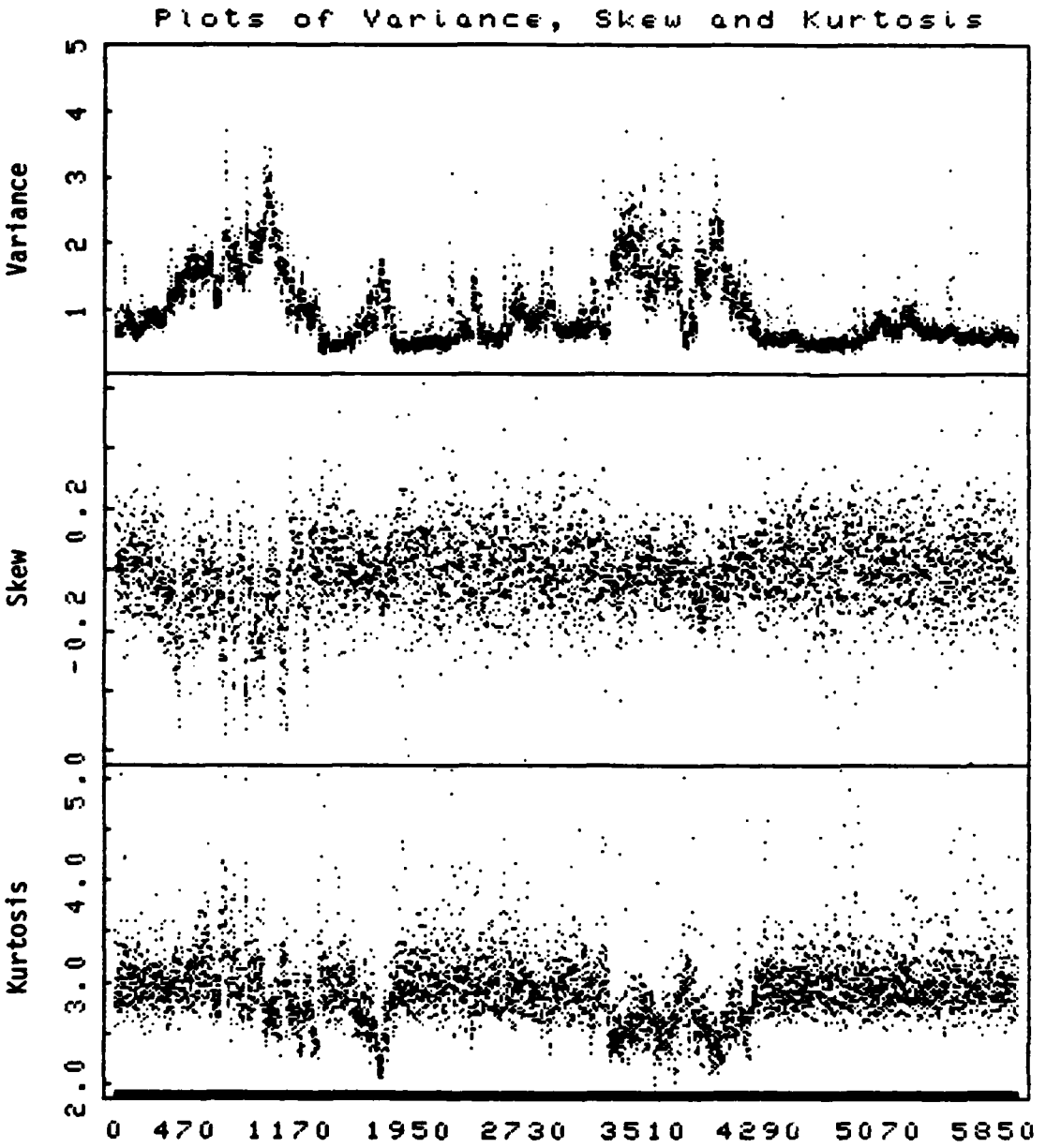

Plots of Variance, Skew and Kurtosis
Variance
1
2
3
4
5
Skew
-0.2
0.2
Kurtosis
2.0
3.0
4.0
5.0
0
470
1170
1950
2730
3510
4290
5070
5850
record number

Exhibit 3.8

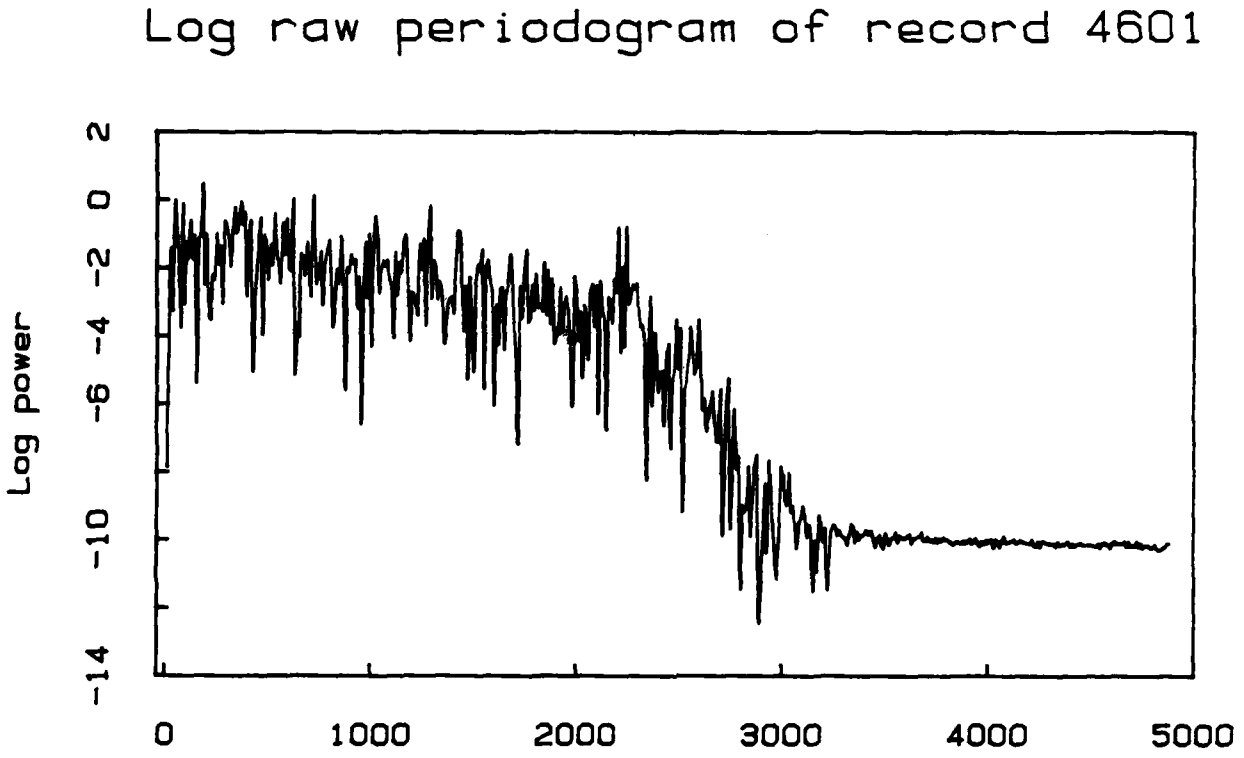

Log raw periodogram of record 4601
Log power
2
0
-2
-4
-6
-10
-14
0
1000
2000
3000
4000
5000
Hz

Exhibit 4.1

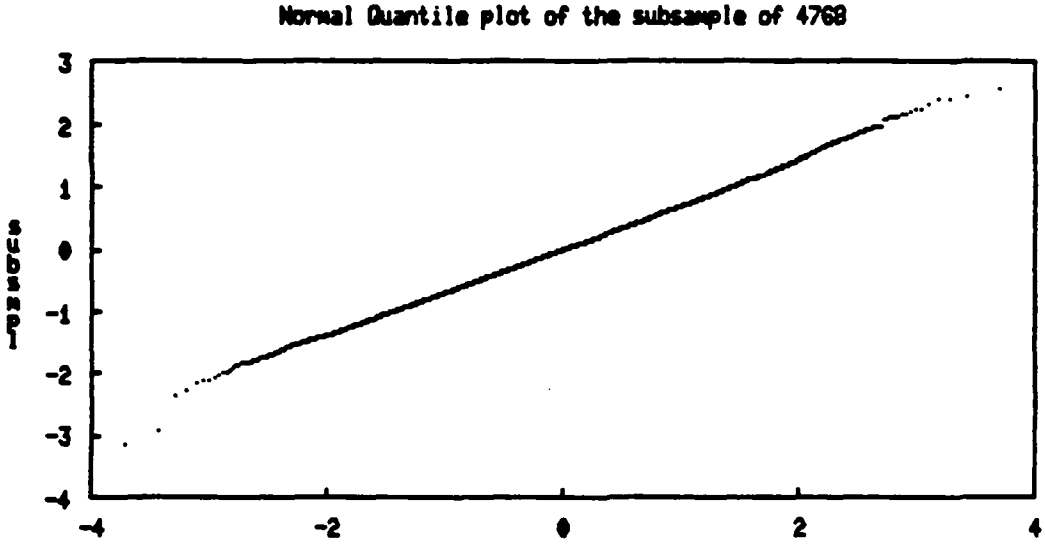


Exhibit 4.2

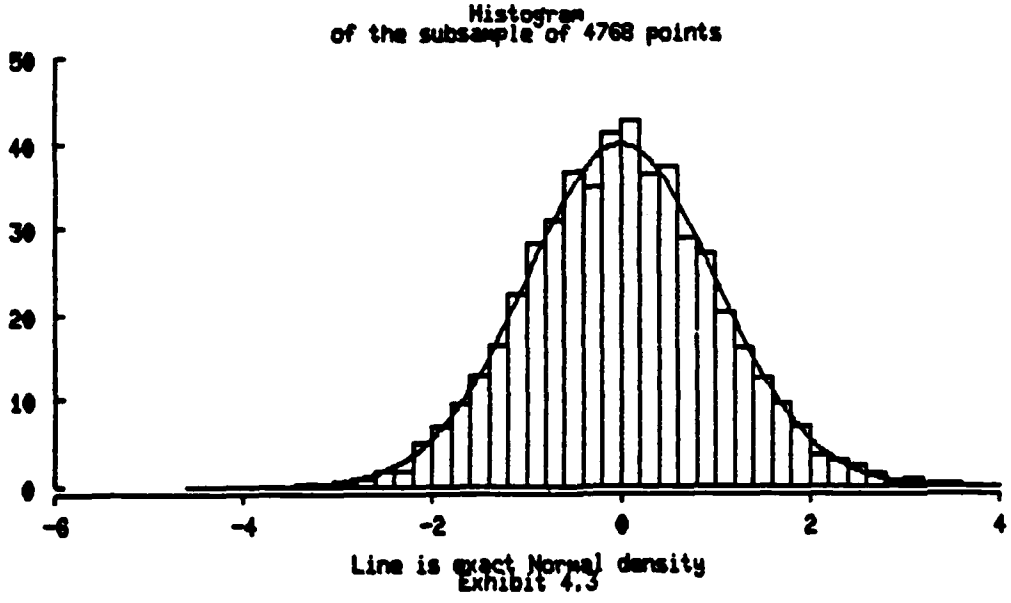


Exhibit 4.3

Spectrum estimate; records 3508-3512

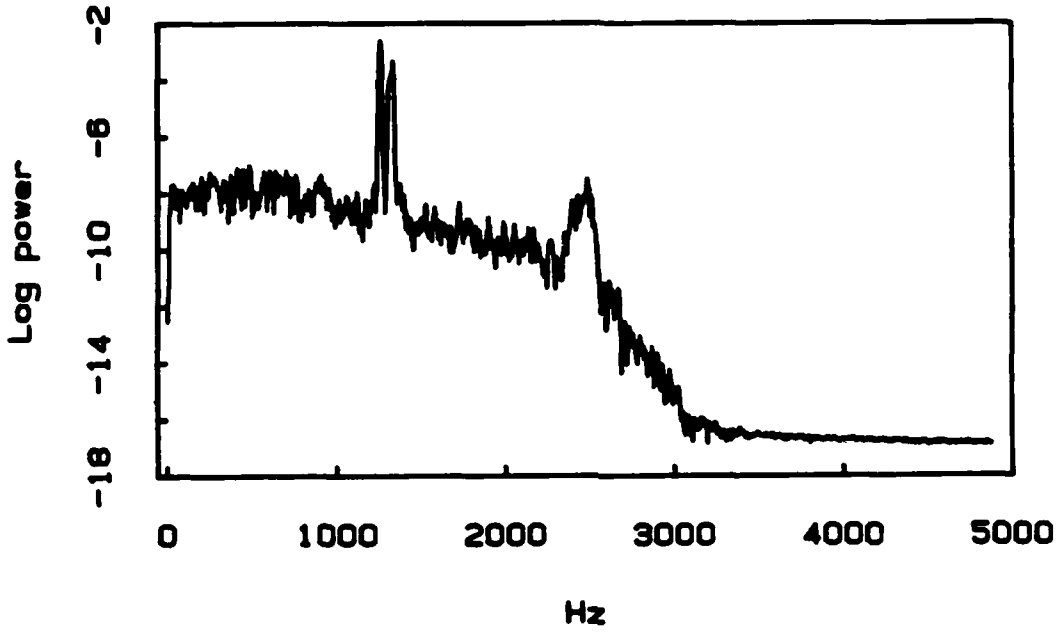


Exhibit 4.4

Log spectrum estimate of record 730

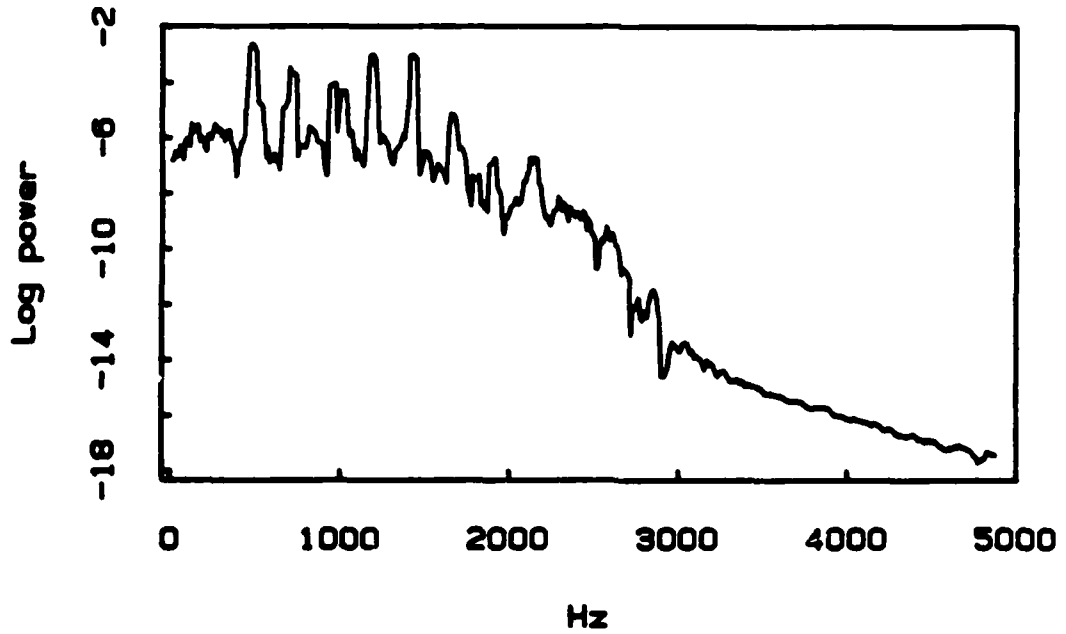


Exhibit 4.5

Comparison of Original Variance and Residual Variance

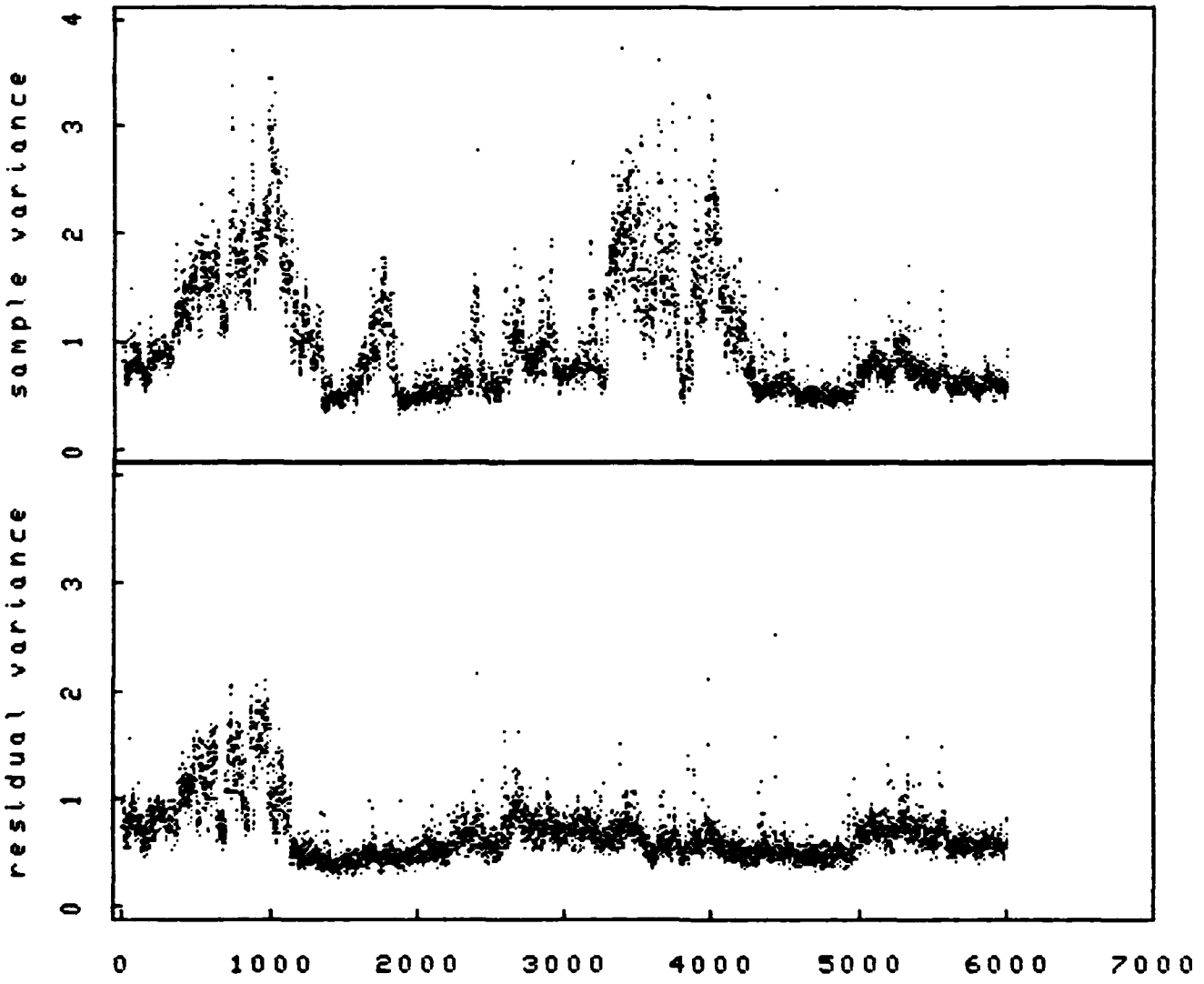


Exhibit 4.6

Plot of spectral spikes in records 1770-1830

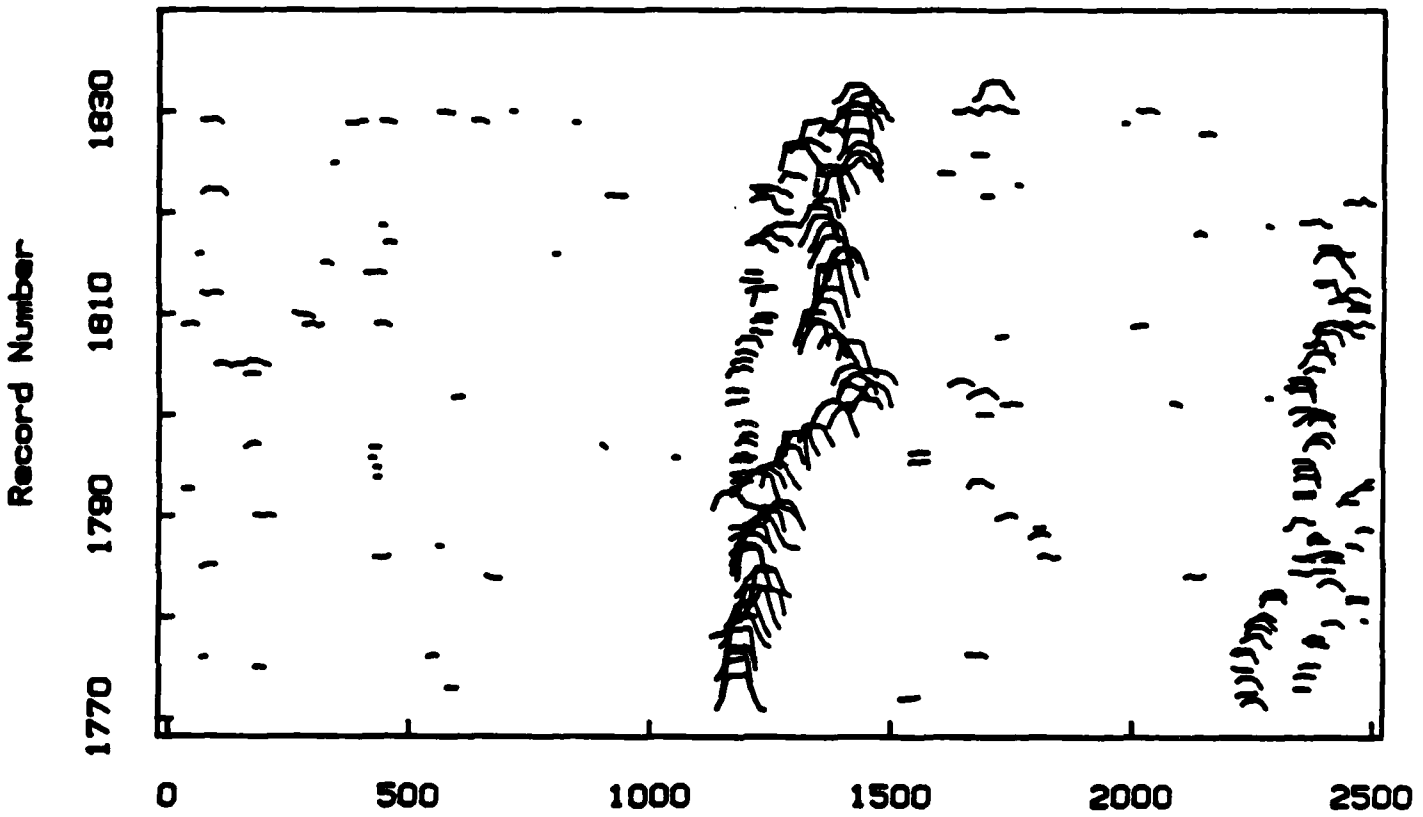


Exhibit 4.7

OFFICE OF NAVAL RESEARCH
STATISTICS AND PROBABILITY PROGRAM

BASIC DISTRIBUTION LIST
FOR
UNCLASSIFIED TECHNICAL REPORTS

FEBRUARY 1982

	Copies
Statistics and Probability Program (Code 411(SP)) Office of Naval Research Arlington, VA 22217	3
Defense Technical Information Center Cameron Station Alexandria, VA 22314	12
Commanding Officer Office of Naval Research Eastern/Central Regional Office Attn: Director for Science Barnes Building 495 Summer Street Boston, MA 02210	1
Commanding Officer Office of Naval Research Western Regional Office Attn: Dr. Richard Lau 1030 East Green Street Pasadena, CA 91101	1
U. S. ONR Liaison Office - Far East Attn: Scientific Director APO San Francisco 96503	1
Applied Mathematics Laboratory David Taylor Naval Ship Research and Development Center Attn: Mr. G. H. Gleissner Bethesda, Maryland 20084	1
Commandant of the Marine Coprs (Code AX) Attn: Dr. A. L. Slafkosky Scientific Advisor Washington, DC 20380	1
Navy Library National Space Technology Laboratory Attn: Navy Librarian Bay St. Louis, MS 39522	1
U. S. Army Research Office P.O. Box 12211 Attn: Dr. J. Chandra Research Triangle Park, NC 27706	1
Director National Security Agency Attn: R51, Dr. Maar Fort Meade, MD 20755	1
ATAA-SL, Library U.S. Army TRADOC Systems Analysis Activity Department of the Army White Sands Missile Range, NM 88002	1
ARI Field Unit-USAREUR Attn: Library c/o ODCSPER HQ USAEREUR & 7th Army APO New York 09403	1
Library, Code 1424 Naval Postgraduate School Monterey, CA 93940	1
Technical Information Division Naval Research Laboratory Washington, DC 20375	1
OASD (I&L), Pentagon Attn: Mr. Charles S. Smith Washington, DC 20301	1

Copies

Director
AMSAA
Attn: DRXSY-MP, H. Cohen
Aberdeen Proving Ground, MD 1
21005

Dr. Gerhard Heiche
Naval Air Systems Command
(NAIR 03)
Jefferson Plaza No. 1
Arlington, VA 20360 1

Dr. Barbara Bailar
Associate Director, Statistical
Standards
Bureau of Census
Washington, DC 20233 1

Leon Slavin
Naval Sea Systems Command
(NSEA 05H)
Crystal Mall #4, Rm. 129
Washington, DC 20036 1

B. E. Clark
RR #2, Box 647-B
Graham, NC 27253 1

Naval Underwater Systems Center
Attn: Dr. Derrill J. Bordelon
Code 601
Newport, Rhode Island 02840 1

Naval Coastal Systems Center
Code 741
Attn: Mr. C. M. Bennett
Panama City, FL 32401 1

Naval Electronic Systems Command
(NELEX 612)
Attn: John Schuster
National Center No. 1
Arlington, VA 20360 1

Defense Logistics Studies
Information Exchange
Army Logistics Management Center
Attn: Mr. J. Dowling
Fort Lee, VA 23801 1

Copies

Reliability Analysis Center (RAC)
RADC/RBRAC
Attn: I. L. Krulac
Data Coordinator/
Government Programs
Griffiss AFB, New York 13441 1

Technical Library
Naval Ordnance Station
Indian Head, MD 20640 1

Library
Naval Ocean Systems Center
San Diego, CA 92152 1

Technical Library
Bureau of Naval Personnel
Department of the Navy
Washington, DC 20370 1

Mr. Dan Leonard
Code 8105
Naval Ocean Systems Center
San Diego, CA 92152 1

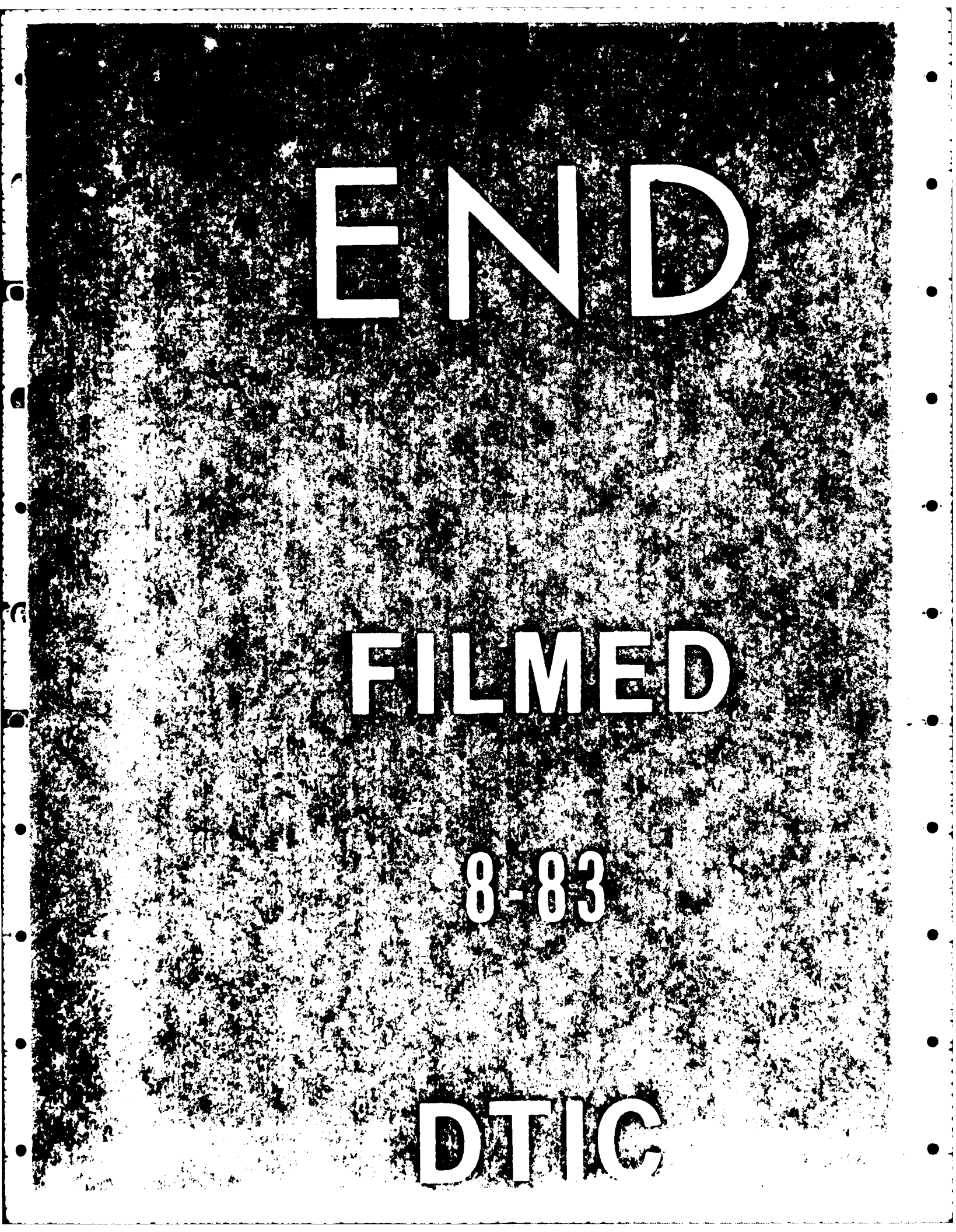
Dr. Alan F. Petty
Code 7930
Naval Research Laboratory
Washington, DC 20375 1

Dr. M. J. Fischer
Defense Communications Agency
Defense Communications Engineering
Center
1860 Wiehle Avenue
Reston, VA 22090 1

Mr. Jim Gates
Code 9211
Fleet Material Support Office
U. S. Navy Supply Center
Mechanicsburg, PA 17055 1

Mr. Ted Tupper
Code M-311C
Military Sealift Command
Department of the Navy
Washington, DC 20390 1

	Copies	Copies
Mr. F. R. Del Priori Code 224 Operational Test and Evaluation Force (OPTEVFOR) Norfolk, VA 23511	1	


END
FILMED
8-83
DTIC