STATISTICAL ASSESSMENT OF A UNIQUE TIME SERIES ANALYSIS TECHNIQUE

Prepared for:
NASA MANNED SPACECRAFT CENTER
Under Contract
NAS 9-11784

Prepared by:
Vernon A. Benignus

BEHAVIORAL TECHNOLOGY CONSULTANTS
Silver Spring, Maryland

January, 1972

INTRODUCTION

It is frequently desirable to detect small changes or shifts of frequency in circadian biological rhythms, especially where there has been some alteration in extrinsic factors which might influence such rhythms. One of the more useful methods used to analyze biological data for the detection and quantification of circadian rhythms is some form of spectrum analysis (Frazier, Rummel, and Lipscomb, 1968). In standard forms of spectrum analysis it is possible to resolve or discriminate between two sinusoidal frequencies separated by Δf where

$$\Delta f = \frac{1}{T} , \qquad [1]$$

and T is the length of time of the time series record being analyzed (Bendat and Piersol, 1966). Among the many problems in biological data acquisition, one of them is that of obtaining records of long duration. This implies that for most circadian rhythm work Δf, the resolution of the analysis program, will be quite large due to short time series records. This report is a preliminary evaluation of a spectrum analysis model which attempts to achieve finer resolution than $\Delta f = 1/T$ by the use of multiple least squares prediction models.

PURPOSE

The specific purposes of this study were to perform empirical tests of a particular least squares multiple prediction program (Rummel, 1966), as well as to relate this particular program to general least squares multiple prediction theory. Empirical evaluation and test of the program involved (1) conversion of the program to run on an IBM 360/44; and replication of test results obtained by NASA, MSC on a Univac 1108 and (2) generation and analysis of Monte Carlo simulated data with the objective of comparison against results theoretically obtainable from such spectrum analysis routines as the FFT (Cooley and Tukey, 1965).

PROCEDURES

General Spectrum Model

The general model for a time series as expressed in the frequency domain is

$$f(t) = K + \sum_{j=1}^{J} \left[a_j \sin(\omega_j t) + b_j \cos(\omega_j t)\right] , \qquad [2]$$

(Bendat and Piersol 1966), where j is the angular frequency index, $0<j\leq J$, j is not necessarily an integer and k is the D.C. component or the mean of the data. The usual approach to the spectrum analysis of f(t) is analogous to discrete Fourier analysis where the coefficients in [2] are estimated by

$$\hat{a}_j = \sum_{t=0}^{T} [f(t)\sin(\omega_j t)] \quad [3]$$

$$\hat{b}_j = \sum_{t=0}^{T} [f(t)\cos(\omega_j t)] \quad , \quad [4]$$

where the carat indicates an estimate. It may be shown that equations [3] and [4] are univariate least squares estimates derived from standard regression theory. These estimates of "real" and "imaginary" amplitude ($\hat{a}_j$ and $\hat{b}_j$ respectively) are usually combined to yield

$$\hat{P}_j = \hat{a}_j^2 + \hat{b}_j^2 \quad , \quad [5]$$

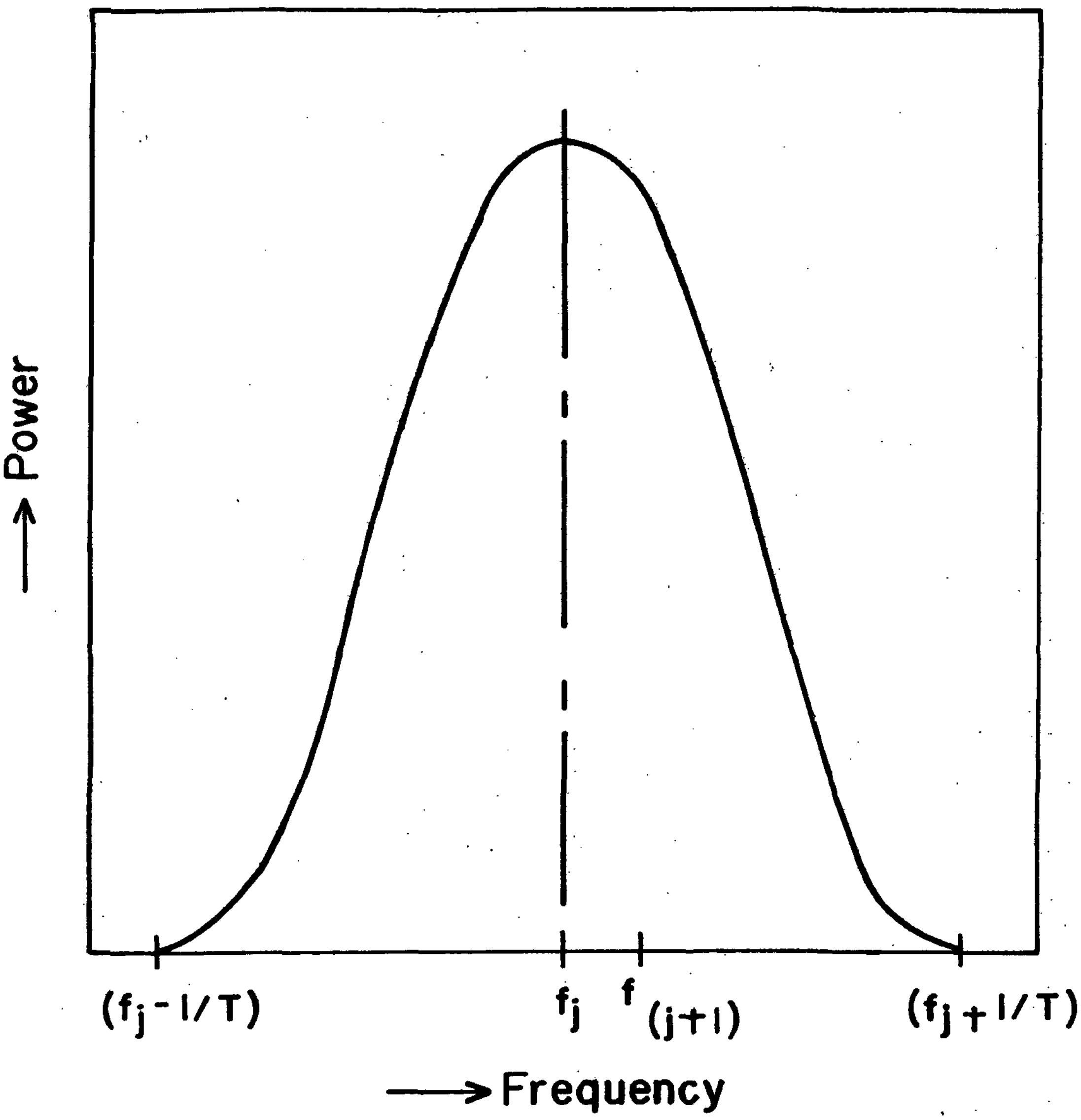
the estimated power in frequency band j or

$$\hat{A}_j = \sqrt{\hat{a}_j^2 + \hat{b}_j^2} \quad , \quad [6]$$

the estimated amplitude in frequency band j (Bendat and Piersol 1966).

It may be shown that two estimates, $\hat{P}_j$ and $\hat{P}_{(j+1)}$ are orthogonal (uncorrelated) if their corresponding frequencies ω_j and $\omega_{(j+1)}$, or of course f_j and $f_{(j+1)}$, are spaced such that $\Delta f = 1/T$. If, in the time series f(t) there exist two signals separated in frequency by much less than Δf, then power estimates at those two frequencies $\hat{P}_j$ and $\hat{P}_{(j+1)}$ can be expressed as a continuous function called a frequency domain or spectrum "window" the main lobe of which is shown in Figure 1. The window function shows that for any estimate, $\hat{P}_j$, if f(t) contains a signal the frequency of which can take on values of $f_j \pm 1/T$ then the value of $\hat{P}_j$ is a function of the true signal frequency. Similarly if two estimates, $\hat{P}_j$ and $\hat{P}_{(j+1)}$ were made

FIGURE 1



APPROXIMATE MAIN LOBE SHAPE OF A SPECTRUM WINDOW at f_j and $f_{(j+1)}$ as shown in Figure 1, and there were two frequencies present in f(t) at f_j and $f_{(j+1)}$, then the two estimates would be nearly equal because data from the signal at f_j are included in $\hat{P}_{(j+1)}$ and vice versa.

Multiple Variable Prediction

Equations [3] and [4] are univariate prediction equations. In usual multiple regression, least squares prediction schemes it is possible to use several predictors simultaneously to estimate the dependent variable. In these cases the several predictors may or may not be correlated. However, when the predictors are highly intercorrelated, estimates of each predictor's contribution are very inaccurate (Draper and Smith, 1966). When k non-independent predictors are used to estimate a dependent variable, the contribution of each is called the "partial regression weight". This regression coefficient is a least squares estimate of the contribution of a given predictor k with the effects of all the other k-1 predictors "accounted for" or "statistically held constant" (Guilford, 1950).

Instead of using univariate predictors such as [3] and [4] to estimate the contribution of a sinusoid of frequency j to f(t), a multiple prediction scheme might be used. In a multiple prediction scheme for estimation of a_j one would not only use a sine wave of frequency j but would include sine waves of several different frequencies in a simultaneous prediction equation. For example if a two predictor scheme were used, then a normalized form of a_j would be estimated using

$$\tilde{a}_j = \frac{R_{dj} - R_{dk}R_{jk}}{1 - R_{jk}} \qquad [7]$$

where the R quantities are Pearson product moment correlation coefficients between the variables indicated in the subscripts and the three variables involved are (1) the dependent variable, f(t) which is identified by the subscript "d"; (2) the sine wave of frequency j, the one whose contribution is estimated as $\tilde{a}_j$ and (3) the sine wave of frequency k, the other simultaneous predictor, the effect of which is to be "controlled" or "accounted" for. Examination of [7] reveals that not only is the relation of a sine wave of frequency j to f(t) considered, R_{dj}, but the relations of the other predictor wave to f(t) and the interpredictor relations are also considered. Thus if $R_{jk} \neq 0$, as it will not be if $\Delta f < 1/T$, then this "overlap" will be considered in estimating a_j. This multiple prediction scheme will hopefully have a higher resolution than equations [3] and [4]. A similar procedure would be used for estimating b_j, the cosine components. When more than two predictors are used in a simultaneous prediction scheme, matrix methods for estimating the contributions of each predictor must be used as shown in [8] and [9]

$$\mathbf{a} = \{ [\sin(\omega t)]^{1} [\sin(\omega t)] \}^{1} [\sin(\omega t)]^{1} f(t) \qquad [8]$$

$$\mathbf{b} = \{ [\cos(\omega t)]^{1} [\cos(\omega t)] \}^{1} [\cos(\omega t)]^{1} f(t) \qquad [9]$$

where **a** and b are vectors of estimates called the real and imaginary amplitude spectra respectively, **sin(ωt)** is a T by J matrix of sines and **cos(ωt)** is a T x J matrix of cosines. When more than two predictor frequencies are used, the contribution of each frequency is made with the contribution of all other included frequencies accounted for.

A Realization of the Multiple Model

The particular program being tested in this study was designed along the lines of a multiple predictor least squares theme as outlined above. The procedure of the program was as follows: (1) compute a spectrum using one frequency at a time as in equations [3] and [4]; (2) examine this spectrum to locate peaks which exceeded a statistical criterion of significance; (3) compute a new spectrum where each frequency's contribution, A_j, was evaluated with the contribution of all other significant peaks held constant by the use of multiple least squares prediction as above; (4) return to step (2) and continue to loop through the procedure until no new peaks are found. In addition to the above procedure, each time step (3) is executed, the frequency value for the significant peaks is moved up and down around the original value and the spectrum is recomputed to guard against the risk of "leakage" from adjacent bands having shifted the original peaks.

Monte Carlo Runs

In order to evaluate the performance of the multiple predictor spectrum analysis program it is necessary to analyze data which approximate that on which the program will be used. Biological signals which are subject to circadian variation can be modeled using a "source of variance" model such as

$$V(\text{Total}) = V(\text{Circadian}) + V(\text{Unaccounted}) \qquad [10]$$

where V(Total) is the total variance (or power) in the wave, V(Circadian) is that portion or component of the wave which is due or correlated with diurnal cycling and $V(\text{Unaccounted})$ is all other variation in the wave. Under the heading of $V(\text{Unaccounted})$ are such sources of variation as short term fluctuations due to stress, homeostatic fluctuation, and in general, any source of variation not related to diurnal cycles. In this discussion $V(\text{Unaccounted})$ will be referred to as either noise or error variance. The general effect of noise in the biological signal is to "mask" the circadian component both with respect to amplitude and frequency. This results in unreliable estimates or variance in the power spectrum since, as with most transformations, the Fourier transformation has as much variance in the resultant as in the original data.

In this paper data were constructed using [10] as a model. The circadian component, $V(\text{Circadian})$, was simulated by generating a sine wave of a particular frequency. The noise, $V(\text{Unaccounted})$, was simulated by a white Gaussian noise generated by sampling from a random number table (Rand Corp., 1955) which was punched onto cards and loaded into a disk file. It is fully realized that biological noise might not have a white spectrum or have a Gaussian distribution. It is true, however, that the assumption of white, Gaussian noise is usually made and it was felt that the program should be evaluated on "fair" theoretical grounds first. Closer approximations to real data can be constructed and tested after the theoretical performance is better understood.

When random noise is involved in data to be analyzed, it is the long-range, average results which are of interest as well as the variation around these averages. The variation around average results is sometimes expressed as variance, error, confidence intervals, failure rates, etc. In order to assess the program's average performance and variance about these averages, a series of records were generated according to [10]. For each of 100 records the noise was obtained by sampling from a unique section of random number table. Each record had a length of 100 observations; the sinusoids which were used as the signals (sine waves) were at period lengths of 23, 24, 26, 27, and 30 samples per cycle. Several SNR (signal to noise ratio) levels were used. Performance on single as well as multiple signal waves was expressed.

Criteria of Performance

These aspects of the performance of the spectrum analysis program were evaluated as follows: (1) finding the correct frequency, (2) finding the correct amplitudes of the components and (3) program failure to find too few or too many peaks. The average performance as well as the variability about the averages was described for (1) and (2) above and a failure probability was computed for (3).

On any given wave the program generated a spectrum which showed significant amplitude peak(s). Due to the noise component there was frequently some error in the frequency of a peak. This error (variability) was described in terms of the relative number of times that the program made various degrees of error. This is the probability of error for a particular degree of error and a graph of error probability vs degrees of error constitutes the probability distribution of frequency errors. Optimally one would want this distribution to be peaked around a mean of zero (high probability of zero error) and to have a narrow width (lower probability of error the greater the degree of error). This probability distribution of frequency errors is analogous to a frequency domain window except that it refers to errors in narrow peaks rather than amplitudes.

On any given analysis, the amplitude of any peak hopefully approximates the correct amplitude of the signal but will frequently be greater or less due to noise. In order to evaluate the variability around the correct amplitude the probability of an amplitude estimate falling into a certain amplitude range or category can be computed. Here a distribution of probabilities can be graphed and it would be desirable for this distribution to be peaked around the correct amplitudes (high probability of finding the correct amplitude) and have a narrow width (lower probabilities of finding amplitudes, the farther the amplitude deviates from the correct amplitude). This is essentially the sampling distribution of amplitudes from which confidence limits would be computed.

RESULTS

Single Frequency Performance

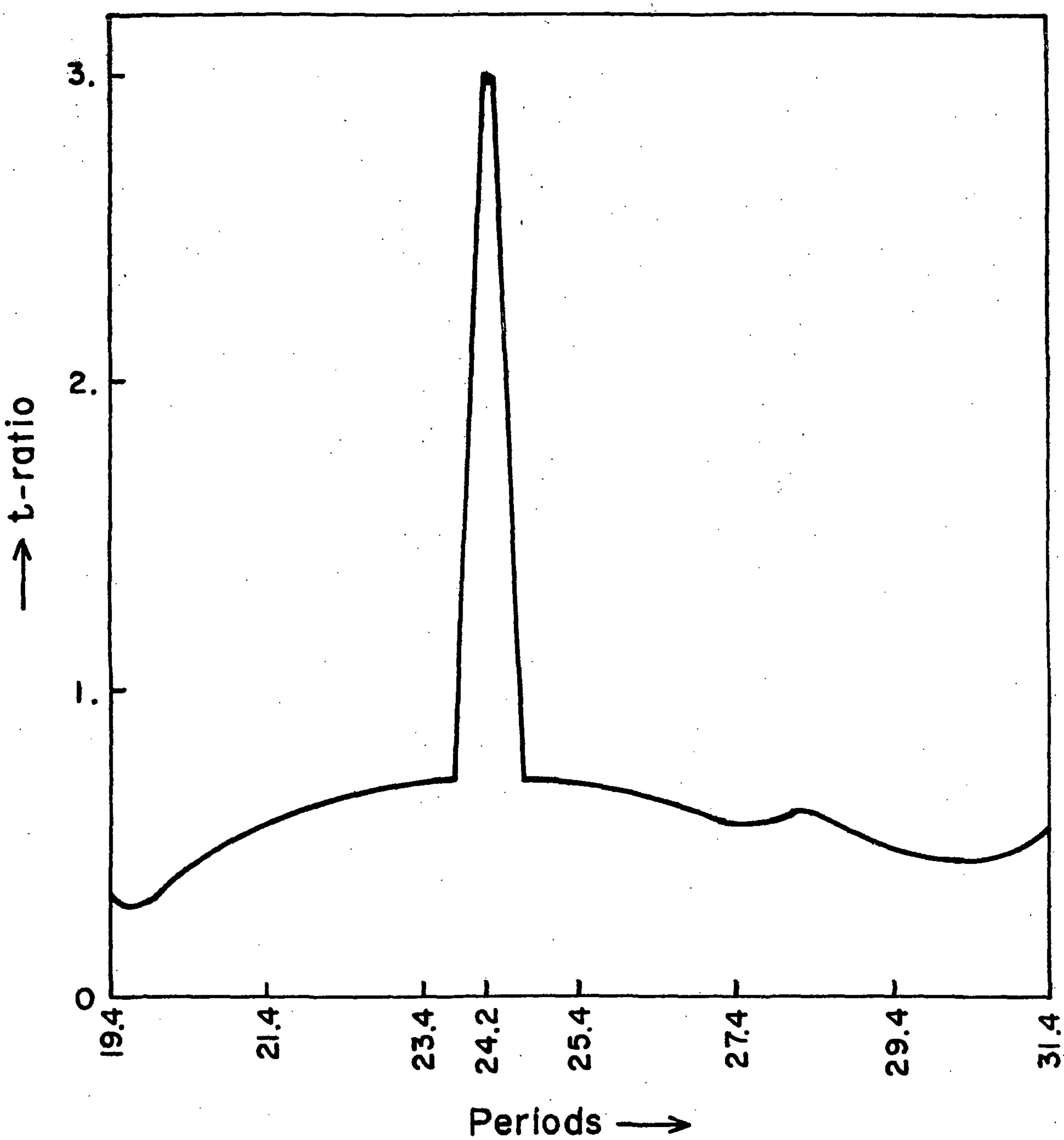
Having verified the fact that the spectrum analysis program was in fact yielding the same results on both the Univac 1108 at NASA/MSC and the IBM 360/44 at Trinity University, the next step of the project was undertaken. This consisted of characterizing the output from the spectrum analysis program for various levels of Gaussian noise, using a single sinusoid as the signal. Statistical criteria of performance were devised and empirically related to SNR.

Three levels of SNR were used as follows: 1/0.5, 1/1 and 1/2. Each time series was 100 points long, having a signal composed of a sine wave, Period=24, and an independently sampled noise. These 100 time series were analyzed using the spectrum analysis subroutine. The spectra resulting from each analysis were stored on disk and an index card for each record was punched so that future access was facilitated.

Two spectrum outputs were stored on disk. The first consisted of the variable called SPER in the program, the t-ratios computed by subroutine PERIOD. The other spectrum was constructed as follows. SPER was tested at each spectral estimate for equal or excess the value of CHEK, the "significance" or reject level. When a value of SPER equalled or exceeded CHEK, the corresponding value of the variable called AMP, the amplitude was stored into the output array. The second output array then consisted of all zeroes except for those spectral estimates whose t-ratio exceeded or equalled CHEK. The first file will be referred to as the t-ratios, the second as the significant amplitudes. These spectra were stored on disk for each of the 100 time series and for each of the three SNRs, thus making a total of 600 spectra. The value of CHEK was 1.96.

The general result of each analysis was as shown in Figure 2. The average baseline level was a function of SNR; the curvature of the baseline was a function of the particular noise sample. Over 100 spectra, the average baseline was virtually flat for all SNRs. The single peak in the spectrum is the estimate of the line element in the time series (the sine wave). The amplitude of this peak varies across spectra and the extent of this variation is a function of SNR. Similarly the single peak occurs at

FIGURE 2



TYPICAL SPECTRAL DISTRIBUTION OF T-RATIOS COMPUTED THROUGH SUBROUTINE PERIOD IN THE ANALYSIS OF A SINUSOIDAL SIGNAL WITH A 24.0 PERIOD SUPERIMPOSED ON A GAUSSIAN NOISE.

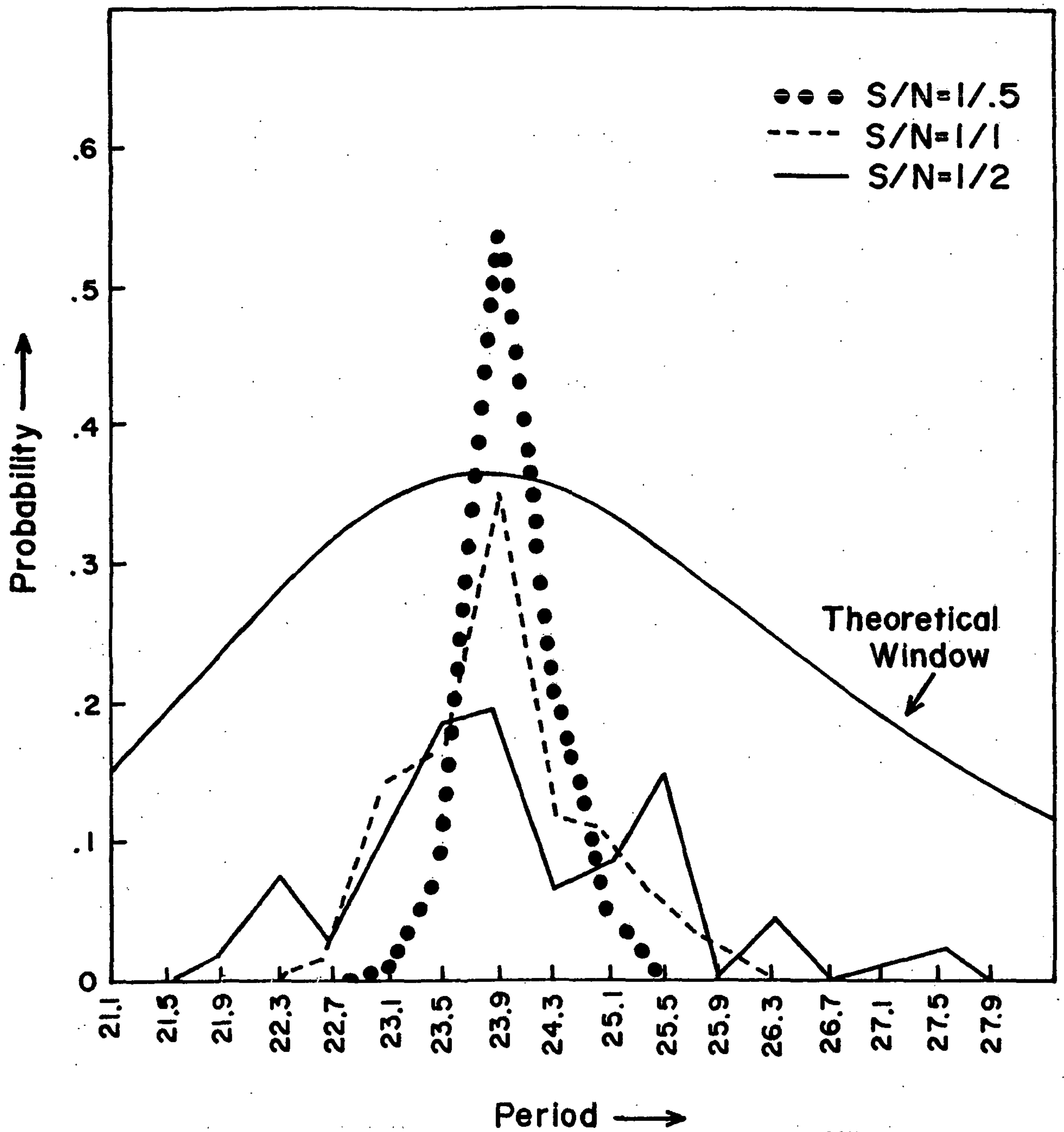
various periods across spectra, not always at Period=24. The extent of this variation is a function of SNR as discussed previously.

Figure 3 shows the plots of the period probabilities for three levels of SNR. Inspection of the figure shows that as proportion of noise increases, so does the variability of the results. Essentially, the more noise, the more widely the period estimates are scattered. The heavy black line shows the shape of the theoretical window for univariate spectrum analysis. It is seen that seven for SNR of 1/2, the most widely deviant result still falls well within the resolution limits of the univariate spectrum window.

Figure 4 shows probability data for the amplitudes of the peak in the spectrum. Again the stability of the estimate decreases as SNR increases. Here, however, the mean also increases. The latter finding is consistent with the notion that the amplitude at the peak is the sum of the sinusoid's amplitude and the amplitude of the noise at that period. There is also noticeable skewness in the amplitude probability distributions.

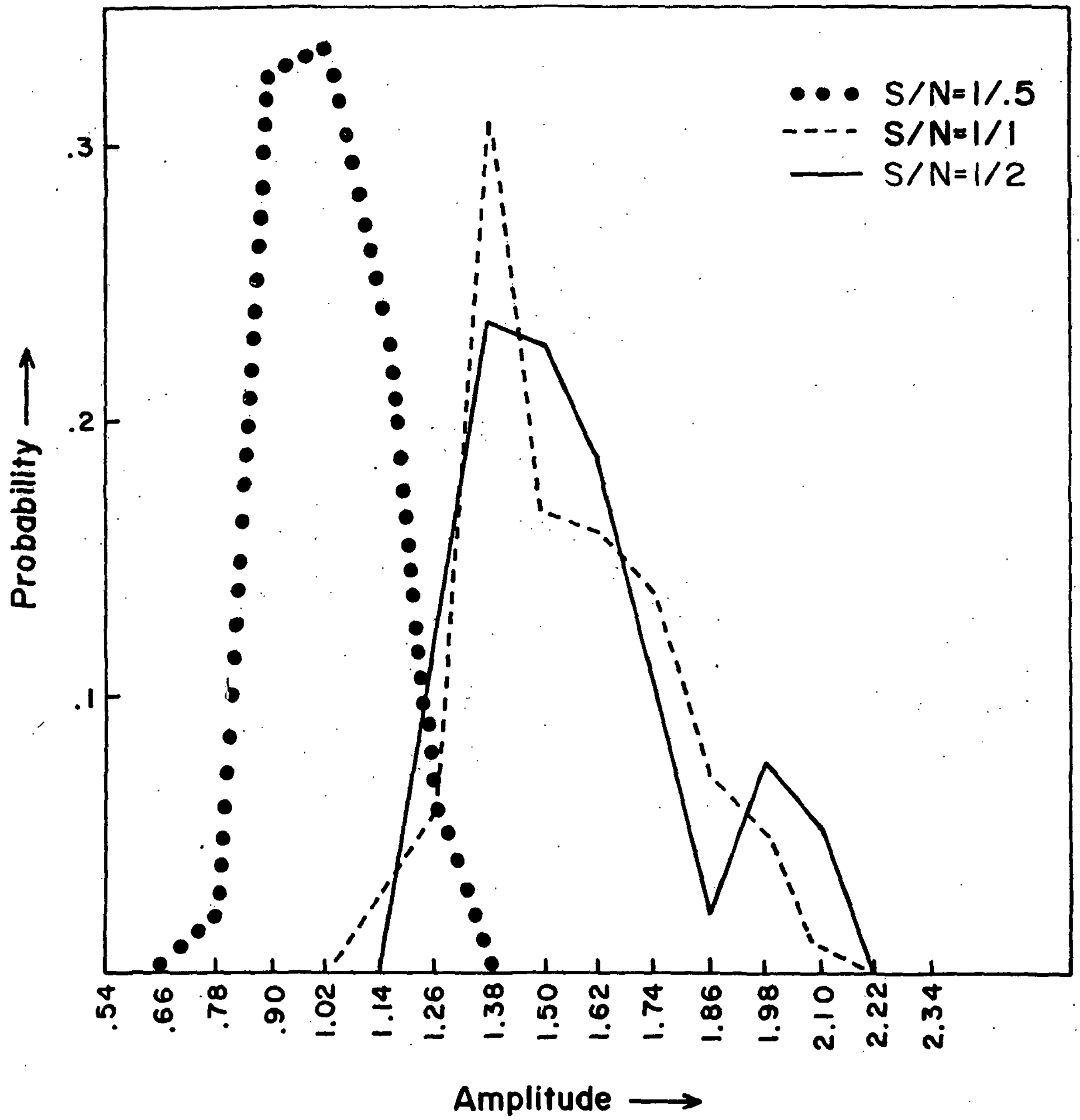
While the probability distributions of the estimates varied in the expected direction, it is not in practical cases known what the SNR figure's value is. An attempt was therefore made to predict the standard deviation of the estimates from the standard error as computed by the spectrum analysis program. Using mean SE values and standard deviations, the curves in Figure 5 were plotted. Lower SE values and lower standard deviation values correspond to lower noise cases. While the relationships are in the expected directions, it is not clear why the "curvature" exists, especially for the standard deviation of the amplitude estimates. The skewness of the probability distributions of amplitudes makes the straight computation of standard

FIGURE 3



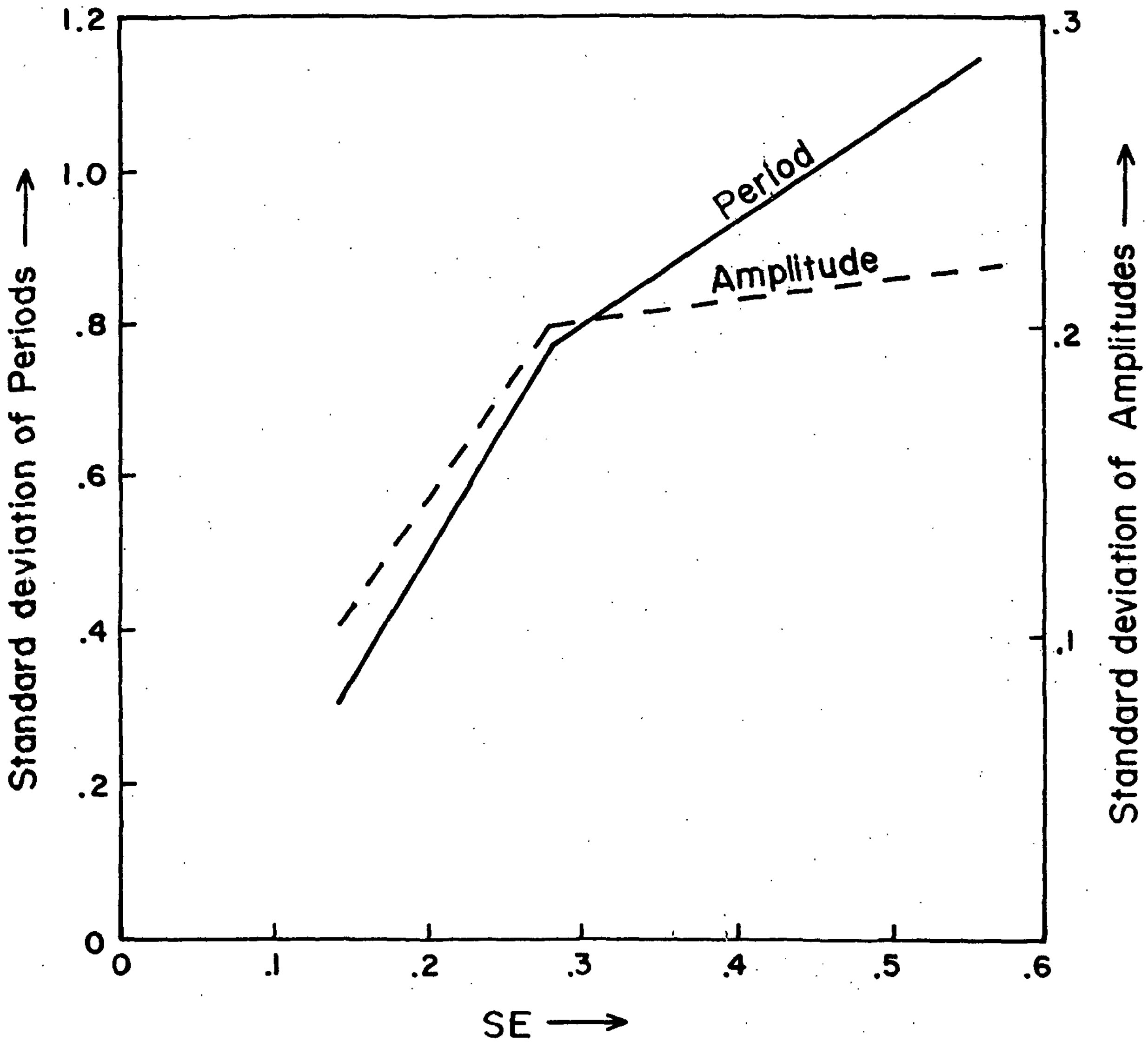
PERIOD PROBABILITIES FOR THREE LEVELS OF SNR

FIGURE 4



PROBABILITY DISTRIBUTION OF PEAK AMPLITUDES AT THREE SNRS

FIGURE 5



RELATIONSHIPS BETWEEN STANDARD DEVIATION ESTIMATES OF PERIOD AND AMPLITUDE AND THE PROGRAM'S STANDARD ERROR ESTIMATES deviations somewhat inappropriate. Transformations on the amplitude data could be investigated to clarify the relationship.

The probability of program failure was calculated for 1/1 signal to noise ratios and for 1/0.5 signal to noise ratios. Table 1 shows the failure probability for various period frequency separations with the 1/1 signal to noise ratio. The left column represents the period separations between the two simulated sinusoids. The middle column represents the frequency separation equivalents of the period separations. The probability of failure at the 1/1 signal to noise ratio is shown in the right column.

PERIOD Separation	FREQ. Separation	P (Failure)
*7	.010	.11
4	.006	.19
3	.005	.46

TABLE 1. Probabilities of failure using a 1/1 signal to noise ratio.

* This is the theoretically resolvable separation.

P (Failure) = Probability of failure.

The theoretically maximal resolution for the data analyzed would correspond to a period separation of 7 units. A separation of 3 period units is twice the expected maximal resolution. It was observed that 11 errors out of 100 estimates were made at the maximal predicted resolution. At twice the expected resolution, the failure rate approached 50 per cent, a prohibitive level.

Most of the errors were of the nature that the output indicated only one frequency peak, usually lying at a position somewhere between the two true frequencies.

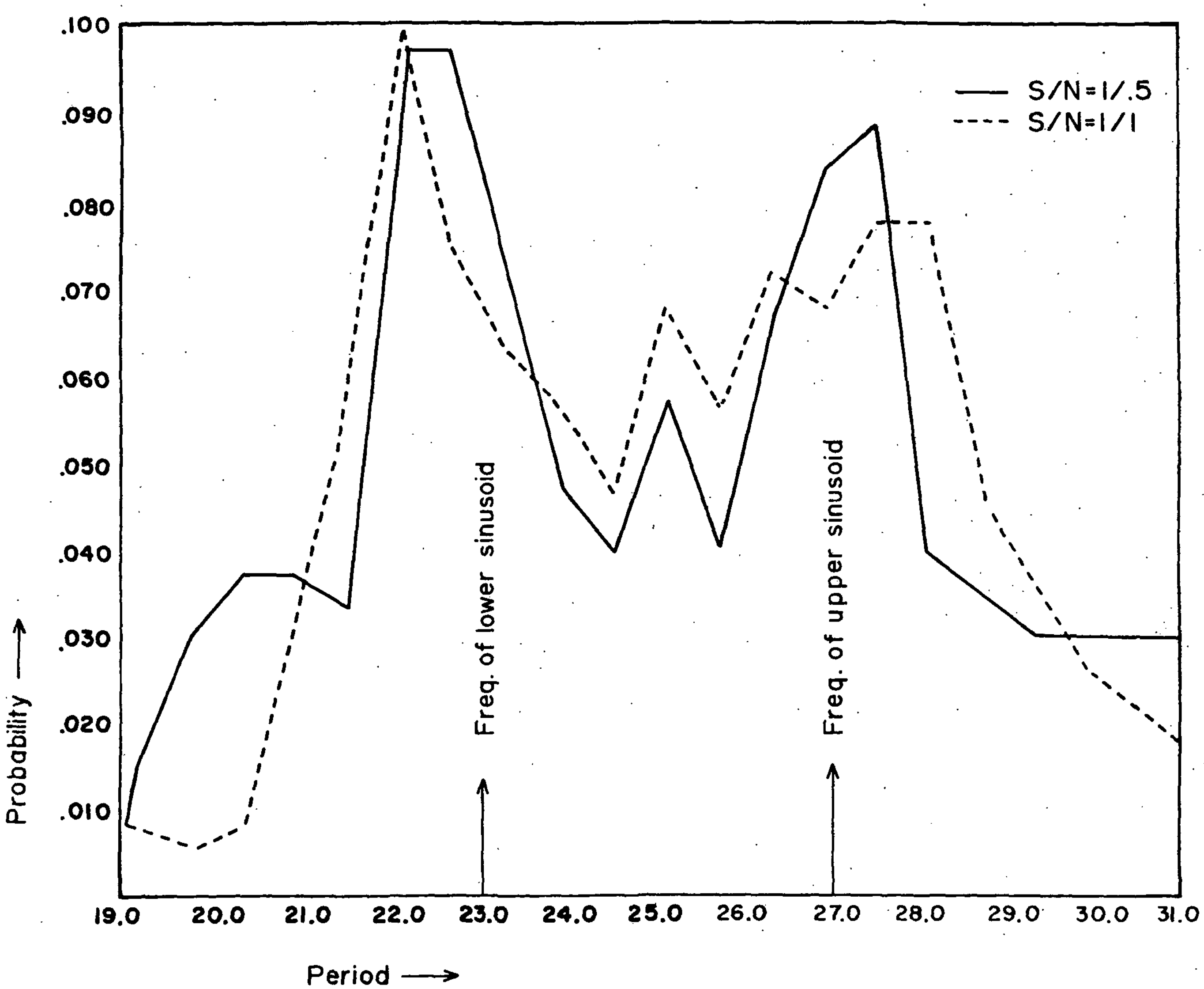
At a signal to noise level of 1/0.5, the failure probability was only .05, or five times per hundred at a period separation of 4 units.

Probability of Peak Detection. Probability distributions were computed to study the distribution of peaks across the various periods comprising the spectrum. Figure 6 shows plots of the obtained results. The two plots shown in this figure refer to period separations of 4 units (not quite twice resolution) with a SNR of 1/0.5 and 1/1.

Examination of Figure 6 shows that there is clearly apparent separation of peaks for the two sinusoid frequencies where the SNR was set at 1/0.5. The degree of separation was clearly less pronounced when signal to noise ratio was changed to 1/1. Probability distributions for separations of twice the expected resolution (3 units) showed even less separated peaks. Even at signal to noise ratios of 1/0.5, the separation between peaks is somewhat "smeared" by error variance across the frequency bands. At ratios of 1/1, the probability of error due to large variability becomes rather large. It would therefore appear that the program has limited utility for signal to noise ratios of poorer than 1/0.5, and even at this value if separation is poor.

The probability distributions for amplitude are not worthy of discussion beyond the mention of the fact that they tended to approximate a chi-square as before. This finding is not unexpected on the basis of distribution theory and previous results.

FIGURE 6



PROBABILITY DISTRIBUTIONS OF TWO-FREQUENCY DISCRIMINATION AT TWO SNRS WITH FOUR-UNIT SEPARATION BETWEEN TEST SIGNALS

Discussion and Conclusions. The single-frequency window performance of the analysis program was sufficiently good that the two-frequency performance was expected to be better than the results revealed. Some factors which are not related to the single frequency window shape are apparently responsible for the "confusion" of adjacent bands. Three possible factors may be implicated as contributors to this problem: (1) problems with the least squares theory as applied to frequency domain data; (2) the statistical method of selecting significant peaks: (3) the algorithm used in sliding periods for best fit; and (4) programming problems which might as yet be undetected. Another possible problem might relate to the fact that when interpredictor correlation is very high, estimation of the contribution of any one predictor becomes less accurate (Draper and Smith, 1966). Inasmuch as one or more of these factors may be involved, questions can be raised as to the generalizability of the analyses reported here with respect to the idea of testing multiple regression analysis as a spectral analysis tool.

If a least squares multiple predictor program were to be created, using all sinusoids in a spectrum as predictors, certain problems such as "significance" and use of peak values as criteria could be eliminated. The procedure of "sliding" periods for optimum fit could also become unnecessary, since all frequencies would be estimated simultaneously. If a "canned" program for multiple regression analysis were to be adapted for testing of the theory, detection of programming errors would be facilitated and all of the standard statistics of regression analysis would be made available. Moreover, techniques such as stepwise regression could be employed for the automatic elimination of predictors (periods).

The procedure of using equally spaced periods to construct a spectrum rather than equally spaced frequencies presents a problem of possible correlation between adjacent frequency/period band estimates. The amount of correlation between adjacent period bands would not be the same at the lower end of the spectrum as at the upper end, since theoretical resolution using estimators equally spaced in frequency is 1/T, where T is the length of the series obtained from measurement. Thus, it would be somewhat easier to interpret broad peaks (high values in more than one adjacent band), when spacing was expressed in constant frequency increments.

Suggestions for Further Research. Considering the various areas of assessment of the least squares analysis over the course of this project, it should be pointed out that the sensitivity of the technique in describing single frequencies clearly exceeded expectations based on time series mathematical theorems. It should also be pointed out that the limitations discovered in its two-frequency discrimination performance should be viewed in the context of its performance with respect to that of other alternative approaches. It is therefore recommended that another series of studies be performed to compare the results obtainable with this least squares analysis program to results obtainable on the same data from alternatives such as general least square multiple regression, standard power spectra computed by the FFT and Halberg's least squares analysis program. Since the assumptions on which much of time series mathematics is based assume "flat" spectral shape, and ultradian spectral distributions are usually not very flat, empirical studies such as those described must be undertaken. Comparisons across the alternative procedures for specifying period, amplitude, and phase information should be made to construct criteria for informed selection among these available alternatives.

It is also suggested that estimates be made of the relative signal to noise ratios to be expected of physiological and behavioral data with respect to the circadian rhythm and other rhythms of interest in the ultradian period range. Since the signal to noise ratio is considered a critical parameter with respect to two-frequency discrimination, the possibility is raised that for some response measures, (with better signal to noise ratios) multiple least squares estimates would be perfectly adequate for unimodal and bimodal spectral distributions.

This same line of research is now seen as basic to the construction of empirical mathematical models of physiological and behavioral functioning, an area in which continuing efforts could yield new knowledge of considerable practical and theoretical significance.

REFERENCES

Bendat, J.S. and Piersol, A.G., Measurement and analysis of random data. New York: John Wiley, 1966.

Cooley, J.W. and Tukey, J.W., An algorithm for the machine calculation of complex Fouriers series. Mathematics of Computation, 1965, 19, 297-301.

Draper, N.R. and Smith, H. Applied Regression Analysis. New York: John Wiley, 1966.

Frazier, T.W., Rummel, J. and Lipscomb, H., Circadian variability in vigilance performance, Aerospace Medicine, 1968, 39, 383-395.

Guilford, J.P. Fundamental Statistics in Psychology and Education. New York: McGraw-Hill, 1950.

RAND Corp., A Million Random Digits with 100,000 Normal Deviates. Glencoe, Ill.: Glencoe Free Press, 1955.

Rummel, John A. A physiological basis for evaluating human circadian rhythms in the extraterestrial environment. Dissertation, Baylor University College of Medicine, Houston, Texas, 1966.