Canadian Studies in Population, Vol. 3, 1976, pp. 1-17.

MAXIMUM ENTROPY SPECTRAL ANALYSIS OF MONTE CARLO SIMULATIONS OF A CLOSED FINITE HUMAN POPULATION

K. Morgan and C. R. Somerville
The University of Alberta, Edmonton, Alberta, Canada

Résumé—Il s'agit dans l'analyse suivante d'une présentation de la méthode entropie maximum (MEM) et son application sur les modèles autorégressifs, "série-temps" convenables. Trois simulations repliées Monte Carlo d'une population humaine close soumises à une variation stochastique en natalité et mortalité, produisent des spectres de puissance similaires en dépit des grandes différences dans les trajectoires des nombres totaux en l'espace des 600 ans d'histoire simulée. Les traits saillants de ressemblance sont attribuables à une "longueur de génération" femelle d'un peu plus de 30 ans et un composant périodique de fécondité plus court amplifié par des oscillations dans un système où le mâle domine et où l'âge nuptial est limité. On a utilisé un critère non-subjectif pour le choix d'un ordre acceptable du modèle AR. On a démontré les avantages de la méthode Burg (MEM) sur les techniques spectrales conventionnelles concernant les courts registres des données avec des intervalles d'échantillonage également espacés.

Abstract—The maximum entropy method (MEM) of spectral analysis and its application to fitting autoregressive (AR) time-series models are presented. Three replicate Monte Carlo simulations of a closed human population subject to stochastic variation in births and deaths yield similar power spectra in spite of large differences in the trajectories of total numbers within the first 600 years of simulated history. Major features of similarity of the power spectra may be referable to a female "generation length" of approximately 30 years and a shorter periodic component of fertility amplified by oscillations in a male-dominant, age-restricted marriage system. A nonsubjective criterion for the choice of sufficient order of an AR model is utilized. The advantages of the Burg method (MEM) over conventional spectral techniques for short data records with equally spaced sampling intervals are demonstrated.

Key words—**spectral analysis, Monte Carlo simulation, maximum entropy method, closed population**

I. Introduction

In spite of recent interest in the application of spectral techniques to the analysis of biological data, for example, Housworth (1972) and Fuller and Tsokos (1971), this approach has not been widely used to examine periodic phenomena which are evident in many biological processes. A major obstacle to the application of conventional methods of spectral analysis to biological time series is the difficulty in obtaining a sufficiently large number of measurements for an adequate representation of the series. Such difficulties may arise, for example, because the sampling interval is fixed by circumstance rather than choice, or because the series is stationary for only a limited time, or both.

Statistical considerations notwithstanding, the data requirements of conventional methods are due largely to a dependence on fixed window functions. Since these functions are independent of the true spectrum, they require a relatively large ratio of record length to lag in order to reduce the fortuitous appearance of spurious peaks or instabilities (bias) in the spectrum which may arise, for example, from leakage through a side lobe in the transfer function of the smoothing window as lag is increased. Jenkins and Watts (1968) have demonstrated empirically that for practical purposes none of the three most commonly used window functions can be considered perferable in this regard. For short data records, the nonadaptive properties of the fixed window function place a restriction on the size of lag such that resolution of peaks in the spectrum may not be possible because of severe smoothing of the spectrum which results from using a small lag window. This is particularly limiting in the

case of biological applications where for most purposes it is the frequency of spectral power that is of primary interest, relative power and slope often being difficult or impossible to interpret.

Burg (for example, see Andersen, 1974, Kanasewich, 1973, and Lacoss, 1971) first proposed a data adaptive method of spectral analysis, called the maximum entropy method (MEM), which largely ameliorates the difficulties associated with windowing. This is accomplished by dispensing with the window function altogether and using Wiener optimum filter theory to design a least-squares prediction filter which transforms the input series to white noise. This is equivalent to finding an m length filter which would most reduce the entropy of the spectrum. Furthermore, this filter can generate a process from white noise which would have the same first m autocorrelations as the data. Therefore, it is possible to obtain the power spectrum of the input series from the response of the prediction filter and the power of the noise series. Thus, MEM is a method of maximizing the information which may be obtained from the data without the assumption of continuation which is inherent in most other spectral techniques.

Another important feature of the MEM is that the coefficients of the prediction filter are efficient estimates of the parameters of the autoregressive model associated with the time series. The autoregressive (AR) model can in turn be used to obtain a forecast of the process under study.

Although certain properties of the MEM are as yet incompletely understood, Lacoss (1971) has examined the MEM in some detail using known autocorrelation functions and concluded that it offers consistently superior resolution to conventional methods of spectral analysis, particularly for short data records. In addition, it is not unusually sensitive to statistical fluctuations in the estimated autocorrelation function.

In this paper we illustrate the use of the MEM in the analysis of three nonstationary time series obtained by Monte Carlo simulation of the growth of a human population. We compare the estimates of the smoothed spectral density using a Tukey spectral window to those of the MEM power spectrum.

II. Methodology

2.1 *A requirement for stationarity*

The MEM has, in common with most other spectral techniques, an absolute requirement for stationarity of the input series. A discrete-time stochastic process is said to be strictly stationary if a joint distribution of any set of observations is unaffected by shifting all times of observation ahead or backward by any integer amount. In practice, this definition is somewhat awkward so that several ad hoc procedures have been employed such as examining the significance of slope by least-squares linear regression (Chatfield and Pepper, 1971) or testing independence with a Kendall rank-order statistic (Fuller and Tsokos, 1971). Visual examination of both the series and its sample autocorrelation function (acf) is universally recommended since a stationary series has no trend and its sample acf damps quickly to zero. In our experience, the detection of trend and subsequent transformation to stationarity are not serious problems. Trend is, in practice, indistinguishable from an unresolved oscillation of very long period. Generally this will be indicated by considerable power at the low frequency end of the spectrum at the expense of resolution elsewhere in the frequency domain. Such low frequency power may then be removed simply by some combination of sum and difference filters or by a multiplicative filter for a series with an exponential trend (see Jenkins and Watts, 1968).

2.2 *Estimation of the spectrum*

In the simplest representation of the MEM, a linear prediction operator A_m of length m, is

defined on the input series x_t with zero mean such that for a prediction distance unity a best estimate of the series at $t + 1$ is obtained as

$$\hat{x}_{t+1} = \sum_{\tau=1}^{m} x_{t+1-\tau} \cdot A_\tau \qquad (t = 1, \ldots, N) \tag{2.2.1}$$

An error series is then obtained as:

$$z_t = \hat{x}_t - x_t \tag{2.2.2}$$

which is seen to represent the nonpredictable part of x_t. The prediction operator A_m is given by the m length series $A_m = a_{m1}, \ldots, a_{mm}$ and may be obtained from the set of relations:

$$\begin{bmatrix} \phi_0 & \phi_1 & \cdot & \cdot & \cdot & \phi_{m-1} \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ & & & & & \phi_1 \\ \phi_{m-1} & & \cdot & \cdot & \cdot & \phi_0 \end{bmatrix} \begin{bmatrix} f_1 \\ \cdot \\ \cdot \\ \cdot \\ f_m \end{bmatrix} = \begin{bmatrix} g_0 \\ \cdot \\ \cdot \\ \cdot \\ g_{m-1} \end{bmatrix} \tag{2.2.3}$$

where for lag τ, ϕ_τ is the autocorrelation of the input series, g_τ is the crosscorrelation between the desired output and the input signals, and f_τ is the least-squares inverse filter or Wiener filter which minimizes the sum of squares of the error series. The Wiener filter minimizes the energy existing in the difference between the desired output and the actual output (Peacock and Treitel, 1969). The sample autocovariance function of the input series for lag τ is estimated as:

$$c_\tau = (1/N) \cdot \sum_t x_t \, x_{t+\tau} \tag{2.2.4}$$

and the crosscovariance between the desired output and input traces for a prediction distance of one as:

$$c_{\tau+1} = (1/N) \cdot \sum_t x_t \, x_{t+\tau+1} \tag{2.2.5}$$

Thus the system of equations (2.2.3) may be rewritten as:

$$\begin{bmatrix} c_0 & c_1 & \cdot & \cdot & \cdot & c_{m-1} \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ & & & & & c_1 \\ c_{m-1} & & \cdot & \cdot & \cdot & c_0 \end{bmatrix} \begin{bmatrix} a_{m1} \\ \cdot \\ \cdot \\ \cdot \\ a_{mm} \end{bmatrix} = \begin{bmatrix} c_1 \\ \cdot \\ \cdot \\ \cdot \\ c_{1+m-1} \end{bmatrix} \tag{2.2.6}$$

The solution of this system of normal equations yields the prediction operator A_m. A prediction error operator with a prediction distance of one is defined as:

$$A'_m = 1, -a_{m1}, \ldots, -a_{mm} \tag{2.2.7}$$

The output power of this filter is estimated as:

$$P_m = \sum_{\tau=0}^{m} c_\tau \, A'_{\tau+1} \tag{2.2.8}$$

The power spectrum may then be obtained over the Nyquist interval $-1/(2\Delta t) \leqslant f \leqslant 1/(2\Delta t)$ as:

$$P(f) = P_m \Delta t \, / \left| 1 - \sum_{n=1}^{m} a_{mn} \, e^{-2\pi i f n \Delta t} \right|^2 \tag{2.2.9}$$

where $i^2 = 1$. For purposes of digital estimation this equation is written in terms of the trigonometric identities as follows:

$$P(f) = P_m \Delta t \Big/ (1 - \sum_{n=1}^{m} a_{mn} \cos 2\pi f n \Delta t)^2 + (\sum_{n=1}^{m} a_{mn} \sin 2\pi f n \Delta t)^2 \qquad (2.2.10)$$

Alternatively, it can be shown that if the filter coefficients $a_{m1}, \ldots, a_{mm}$ and the output power P_m were known independently of the autocorrelation function, we could obtain an improved estimate of the sample autocorrelation function (acf) by solving the set of normal equations:

$$\begin{bmatrix} \phi_0 & \phi_1 & \cdot & \cdot & \cdot & \phi_m \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ & & & & & \phi_1 \\ \phi_m & & \cdot & \cdot & \cdot & \phi_0 \end{bmatrix} \begin{bmatrix} 1 \\ -a_{m1} \\ \cdot \\ \cdot \\ \cdot \\ -a_{mm} \end{bmatrix} = \begin{bmatrix} P_m \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \qquad (2.2.11)$$

Burg first outlined the procedure for obtaining the filter coefficients and the output power independently of the acf and Andersen (1974) has more recently presented the following simple algorithm for digital computation.

For the general case, the average forward and reverse power of the m length prediction error operator is given by:

$$\Pi_m = \frac{1}{2(N-m)} \sum_{t=1}^{N} [(x_t - \sum_{k=1}^{m} a_{mk} \cdot x_{t+k})^2 + (x_{t+m} - \sum_{k=1}^{m} a_{mk} \cdot x_{t+m-k})^2] \qquad (2.2.12)$$

which can be minimized with respect to a_{mm} to give the correct solution of the highest order filter coefficients as:

$$a_{mm} = 2 \sum_{t=1}^{N-m} b_{mt} \cdot b'_{mt} \Big/ \sum_{t=1}^{N-m} (b^2_{mt} + b'_{mt})^2 \qquad (2.2.13)$$

where:

$$b_{mt} = b_{m-1,t} - a_{m-1,m-1} \cdot b'_{m-1,t} \qquad (2.2.14)$$

$$b'_{mt} = b'_{m-1,t+1} - a_{m-1,m-1} \cdot b_{m-1,t+1} \qquad (2.2.15)$$

The remaining filter coefficients are then obtained as:

$$a_{mk} = a_{m-1,k} - a_{mm} \cdot a_{m-1,m-k} \qquad (2.2.16)$$

and P_m is estimated by:

$$P_m = P_{m-1} \cdot (1 - a^2_{mm}) \qquad (2.2.17)$$

Thus the power spectrum can be obtained directly without actually calculating the sample acf by starting with the two-point filter $(1, -a_{11})$ which is determined by the starting values $b_{1t} = x_t$, $b'_{1t} = x_{t+1}$ and $P_0 = (1/N) \cdot \sum_t x_t^2$ and building up the $m + 1$ point filter recursively. It is convenient to normalize the zero-mean input series as follows:

$$x_t \cdot \sqrt{N} \Big/ \sqrt{\sum_t x_t^2} \qquad (2.2.18)$$

This transformation gives a power scale which is comparable to that obtained by conventional methods. If the series is not normalized, solution of equation (2.2.11) provides an estimate of the sample autocovariance rather than the sample autocorrelation function.

Criteria for choosing an optimal prediction filter length have not been explicitly defined. Peacock and Treitel (1969) suggested, on the basis of empirical results, that the filter length should span two or more orders of any multiple pattern in the autocorrelation function. However, in some applications multiple patterns may not be sufficiently evident. Andersen

(1974) has noted that the choice of filter length will depend on the kind of information one wants to obtain; for example, identification of major frequency components, the signal-to-noise ratio, etc. An objective procedure might be defined as follows: If the true order *(m)* of the AR or ARMA (autoregressive and moving average) process generating the series were known, then the m length prediction filter would be a sufficient description of the process. Since the MEM spectrum is simply the inverse of the squared response of the filter, then it is reasonable to assume that the MEM spectrum obtained from the m length prediction filter is an adequate representation of the true spectrum. A recently developed criterion for determining the order of an AR or ARMA process is utilized in a later section of this paper.

2.3 *Interpretation of the spectrum*

In contrast to conventional techniques which estimate relative power by peak values, the power in the MEM spectrum is proportional to the area under the peak and approximately proportional to the square of the peak values. In view of the efficiency of numerical integration techniques, this property of the MEM, although inconvenient, should not be a serious drawback to widespread application. Moreover, Lacoss (1971) has concluded from empirical studies that the MEM is often a good pointwise approximation to the true power, particularly if the peaks are broad.

R. E. Kromer (cited in Lacoss, 1971) has shown that the estimate of power of the MEM spectrum is asymptotically normally distributed with approximate degrees of freedom N/L where L is filter length and N is record length, so that confidence intervals for the power may be obtained as for conventional methods from a t-distribution. Because the MEM does not have a window function proper, the concept of standardized bandwidth is not applicable to the MEM spectrum. However, this must not be considered limiting since the resolution of the MEM is such that the actual bandwidth of spectral peaks will generally be very narrow.

2.4 *Autoregressive model building*

For a stationary series with mean μ the general autoregressive process (AR) is given by:

$$X_t - \mu = \alpha_{m1} \cdot (X_{t-1} - \mu) + \ldots + \alpha_{mm} \cdot (X_{t-m} - \mu) + Z_t \qquad (2.4.1)$$

where Z_t is white noise and m is the order of the process. From equation (2.2.1) it is seen that the coefficients of the AR model $\alpha_{m1}, \ldots, \alpha_{mm}$ are estimated by the m length prediction filter $\alpha_{m1}, \ldots, \alpha_{mm}$ and Z_t is the error series (2.2.2). Since the prediction filter is obtained so as to maximize the "whiteness" of Z_t, the filter coefficients are best estimates for the AR model.

The residual sum of squares of the AR model of order m may be expressed as:

$$S(\hat{\mu}, \hat{\alpha}_{m1}, \ldots, \hat{\alpha}_{mm}) \simeq (N-m) \cdot [C_{xx}(0) - \hat{\alpha}_{m1} C_{xx}(1) - \ldots - \hat{\alpha}_{mm} C_{xx}(m)] \qquad (2.4.2)$$

where $C_{xx}(\tau)$ is the sample autocovariance for lag τ. The residual variance is then obtained as:

$$s^2 = S(\hat{\mu}, \hat{\alpha}_{m1}, \ldots, \hat{\alpha}_{mm}) / (N-2m-1) \qquad (2.4.3)$$

Recently Akaike (1974) and Parzen (1974) have formulated criteria for determining the order of the ARMA model which best approximates the stationary process. The simplest and to date most widely used of these is Akaike's final prediction error (FPE) criterion. The FPE for an AR model of order m is defined as:

$$FPE(m) = [(N+m) / (N-m)^2] \cdot S(\hat{\mu}, \hat{\alpha}_{m1}, \ldots, \hat{\alpha}_{mm}) \qquad (2.4.4)$$

By scanning AR models successively from order $m = 0$ to some upper limit $m = M$, the model of choice is given by that order m and the associated $[\hat{\alpha}_{m\tau};\ \tau = 1, \ldots, m]$ which determine the minimum $FPE(m)$ $[m = 0, 1, \ldots, M]$. Although in this procedure the

choice of the correct model is straightforward, the determination of the upper limit M requires judgment.

If the original series required filtering prior to estimation of the prediction operator, it is necessary to backfilter the AR model in order to obtain a representation of the original series. For example, the $m + 2$ order AR model is obtained from the m filter coefficients $\alpha_{m1}, \ldots, \alpha_{mm}$ obtained on the double-differenced series as:

$$X_t = \hat{\mu}(1 - \sum_{\tau=1}^{m} a_{m\tau}) + \sum_{\tau=1}^{m} X_{t-\tau}(C_\tau - 2C_{\tau+1} + C_{\tau+2}) + Z_t \tag{2.4.5}$$

where $\hat{\mu}$ is the sample mean of the double-differenced series and the C_τ are obtained from the filter coefficients as:

$$[C] = [0, -1, a_{m1}, \ldots, a_{mm}, 0, 0] \tag{2.4.6}$$

After backfiltering, the AR model has increased in order from m to $m+2$ and may then be used to predict the behaviour of the original process for some relatively short lead time l where l is dependent on the order of the process (see Box and Jenkins, 1970).

III. Application to Simulated Time Series

3.1 *Structure of the simulation*

The maximum entropy method (MEM) is illustrated here with a subset of three runs (see Figure 1) from a total of forty Monte Carlo microsimulations of a closed human population. The simulation study comprises eight combinations of incest prohibition, clan exogamy, and remarriage rules (Morgan, 1974). The computer programme simulates a stochastic age-and-sex dependent, birth-and-death process with a male-dominant marriage market. Age-and-sex specific mortality schedules and female marital fertility patterns were chosen to approximate an intrinsic rate of increase of 0.5 per cent per cycle for a stable population. Each run began with the same initial population of one hundred unmarried and unrelated individuals of each sex. The initial population was age structured to approximate the equilibrium distribution which would eventually be reached in the stable population. The programme provided summary statistics every ten cycles and was terminated in most cases prior to 600 cycles if the population decreased below a certain level or would have exceeded storage for a population of about 1,000 individuals. A cycle is considered nominally equivalent to one year.

The effect of the starting conditions was to introduce an early pulse of births subsequent to the initial pulse of some forty or so marriages. After an initial period of growth, many of the populations experienced a temporary stationary phase. The result of the starting conditions is to introduce one or more waves into the population numbers, which are expected to reverberate with progressively damped amplitudes until the population eventually reaches age-and-sex structure equilibrium. It should be noted that although spectral techniques are not affected by a nonconstant amplitude in the signal, the analysis requires the assumption that within the record length the period remains approximately constant.

In the three series chosen for illustration here (Figure 1), all matings are monogamous and remarriage is not permitted upon the death of a spouse. The only input difference among the three simulations is in the use of a unique seed number for the pseudo-random number generator. The number of data points for the original series are 58, 61, and 56 for runs 1, 2, and 3, respectively. A point represents a census of the population at the beginning of each ten-year interval.

3.2 *Calculation of the spectrum*

Since the three series were not stationary, we examined the effectiveness of various digital filtering schemes in transforming the original series to stationarity. Kendall rank order

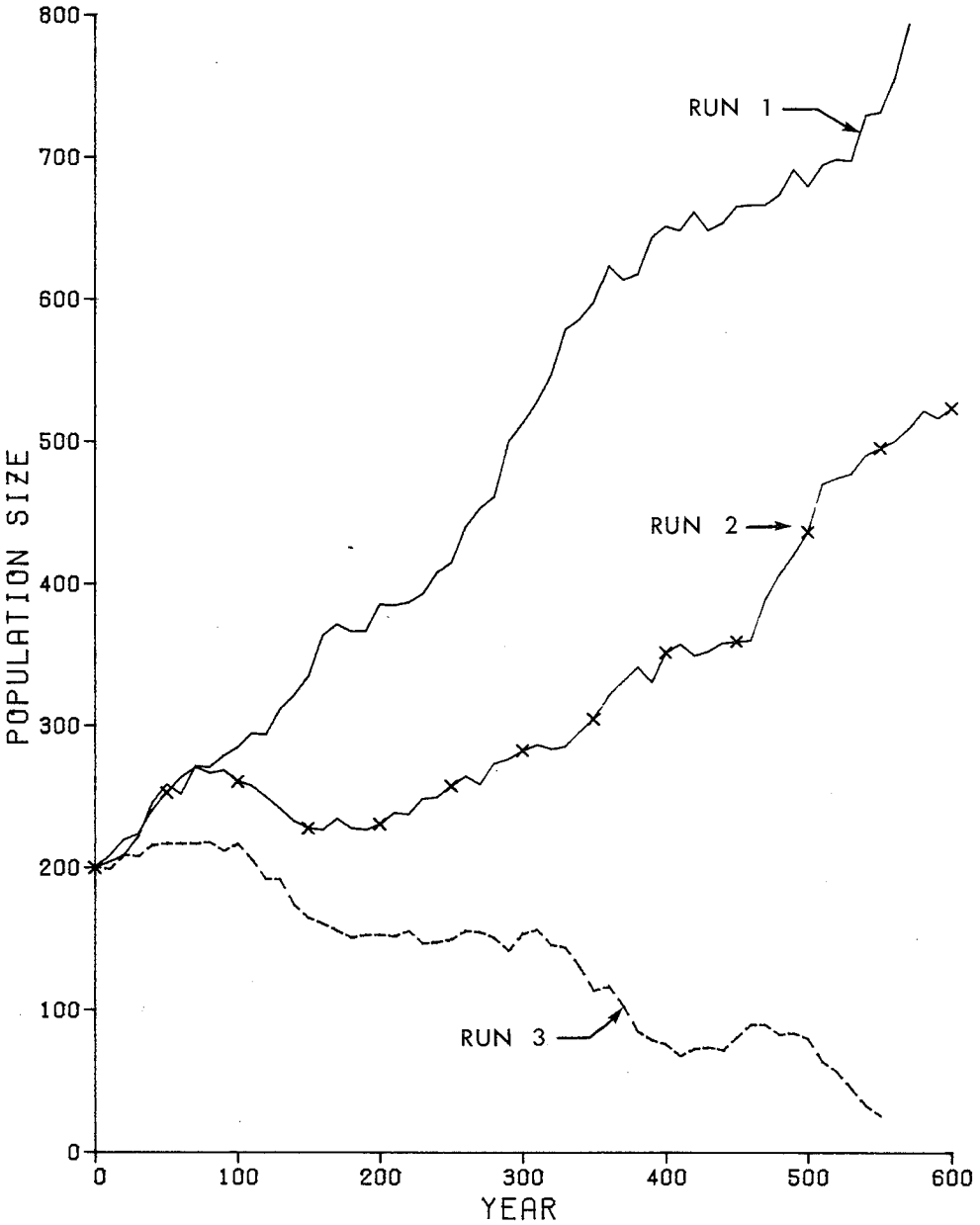

FIGURE 1. TOTAL SIZE OF THREE REPLICATES OF A SIMULATED HUMAN POPULATION INITIALLY AND AT EVERY TENTH CYCLE (YEAR)

correlation coefficients for the series of first differences were not significantly different from zero at the 5 per cent level on a two-tailed test. (However, the first differences for run 2 do exhibit a statistically significant trend on a one-tailed test: $\tau = 0.164$, $z = 1.83$.) White noise tests of the series of first differences utilizing the sample integrated spectrum (Jenkins and Watts, 1968) demonstrate periodic effects for all three runs. Furthermore, the MEM spectra of the series of first differences show appreciable power at low frequencies (Figures 2, 3, 4). Determination of the prediction filter length is given below.

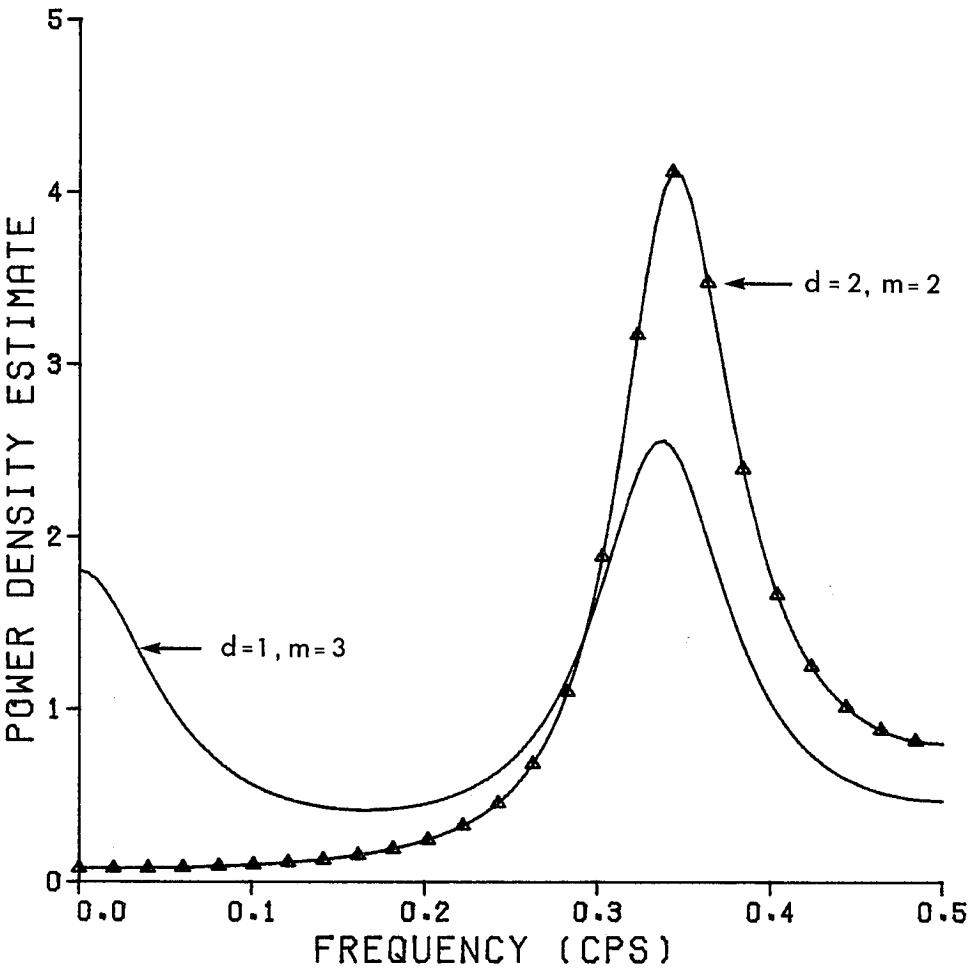

FIGURE 2. MEM POWER SPECTRUM ESTIMATES OF RUN 1 FOR SERIES OF FIRST AND SECOND DIFFERENCES

Since the presence of power at very low frequencies in the MEM spectrum could be the result of lack of stationarity, and because we are interested primarily in spectral components of higher frequency, a second-order difference filter was applied to the original series as follows:

$$Y_t = X_t - 2X_{t-1} + X_{t-2} \tag{3.2.1}$$

We thereby preferentially sacrifice information at low frequencies in order to obtain better resolution at higher frequencies.

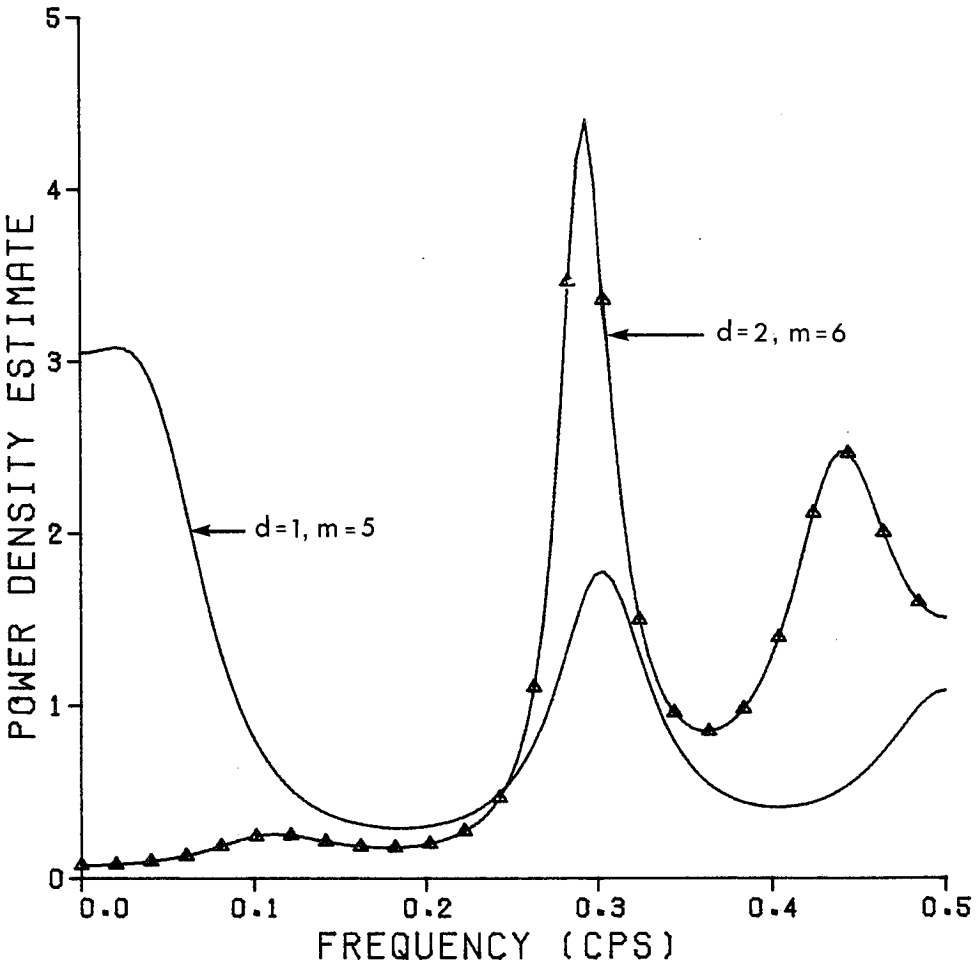

FIGURE 3. MEM POWER SPECTRUM ESTIMATES OF RUN 2 FOR SERIES OF FIRST AND SECOND DIFFERENCES

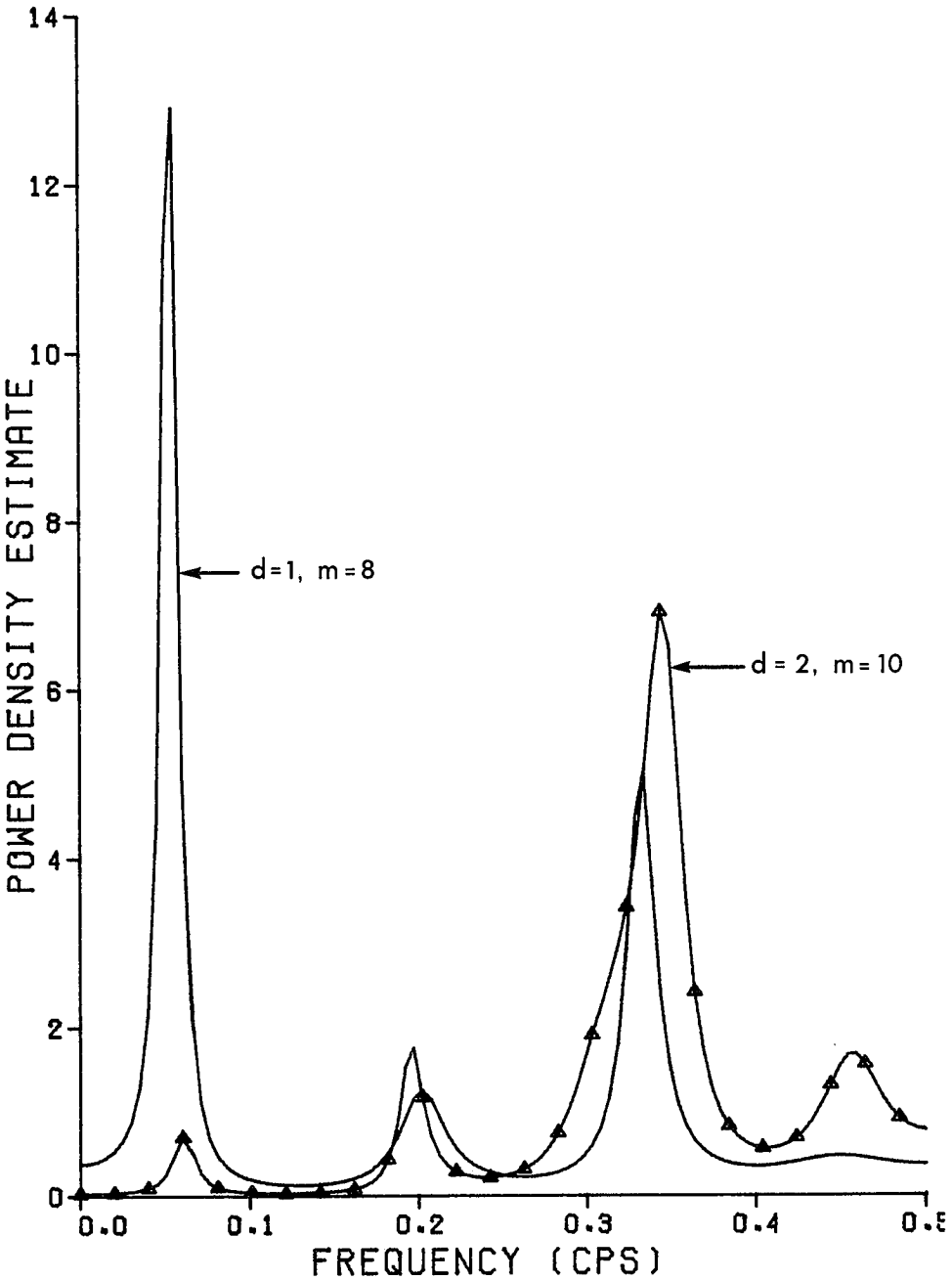

FIGURE 4. MEM POWER SPECTRUM ESTIMATES OF RUN 3 FOR SERIES OF FIRST AND SECOND DIFFERENCES.

The conventional spectral density estimates for the three double-differenced series were computed using the Tukey window for an increasing series of lag lengths (Figures 5, 6, 7). By following the window closing procedure suggested by Jenkins and Watts (1968), apparent qualitative stability of the smoothed spectral density is obtained with lags of twelve or less, as can be judged from Figures 5, 6, and 7.

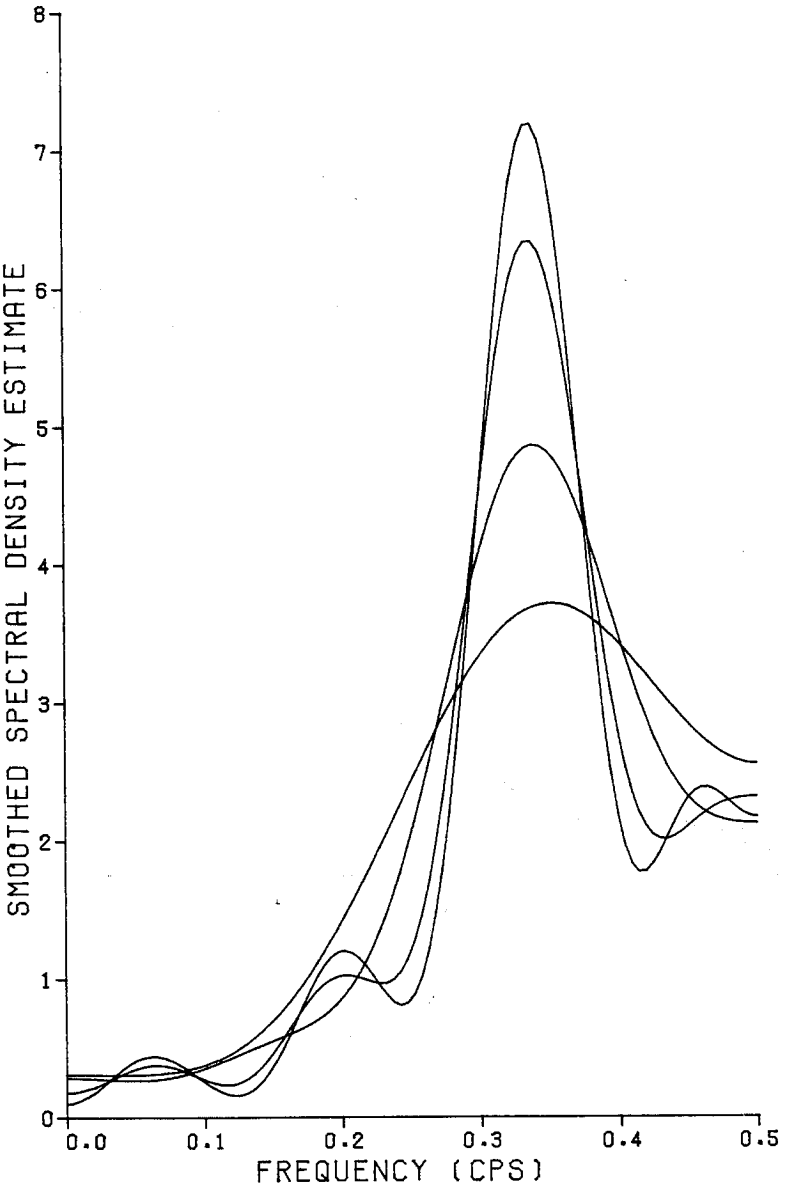

FIGURE 5. SMOOTH SPECTRAL DENSITY ESTIMATES FOR RUN 1 USING THE TUKEY SPECTRAL WINDOW WITH LAGS 5, 8, 12 AND 15

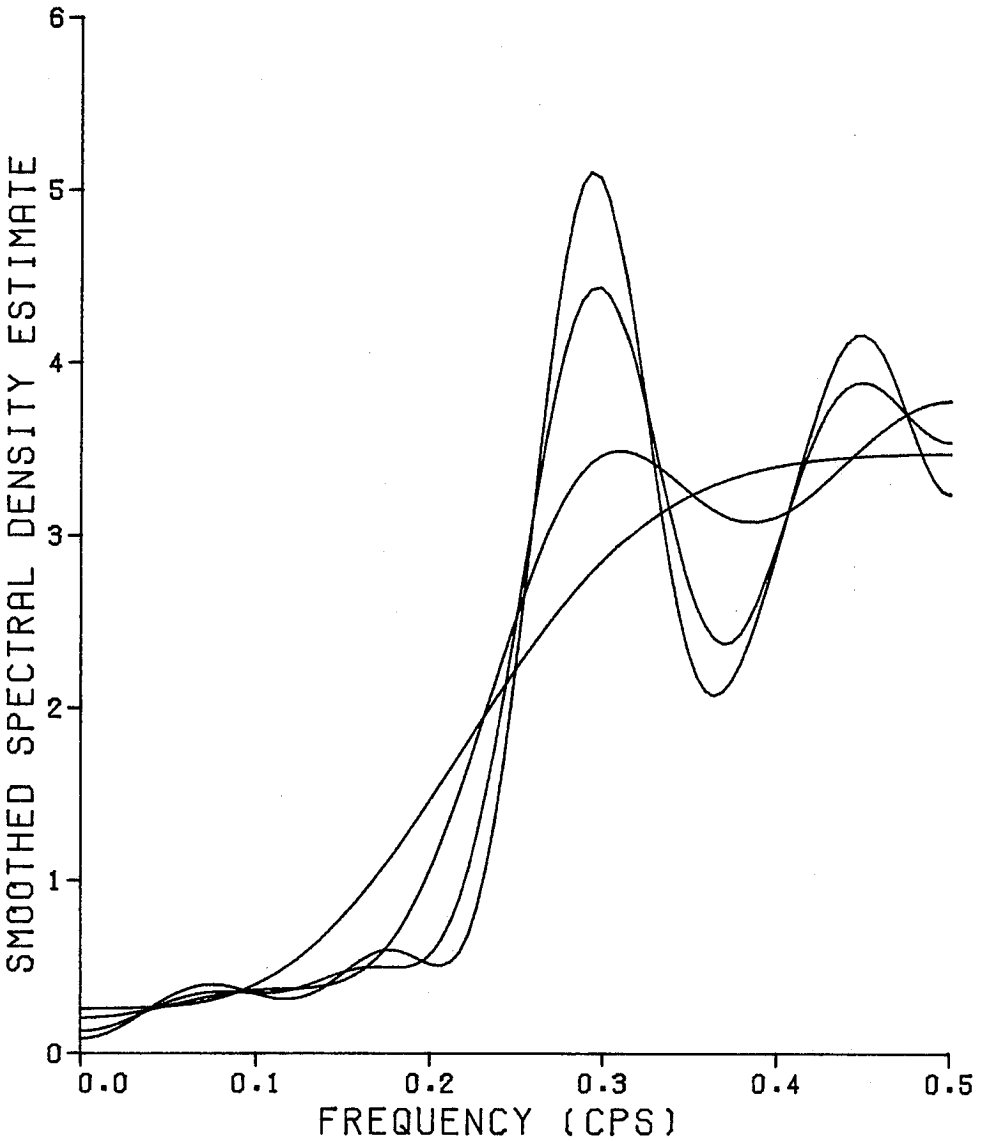

FIGURE 6. SMOOTH SPECTRAL DENSITY ESTIMATES FOR RUN 2 USING THE TUKEY SPECTRAL WINDOW WITH LAGS 5, 8, 12, AND 15

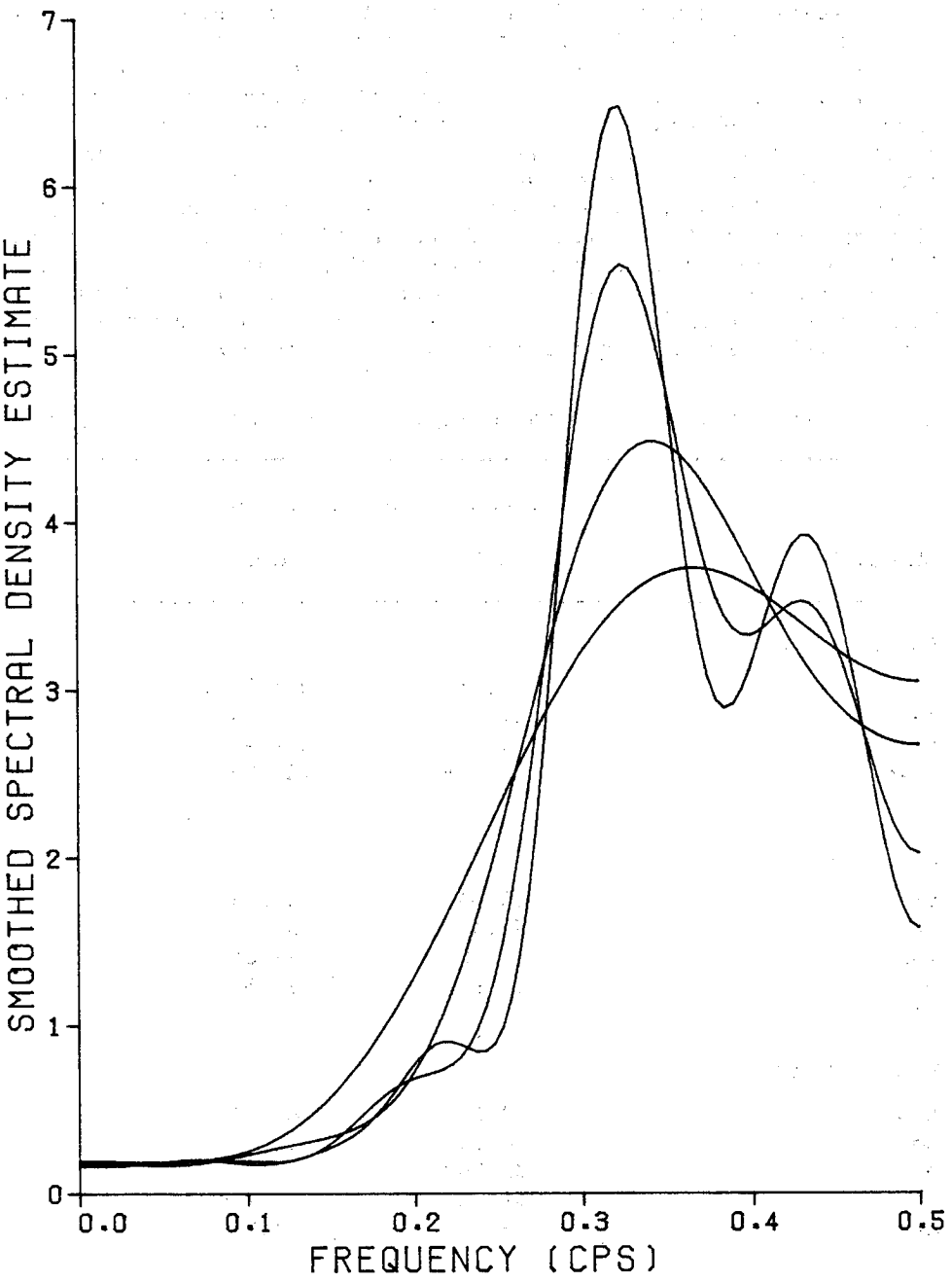

FIGURE 7. SMOOTH SPECTRAL DENSITY ESTIMATES FOR RUN 3 USING THE TUKEY SPECTRAL WINDOW WITH LAGS 5, 8, 12 AND 15

For the calculation of the MEM spectrum, we first obtained an estimate of the correct prediction filter length. The *FPE(m)* for $m = 0, \ldots, 15$ were calculated on the series of first and second differences. The m filter coefficients for each MEM model were calculated by an algorithm published by Andersen (1974). For the series of first differences ($d = 1$), local minima of the *FPE(m)* were obtained at $m = 3$ for run 1, $m = 5$ for run 2, and $m = 8$ for run 3, where M was restricted to 10 in the interest of parsimony. (For $M = 15$, the local minimum of the *FPE(m)* for $d = 1$ of run 1 was at $m = 14$.) For the series of second differences ($d = 2$), local minima of the *FPE(m)* were obtained at $m = 2$ for run 1, $m = 6$ for run 2, and $m = 10$ for run 3, where again M was 10. Values of the *FPE(m)* for various MEM models are provided in Table 1. Utilization of the minimum information theoretical (AIC) estimate (MAICE) leads to the same conclusion as that of the minimum FPE criterion except for run 3, $d = 2$, where the MAICE is for the model of order $m = 13$ (then $m = 10$). The definition of these measures and their relationship are provided by Akaike (1974) (see also Parzen, 1974).

TABLE 1. VALUES OF THE FINAL PREDICTION ERROR (FPE) FOR MEM AR MODELS

	Degree of Differencing	
AR Model	d=1	d=2
Run 1 (N=58)		
m=0	137.23	285.02
m=2	146.10	141.14*
m=3	128.26†	142.46
m=14	119.72*	144.62
Run 2 (N=61)		
m=0	80.04	120.64
m=3	73.43	80.00
m=5	71.78*	83.01
m=6	72.90	79.50*
Run 3 (N=56)		
m=0	54.21	83.54
m=2	54.06	50.83
m=8	37.78†	54.50
m=10	39.94	46.23*
m=15	37.61*	47.69

† Minimum for $0 \leq m \leq 10$

* Minimum for $0 \leq m \leq 15$

The MEM power spectra calculated using the shorter prediction filter lengths are illustrated in Figures 2, 3, and 4. By comparison of these figures with Figures 5, 6, and 7 it can be seen that the MEM spectra obtained in this manner correctly resolve the major feature evident in the Tukey spectra of the series of second differences.

Estimates of the frequencies of the peak power of the major spectral feature of each of the runs can be obtained from the corresponding MEM power spectra as shown in Figures 2, 3,

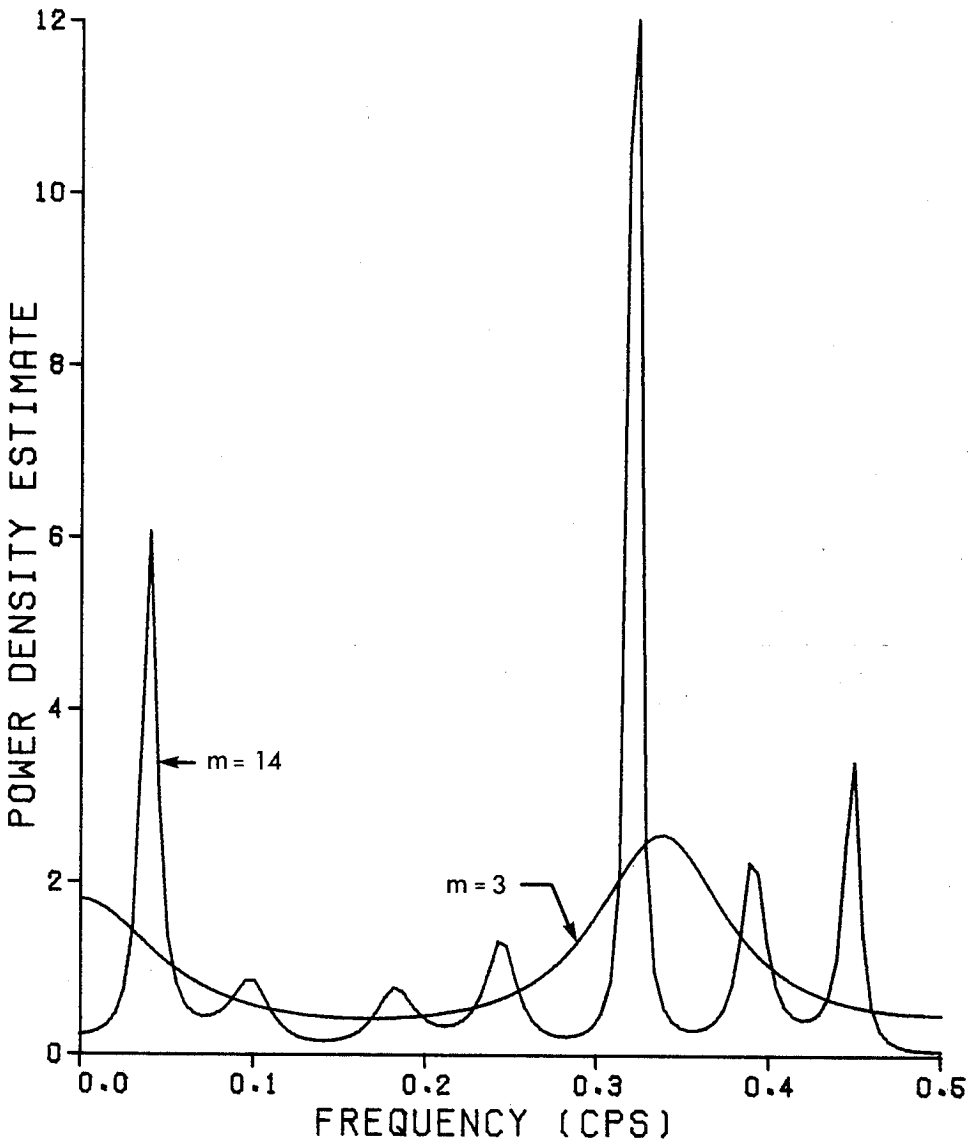

FIGURE 8. MEM POWER SPECTRUM ESTIMATES OF RUN 1 FOR SERIES OF FIRST DIFFERENCES

and 4 for the series of second differences. These values are approximately 29, 34, and 29 years for runs 1, 2, and 3, respectively. We interpret this primary short period as the average generation time, since the input fertility schedule for married females has a mean of 29 years. The initial "driver" of this oscillation is believed to be the "pulse" of births which occur at the beginning of the simulation as noted previously.

We suggest that the secondary peak of shorter period, which is especially apparent for run 2 (Figures 3 and 6), represents a fertility cycle induced by fluctuation in the number of marriages. This fluctuation may be the result of an oscillation in the sex ratio of births. The marriage subsystem of the simulation may be behaving as a feedback amplifier (see Cole, 1972, Chapter 6). The MEM spectrum of second differences of run 3 (Figure 4) provides superior resolution of the secondary peak of short period compared to the Tukey spectrum (Figure 7). The resolution of the MEM spectrum for run 1 can be increased by extending the prediction error filter length to $m = 14$ for the series of first differences (Figure 8). In this case, the long period (approximately 250 years) as well as short period oscillations are clearly resolved. For all three runs the high frequency peak in the MEM spectra corresponds to a period of about 22 to 23 years.

Comparison of AR models for series with $d = 1$ and $d = 2$ shows that for all three runs a smaller final prediction error is obtained for the appropriate MEM model of series of first differences than for second differences. These results suggest that the AR models of first differences may be preferred to those of second differences for the purposes of forecasting. However, it is not obvious that this procedure for AR model selection is valid in general for nonstationary time series. Furthermore, difference filtering may not be the optimal procedure for the purposes of removing trend from these data.

Acknowledgments

This research was supported in part by grants from National Research Council (A6873) and Canada Council (S73-0357) to the first author, and by computing funds provided by the University of Alberta. The spectral analyses were performed in APL operating under the Michigan Terminal System.

We wish to thank Drs. Kanasewich and Cummings, Department of Physics, University of Alberta, for suggesting the MEM; and a reviewer for helpful comments which materially improved an earlier version of this paper.

References

Akaike, H. 1974. A New Look at the Statistical Model Identification. IEEE Transactions on Automatic Control AC-19:716-723.

Andersen, N. 1974. On the Calculation of Filter Coefficients for Maximum Entropy Spectral Analysis. Geophysics 39:69-72.

Box, G. E. P. and G. M. Jenkins. 1970. Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day (revised edition 1976).

Burg, J. P. 1972. The Relationship Between Maximum Entropy Spectra and Maximum Likelihood Spectra. Geophysics 37:375-376.

Chatfield, C. and M. Pepper. 1971. Time Series Analysis: An Example From Geophysical Data. Applied Statistics 20:217-238.

Coale, A. J. 1972. The Growth and Structure of Human Populations. Princeton: Princeton University Press.

Fuller, F. C. and C. P. Tsokos. 1971. Time Series Analysis of Water Pollution Data. Biometrics 27:1017-1034.

Housworth, W. J. 1972. Hybrid Polynomial and Periodic Regression with and without Missing Observations. Biometrics 28:1025-1042.

Jenkins, G. M. and D. G. Watts. 1968. Spectral Analysis and Its Applications. San Francisco: Holden-Day.

Kanasewich, E. R. 1973. Time Sequence Analysis in Geophysics. Edmonton: University of Alberta Press.

Lacoss, R. T. 1971. Data Adaptive Spectral Analysis Methods. Geophysics 36:661-675.

Morgan, K. 1974. Computer Simulation of Incest Prohibition and Clan Proscription Rules in Closed Finite Populations. In B. Dyke and J. W. MacCluer (eds.), Computer Simulation in Human Population Studies. New York: Academic Press.

Parzen, E. 1974. Some Recent Advances in Time Series Modeling. IEEE Transactions on Automatic Control AC-19:723-730.

Peacock, K. L. and S. Treitel. 1969. Predictive Deconvolution: Theory and Practice. Geophysics 34:155-169.

Received June, 1975; Revised October, 1976.