


1. Classificação <i>INPE-COM.10/PE</i> <i>CDU:551.510.535</i>		2. Período <i>Junho de 1977</i>	4. Critério de Distribuição: interna <input type="checkbox"/> externa <input checked="" type="checkbox"/>
3. Palavras Chaves (selecionadas pelo autor) <i>Método da Entropia Máxima, Espectro de Potência, Séries de Tempo, Predição.</i>			
5. Relatório nº <i>INPE-1082-PE/068</i>	6. Data <i>Julho 1977</i>	7. Revisado por <i>J.H.</i> <i>José H.A. Sobral</i>	
8. Título e Sub-Título <i>Espectro de Potência do Campo Geomagnético pelo Método da Máxima Entropia</i>		9. Autorizado por <i>Parada</i> <i>Nelson de Jesus Parada</i> <i>Diretor</i>	
10. Setor <i>CEA/IONS</i>	Código <i>412</i>	11. Nº de cópias <i>06</i>	
12. Autoria <i>I.J. Kantor</i> <i>N.B. Trivedi</i>		14. Nº de páginas <i>39</i>	
13. Assinatura Responsável 		15. Preço	
16. Sumário/Notas <i>Médias mensais do campo geomagnético em Vassouras são analisadas pelo método da máxima entropia. O método é descrito e explicado. É feita uma comparação deste método com outros métodos de obtenção de espectro de potência. São apresentadas as vantagens e desvantagens do uso deste método.</i>			
17. Observações <i>Este trabalho foi parcialmente subvencionado pelo Fundo Nacional de Desenvolvimento Científico e Tecnológico - FNDCT, através de convênio FINEP-271-CT.</i>			

POWER SPECTRUM OF THE GEOMAGNETIC FIELD

BY THE MAXIMUM ENTROPY METHOD

by

I.J. Kantor and N.B. Trivedi

Instituto de Pesquisas Espaciais - INPE

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

São José dos Campos - SP - Brasil

ABSTRACT

Monthly mean values of Vassouras H geomagnetic field is analysed using the maximum entropy method, MEM. The method is described, explained and compared with other methods of spectral analysis. Advantages and disadvantages of the use of MEM are presented.

ESPECTRO DE POTÊNCIA DO CAMPO GEOMAGNÉTICO PELO MÉTODO DA MÁXIMA ENTROPIA

por

Ivan J. Kantor, Nalin B. Trivedi
Instituto de Pesquisas Espaciais - INPE
Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq
São José dos Campos, São Paulo

RESUMO

Médias mensais do campo geomagnético em Vassouras são analisados pelo método da máxima entropia. O método é descrito e explicado. É feita uma comparação deste método com outros métodos de obtenção de espectro de potência. São apresentadas as vantagens e desvantagens do uso deste método.

Introdução

Este trabalho visa esclarecer detalhes do cálculo do espectro de potência do campo magnético terrestre usado no trabalho apresentado na SBPC¹ com o título: "Variação anual do Campo Magnético Terrestre" por N.B. Trivedi, I.J. Kantor e J.M. da Costa.

A estimativa espectral pelo método da máxima entropia, é relativamente nova e foi primeiramente apresentada por Burg² em 1967. Esse método é importante não só para a análise de variação de campos magnéticos e índices de atividade magnética como, também, para qualquer série de tempo.

A teoria do método abrange diversos campos, que permitem muitas outras aplicações. O "Maximum Entropy Method", doravante designado por MEM tem aplicações em espectro de potência, espectro cruzado, cálculo da função de auto-correlação, predição linear, filtragem de dados, transmissão de sinais e cancelamento de ruído.

Um estudo completo do MEM envolve campos bastante diversos como entropia, teoria linear de predição de Wiener, matrizes positivas definidas, processos auto-regressivos, matrizes Toeplitz, espaços de Hilbert e transformada Z, além de um conhecimento de processos estocásticos. Aqui será apresentada uma maneira de chegar ao MEM evitando conhecimentos muito especializados, deduzindo-se o método da maneira mais direta possível. O leitor interessado nos detalhes e variantes teóricas deve consultar Smylie et al³, Makhoul⁴, Radoski et al⁵,

Ulrych e Bishop⁶ ou Ball⁷.

Espectro de Potência

Em geomagnetismo, como em muitos outros campos, os dados medidos são, geralmente, uma série de números tomados em tempos e quidistantes. Vamos representar esta série de tempo por

$$X_1, X_2, \dots, X_N$$

Existem vários métodos para se obter o espectro de potência desta série.

O método da auto-correlação, desenvolvido por Blackman-Tukey⁸, calcula a auto-correlação para vários "lags", fazendo em seguida a transformada de Fourier desta auto-correlação. Se N é o número total de amostras e M o "lag" da auto-correlação, vem:

$$\phi_M = \frac{1}{N-M} \sum_{i=1}^{N-M} X_i X_{i+M} \quad P(f) = \frac{1}{2\Delta t} \sum_{M=-N+1}^{N-1} \phi_M \cos 2\pi f_M \Delta t$$

O método do periodograma desenvolvido por Cooley-Tukey calcula a transformada de Fourier da série de tempo. O espectro de potência é simplesmente o quadrado do valor da transformada.

$$X_k = \sum_{j=0}^{N-1} X_j e^{-2\pi i j k / N} \quad P_k = |X_k|^2$$

O fenômeno do qual foi medida a série de tempo, no nosso caso o campo H, na verdade não é nem discreto nem limitado no tempo. Portanto o espectro calculado não pode representar com exatidão

o espectro real de potência do sinal original do qual a série de tempo foi medida. O fato do tempo total de medida ser limitado, T por exemplo, implica numa resolução de frequência $1/T$. Se medirmos durante muito pouco tempo o sinal, perdemos resolução. O fato de não conhecermos o sinal fora do intervalo de tempo das medidas provoca esta perda de resolução. Cada um dos métodos citados contem intrinsicamente uma condição para os valores da função fora do intervalo da série de tempo. O método da autocorrelação supõe que os pontos fora do intervalo sejam nulos. O periodograma assume uma repetição periodica dos pontos. Estas extensões dos dados certamente não são as mais naturais, o correto seria estender a sequência dos dados através de um previsão.

O método da máxima entropia, MEM, foi desenvolvida por Burg². É construido um filtro preditor linear, que dá o mínimo erro quadrático ($\min E(\hat{x}_t - x_t)^2$) entre os dados e o valor previsto \hat{x}_t destes mesmos dados. O espectro de potência $P(f)$ pode ser calculado usando-se os valores dos coeficientes deste filtro

$$\hat{x}_t = \sum_{k=1}^M g_k x_{t-k} \quad P(f) = \frac{P_M \Delta t}{|1 - \sum_k g_k \exp(-2\pi i f \Delta t k)|^2}$$

Será visto mais tarde que a série de tempo de previsão usa apenas a informação dada pela série de tempo original, daí a maximização da entropia. Na verdade o espectro de potência é o da série de previsão que tem extensão infinita e, portanto, a resolução não depende mais do tempo total da série dada, mas sim da previsibilidade dela.

Processos auto-regressivos (AR)

Para a compreensão do filtro predictor devemos introduzir o conceito de processo auto-regressivo, comumente abreviado por AR.

Vamos imaginar um carro deslocando-se ao longo de uma estrada. Podemos medir sua posição em cada instante de tempo (cada minuto p.e.). Teremos assim uma série de tempo.

$$x_0, x_1, x_2, x_3, x_4 \dots$$

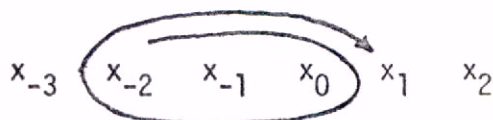
Se a velocidade do carro for constante, a diferença entre cada x_t e seu anterior, x_{t-1} , será constante e proporcional à velocidade, e cada valor x_t pode ser previsto (\hat{x}_t valor de x_t predito)

$$\hat{x}_t = x_{t-1} + (x_{t-1} - x_{t-2}) = 2x_{t-1} - x_{t-2}$$

Se o carro tiver aceleração constante serão necessários 3 valores de x para fazer a predição.

$$\hat{x}_1 = 2,5x_0 - 2x_{-1} + 0,5x_{-2}$$

A cada 3 valores consecutivos da posição podemos prever o próximo. Isto pode ser representado pelo diagrama abaixo.



Mesmo que a velocidade ou aceleração do carro não sejam constantes, a sua posição pode ser aproximadamente predita pelo uso dos 3 pontos an

teriores. Neste caso estaríamos usando um filtro preditor com três pesos ou coeficientes, embora, é claro, o número de coeficientes possa ser qualquer.

Para uma expansão em série de Taylor no ponto $x(t=0)$ de uma função $x(t)$, podemos usar os valores de $t=-1, -2, -3 \dots$ para estimar as derivadas em $t=0$. A primeira derivada seria dada pela diferença $x_0 - x_{-1}$. A segunda derivada depende dos 3 últimos pontos, a terceira derivada dos quatro últimos pontos, etc. Portanto a expansão em série de Taylor, para calcular o valor de x_1 , daria a estimativa \hat{x}_1 em termos de uma função linear dos pontos anteriores. Se a série for truncada em $N-1$ termos teremos.

$$\hat{x}_1 = \sum_{k=0}^N \frac{x^{(k)}(0)}{k!} = \sum_{k=0}^N g_k x_{-k}$$

Este preâmbulo tenta mostrar que, tomar cada ponto futuro como uma combinação linear dos pontos do passado, não é tão restritivo como poderia parecer à primeira vista. É o mesmo que expandir a função em série de Taylor, truncada na n -ésima derivada, correspondendo à ordem do filtro previsor.

Em 1927, Yule⁹ mostrou que qualquer processo estacionário, no qual a média é constante e, em particular, nula, e a autocorrelação é estacionária no sentido amplo, ou seja

$$E x_j = 0$$

$$E x_j x_k = \phi(j-k) \quad (\text{auto-correlação})$$

A sequência x_t pode ser expressa por

$$x_t = \sum_{k=1}^{\infty} g_k x_{t-k} + \varepsilon_k$$

onde ε_k é o resíduo ou inovação, ou seja, toda série, resultante de um processo estacionário, pode ser decomposto numa parte previsível ou de terminística e numa parte aleatória. Isto corresponde à decomposição de Wold¹⁰.

Predição de um passo (One-step prediction)

O problema da predição foi atacado, pela primeira vez, por Kolmogorov¹¹ em 1941 e, em 1942, por Norbert Wiener¹². Sua importância era prever a posição de um avião inimigo para poder derrubá-lo. As coordenadas do avião, em instantes sucessivos, podem ser consideradas um processo estocástico e, talvez, estacionário.

Vamos assumir que o ponto previsto, \hat{x}_t , seja uma função linear dos M pontos do passado

$$\hat{x}_t = \sum_{k=1}^M g_k x_{t-k}$$

Como otimizar os coeficientes g_k para que a predição seja a melhor possível? Vamos impor que a esperança do quadrado da diferença entre o predito e o real seja mínimo. Este é o erro quadrático médio da predição.

$$E^2 = E(x_t - \hat{x}_t)^2 = \underbrace{E(x_t x_t)}_{\phi(0)} - 2 \sum_k g_k \underbrace{E(x_t x_{t-k})}_{\phi(k)} + \sum_{kl} g_k g_l \underbrace{E(x_{t-k} x_{t-l})}_{\phi(k-l)}$$

O m̄nimo pode ser encontrado igualando a zero as derivadas parciais em relação a cada coeficiente, g_k , isto ē

$$\frac{\partial E^2}{\partial g_m} = 0 \quad (m = 1, 2, \dots, M)$$

Encontramos assim um sistema de equaçães chamado de Wiener-Hopf

$$\sum_{k=1}^M g_k \phi(k-m) = \phi(m) \quad (m = 1 \dots M)$$

Tal sistema pode ser obtido por outro raciocı́nio, devido a Yule⁹ e Walker¹³.

Dada uma s̄erie de tempo representada por

$$x_t = \sum_{k=1}^M g_k x_{t-k} + \epsilon_j$$

multiplicamos por x_{t-k} e tomamos a esperança. Tomando os v̄arios valo res do indice k, recaimos no sistema de Wiener-Hopf.

Conhecendo a auto-correlação para diversos "lags", pode mos resolver o sistema e obter os coeficientes g_k . As auto-correlaçães podem ser calculadas a partir dos dados, como no m̄todo de Blackman-Tu key. Este m̄todo de calcular o espectro de pot̄ncia, usando os coefi cientes g_k , calculados a partir das auto-correlaçães obtidas dos da

dos, chama-se Método de Levinson. Entretanto, o método de Burg, que será exposto a seguir, apresenta mais vantagens, como será visto.

Conhecidos os g_k , que minimizam o erro, podemos calcular este erro mínimo quadrático, que chamaremos de potência do erro de previsão, P_M . Substituindo o sistema de Wiener-Hopf na expressão do erro temos

$$P_M = E^2_{\text{MIN}} = \phi(0) - \sum_{k=1}^M g_k \phi(k)$$

Vamos, agora, transformar a notação, passando a usar matrizes e vetores, o que permitirá uma visão melhor das equações, pois simplifica a notação, permitindo escrever as equações de forma mais compacta.

Inicialmente, construiremos o vetor "filtro do erro de predição", de dimensão $M + 1$

$$\tilde{\Gamma}_M = \begin{bmatrix} \Gamma_0 \\ \Gamma_1 \\ \vdots \\ \Gamma_M \end{bmatrix} \quad \text{onde} \quad \begin{cases} \Gamma_0 = 1 \\ \Gamma_k = -g_k \quad k = 1, 2 \dots M \end{cases}$$

A matriz de auto-correlação $\tilde{\Phi}$ é composta dos $\phi(k)$, e é simétrica em relação à diagonal.

$$\tilde{\Phi}_M = \begin{bmatrix} \phi(0) & \phi(1) & \dots & \phi(M) \\ \phi(1) & \phi(0) & \dots & \phi(M-1) \\ \vdots & \vdots & \ddots & \vdots \\ \phi(M) & \phi(M-1) & \dots & \phi(0) \end{bmatrix}$$

As matrizes deste tipo são chamadas matrizes Toeplitz, cujas várias propriedades serão apresentadas em seguida.

Finalmente, temos o vetor "potência do erro de previsão", dado por

$$\tilde{P}_M = \begin{bmatrix} P_M \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Substituindo os novos coeficientes Γ_k , no sistema de Wiener-Hopf e na fórmula de P_M , obtemos

$$\sum_{k=0}^M \Gamma_k \phi(k-m) = 0$$

$$\sum_{k=0}^M \Gamma_k \phi(k) = P_M$$

Este sistema pode se escrito na forma matricial

$$\begin{bmatrix} \phi(0) & \phi(1) & \dots & \phi(M) \\ \phi(1) & \phi(0) & & \vdots \\ \vdots & & \ddots & \vdots \\ \phi(M) & \dots & \dots & \phi(0) \end{bmatrix} \begin{bmatrix} \Gamma_0 \\ \Gamma_1 \\ \vdots \\ \Gamma_M \end{bmatrix} = \begin{bmatrix} P_M \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

e expresso, de forma compacta

$$\tilde{P}_M \tilde{\Gamma}_M = \tilde{P}_M$$

Ele permite encontrar os coeficientes do erro de previsão e o valor da potência do erro de previsão, sendo conhecidas as au

to-correlações. O cálculo de um sistema de equações cuja matriz é Toeplitz, pode ser bastante simplificado se for adotado um método especial, o de Recursão de Levinson.

Método de Recursão de Levinson

Como a matriz Toeplitz é simétrica em relação à diagonal, se invertermos o vetor $\begin{bmatrix} 0 \\ \vdots \\ M \end{bmatrix}$ da equação matricial, ficará invertido o vetor $\begin{bmatrix} P_0 \\ \vdots \\ P_M \end{bmatrix}$:

$$\begin{bmatrix} \phi_0 & & & \phi_M \\ & \ddots & & \\ & & \ddots & \\ \phi_M & & & \phi_0 \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ M \end{bmatrix} = \begin{bmatrix} P_0 \\ \vdots \\ P_M \end{bmatrix} \Rightarrow \begin{bmatrix} \phi_0 & & & \phi_M \\ & \ddots & & \\ & & \ddots & \\ \phi_M & & & \phi_0 \end{bmatrix} \begin{bmatrix} M \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \vdots \\ 0 \\ P_M \end{bmatrix}$$

Isto corresponde à previsão para frente (forward), caso da primeira equação matricial, e para atrás (backward), caso da segunda equação matricial. Por exemplo, com os pontos x_1, x_2, x_3 podemos prever "forward" encontrando \hat{x}_4 , ou "backward" encontrando \hat{x}_0 .



Esta propriedade será aproveitada agora. Em vez de resolver diretamente o sistema, ele será resolvido passo a passo. Admitindo-o resolvido para a dimensão M , existe uma maneira de calcular diretamente a solução, para a dimensão $M + 1$.

O sistema de ordem $M + 1$ pode ser escrito

$$\begin{bmatrix} \mathbb{1}_{M+1} \\ \vdots \\ \vdots \\ \vdots \end{bmatrix} \begin{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ \vdots \end{bmatrix} \\ \begin{bmatrix} M \\ \vdots \\ \vdots \\ \vdots \end{bmatrix} \\ \begin{bmatrix} M+1 \\ \vdots \\ \vdots \\ \vdots \end{bmatrix} \end{bmatrix} = \begin{bmatrix} P_{M+1} \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

O sistema de ordem M é da forma

$$\begin{bmatrix} \mathbb{1}_M \\ \vdots \\ \vdots \\ \vdots \end{bmatrix} \begin{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ \vdots \end{bmatrix} \\ \begin{bmatrix} M \\ \vdots \\ \vdots \\ \vdots \end{bmatrix} \end{bmatrix} = \begin{bmatrix} P_M \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Vamos tomar agora um vetor constituído pelo vetor filtro do erro de previsão, de ordem M , acrescido de um zero. Fazemos agora dois sistemas:

$\mathbb{1}_{M+1}$ multiplicada por $\begin{bmatrix} \sim \\ M \end{bmatrix}$ acrescido de um zero, e pelo mesmo $\begin{bmatrix} \sim \\ M \end{bmatrix}$ invertido, neste caso depois de convenientemente particionada a matriz

$\mathbb{1}_{M+1}$

$$\begin{bmatrix} \mathbb{1}_M & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ \vdots \end{bmatrix} \\ \begin{bmatrix} M \\ \vdots \\ \vdots \\ \vdots \end{bmatrix} \\ 0 \end{bmatrix} = \begin{bmatrix} P_M \\ 0 \\ \vdots \\ 0 \\ \Delta \end{bmatrix}$$

$$\begin{bmatrix} \phi_0 & \vdots & \phi_1 & \dots & \phi_{M+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ \begin{bmatrix} M \\ \vdots \\ \vdots \\ \vdots \end{bmatrix} \\ 0 \end{bmatrix} = \begin{bmatrix} \Delta \\ \vdots \\ 0 \\ \vdots \\ P_M \end{bmatrix}$$

Se do primeiro sistema subtrairmos o segundo multiplicado por Δ/P_M , encontraremos um sistema da mesma forma de ordem $M + 1$.

Temos

$$\Gamma_{M+1}^i = - \Delta/P_M$$

$$P_{M+1} = P_M (1 - \Gamma_{M+1}^i{}^2)$$

$$\tilde{\Gamma}_{M+1} = \begin{bmatrix} \Gamma_0 \\ \vdots \\ \Gamma_M \\ \vdots \\ 0 \end{bmatrix} + \Gamma_{M+1}^i \begin{bmatrix} 0 \\ \vdots \\ \Gamma_M \\ \vdots \\ 0 \end{bmatrix}$$

A última fórmula expressa o algoritmo de Levinson. O coeficiente Γ_{M+1}^i é chamado "coeficiente de reflexão" ou "coeficiente de correlação parcial".

Então, cada vetor $\tilde{\Gamma}_{M+1}$ pode ser calculado a partir de $\tilde{\Gamma}_M$, se for conhecido Γ_{M+1}^i . Assim, podemos sucessivamente calcular $\tilde{\Gamma}_1$, $\tilde{\Gamma}_2$, $\tilde{\Gamma}_3 \dots \tilde{\Gamma}_M$, $\tilde{\Gamma}_{M+1}$. Observe que a potência do erro, P_M , decresce a cada ordem sucessiva. Como $\Gamma_{M+1}^i{}^2$ é necessariamente positivo ou nulo, então $P_{M+1} \leq P_M$. Isto é intuitivo, pois com um maior número de coeficientes do filtro de previsão, o erro médio é menor.

Resta um problema. Como obter os coeficientes de reflexão Γ_{M+1}^i ? No método de Levinson, em que são calculados os $\phi(K)$, cada coeficiente de reflexão pode ser calculado por

$$\Gamma_{M+1}' = - \frac{\Delta}{P_M} = - \frac{1}{P_M} (\phi_1 \Gamma_M + \phi_2 \Gamma_{M-1} + \dots + \phi_{M+1} \Gamma_0)$$

Burg, porém, idealizou um maneira de achar os Γ' que prescinde dos $\phi(k)$, e o cálculo pode ser feito diretamente dos dados. E, além disso, otimiza o resultado, maximizando a plausibilidade ("likelihood"), como demonstrou Lacoss¹⁴.

Método de Burg (1967)

Em cada calculo, em que \bar{e} é acrescida de um a ordem do filtro de predição, iremos minimizar P_{M+1} em relação ao Γ_{M+1}' ou seja

$$\frac{\partial P_{M+1}}{\partial \Gamma_{M+1}'} = 0$$

onde vamos tomar a potência do erro como a média entre o erro de previsão para frente, F (forward) e o da previsão para atrás, B (backward)

$$P_M = \frac{1}{N-M} \sum_{k=1}^{N-M} \frac{1}{2} \left[\underbrace{\left(\sum_{\ell=0}^M \Gamma_{\ell} x_{k+\ell} \right)^2}_{F_k} + \underbrace{\left(\sum_{\ell=0}^M \Gamma_{N-\ell} x_{k+\ell} \right)^2}_{B_k} \right]$$

$$\Gamma_{M+1}' = - 2 \frac{\sum F_k B_k}{\sum (F_k^2 + B_k^2)}$$

O coeficiente de reflexão, assim calculado, garante que o erro de previsão no passo seguinte é mínimo. Observar que só usamos os pontos da série de tempo que já temos (isto é os dados).

Lacoss¹⁴ demonstrou que a minimização de P_M é equivalente a maximizar a probabilidade conjunta dos dados, desde que conhecidos os g_k e a variância dos resíduos em relação aos g_k .

Espectro de potência

Obtidos os coeficientes que descrevem o processo AR,

$$x_t = \sum g_k x_{t-k} + \epsilon_t$$

podemos agora obter o espectro de potência.

Tomando a transformada Z da expressão acima, temos

$$X(z) = G(z)X(z) + E(z)$$

$$X(z) = \frac{E(z)}{1-G(z)} = \frac{E(z)}{\Gamma(z)}$$

Para obter o espectro de potência, podemos multiplicar pelo conjugado e substituir z por $\exp(-2\pi i f \Delta t)$. Com isso, obtemos a fórmula de Fejer-Riesz

$$P(f) = \frac{P_M \cdot \Delta t}{\left| \sum_{k=0}^M \Gamma_k e^{-2\pi i f \Delta t k} \right|^2}$$

onde Δt é o tempo de amostragem da série de tempo.

Obtivemos, assim, o espectro de potência pelo método da entropia máxima. Ele é o espectro da previsão da série toda. Assim, mesmo que seja conhecido um número pequeno de pontos da série de tempo, se eles forem tais que permitam conhecer ou prever o resto da série, o espectro resultante terá melhor resolução do que o espectro obtido pelos métodos de auto-correlação ou pelo periodograma, pois estes usam apenas os valores da série de tempo conhecidos, sem predizerem outros pontos.

Auto-correlação

Uma série de tempo, com N pontos, permite estimar a auto-correlação para um "lag" máximo, correspondente ao número de pontos da série de tempo. A série de previsão, como um processo AR, permite estimar a auto-correlação para "lags" maiores do que o número de pontos. Multiplicando a expressão do processo AR por x_{t-k} , e tomando a esperança da expressão toda, obtemos

$$\phi(N+l) = - \sum_{j=1}^N \Gamma_j \phi(N+l-j)$$

o que permite extrapolar os $\phi(k)$ além do lag N-1.

Roteiro para o método da entropia máxima

Para calcular o espectro de potência de uma série de tempo devemos seguir os passos indicados no diagrama 1.

Inicialmente calculamos a potência do sinal P_0 e o vetor Γ_0 , de dimensão 1, que tem valor [1]. Esta parte representa a inicialização do algoritmo de Levinson. Entramos depois em um "loop", para o cálculo do vetor do filtro de previsão, indo para ordens progressivamente maiores. Em seguida, calculamos o coeficiente de reflexão Γ^1 que é introduzido no algoritmo de Levinson para acharmos os coeficientes do filtro de ordem $m+1$. Então a potência do erro de previsão pode ser calculada a partir do P_{m-1} . A ordem do vetor filtro de previsão pode ser qualquer, e existem métodos para sua otimização. Em geral podemos tomar um filtro de ordem $M=N/2$, onde N é o número de pontos da série de tempo. Tomar $M=N/2$ não é o melhor, mas é bastante conveniente. Com os Γ_k , pela fórmula de Fejer-Riesz é possível calcular $P(f)$.

Como o processo necessariamente calcula os P_m sucessivos, pela observação dos valores de P_m para cada m , podemos ver quando um aumento de m não diminui sensivelmente o valor da potência do erro, ou então, se P_m já é suficientemente próximo do $\lim_{m \rightarrow \infty} P_m$. A relação P_m/P_0 também indica qual m dá suficiente precisão na previsão. Akaike¹⁵ desenvolveu uma maneira de calcular o melhor m . Este corresponderia ao mínimo do "Final Prediction Error" (FPE)

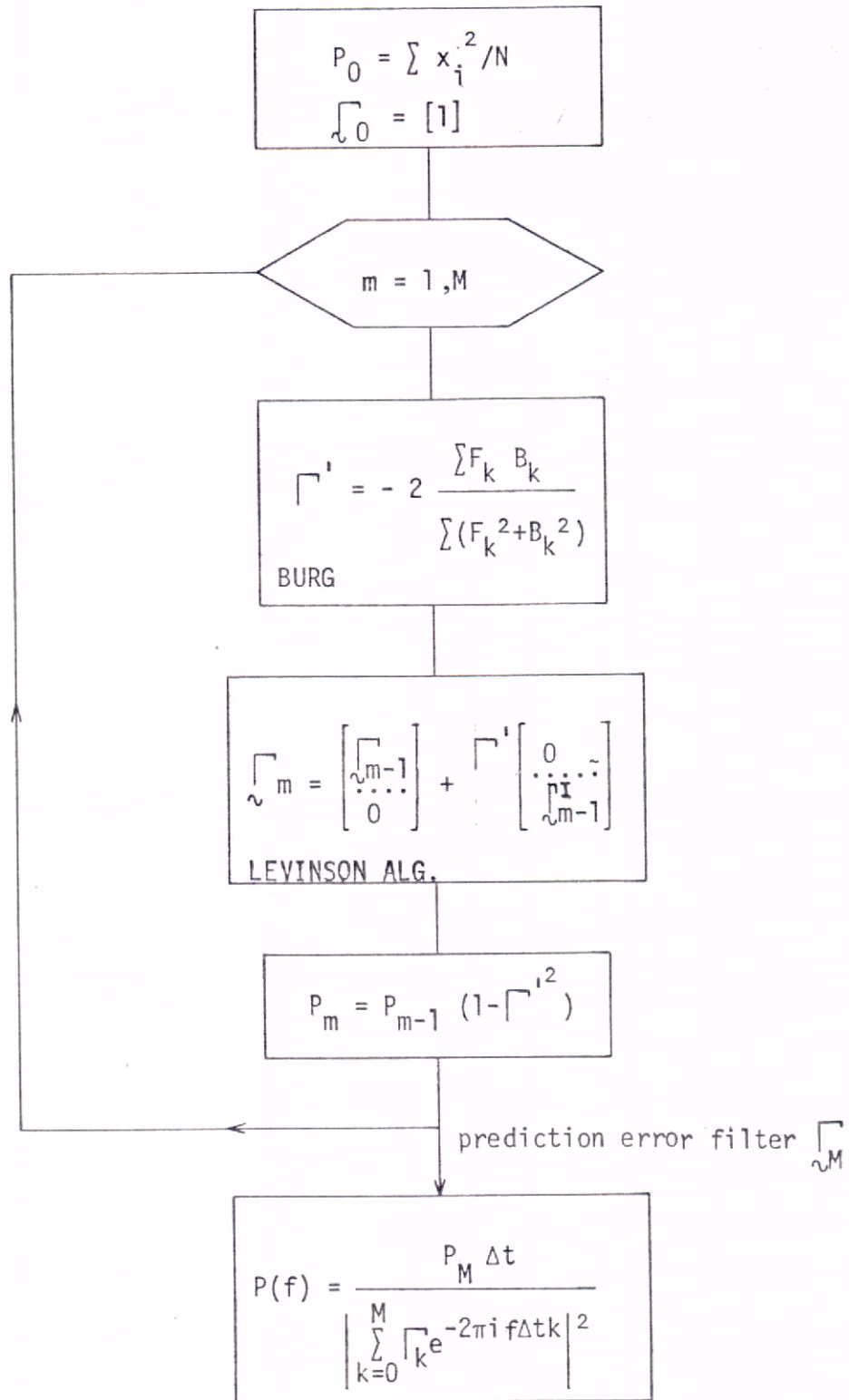


Diagrama 1 - Roteiro para o cálculo do espectro de potência pelo MEM.

$$\text{FPE} = \frac{N + (M + 1)}{N - (M + 1)} P_{11}^2$$

o FPE pode ser calculado durante o "loop".

No apêndice encontra-se listada uma sub-rotina nas linguagens ALGOL e FORTRAN que, a partir da série de tempo, calcula os coeficientes de reflexão, as potências do erro e o vetor dos coeficientes do filtro de predição.

Aplicações

O diagrama 2 mostra em que o método da máxima entropia pode ser empregado.

O método permite calcular os coeficientes do filtro de predição. Estes coeficientes podem ser usados em diversas aplicações. *Espectro de Potência P(f)*. Permite melhor resolução do que os métodos convencionais. Permite obter melhor resultado quando o período da componente em que se está interessado é próximo do intervalo total de tempo dos dados. Permite obter um melhor resultado quando temos poucos pontos da série de tempo. Para bons resultados não é necessário a aplicação de filtros, como o "hamming", pois o método dispensa pré-filtragem. Uma pequena adaptação permite calcular o *espectro cruzado* ("cross spectrum"). O vetor $\hat{\alpha}$ pode ser usado para *prever* valores de uma série de tempo. Por exemplo a série das manchas solares permite calcular a previsão para anos futuros com bastante precisão. A predictabilidade de uma série de tempo está ligada ao limite P_{∞} . Ou seja, numa *decomposição de Wold*,

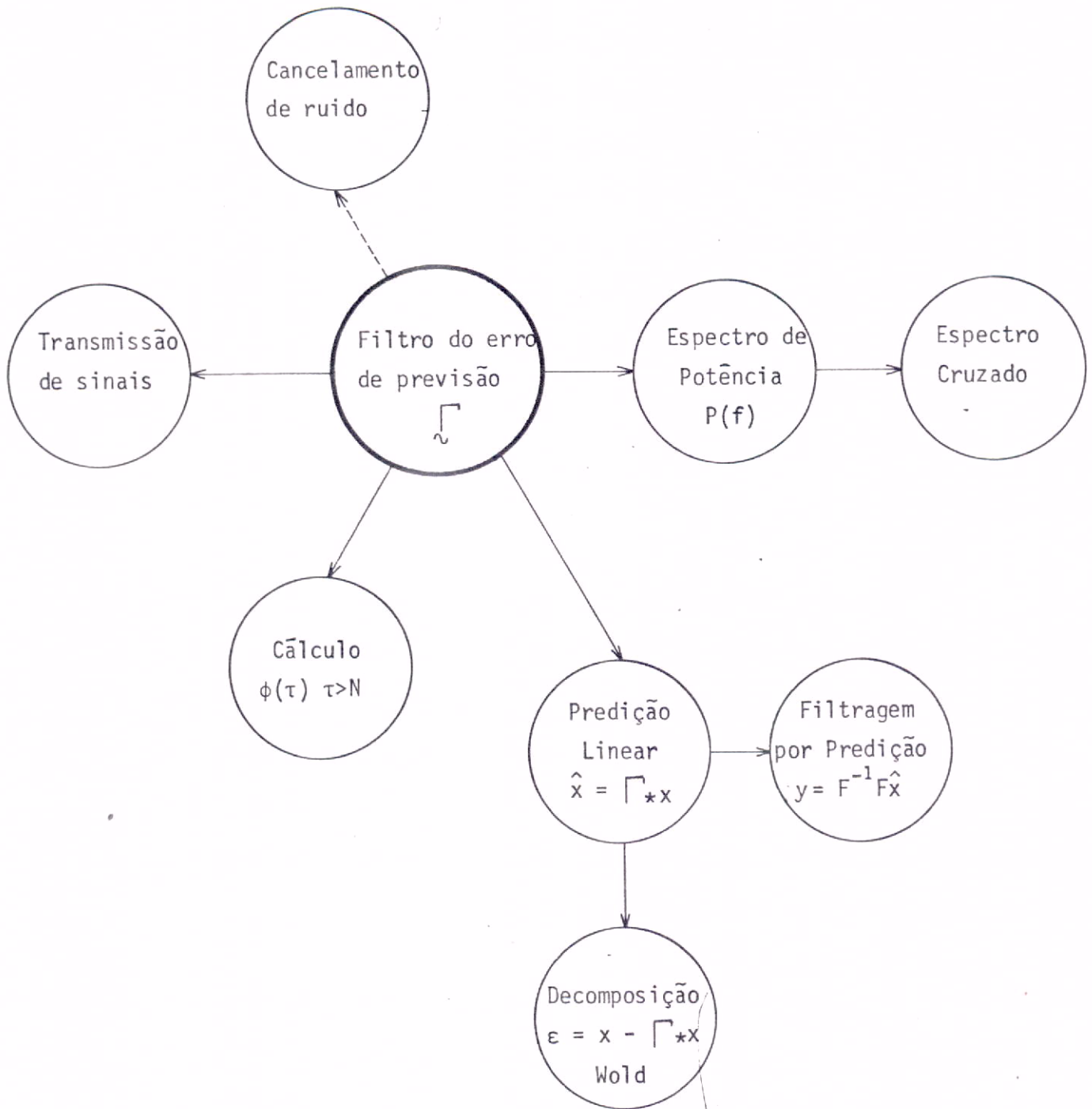


Diagrama 2 - Aplicações do MEM

qual a potência do sinal determinístico em relação ao ruído. Uma filtragem de uma série de tempo com poucos dados pode ser difícil. Porém, podemos aplicar a transformada de Fourier na série estendida pela predição, filtrar este espectro e reconstruir a série aplicando a transformação inversa. $y = F^{-1}F\hat{x}$. Como já vimos, podemos também calcular os coeficientes de auto-correlação $\phi(\tau)$ para $\tau > N$. Podemos eliminar o ruído de uma série de tempo tomando a parte determinística. A transmissão de sinais pode então ser feita, transmitindo-se apenas os têrmos de resíduo ("Linear Prediction Vocoder").

Exemplos Práticos

Os exemplos de aplicação mais espetaculares são quando o MEM é aplicado em trechos de senoide com extensão menor do que o período. Na primeira coluna da Figura 1 temos o sinal original, na coluna seguinte o espectro pelo método Cooley-Tukey, sem pré-filtragem, e em seguida o espectro pelo MEM¹⁷. Mesmo 57% do período permite obter uma boa curva espectral com o MEM. O pequeno trecho da senoide permite prever o resto da curva, da qual é calculado o espectro.

Na Figura 2 temos a comparação entre o Blackman-Tukey e o MEM⁵. A série de tempo foi gerada pela soma de 4 senoides acrescidas por um ruído de amplitude 1/2. As frequências são respectivamente 4; 64; 65 e 124,5 Hz. A série tem 257 pontos, foram tomados 100 "lags" e um filtro com 100 coeficientes. Compare os dois métodos para as altas e baixas frequências e observe que a Blackman-Tukey não consegue separar as duas frequências próximas.

O trabalho de Radoski et al⁵ apresenta extensa comparação entre os vários métodos.

Na Figura 3 temos cálculo do espectro de potência pelo MEM, para uma série de 16 pontos dados no gráfico. Para comparação é apresentado o espectro pelo Cooley-Tukey, em estrelinhas e, em linha contínua, o calculado pelo MEM. Abaixo temos uma tabela dos coeficientes de reflexão representados por (G) e o respectivo P_M e a porcentagem do P_M em relação ao P_0 . Mais abaixo temos a listagem dos Γ_k representados por GAMMA. Pela tabela P%, um filtro com mais de 4 termos, e erro médio quadrático da previsão em relação à potência do sinal é menor do que 0,003%. Portanto, um filtro com 4 coeficientes já é suficiente para uma boa representação do sinal. Observe como o pico do sinal é bem mais estreito com o MEM. O sinal simulado é uma cosenoide truncada. Se desejarmos interpolar pontos do espectro quando o cálculo da fórmula de Fejer-Riesz é feito usando FFT, basta acrescentar zeros aos coeficientes Γ_k .

Uma das séries mais estudadas pelo MEM é a série dos números de manchas solares. Pode-se encontrar referências em Cohen e Lintz¹⁶ e Radoski et al⁵.

Em nosso caso particular, o método de análise foi aplicado aos dados citados no trabalho¹ "Variação anual do campo Magnético Terrestre" por N.B. Trivedi, I.J. Kantor e J.M. da Costa. As Figuras 4 e 5 apresentam os resultados. Foram analisados as médias mensais do campo geomagnético H em Vassouras no período de 1957-1968 (114 meses).

Os dados apresentam uma variação linear de aproximadamente 60 gammas/ano. Esta tendência secular foi removida dos dados por ajuste dos mínimos quadrados a uma linha reta. Foram usados dois métodos para análise espectral dos dados. O primeiro foi o da auto-correlação e, o segundo, foi o MEM. Para as nossas finalidades não estão interessando as amplitudes, por isso podemos comparar apenas a resolução, que é observada pela superposição de ambos resultados.

Comentários Finais

Um estudo mais rigoroso do MEM envolveria conceitos sobre processos estocásticos estacionários, entropia, processos AR e MA ("moving average" isto é, a média móvel), modelos "all-pole", espaços de Hilbert etc.

Curiosamente, durante a dedução do MEM para o cálculo do espectro de potência não foi mencionado o conceito de entropia, apesar do nome Método da Máxima Entropia. É que a dedução das fórmulas, a partir do conceito da entropia, é bem mais complexa do que a que foi aqui apresentada. O problema é formulado da seguinte maneira: Se $P(f)$ é o espectro de potência de um processo estacionário, a entropia é dada por

$$H = \int_{-f_N}^{f_N} \log P(f) df \quad f_N = \text{frequência de Nyquist.}$$

que deve ser maximizada sob a condição que tenha as auto-correlações expressas por

$$\int_{-f_N}^{f_N} P(f) e^{-2\pi i f \Delta t n} df = \phi_n \quad - N \leq n < N$$

Então queremos saber qual o $P(f)$ que maximiza H , nas condições dadas. Usando cálculo variacional obtemos o sistema de Wiener-Hopf. Vale notar que, pelo uso dos conceitos de entropia, as razões lógicas do método ficam mais claras.

O MEM é bem recente pelo que existem muitos assuntos ainda não esclarecidos como, por exemplo: Qual o tamanho ideal para o filtro do erro de predição? Qual a variância e erro do espectro de potência obtido? Qual o erro na posição dos máximos?

Apêndice A

Apresentamos um diagrama de bloco para cálculo dos coeficientes gamma, coeficientes de reflexão e potência do erro de predição¹⁸. O programa é sob a forma de sub-rotina. Em FORTRAN foi testado num computador CDC3300 e em ALGOL, foi testado num Burroughs B-6700.

Os dados de entrada são o "array" X, de dimensão N, e o MM, que é a dimensão do "array" G.

O resultado de saída é o "array" G e o valor de P.

X = série de tempo

G = filtro do erro de predição

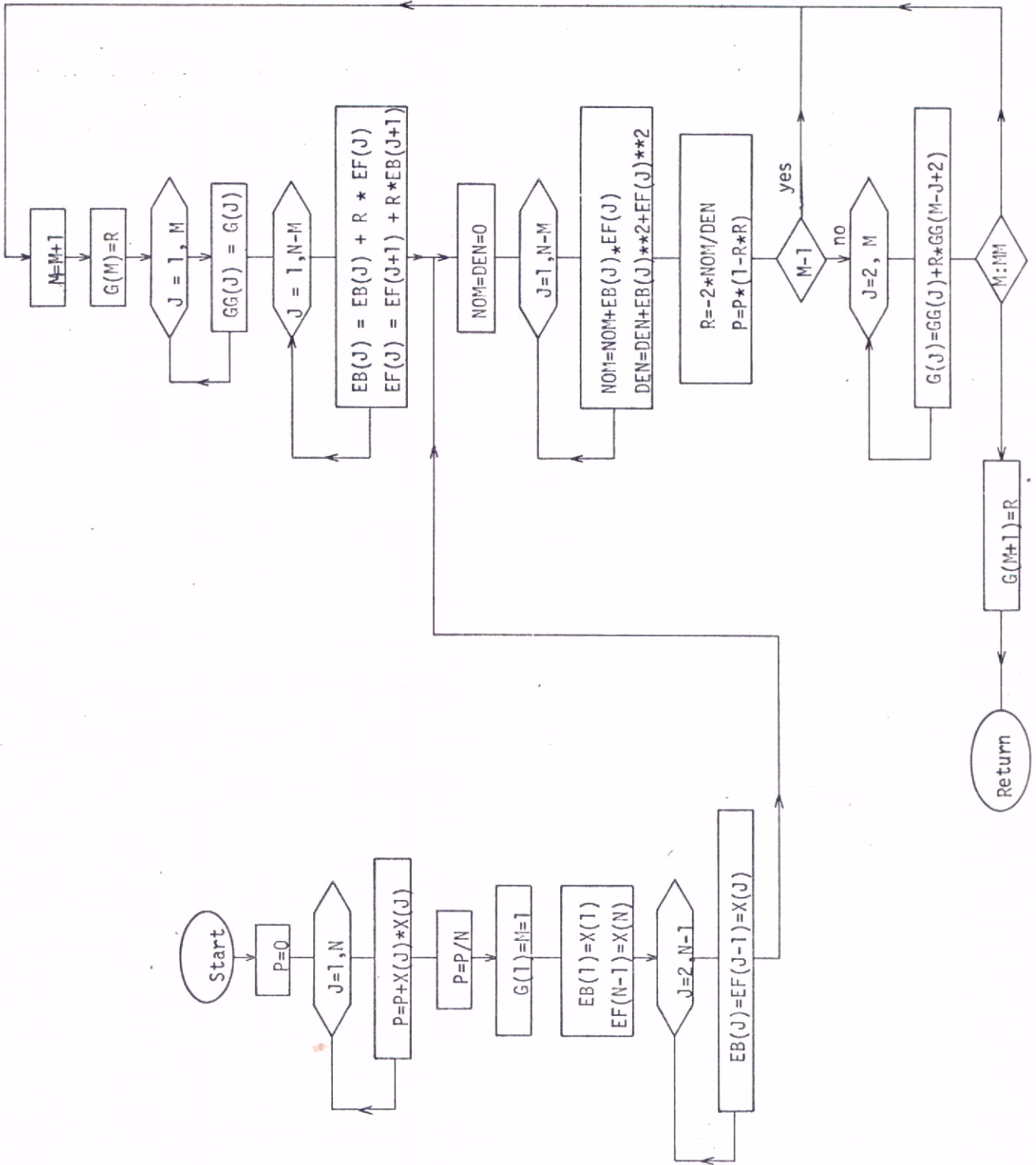
P = corresponde ao P_M .

Nos cálculos, ER e EF, são os erros "forward" e "backward". R é o coeficiente de reflexão. DEN e NOM são o denominador e o numerador da fórmula que dá o coeficiente de reflexão.

Num computador CDC-3300, usando o MEM para séries de tempo complexas, foi obtido o seguinte tempo de execução da sub-rotina:

$$\text{Tempo (seg)} \approx 119 \cdot N^{2,03} \cdot 10^{-6} \quad (10 \leq N \leq 128)$$

onde N é o número de pontos da série de tempo e MM é N/2. Para N = 20 o tempo é 0,05 seg e para N = 128 é 2,3 seg.




```

FILE SAIDA (KIND = PRINTER);
%--PREDICTION ERROR FILTER - BURG
PROCEDURE MEM (G,MM,X,N,P);
VALUE MM, N; INTEGER MM,N; REAL P; ARRAY G[0], X[1];
BEGIN INTEGER I,J,NM; LABEL LØØP, IN; ARRAY EB,EF[1:N], GG[0:M];
    DEFINE ATE = STEP 1 UNTIL *;
    P:=0; FØR J:=1 ATE N DØ P:=*+X[J]*X[J]; P:=*/N;
    WRITE (SAIDA, <"bM=bb0", X17,"bP=", E14.6>, P);
    G[0]:=0; M:=1; EB[1]:=X[1]; EF[N-1]:=X[N];
    FØR J:=2 ATE N-1 DØ EB[J]:=EF[J-1]:=X[J];
    NM:=N-M; GØ TØ IN;
LØØP: G[M:=*+1]:=R; FØR J:=1 ATE M DØ GG[J]:=G[J];
    FØR J:=1 ATE NM DØ BEGIN
        EB[J]:=*+R*EF[J];
        EF[J]:=EF[J+1]+R*EB[J+1]; END;
IN: NØM:=DEN:=0;
    FØR J:=1 ATE NM DØ BEGIN
        NØM:=*+EB[J]*EF[J];
        DEN:=*+EB[J]*EB[J]+EF[J]*EF[J]; END;
    R:=-2*NØM/DEN; P:=*(1-R*R).
    WRITE (SAIDA, <"bM=",I3,"bR=",E14.6,"bP=",E14.6>,M,R,P);
    IF M=1 THEN GØ TØ LØØP;
    FØR J:=1 ATE M DØ G[J]:=GG[J]+R*GG[M-J];
    IF M<MM THEN GØ TØ LØØP;
    G[M+1]:=R
END MEM;

```

```

SUBROUTINE MEM (G,MM,X,N,P)
C  MAXIMUM ENTROPY METHOD - I.J. KANTOR - 040776 - CDC 3300
C  N IS LIMITED TO 128
REAL NØM
DIMENSION G(1),X(1),EB(128),FF(128),GG(128)
N1 = N-1
P = 0
DØ 10 J = 1,N
10  P = P + X(J)*X(J)
P = P/N
M = 1
G(1) = 1
EB(1) = X(1)
EF(N-1) = X(N)
WRITE (61,15)P
15  FØRMAT (6HbM = bb0,17X,3HbP = E14.6)
DØ 20 J = 2,N1
20  EB(J) = EF(J-1) = X(J)
NM = N-M
GØ TØ 50
30  M = M+1
NM = N-M
G(M) = R
DØ 35 J = 1,M
35  GG(J) = G(J)
DØ 40 J = 1,NM
EB(J) = EB(J) + R*EF(J)
40  EF(J) = EF(J+1) + R*EB(J+1)
50  DEN = NØM = 0
DØ 60 J = 1,NM
NØM = NØM + EB(J)*EF(J)
60  DEN = DEN + EB(J)*EB(J) + EF(J)*EF(J)
R = - 2*NØM/DEN
P = P*(1 - R*R)

```

```
WRITE (61,70) M,R,P
70  FORMAT (3HbM = I3, 3HbR = E14.6, 3HbP = E14.6)
    IF (M.EQ:1) GØ TØ 30
    DØ 80 J = 2,M
    I = M - J + 2
80  G(J) = GG(J) + R*GG(I)
    IF(M.LT: MM) GØ TØ 30
    G(M+1) = R
    RETURN
    END
```


Referências

- 1 - Trivedi, N.B., I.J. Kantor e J.M. da Costa, Variação Anual do Campo Magnético Terrestre, trabalho 12 F.2, SBPC, Fortaleza, 1977.
- 2 - Burg, J.P., Maximum entropy spectral analysis, trabalho apresentado na 37^a Reunião, Soc. of Explor. Geophys., Oklahoma City, Okla., Oct. 1967.
- 3 - Smylie, D.E., G.K.C. Clarke and T.J. Ulrych, Analysis of Irregularities in the Earth's Rotation, in Methods in Computational Physics, vol.13, 391-430, B.A. Bolt., ed, New York, Academic Press 1973.
- 4 - Makhoul, J., Linear Prediction: A Tutorial Review, Proc. IEEE, 63, N.4, 561-580, April 1975.
- 5 - Radoski, H.R., P.F. Fougere and E.J. Zawalick, A Comparison of Power Spectral Estimates and Application of the Maximum Entropy Method, J. Geophys. Res, 80, N^o4, 619-625, Feb 1975.
- 6 - Ulrych, T.J. and T.N. Bishop, Maximum Entropy Spectral Analysis and Autoregressive Decomposition, Reviews of Geophysics and Space Physics, 13, N^o1, 183-200, Feb 1975.
- 7 - Ball, J.A., Computations in Radio-Frequency Spectroscopy, in Computational Physics, vol.14, B. Alder. ed, New York, Academic Press. 1975.
- 8 - Blackman, R.B. and J.W. Tukey, The Measurement of Power Spectra, Dover Publications, Inc., New York, 1958.

- 9 - Yule, G.U., On a method of investigating periodicities in disturbed series, with special reference to wolfer's sunspot numbers, Phil. Trans. Roy. Soc. London, Ser. A, 226, 267-298, 1927.
- 10 - Wold, H., A Study in the Analysis of Stationary Time Series, Uppsala: Almqvist and Wihells, 1938, pp 214.
- 11 - Kolmogorov, A, "Interpolation und Extrapolation Stationären Zufälligen Folgen", Bull. Acad. Sci, U.S.S.R., Ser. Math, vol.5, pp.3-14, 1941.
- 12 - Wiener, N., Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications. Cambridge, Mass.: MIT. Press, 1949.
- 13 - Walker, G., On the periodicity in series of related terms, Proc. Roy. Soc. London, Ser. A, 131, 518-532, 1931.
- 14 - Lacoss, R.T., Data adaptive spectral analysis methods, Geophysics, 36, 661-675, August 1971.
- 15 - Akaike, H., Fitting autoregressions for prediction, Ann. Inst. Statist. Math., 21, 243-247, 1969.
- 16 - Cohen, T.J. and P.R. Lintz, Long term periodicities in the sunspot cycle, Nature 250, 398-399, 1974.
- 17 - Ulrych, T.J., Maximum Entropy Power Spectrum of Truncated Sinusoids, J. Geophys. Res, 77, 1396-1400, 1972.
- 18 - Andersen, N., On the calculation of filter coefficients for Maximum Entropy Spectral Analysis, Geophysics, 39, 69-72, 1974.

Lista de Figuras

- Fig. 1 - Na primeira coluna temos o sinal original, na segunda coluna o espectro pelo método Cooley-Tukey, sem pre-filtragem e, na terceira coluna, o espectro pelo MEM. A primeira linha representa 1 segundo de uma senoide de 1Hz em 10% de ruído. A segunda linha, 0,57 segundos do mesmo dado¹⁷.
- Fig. 2 - Comparação entre o Blackman-Tukey com 100 "lags" (linha tracejada) o MEM com 100 coeficientes (linha contínua). A série de tempo com 257 pontos foi gerada pela soma de 4 senoides de frequências, 4, 64, 65 e 124,5 Hz, afetadas por um ruído de amplitude $1/2^5$.
- Fig. 3 - Cálculo do espectro de potência pelo MEM, para uma série de 16 pontos do gráfico. As estrelas representam o espectro calculado pelo método Cooley-Tukey e a linha contínua, pelo MEM. Abaixo temos uma tabela dos coeficientes de reflexão (G) e o respectivo P_M e $P_M/P_0 \cdot 100$. Mais abaixo temos a listagem dos Γ_k representados por GAMMA.
- Fig. 4 - Análise espectral da série de tempo dos dados mensais dos valores médios de H, em Vassouras, para o período 1857-1968. Método da auto-correlação (linha tracejada) e MEM (linha contínua) com 50 coeficientes.
- Fig. 5 - Análise espectral de dados do D_{st} equatorial para o período 1957-1968. Método da auto-correlação (linha tracejada) e MEM (linha contínua) com 50 coeficientes.

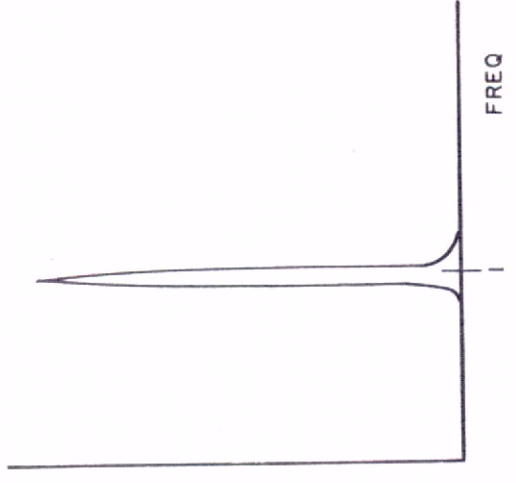
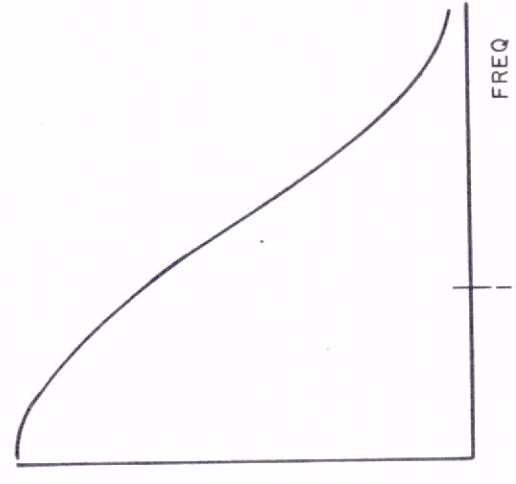
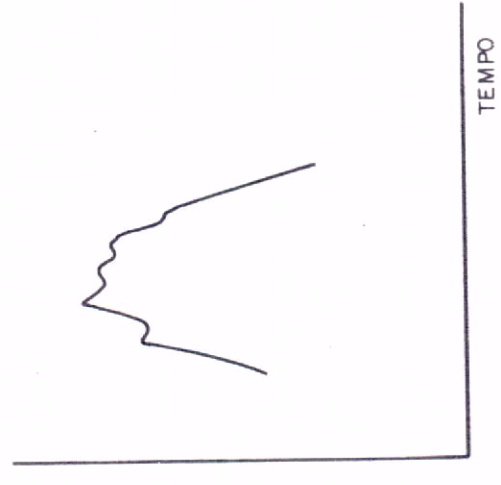
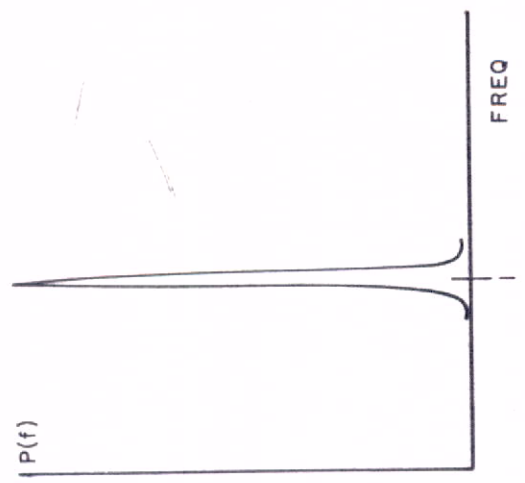
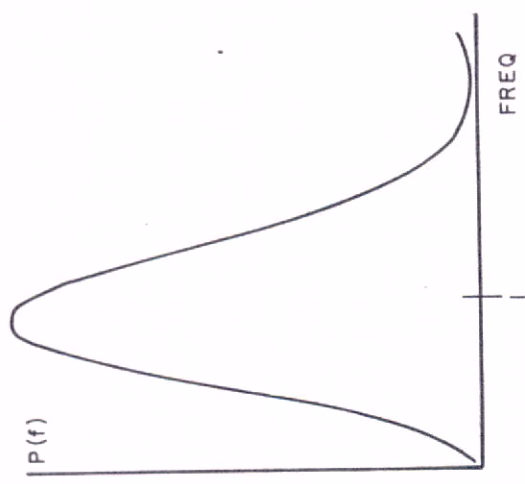
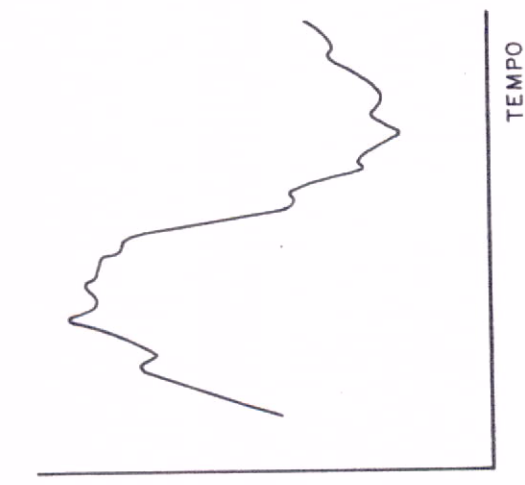


Fig. 1

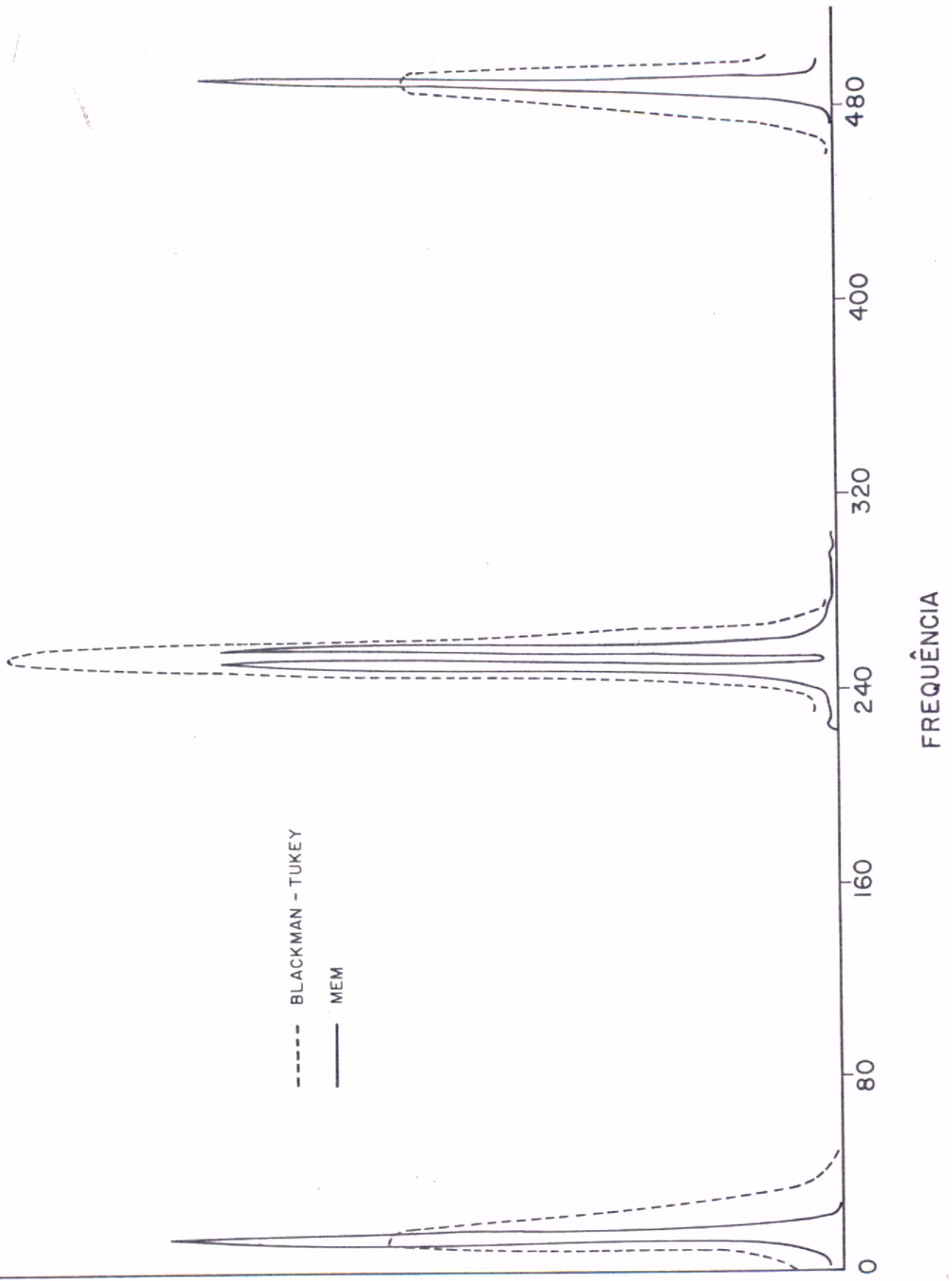
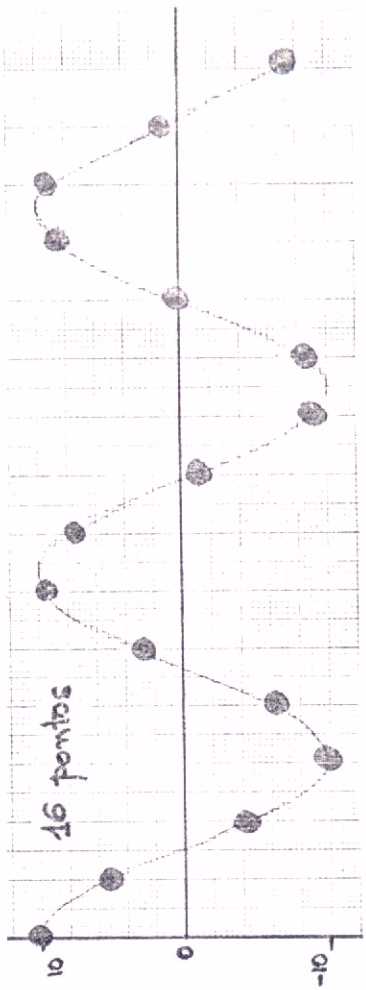


Fig. 2

FFT
DC= 1.12E-01

N FOURIER PERIODO FREQ

1	9.55E-02	63.99	1.56E-02	*****
2	4.63E-02	31.99	3.13E-02	*****
3	8.78E-02	21.32	4.69E-02	*****
4	1.84E-01	15.99	6.25E-02	*****
5	2.02E-01	12.79	7.81E-02	*****
6	6.86E-02	10.66	9.38E-02	*****
7	2.42E-01	9.13	1.09E-01	*****
8	6.50E-01	7.99	1.25E-01	*****
9	1.03E 00	7.10	1.41E-01	*****
10	1.26E 00	6.39	1.56E-01	*****
11	1.22E 00	5.81	1.72E-01	*****
12	9.37E-01	5.32	1.88E-01	*****
13	4.95E-01	4.91	2.03E-01	*****
14	5.23E-02	4.56	2.19E-01	*****
15	2.53E-01	4.26	2.34E-01	*****
16	3.40E-01	3.99	2.50E-01	*****
17	2.29E-01	3.75	2.66E-01	*****
18	2.29E-02	3.55	2.81E-01	*****
19	1.65E-01	3.36	2.97E-01	*****
20	2.31E-01	3.19	3.13E-01	*****
21	1.62E-01	3.04	3.28E-01	*****
22	1.41E-02	2.90	3.44E-01	*****
23	1.33E-01	2.77	3.59E-01	*****
24	1.89E-01	2.66	3.75E-01	*****
25	1.34E-01	2.55	3.91E-01	*****
26	1.03E-02	2.45	4.06E-01	*****
27	1.21E-01	2.36	4.22E-01	*****
28	1.71E-01	2.28	4.38E-01	*****
29	1.21E-01	2.20	4.53E-01	*****
30	9.36E-03	2.12	4.69E-01	*****
31	1.17E-01	2.05	4.84E-01	*****
32	1.65E-01	1.99	4.00E-01	*****
33	1.17E-01	1.93	5.16E-01	*****



G(2)	=	5.119651730E-01	P(1)	=	5.034781333E 01	P(7)	=	73.788
G(3)	=	9.980007437E-01	P(2)	=	3.715123164E 01	P(8)	=	0.000
G(4)	=	4.632244841E-01	P(3)	=	1.484011760E-01			
G(5)	=	9.912541461E-01	P(4)	=	1.165577084E-01			
G(6)	=	2.347842464E-01	P(5)	=	2.029877871E-03			
G(7)	=	5.255519409E-01	P(6)	=	1.917983609E-03			
G(8)	=	8.058265704E-01	P(7)	=	1.388227249E-03			
			P(8)	=	4.867729145E-04			

GAMMA

1	0.00000000E 00	2	-1.876996822E 00	3	3.034876566E 00	4	-1.402059868E 00	5	1.663209638E-03
6	1.962057878E 00	7	-1.328252520E 00	8	8.058265704E-01				

Fig. 3

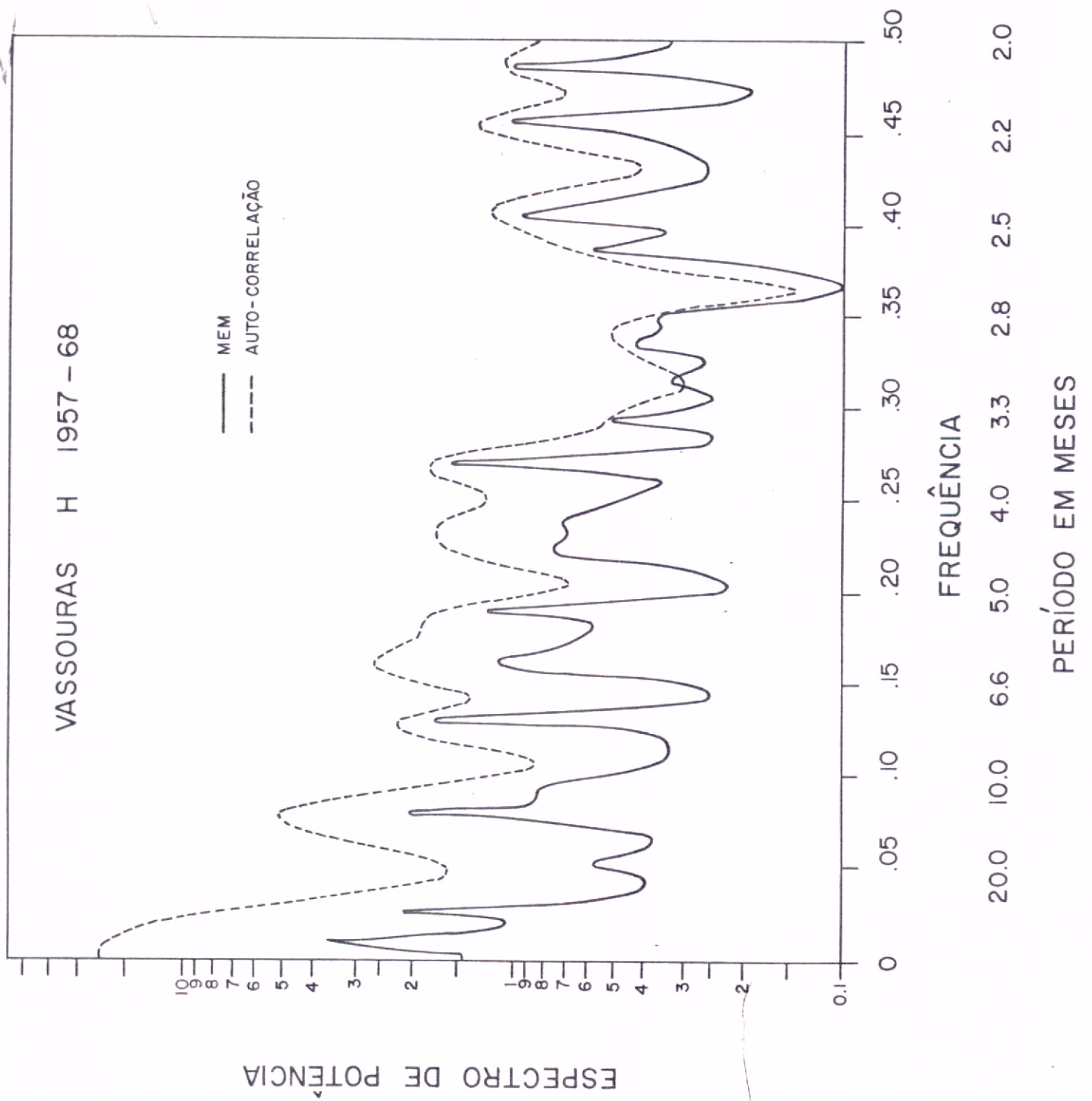


Fig. 4

EQUATORIAL DSt 1957 - 70

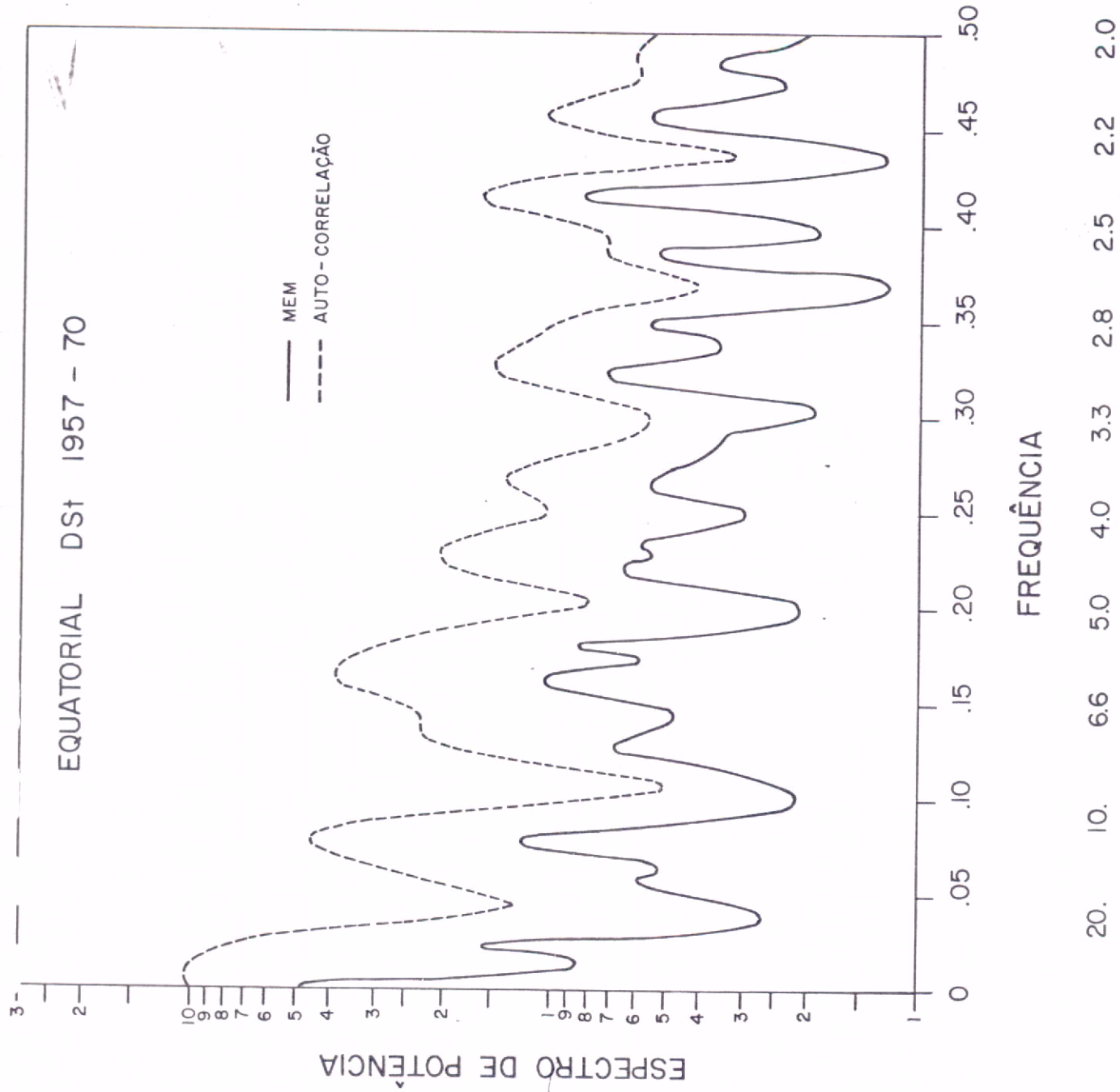


Fig. 5

