H. DRÄGERT, Göttingen

"Maximum Entropy spectral analysis for geomagnetic variations"

Mittwoch, den 13.03.1974

Introduction

Practically all branches of geophysics presently employ frequency analysis of time-varying phenomena to such an extent that the development of frequency analysis techniques has itself become a research area in geophysics. The value of the representation of geophysical data in the frequency domain stems from the cyclical or resonance properties possessed by many geophysical phenomena, and from the simplification of analytical development in this domain.

To allow the application of spectral analysis to empirical data, certain basic assumptions must be made. Suppose that x(t) is a real, discrete data set with samples equally spaced throughout some range of the parameter t. In order to obtain a spectral estimate for the process sampled by x(t), the assumptions shown in Fig.1 must be made. The first two assumptions are necessary in order that the data sample can be related to a stochastic process with a spectral representation, and that this spectrum may be estimated through time averaging. The validity of these assumptions is usually qualitatively judged by the credibility of the results of the analysis. The last two assumptions allow the autocorrelation function and the spectral density function to be considered as a Fourier Transform Pair, and they assure us that the process has finite and non-zero power at all frequencies.

1. Conventional Spectral Analysis

Most of the conventional computational methods for estimating the power spectral density are based on one of two general approaches:

(1.) the AUTOCORRELATION METHOD

(2.) the PERIODOGRAM METHOD

With reference to the summary of equations shown in Fig.2, assume that the finite sample x(t) is sampled at intervals $\Delta t$

## BASIC ASSUMPTIONS FOR SPECTRAL ANALYSIS

1.) The set x(t) is a finite subset of some infinite set y(t) which constitutes one realization of the weakly stationary random process $Y_t$.

2.) $Y_t$ is ergodic with respect to the mean and with respect to the autocorrelation.

3.) $Y_t$ has an absolutely continuous spectral distribution given by:

$$S(f) = \int_{-\infty}^{f} s(f)\, df$$

where s(f) = the spectral density of $Y_t$

4.) The PALEY-WIENER CRITERION is satisfied:

$$\int_{-\infty}^{+\infty} |\log s(f)| df < \infty$$

Fig.1: A summary of basic assumptions used in the spectral analysis of time series.

## AUTOCORRELATION APPROACH

Autocorrelation Estimate:

$$R_x(r\Delta t) = \frac{1}{N-r} \sum_{n=1}^{N-r} x_n x_{n+r} \qquad r = 0,1,2,\ldots m \qquad (1.1)$$

Raw Power Spectral Density Estimate:

$$G_x(f) = 2\Delta t \left\{ R_o + 2 \sum_{r=1}^{m-1} R_r \cos\left(\frac{\pi r f}{f_c}\right) + R_m \cos\left(\frac{\pi m f}{f_c}\right) \right\} \qquad (1.2)$$

Example of Tapering Window:

$$D_r(r\Delta t) = \frac{1}{2}\left(1 + \cos\frac{\pi r}{m}\right) \qquad r = 0,1,2,\ldots m$$
$$= 0 \qquad r > m \qquad (1.3)$$

Smoothed Power Spectral Density Estimate:

$$\bar{G}_x(f) = 2\Delta t \left\{ R_o D_o + 2 \sum_{r=1}^{m-1} D_r R_r \cos\left(\frac{\pi r f}{f_c}\right) \right\} \qquad (1.4)$$

## PERIODOGRAM APPROACH

Fourier Coefficients:

$$A_n = \frac{1}{N} \sum_{k=0}^{N-1} x_k \, e^{-2\pi i k n/T} \qquad n = 0,1,2,\ldots N-1 \qquad (2.1)$$
$$T = N\Delta t$$

Raw Power Spectral Density Estimate:

$$I_x(\omega_n) = \frac{N}{2\pi} |A_n|^2 \qquad \text{where} \quad \omega_n = \frac{2\pi n}{N} \qquad (2.2)$$
$$n = 0,1,2,\ldots N/2$$

Example of Smoothing Window:

$$W(\omega_n, M) = \frac{3M}{4\pi} \left\{ \frac{\sin(\omega_n M/4)}{\omega_n M/4} \right\}^4 \qquad (2.3)$$

Smoothed Power Spectral Density Estimate:

$$S_x(\omega_n) = \sum_{m=-2T/M}^{m=+2T/M} W(\omega_m, M) \cdot I_x(\omega_{n-m}) \qquad (2.4)$$

Fig.2: Summary of equations used in conventional spectral analysis.

giving N sample points over a record length T. Without loss of generality, a mean of zero can be assumed; hence, the autocorrelation estimate at a displacement lag of $r\Delta t$ is given by equation (1.1). Here, r represents the lag number, and m is the maximum lag number. The raw estimate of the power spectral density, $G_x(f)$, defined for a frequency band $(0,f_c)$, is given by the cosine transform of the autocorrelations as shown by equation (1.2). Here, $f_c$ is the Nyquist or cut-off frequency, and $G_x(f)$ is evaluated at the discrete frequencies $f=kf_c/m$.

It should be noted that in applying equations (1.1) and (1.2), the infinite realization y(t) has been in effect redefined outside of the sample x(t); i.e.

$$y(t) = x(t) \quad \text{for } 0 \leq t \leq T$$
$$= 0 \quad \text{otherwise}$$

It is often this unrealistic zero-extension which is a major source of error in the spectral estimate (1.2).

Furthermore, to increase stability and reduce the variance of the spectral estimate, the autocorrelation estimates are often tapered by a lag window such as, for example, the Hann Lag Window defined by equation (1.3). The resulting smoothed power spectral density estimate is given by equation (1.4). It is obvious that the window $D_r$ is independent of the properties of the data; therefore, $\overline{G}_x(f)$ tends to give not an estimate of the true spectrum, but an estimate of a spectrum biased by the convolution of the transform of $D_r$.

In the periodogram approach, the Fourier coefficients, $A_n$, for the discrete time series x(t) can be calculated by use of equation (2.1). Here, the data sample length T is considered to be the fundamental data period. The periodogram or raw power spectral density estimate is then calculated by relation (2.2). This raw estimate, $I_x$, is defined over the frequency band $(-f_c, +f_c)$, and evaluated at the discrete frequencies $\omega_n=2\pi n/N$, where $n=0,1,2,\ldots N/2$. Again it should be noted that through the use of equations (2.1) and (2.2), the infinite realization y(t) has been redefined outside the given sample x(t); i.e. y(t) is defined to be perfectly periodic in terms

of x(t). Needless to say, such a periodic extension of the data can insert periodicities in the spectral estimate which do not exist in the data.

As before, to ensure that the variance of the periodogram estimates converge with increasing N, periodograms must be averaged for different data samples, or equivalently, $I_x$ must be convolved with a smoothing window in the frequency domain. As an example, equation (2.3) defines a Parzen Window, which, when applied to $I_x$, results in the smoothed power spectral density, $S_x$, given by equation (2.4). Again, such a smoothing operation will bias the spectral estimate.

## 2. The Maximum Entropy Approach

It is well known that the conventional methods of spectral estimation result in questionable accuracy whenever periods analyzed are comparable to data sample lengths. Not only do the necessary assumptions of either data-length periodicity or zero-extension of data become suspect, but also the window functions required to obtain a stable spectral estimate from the finite data sets introduce spectral shifts and spectral leakage, as well as smearing resolution. Consequently, whenever only limited data are available, or, whenever the stationarity of the data is limited to a time interval of the order of the periods of interest, a different analysis method is required.

A newer method appropriate to the analysis of short records, especially those containing narrow spectral peaks, is the MAXIMUM ENTROPY METHOD (MEM), first proposed by BURG (1967). In general terms, it may be described as a data-adaptive or data-regression method in that it assumes neither a null nor a periodic extension of the data, but adapts itself to the sample of the process being analyzed in such a way that the spectral estimate displays maximum entropy or maximum information content for the sample, while still fully agreeing with the available data.

The rationale behind this approach can also be shown by the following argument (McGEE,1969): The finite sample x(t)

can be considered to be a partial realization of an infinite number of random processes, of which "zero-extension of x(t)" and "periodic extension of x(t)" are naturally two possibilities. Since there is no prior knowledge of the amount of information contained in x(t), the most reasonable choice for the random process is that process whose spectral estimate displays as much information as is fully consistent with the data and no more.

## 3. A Measure of Information

To allow an analytical formalation of this maximum-entropy or maximum-information condition, a measure of information for a stochastic process is required. An intuitive definition can be illustrated by the following argument. Consider the occurrence of the event:

" Income tax will be increased next year."

This event is almost certain with a probability close to 1; hence, the statement conveys very little information. However, the following event:

" Income tax will be abolished next year."

is an almost impossible event with a probability of close to 0; hence, this statement conveys a great deal of information. Consequently, the following relationship between the probability of an event and its information content is suggested:

$$\text{INFORMATION} \propto -\log(\text{PROBABILITY})$$

The same relationship can be deduced from the communication theory definition for information, which states that the information contained in a message is proportional to the time required to transmit the message. For example, consider the binary encoding of equiprobable messages as outlined in Table 1. For two equiprobable messages, $a_1$ and $a_2$, only one binary digit is required to convey a message. For n equiprobable messages, $\log_2 n$ binary digits per message are required. Since the time required to transmit such messages is directly proportional to the number of digits defining the message, it follows that the information transmitted is proportional to $\log_2 n$; i.e.

$$I \propto \log_2 n$$

| EQUIPROBABLE MESSAGES | BINARY ENCODING | BINARY DIGITS PER MESSAGE |
|---|---|---|
| $(a_1,a_2)$ | 0,1 | 1 |
| $(a_1,a_2,a_3,a_4)$ | 00,01,10,11 | 2 |
| $(a_1,a_2,\ldots a_8)$ | 000,001,...111 | 3 |
| . | . | . |
| . | . | . |
| . | . | . |
| $(a_1,a_2,\ldots a_n)$ | 00..0,00..1,etc. | $\log_2 n$ |

Table 1: Binary encoding of equiprobable messages.

But since the probability of each message is given by $P=1/n$, a proportionality similar to the previously intuitively derived one results; that is,

$$I \propto -\log_2(P)$$

The outlined argument may be generalized for any arbitrary logarithmic base and for messages with arbitrary probabilities.

When not all events are equally probable, a weighted average is used to express the average information per message. Consider a source $\Sigma$ emitting n messages $\{s_1,s_2,\ldots s_n\}$ randomly, each message with its respective probability $P_1,P_2,\ldots P_n$. The information associated with the message $s_j$ is given by

$$I_j = -k \log P_j \qquad (3.0)$$

where the proportionality constant k includes the factor for the change in the logarithmic base. The average information emitted by the source $\Sigma$ is given by

$$H(\Sigma) = \sum_{j=1}^{n} P_j I_j = -k \sum_{j=1}^{n} P_j \log P_j \qquad (3.1)$$

The quantity H given by (3.1) is defined as the ENTROPY of the source $\Sigma$. Obviously, the function H can also be associated with a stochastic process $Y_t$ having the n realizations $y_1,\ldots y_n$ and the joint probability density $p(y_1,y_2,\ldots y_n)$.

The resulting general expression for the entropy of $Y_t$ is given by equation (3.2).

$$H(Y_t) = -k \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} p(y_1,\ldots y_n)\cdot\log\{p(y_1,\ldots y_n)\}dy_1\ldots dy_n \quad (3.2)$$

In the case that $Y_t$ is a process of infinite duration, the absolute entropy will diverge; hence, the ENTROPY DENSITY or ENTROPY RATE, defined by equation (3.3), is generally used as a measure of information.

$$h(Y_t) = \lim_{n\to\infty} \frac{H(Y_t)}{n+1} \quad (3.3)$$

Furthermore, in spectral estimation, the analysis is usually limited to second-order statistics. This implies that it is not possible to distinguish a given time series from one characterized only by its first- and second-order statistics, namely a GAUSSIAN PROCESS. It can be shown (SMYLIE et al.,1973) that the entropy rate for a zero-mean, stationary Gaussian process can be expressed in terms of the spectral density function or the autocorrelation functions as follows:

$$h = \frac{1}{2}\log(2f_c) + \frac{1}{4f_c}\int_{-f_c}^{+f_c} \log\{S(f)\}df \quad (3.4)$$

$$\text{or,}\quad h = \frac{1}{4f_c}\int_{-f_c}^{+f_c} \log\left\{\sum_{k=-\infty}^{\infty} \phi(k)\cdot e^{-i2\pi fk\Delta t}\right\}df$$

where $f_c = 1/2\Delta t$, the cut-off frequency

$S(f) = \Delta t \sum_{k=-\infty}^{\infty} \phi(k)\cdot e^{-i2\pi fk\Delta t}$, the spectral density

$\phi(k)$ = the autocorrelation function at lag k

## 4. The Burg Maximum Entropy Method

With reference to equation (3.4), the maximum entropy condition is now formulated as follows. For finite data samples, only a small number (N) of autocorrelations can be estimated reliably. However, for an accurate estimate of the spectral density, a much larger number of autocorrelations are required. As mentioned before, the standard methods of spectral analysis assume a zero-extension or a periodic extension for the autocorrelation functions, resulting in limited resolution or in spurious periodicities respectively. A more reasonable choice of the unknown autocorrelations is the one which adds no information or entropy; i.e., the entropy rate evaluated from the known data must not be increased by the added autocorrelations, but remain as a maximum entropy estimate.

This condition is expressed analytically by

$$\frac{\partial h}{\partial \phi(k)} = 0 \ , \quad |k| \geq N+1$$

and leads to the equation

$$\int_{-f_c}^{+f_c} \frac{e^{-2i\pi fk\Delta t}}{S(f)}\, df = 0 \ , \quad |k| \geq N+1 \tag{4.1}$$

Naturally, a secondary condition is that the spectral density evaluated from this extended set of autocorrelations must be consistent with the known autocorrelations.

BURG (1967) has shown that the spectral density which gives an entropy rate that is stationary with respect to the unknown autocorrelations and which is also consistent with the N known autocorrelations is given by

$$S(f) = \frac{P_{N+1}}{2f_c \left| 1 + \sum_{j=1}^{N} \gamma_j \cdot e^{-i2\pi fj\Delta t} \right|^2} \tag{4.2}$$

where $P_{N+1}$ and $\gamma_j$ satisfy the N+1 normal equations:

$$\begin{pmatrix} \phi(0) & \phi(-1) & \cdots & \phi(-N) \\ \phi(1) & \phi(0) & & \phi(1-N) \\ \vdots & & & \vdots \\ \phi(N) & \phi(N-1) & \cdots & \phi(0) \end{pmatrix} \begin{pmatrix} 1 \\ \gamma_1 \\ \vdots \\ \gamma_N \end{pmatrix} = \begin{pmatrix} P_{N+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{4.3}$$

From Wiener filter theory, this system of N+1 equations can be recognized as the system of normal equations generating the (N+1)-point optimum PREDICTION-ERROR FILTER, given here by

$$\Gamma = \{ 1, \gamma_1, \gamma_2 \cdots, \gamma_N \}$$

Correspondingly, $P_{N+1}$ is the power of the error series produced by the convolution of $\Gamma$ with the data series:

$$P_{N+1} = \sum_{k=0}^{N} \gamma_k \phi(-k) \qquad (4.4)$$

From equation (4.2), it can be seen that the determination of the maximum entropy spectral density has been reduced to the calculation of the prediction-error filter and the corresponding error power; that is, the solution of system (4.3).

## 5. The Burg Maximum Entropy Algorithm

BURG (1968) has also demonstrated a recursive algorithm for solving the prediction-error system, an outline of which is given below.

(i) For N=0, the trivial case of the one-point filter, the normal system of equations reduces to equation (5.1), where the estimate for the zero-lag autocorrelation is given by the usual summation of squares.

$$\phi(0) = P_1 \qquad (5.1)$$

$$\text{where} \quad \phi(0) = \frac{1}{M} \sum_{j=1}^{M} x_j^2 \,.$$

(ii) For N=1, the two-point filter, the normal system reduces to equation (5.2).

$$\begin{vmatrix} \phi(0) & \phi(-1) \\ \phi(1) & \phi(0) \end{vmatrix} \begin{vmatrix} 1 \\ \gamma_{11} \end{vmatrix} = \begin{vmatrix} P_2 \\ 0 \end{vmatrix} \qquad (5.2)$$

Note that the second subscript for $\gamma$ is used to show the number of the recursion. The Levinson Algorithm (WIGGINS and ROBINSON, 1965), which generates the (K+1)-point prediction-error filter from the K-point filter, is then applied to equation (5.2)

which becomes

$$\begin{vmatrix} \phi(0) & \phi(-1) \\ \phi(1) & \phi(0) \end{vmatrix} \cdot \left| \begin{vmatrix} 1 \\ 0 \end{vmatrix} + \gamma_{11} \begin{vmatrix} 0 \\ 1 \end{vmatrix} \right| = \left| \begin{vmatrix} P_1 \\ \Delta_1 \end{vmatrix} + \gamma_{11} \begin{vmatrix} \Delta_1 \\ P_1 \end{vmatrix} \right| = \begin{vmatrix} P_2 \\ 0 \end{vmatrix} \qquad (5.3)$$

where $P_2 = P_1 + \gamma_{11}\Delta_1$

$\Delta_1 = -\gamma_{11}P_1$

The application of the Levinson Algorithm ensures that the (K+1)-estimate is consistent with the K-estimate, and that the correlation matrix remains non-negative definite. To solve (5.3), the coefficient $\gamma_{11}$ is estimated by minimizing the mean output power of the two-point filter with respect to $\gamma_{11}$. The mean output power is determined by running the filter over the data in both a forward and a backward direction; i.e., $\gamma_{11}$ is determined from the equation

$$\frac{\partial}{\partial \gamma_{11}} \left\{ \sum_{j=1}^{M-1} (x_j + \gamma_{11} x_{j+1})^2 + \sum_{j=1}^{M-1} (x_{j+1} + \gamma_{11} x_j)^2 \right\} = 0 \qquad (5.4)$$

The resulting value of $\gamma_{11}$ is then used in (5.3) to evaluate the autocorrelation at lag 1 and the new error power $P_2$, which yields the expressions

$$\phi(1) = -\gamma_{11}\phi(0)$$

$$P_2 = (1-\gamma_{11}^2)P_1$$

(iii) For N=2, the three-point filter, the normal equations are given by

$$\begin{vmatrix} \phi(0) & \phi(-1) & \phi(-2) \\ \phi(1) & \phi(0) & \phi(-1) \\ \phi(2) & \phi(1) & \phi(0) \end{vmatrix} \begin{vmatrix} 1 \\ \gamma_{12} \\ \gamma_{22} \end{vmatrix} = \begin{vmatrix} P_3 \\ 0 \\ 0 \end{vmatrix} \qquad (5.5)$$

Again, to ensure consistency with the previous estimates and to maintain a non-negative definite correlation matrix, the Levinson Algorithm is applied and (5.5) becomes

$$\begin{vmatrix} \phi(0) & \phi(-1) & \phi(-2) \\ \phi(1) & \phi(0) & \phi(-1) \\ \phi(2) & \phi(1) & \phi(0) \end{vmatrix} \cdot \left| \begin{vmatrix} 1 \\ \gamma_{11} \\ 0 \end{vmatrix} + \gamma_{22} \begin{vmatrix} 0 \\ \gamma_{11} \\ 1 \end{vmatrix} \right| = \left| \begin{vmatrix} P_2 \\ 0 \\ \Delta_2 \end{vmatrix} + \gamma_{22} \begin{vmatrix} \Delta_2 \\ 0 \\ P_2 \end{vmatrix} \right| = \begin{vmatrix} P_3 \\ 0 \\ 0 \end{vmatrix} \qquad (5.6)$$

where $P_3 = P_2 + \gamma_{22}\Delta_2$

$\Delta_2 = -\gamma_{22}P_2$

$\gamma_{12} = \gamma_{11} + \gamma_{22}\gamma_{11}$

As before, the new mean output power of the error series for the filter $(1,\gamma_{11}+\gamma_{22}\gamma_{11},\gamma_{22})$ is determined by forward and backward convolution with the data, and a value for $\gamma_{22}$ minimizing the error power is calculated. Using equation (5.6), $\phi(2)$ and $P_3$ can then be evaluated as follows:

$$\phi(2) = -\gamma_{12}\phi(1) - \gamma_{22}\phi(0)$$

$$P_3 = (1 - \gamma_{22}^2)P_2$$

This procedure generalizes, and after N+1 recursions yields the (N+1)-point prediction-error filter, the error power, and the corresponding autocorrelation matrix, ALL BASED ON ONLY THE AVAILABLE DATA.

## 6. The Maximum Entropy Prediction Method (MEP)

From equation (4.2), it is obvious that a spectral density estimate can be evaluated directly from the transform of the prediction-error filter. However, such an estimate will reveal no phase information. To provide both amplitude and phase information, a simple extension of the maximum entropy method suggested by ULRYCH et al. (1973) can be utilized.

From the prediction-error filter $\Gamma = \{1,\gamma_1,\gamma_2,\ldots\gamma_N\}$, a unit-distance prediction filter $A = \{\alpha_1,\alpha_2,\ldots\alpha_N\}$ can be easily determined since $\alpha_j = -\gamma_j$ for $j=1,2,\ldots N$. The N-point prediction filter A can then be convolved with the last N points of the original M-point data series $x_t$ in order to obtain $x_{M+1}$. Furthermore, Ulrych has shown that the prediction filter for prediction distance L may be derived from the prediction filter for prediction distance L-1. This means that the point $x_{M+L}$ can be obtained by L successive applications of the unit-distance prediction filter A, and a stepwise-extended series, $\hat{x}_t$, as shown in the expression (6.1) can be generated.

$$\{x_1, x_2, \ldots x_M, \sum_{s=1}^{N} x_{M+1-s}\alpha_s, \sum_{s=1}^{N} x_{M+2-s}\alpha_s, \ldots \sum_{s=1}^{N} x_{M+L-s}\alpha_s \} \qquad (6.1)$$

where M = original data length
N = prediction filter length
L = extension length

From (6.1) it can be seen that each successive prediction incorporates the previously predicted points. Furthermore, the original series $x_t$ is usually predicted in a forward and a backward direction, resulting in an extended series length of M+2L. By applying conventional Fourier transform techniques to this extended series, the excellent resolution of the maximum entropy method is almost fully retained, and relative phase information can be recovered through the calculation of standard cross-spectra or cross-correlations.

## 7. Examples of Application

### (i) Comparison of Techniques using Synthetic Data

Fig.3 shows a comparison of spectral estimates for a 48-point sample of synthetic data having a 10% white noise level. The input signal contains three distinct periods which are given in terms of the sample length T, and which are such that the sample T does not contain an integral number of cycles. Furthermore, the relative phases are arbitrary , and hence the sample does not conveniently start and end at zero.

The periodic extension analysis yields three significant harmonics whose relative amplitudes and phases are inexact . since the input frequencies are not equivalent to the fixed sample harmonics of k/T. The zero-extension or transient analysis yields a smoothed amplitude spectrum which does not resolve the two higher frequencies. The Maximum Entropy Prediction (MEP) spectral estimate resolves all three components with minimal frequency shifts and reasonable relative phases and amplitudes. (The MEP spectral estimate illustrated in Fig.3 was generated using a 24-point prediction-error filter, an extension length of four in a forward and backward direction, and a subsequent periodogram of the extended series by means of a Fast Fourier Transform.)

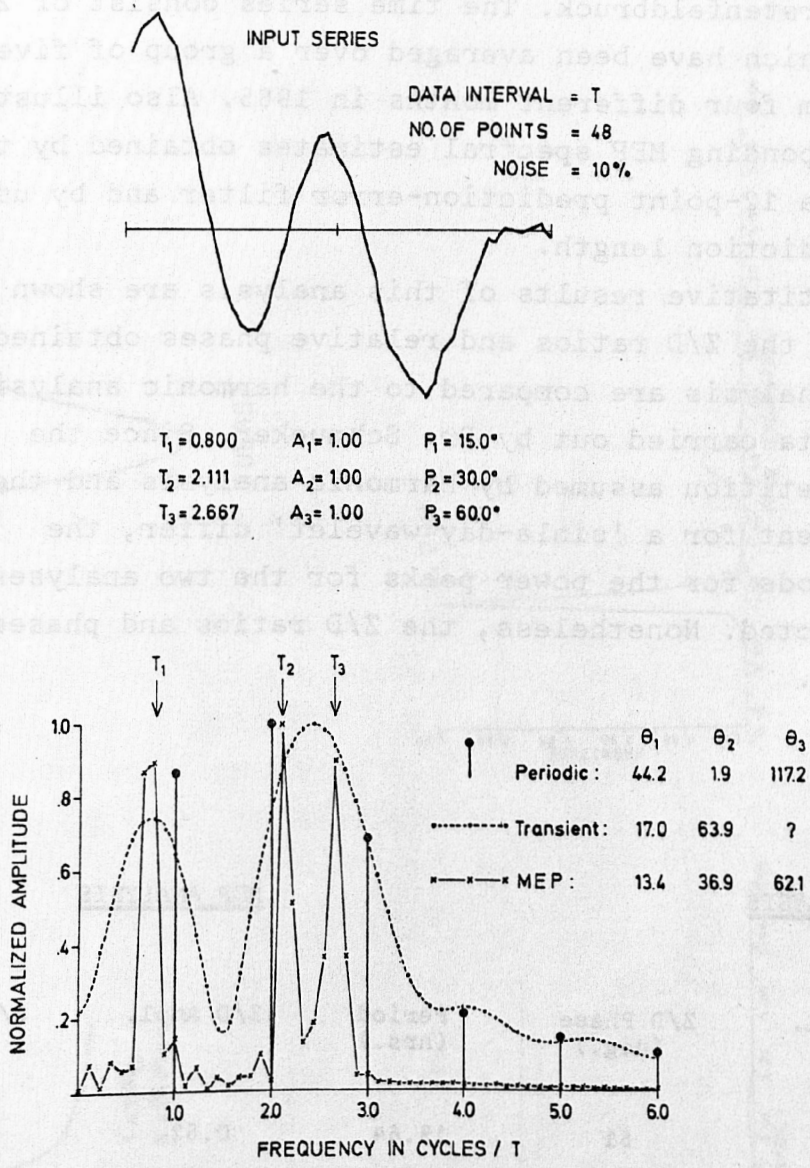


Fig.3: A comparison of amplitude and phase recovery for a synthetic data sample using conventional spectral estimates and the maximum entropy prediction method.

(ii) Application to SQ Data

Fig.4 shows averaged SQ data for the D and Z components measured at Fürstenfeldbruck. The time series consist of 24 hourly means which have been averaged over a group of five quiet days from four different months in 1965. Also illustrated are the corresponding MEP spectral estimates obtained by the generation of a 12-point prediction-error filter and by using a four-day prediction length.

The quantitative results of this analysis are shown in Table 2, where the Z/D ratios and relative phases obtained from the MEP analysis are compared to the harmonic analysis of the same data carried out by Dr. Schmucker. Since the periods of repetition assumed by harmonic analysis and the frequency content for a 'sinle-day wavelet' differ, the different periods for the power peaks for the two analyses are to be expected. Nonetheless, the Z/D ratios and phases are comparable.

| HARMONIC ANALYSIS | | | MEP ANALYSIS | | |
|---|---|---|---|---|---|
| Period (hrs.) | Z/D Ampl. | Z/D Phase (deg.) | Period (hrs.) | Z/D Ampl. | Z/D Phase (deg.) |
| 24.0 | 0.41 | 51 | 19.64 | 0.52 | 55 |
| 12.0 | 0.39 | 54 | 9.39 | 0.41 | 59 |
| 8.0 | 0.39 | 48 | 6.17 | 0.48 | 69 |
| 6.0 | 0.46 | 52 | | | |

Table 2: Comparison of the Z/D relationships for SQ data at Fürstenfeldbruck using harmonic analysis and MEP analysis.

It should be pointed out that this sample analysis of SQ data using the MEP method was simply carried out for interest's sake, and perhaps to show that reasonable results are obtained. Because of the A PRIORI knowledge of the periodicity of SQ data, harmonic analysis is usually the better approach.

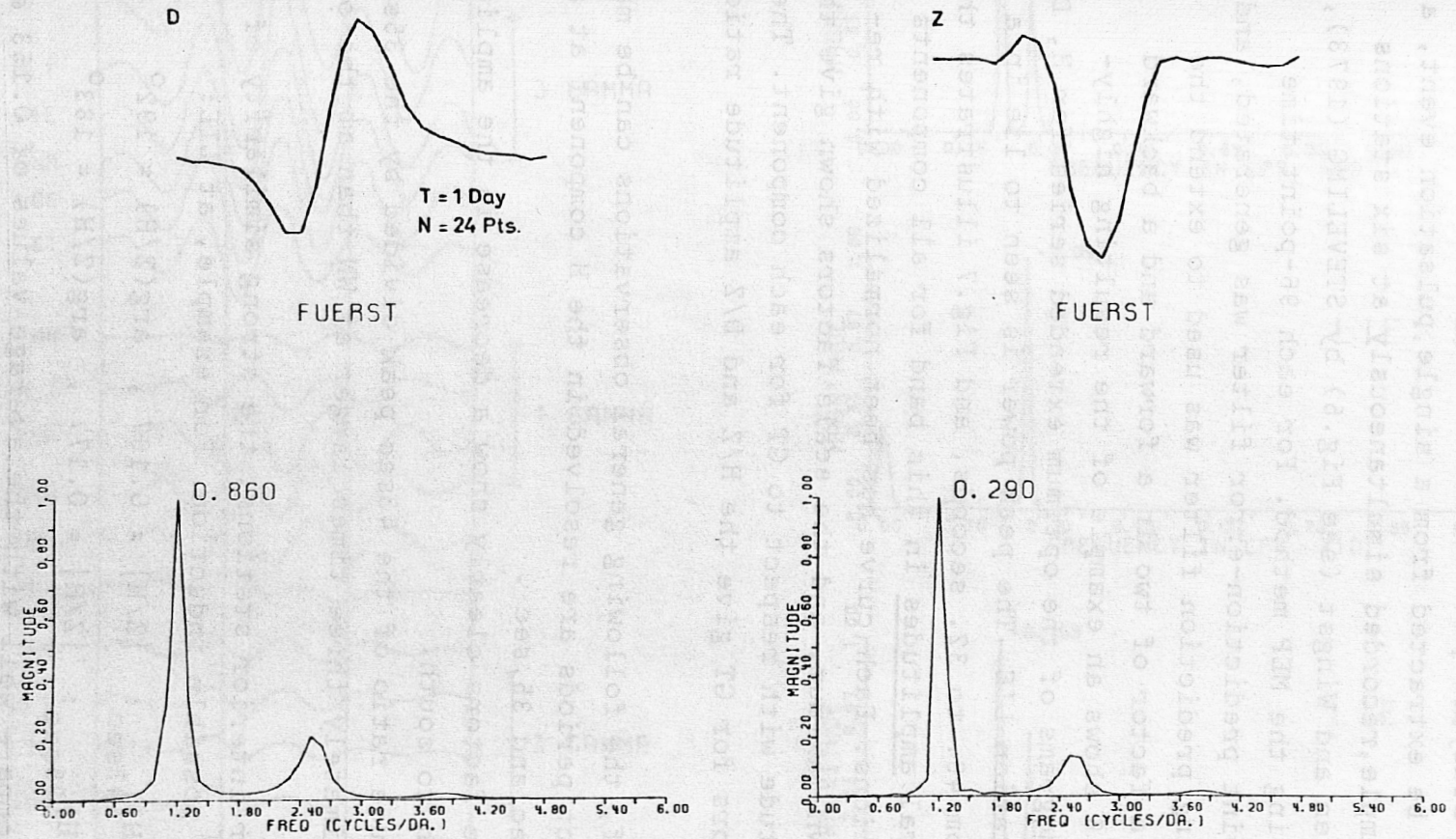


Fig.4: Averaged SQ data and their MEP spectral estimates for D and Z measured on quiet days at Fürstenfeldbruck.

(iii) Analysis of Pulsations

A more fruitful area of MEP analysis application is probably in the study of pulsations. To illustrate that ample information can be extracted from a single pulsation event, a three-minute sample, recorded simultaneously at six stations between Göttingen and Wingst (see Fig.5) by STEVELING (1973), was analyzed using the MEP method. For each 96-point time series, a 40-point prediction-error filter was generated, and the corresponding prediction filter was used to extend the time series by a factor of two in a forward and a backward direction. Fig.6 shows an example of the resulting highly-resolved periodograms of the optimum extended series for H, D, and Z at the station LIE. The peak power is seen to lie in a period band from 48. to 32. seconds, and Fig.7 illustrates the relative spectral amplitudes in this band for all components at the six stations. Each curve has been normalized with respect to its own maximum, and the scale factors shown give the relative magnitude with respect to GT for each component. The bracketed factors for GT give the H/Z and D/Z amplitude ratios at GT.

In brief, the following general observations can be made:

1.) Two distinct periods are resolved in the H component at all stations: 43.sec and 35.sec .

2.) The H scale factors clearly show a decrease in the amplitudes of H from north to south.

3.) The amplitude ratio of the 43sec peak divided by the 35sec peak is approximately three times larger at WN than at the other stations.

4.) At the four interior stations, the strong similarity of H and Z suggests possible induction. For example, at LIE:

$$(Z/H)_{43sec}: \quad |Z/H| = 0.167 , \quad \arg(Z/H) = 192^{\circ}$$

$$(Z/H)_{35sec}: \quad |Z/H| = 0.177 , \quad \arg(Z/H) = 183^{\circ}$$

These values compare well with the average values of 0.153 and $178^{\circ}$ obtained by Steveling for the period band 31 to 51 sec in his transfer function analyses.

5.) Although the spectral distribution pattern for the D components is unclear, its character is quite obviously different

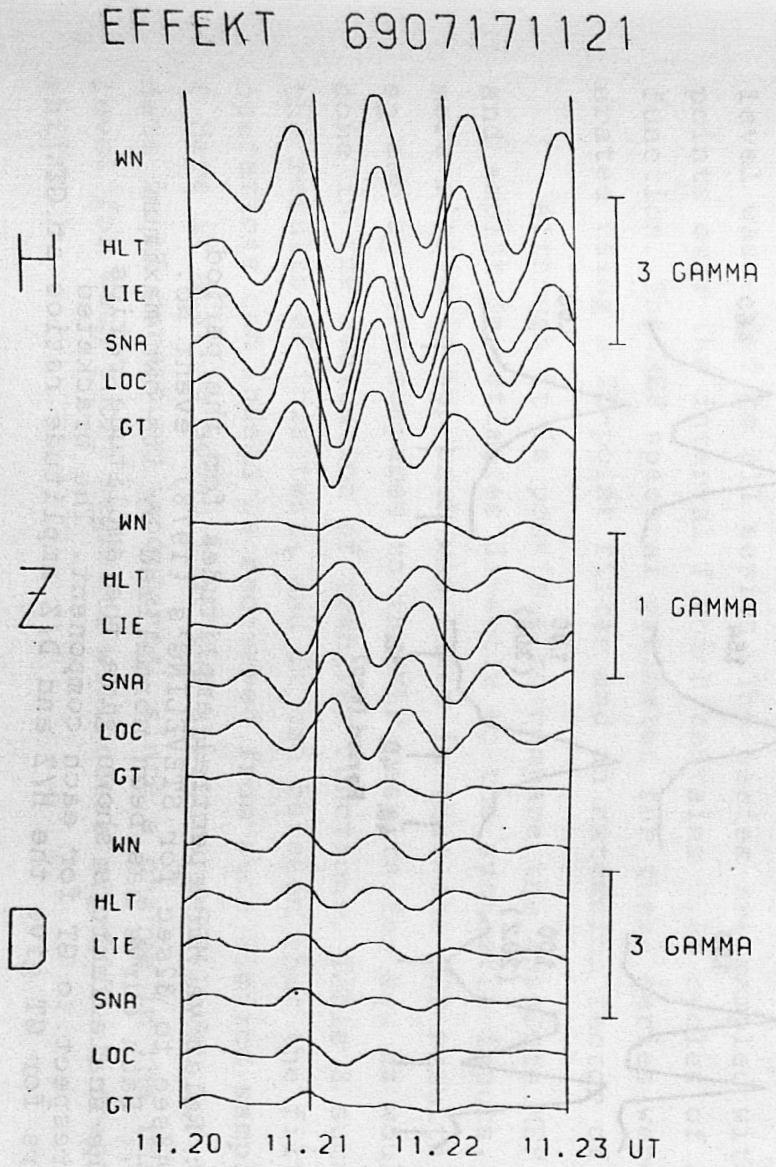


Fig.5: Sample pulsation event recorded at six stations in northern Germany. (after STEVELING,1973)

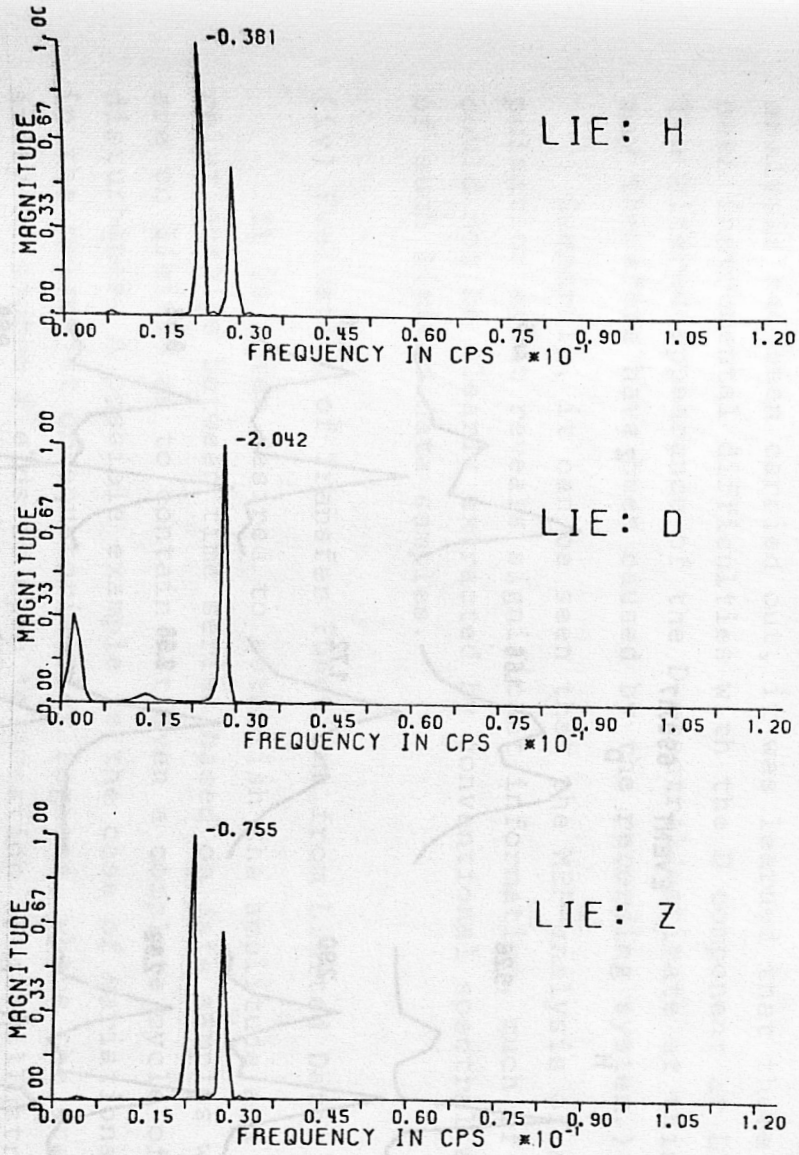


Fig.6: The MEP power spectral estimates for H, D, and Z at LIE. Peak values give the logarithm of the approximate relative power.

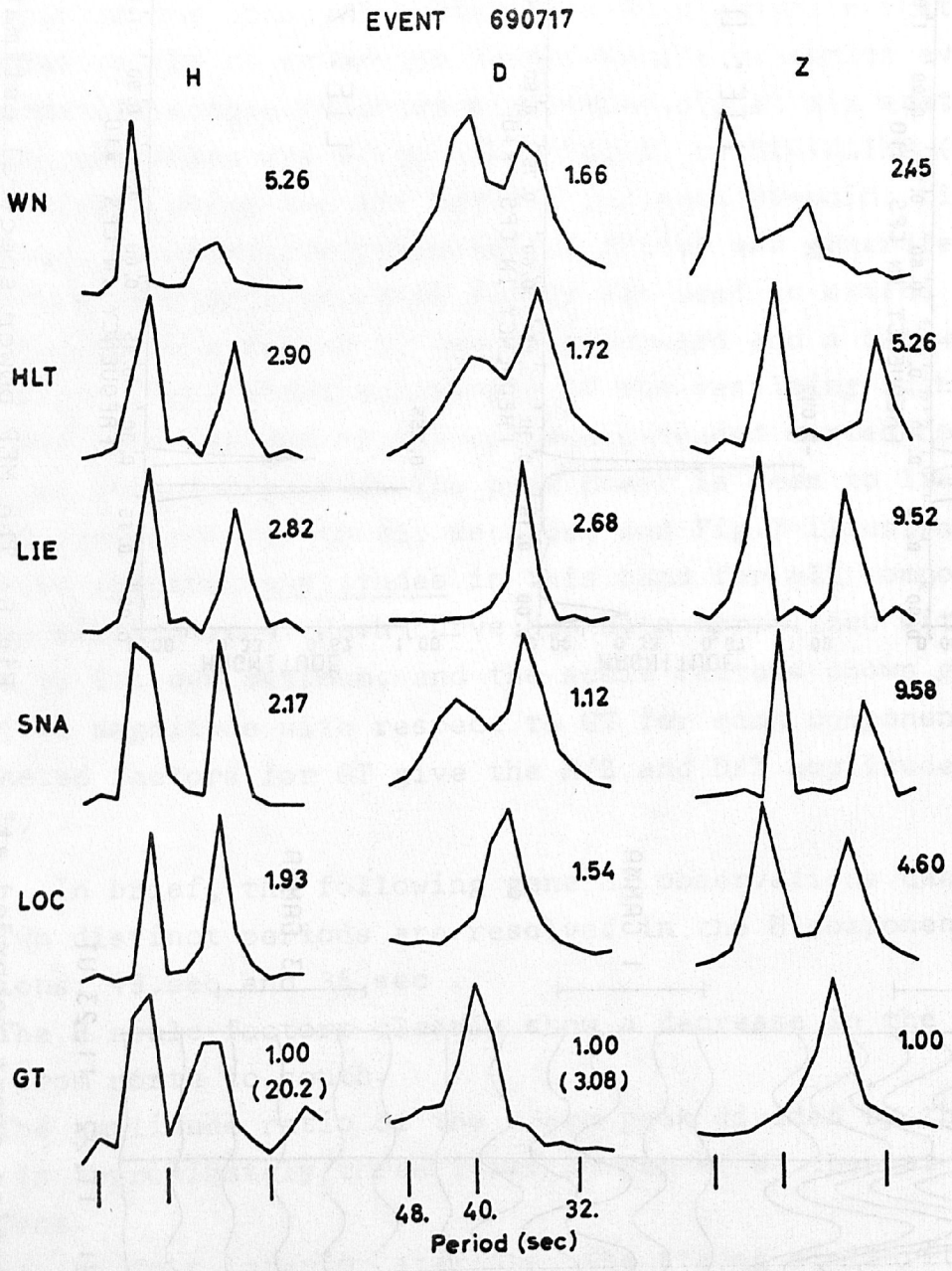


Fig.7: Relative MEP spectral amplitudes for the period band 48sec to 32sec for STEVELING's (1973) event No. 690717. Each curve has been normalized by its own maximum and the scale factors shown give the amplitude ratios with respect to GT for each component. The bracketed factors for GT give the H/Z and D/Z amplitude ratios at GT.

from the H components. ( It should be added that after this analysis had been carried out, it was learned that there had been instrumental difficulties with the D component at LIE. The altered appearance of the D spectral estimate at this site may therefore have been caused by the recording system.)

Summarily, it can be seen that the MEP-analysis of a single pulsation event reveals significant information, much of which could not be clearly extracted by conventional spectral analysis of such limited data samples.

(iv) Evaluation of Transfer Functions from Limited Data

It is often desired to establish the amplitude and phase relationships between time series based on data samples which are so limited as to contain not even a complete cycle of a disturbance. A possible example is the case of variational data in the polar-jet or equatorial-jet regions, where the stationarity of source fields is open to question. Fig.8 illustrates three synthetic series, where the series $s_2$ and $s_3$ have been generated from $s_1$ by simple filter operations altering the phase and the amplitude of the incomplete sinusoid of $s_1$. A 5% noise level was added to all series. The series are sampled with 40 points over the interval T which contains 0.55 cycles of a sine function. The MEP spectral estimates for these series were generated using a 20-point filter and an extension factor of four.

Although it is possible to compare the absolute MEP phase and amplitude estimates directly to the synthetic inputs, a more realistic approach was taken in that the MEP results for $s_2$ and $s_3$ were compared to the MEP results for $s_1$, as would be done in the evaluation of transfer functions. Table 3 summarizes the results of this test, and it can be seen that the filter operations can indeed be recovered from such limited samples. ( Here it must be added that the absolute zero-level for the data is of paramount importance, since a change in the reference level for such short segments will cause significant frequency shifts.)

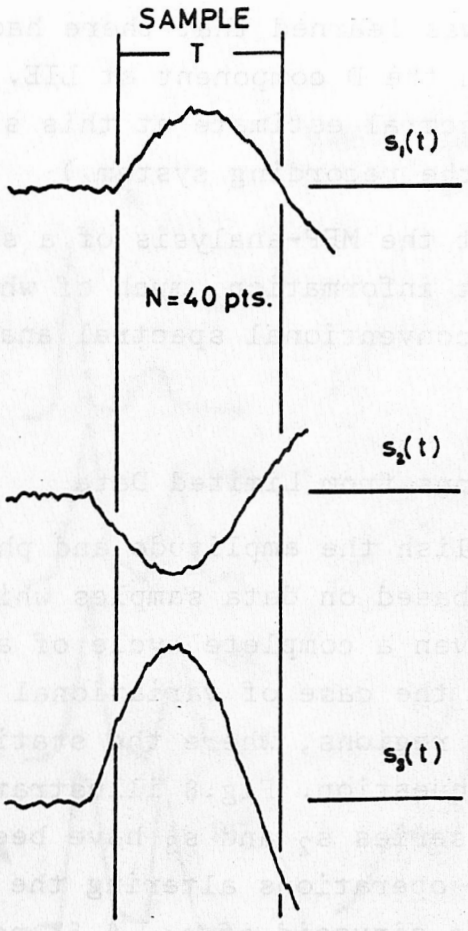


Fig.8: Truncated sinusoidal signals with 5% white noise. $s_2$ and $s_3$ are generated from $s_1$, and all series are sampled only over the interval T.

| | APPLIED FILTER OPERATION | | RECOVERED FILTER OPERATION | |
|---|---|---|---|---|
| | Amplitude Factor | Phase Change | Amplitude Factor | Phase Change |
| $s_1$: | 1.00 | 0.0° | 1.00 | 0.0° |
| $s_2$: | 1.00 | -150.0° | 0.96 | -149.7° |
| $s_3$: | 1.99 | 30.0° | 1.91 | 32.5° |

Table 3: Recovery of filter operations using MEP spectral estimates of extremely limited samples of synthetic data.

(v) Time-Gate Spectral Analysis

The short sample lengths allowed by the maximum entropy method make this technique well suited to the analysis of piecewise-stationary processes. To demonstrate the resolution of temporal changes in the spectral estimates of pulsations, a 12-minute $P_c$ event, recorded at ALLoluokta and ARJeplog in northern Scandinavia, was subdivided into four consecutive 3-minute events. These shorter events of 60 points each were then individually analyzed using 24-point filters and extension factors of two.

To illustrate again the resolution capabilities of the MEP method, Fig.9 shows the third and fourth 3-minute sections for all components as recorded at ALL,along with the corresponding conventional autocorrelation-transform spectral density estimates and the MEP estimates. The time dependence of the spectral estimates at both stations is clearly indicated in Fig.10. All three components show the arrival of a longer period pulsation which dominates during the $t_3$ interval. Furthermore, the similarity of the $t_1$ spectra at ARJ to the $t_2$ spectra at ALL reveals a later arrival of this longer period disturbance at ALL. Needless to say, these detailed, time-dependent spectral characteristics could not be resolved by conventional spectral analysis.

Summary

The basic assumptions, formulations, and limitations for conventional spectral analysis have been summarized, whereas the formulations of the Burg Maximum Entropy Method and the Ulrych Maximum Entropy Prediction Method for spectral estimation have been outlined in greater detail. The examples of application of these newer, data-adaptive methods to geomagnetic variational data show these methods to be extremely effective in the analysis of short data samples and in the analysis of piecewise-stationary magnetic disturbances.

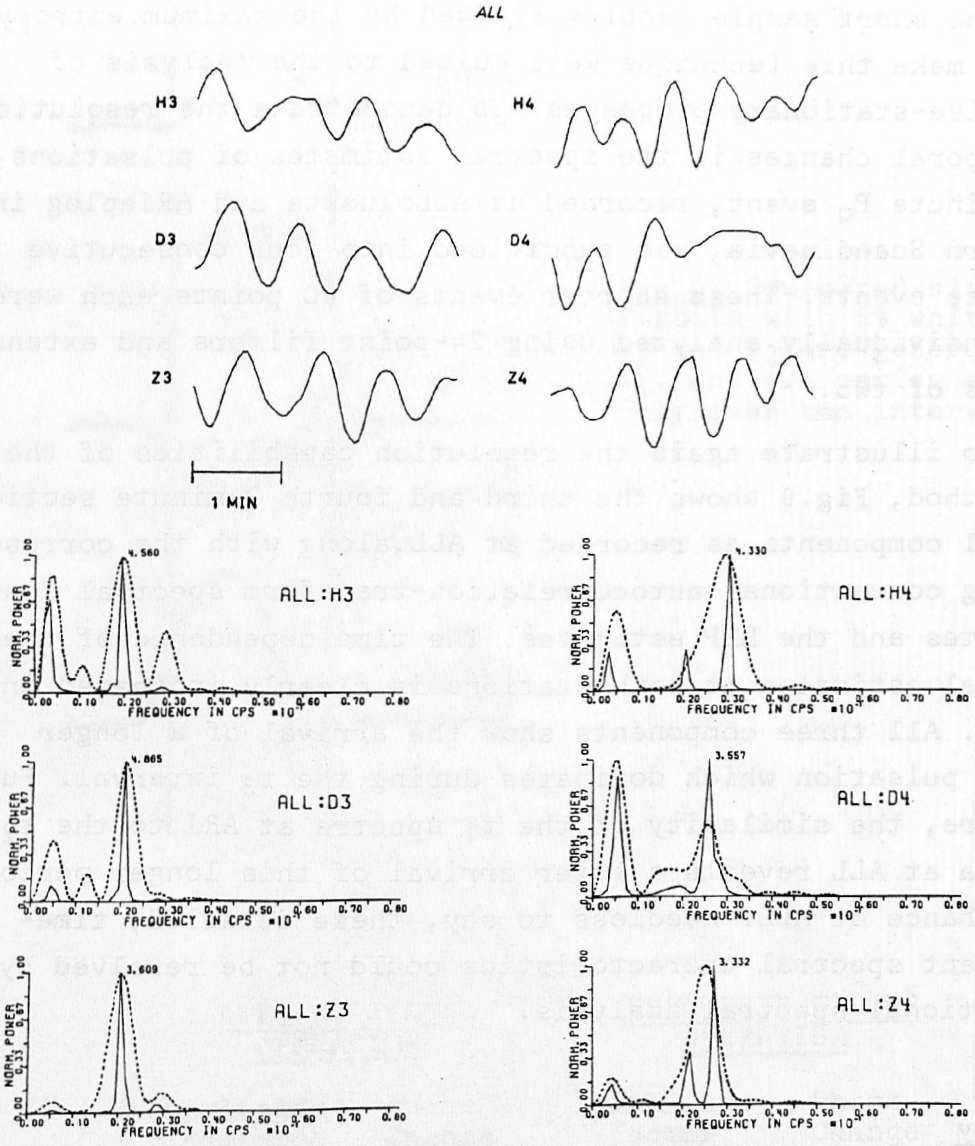


Fig.9: Two consecutive 3-minute record sections for a Pc3 disturbance at Alloluokta, Scandinavia. The corresponding 'zero-extension' spectral estimates (dashed lines) and the MEP spectral estimates are also shown. ( The data is courtesy of M. Palandt, Göttingen.)

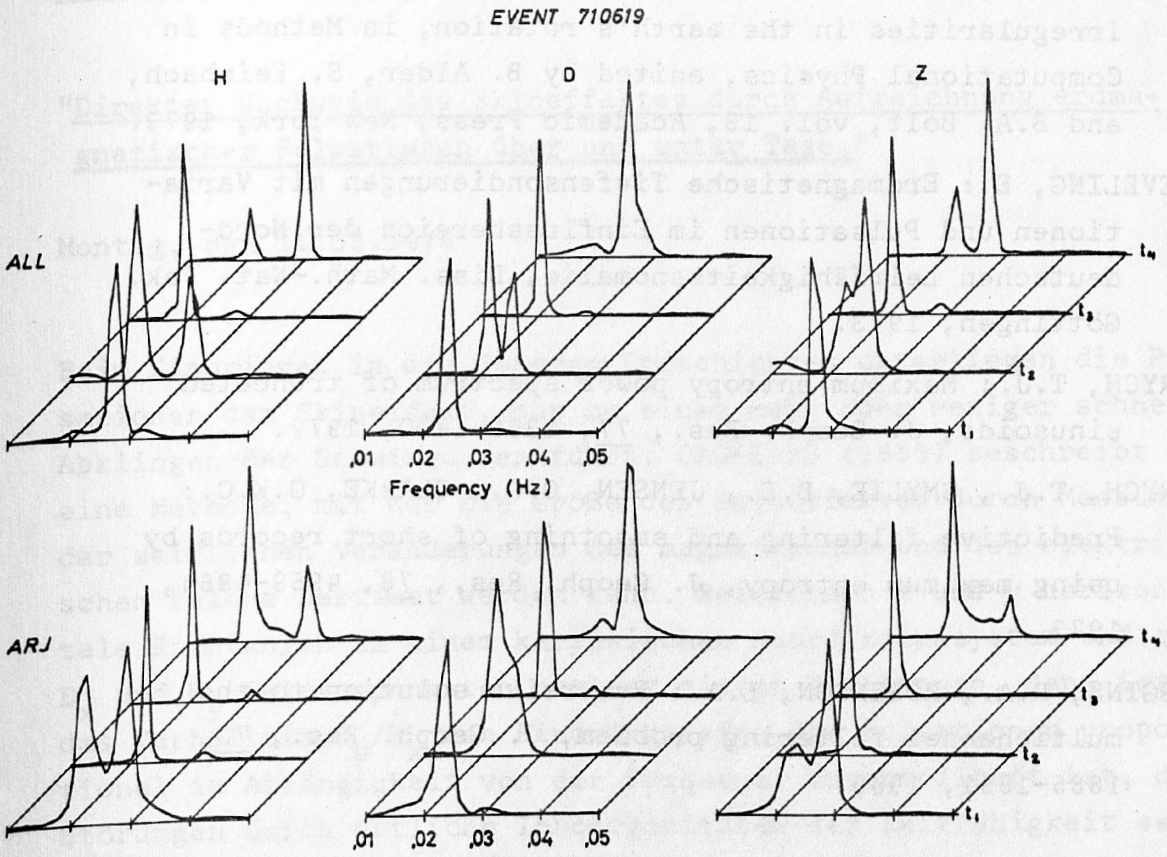


Fig.10: MEP spectral estimates of four consecutive 3-minute record sections at Alloluokta and Arjeplog for a Pc3 event.

References

BURG, J.P.: Maximum entropy spectral analysis, paper presented at the 37th meeting, Soc. Explor. Geophys., Oklahoma City, Okla., Oct.31, 1967.

BURG, J.P.: A new analysis technique for time series data, paper presented at NATO Advanced Study Institute on Signal Processing, Enschede, Netherlands, August 1968.

LACOSS, R.T.: Data adaptive spectral analysis methods, Geophysics, 36, 661-675, 1971.

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

SMYLIE. D.E., CLARKE, G.K.C., ULRYCH, T.J.: Analysis of irregularities in the earth's rotation, in Methods in Computational Physics, edited by B. Alder, S. Feinbach, and B.A. Bolt, vol. 13, Academic Press, New York, 1973.

STEVELING, E.: Erdmagnetische Tiefensondierungen mit Variationen und Pulsationen im Einflussbereich der Norddeutschen Leitfähigkeitsanomalie, Diss. Math.-Nat. Fak. Göttingen, 1973.

ULRYCH, T.J.: Maximum entropy power spectrum of truncated sinusoids, J. Geoph. Res., 77, 1396-1400, 1972.

ULRYCH, T.J., SMYLIE, D.E., JENSEN, O.G., CLARKE, G.K.C.: Predictive filtering and smoothing of short records by using maximum entropy, J. Geoph. Res., 78, 4959-4964, 1973.

WIGGINS, R.A., ROBINSON, E.A.: Recursive solution to the multichannel filtering problem, J. Geoph. Res., 70, 1885-1891, 1965.