

## POWER BIAS IN MAXIMUM ENTROPY SPECTRAL ANALYSIS

JUAN J. EGOZCUE (\*), JAUME PAGÈS (+)

*Several algorithms have been used in maximum entropy spectral analysis. Among them, the standard Burg procedure and the forward-backward least squares method are considered.*

*When the autoregressive model, which is implicit in these estimation methods, is used to simulate the analyzed process, the power or variance of the simulation can differ from power estimated from the signal in several orders of magnitude. This is specially dangerous in simulated studies about the maxima of certain parameters.*

*Burg's method, although not optimal in the least squares sense, produces autoregressive models whose estimated prediction error power is consistent with the estimated total power, while least squares method sometimes do not.*

*Suitable corrections to power bias are described and two numerical examples clear up different situations.*

Keywords: MAXIMUM ENTROPY. AUTOREGRESSIVE. VARIANCE ESTIMATION. SIMULATION. MAXIMUM LIKELIHOOD.

### 1. INTRODUCTION.

The spectral analysis of autoregressive (AR) processes has been the subject of a great number of studies and papers, and has motivated a number of different estimation algorithms, such as Burg method /3/, which we refer as Burg Forward-Backward method (BFB), and the method introduced by Nuttall /12/ and independently by Ulyrch and Clayton /13/, which we will refer as Forward-Backward Least Squares method (FBLS).

The main goal of these studies have been the detection of dominant frequencies and resolution of spectral maxima. The evaluation of the power associated to each spectral peak and the total power or variance of the process are secondary problems. This is possibly because power estimate bias can be corrected by a scale factor, and also because in a large number of applications, the estimation of the power is not an important matter.

However, in simulation problems, as often stated in engineering, the correct estimation

of the power is of central importance. Indeed the extrema of a simulated process are, in many cases, the parameters that the engineer is controlling, those extrema being approximately proportional to the standard deviation of the process.

Clearly we could set up ourselves within the general framework of variance estimation problems, so that, in many situations it would be enough to use for a zero mean stationary process, the unbiased variance estimator given by

$$(1) \hat{P}_0 = \frac{1}{N} \sum_{k=1}^N x_k^2 = \frac{1}{N} X^t X$$

where  $X = (x_1, \dots, x_N)^t$  stands for real data vector, and the superscript  $t$  indicates matrix transposition.

The value of  $\hat{P}_0$  can differ significantly from the true value of the AR process variance  $\sigma_0^2$ . Thus, a scale factor correction may

- (\*) Juan J. Egozcue - Escola T.S. Engenyiers de Camins, Canals i Ports - Universitat Politècnica de Barcelona - Jordi Girona Salgado, 31 - Barcelona 34.

- (+) Jaume Pagès - Facultat d'Informàtica - Universitat Politècnica de Barcelona - Jordi Girona Salgado - Barcelona

- Article rebut el Setembre del 1983.

be necessary.

We will disting at least between two -- common cases when we try to analyze a data - vector X:

A.- The signal X is large with respect to -- the periods we must study. Then,  $\hat{P}_0$  is a good estimate of  $\sigma_0^2$  and the hypothesis -  $\sigma_0^2 = \hat{P}_0$  can be assumed.

B.- The signal contain low frequencies that we do not want to reject and it have in complete cycles of these frequencies. -- Perhaps then,  $\hat{P}_0$  do not approximate well the true value of  $\sigma_0^2$ .

From now on we will discuss the power estima- tion capabilities of the BFB method and the FBLs method.

## 2. BURG MAXIMUM ENTROPY SPECTRAL ANALYSIS.

This estimation algorithm (Burg, /3/) has be- come a standard procedure in many applica--- tions and we shall be interested in the stu- dy of some of its power properties.

We shall be concerned with m-th order autor- regressive processes, shortly AR(m), describ- ed by the well known equations

$$(2) \sum_{k=0}^m g_{mk} x_{j-k} = \epsilon_j \quad g_{m0} = 1$$

$$\sigma_m^2 = E(\epsilon_j^2)$$

where E(.) is the expectation operator and -  $\epsilon_j$  is a white noise process, assumed normal and of zero mean. The coefficients  $g_{mk}$  ---  $k=1,2,\dots,m$  and the noise variance  $\sigma_m^2$  are to be estimated from the data.

The power spectral density of the AR process (2) is given by

$$(3) S_m(f) = \frac{\sigma_m^2 T}{\left| \sum_{k=0}^m g_{mk} e^{-i2\pi fT} \right|^2}$$

where T is the time sampling interval.

The so called normal equations

$$(4) R_s G_s = \sigma_s^2 I_s$$

$$(R_s)_{jk} = R(|j-k|) = E(x_j x_k) \quad j,k=0,1,\dots,s$$

$$G_s^t = (1, S_1, \dots, S_s) \quad I_s^t = (1, 0, \dots, 0)$$

are to be fulfilled by the coefficients of an AR(s) model.

The BFB procedure assumes that  $\sigma_0^2 = \hat{P}_0$  and --- uses Levinson's recurrence

$$(5) g_{sk} = g_{s-1,k} + g_{ss} g_{s-1,s-k}$$

to solve the normal equations (4) for subse--- quent orders  $s=1,2,\dots,m$ . At each step, we -- consider the unbiased estimator of the noise - variance (prediction error power given by

$$(6) \hat{P}_s = \frac{1}{2(N-s)} \sum_{k=1}^{N-s} e_{1s}^2(k) + e_{2s}^2(k)$$

where for any  $j=1,2,\dots,N-s$

$$(7) e_{1s}(j) = \sum_{k=0}^s x_{j+s-k} g_{sk} \quad e_{2s}(j) = \sum_{k=0}^s x_{j+k} g_{sk}$$

and  $\hat{P}_s$  is to be minimized as a function of  $g_{ss}$  at such an step. The errors (7) are often cal- led forward and backward prediction errors.

The minimization of (6) leads to the following expression of  $g_{ss}$

$$(8) g_{ss} = \frac{-2 \sum_{k=1}^{N-s} e_{1,s-1}(k+1) e_{2,s-1}(k)}{\sum_{k=1}^{N-s} \left[ e_{1,s-1}^2(k+1) + e_{2,s-1}^2(k) \right]}$$

which together with (5) complete the Burg ite- rative procedure.

In this fashion, we subject  $\hat{P}_m$  to a minimiza- tion which is constrained by the previous mini- mizations of all  $\hat{P}_s$   $s=1,2,\dots,m-1$ . Thus, BFB is a suboptimal least squares method.

We now can ask about the estimate of  $\sigma_m^2$  that - we should take in (2) in order to simulate -- the process, or in (3) in order to perform a - power analysis. It is a common practice to ta- ke the standard estimate  $P_m$  based in the Levin- son's recurrence as follows:

$$(9) P_m = \hat{P}_0 \prod_{j=1}^m (1 - g_{jj}^2)$$

By doing so we are placing ourselves within case A.

On the other hand BFB has the property that the expected value of  $\hat{P}_m$  is  $P_m$ , when the adequate AR model coefficients are fixed and the hypothesis  $\sigma_0^2 = \hat{P}_0$  is assumed (see appendix). This is to say

$$(10) E(\hat{P}_m | G_m(\text{BFB}), \sigma_0^2 = \hat{P}_0) = P_m$$

so that, if we define

$$(11) \tilde{P}_0 = \hat{P}_m \prod_{j=1}^m [(1 - g_{jj}^2)]^{-1}$$

as another estimator of  $\sigma_0^2$ , then

$$(12) E(\tilde{P}_0 | G_m(\text{BFB}), \sigma_0^2 = \hat{P}_0) = \hat{P}_0$$

Numerical experience shows that  $\tilde{P}_0$  fluctuations about the value  $\hat{P}_0$  can be ignored (see numerical examples).

Relations (10) and (12) do not hold if the AR model is not adequate, specially for orders - higher than 2/3 of the data length.

Therefore, we conclude that the use of  $P_m$  (9) as an estimator of  $\sigma_m^2$  is equivalent to the hypothesis  $\hat{P}_0 = \sigma_0^2$ , often used in case A. -- However no significative difference is obtained if we take  $\hat{P}_m$  as an alternative estimator of  $\sigma_m^2$  for moderate m in BFB.

In case B,  $\hat{P}_0 = \sigma_0^2$  can be a wrong hypothesis and, as a result,  $P_m$  should be a bad estimator of  $\sigma_m^2$ . Also  $\hat{P}_m$  is not a reliable estimator of  $\sigma_m^2$ , because its value is controlled - by (10). Moreover, in section 4 we shall --- show that we cannot correct the estimated -- power  $\hat{P}_0$  in BFB, by using a scale factor, -- without decreasing the likelihood of the --- estimated model. Hence, BFB method should be avoided in case B to study the signal power.

### 3. FORWARD-BACKWARD LEAST SQUARES METHOD.

This method was introduced by Nuttall /12/, -- and also by Ulrych and Clayton /13/. The algorithm presented by Fouguere /9/ added some -- new features, while a recursive procedure designed by Marple /11/ improved some computational aspects of the FBLs method.

The main idea behind FBLs is to minimize  $\hat{P}_m$  - as a function of the AR(m) model coefficients (Nuttall /12/, Ulrych and Clayton /13/) or as a function of the reflection coefficients ---

$g_{jj} \quad j=1,2,\dots,m$  /9/. The latter alternative leads to a cumbersome non-linear problem, -- but having the advantage of guaranteeing the stability of the AR(m) model. The Nuttall -- method behaves computationally better, but - may give a non-stable prediction filter. In the opposite case, the stable filter obtained by Nuttall method equals Fouguere's filter.

The minimization of  $\hat{P}_m$  as a function of ----  $g_{m1}, g_{m2}, \dots, g_{mm}$  is reduced to solve the linear system of equations

$$(13) Q_m G_m = \hat{P}_m I_m$$

where the matrix  $Q_m$  is defined as

$$(14) (Q_m)_{jk} = \sum_{r=m+1}^N (x_{r-j} x_{r-k} + x_{r-m+j} x_{r-m+k})$$

$j, k = 0, 1, \dots, m$

Like in BFB method we have some freedom to - choose the estimator of  $\sigma_m^2$ . For instance, we could use  $\hat{P}_m$  as estimator, but there is an alternative option using  $P_m$  (9).

If we select  $P_m$  as the adequate estimator, - Nuttall's FBLs method requires the step down Levinson's procedure

$$(15) g_{s-1,k} = \frac{g_{sk} - g_{ss} g_{s,s-k}}{1 - g_{ss}^2} \quad k=1,2,\dots,s-1$$

to find the reflection coefficients -----  $g_{jj} \quad j=1,\dots,m-1$ , (otherwise necessary to - calculate  $P_m$ ), because those were not explicitly obtained. This step can be omitted in

Fougere FBLs method, because the reflection coefficients are the variables of the minimization problem.

In any FBLs method relations like (10) and (12) do not hold, and it may happen that  $P_m$  and  $\hat{P}_m$  differ in one or even two orders of magnitude. The same applies to  $\hat{P}_0$  and  $P_0$ . An example of this is given latter on.

In FBLs we can choose  $P_m$  as estimator of  $\sigma_m^2$  which amounts to assume case A. (like in the BFB), while if we select  $\hat{P}_m$  as estimator we are in case B. In fact, in this latter option it seems preferable to assume  $\sigma_m^2 = \hat{P}_m$  rather than the alternative hypothesis  $\sigma_m^2 = P_m$ . These two assumptions are not incompatible but often lead to different results in evaluating total power.

The hypothesis  $\sigma_m^2 = \hat{P}_m$  can be justified when we analyze sinusoids with additive white noise: if the order of the AR model is greater than twice the number of sinusoids, we expect  $\hat{P}_m$  to be a good estimator of  $\sigma_m^2$ , and, on the other hand, to estimate  $\sigma_0^2$  by  $\hat{P}_0$  may give substantial errors when incomplete cycles are present. However sinusoids in white noise are not regular processes and stationary AR models are not adequate to analyze these signals.

When FBLs method is applied to fit an AR(m) model, with m greater than 2/3 of the data length, there are too many parameters to be fitted given the number of sample prediction errors. This causes drastic decrease of  $\hat{P}_m$ , while the model is brought close to singularity and  $\hat{P}_0$  underestimates  $\sigma_0^2$  in several orders of magnitude. Moreover, these overfitting effects introduce numerical problems in the FBLs algorithms.

#### 4. MAXIMUM LIKELIHOOD CORRECTIONS FOR LEAST SQUARES METHODS.

The Maximum Likelihood Principle can be used to partially correct the estimated power, by maximizing the likelihood of the data sample X with respect to a scale factor of variance.

Assume X to be a gaussian vector sampled from an AR(m) process. The sample likelihood is

$$(16) L(\sigma_0^2, G_m | X) = (2\pi)^{-N/2} |R_{N-1}|^{-1/2} \exp(-\frac{1}{2} X^t R_{N-1}^{-1} X)$$

where  $R_{N-1}$  is the autocovariance matrix of  $N \times N$  order and  $|\cdot|$  stands for matrix determinant.

For a fixed prediction filter  $G_m$  we can define a new power estimator by  $\hat{\sigma}_0^2 = a \hat{P}_0$  with  $\hat{P}_0$  fixed and a as a variable. It easily seen (Burg et al., /3/, Kay, /10/) that the value of a which maximizes L is

$$(17) a = \frac{X^t T_{N-1}^{-1} X}{N}$$

where  $T_{N-1}$  is the  $N \times N$  autocovariance matrix determined by  $G_m$  with its main diagonal being equal to  $\hat{P}_0$ . Then, an additional amount of likelihood would be obtained correcting  $\hat{P}_0$ . The new estimates are

$$(18) \sigma_0^2 \approx \hat{\sigma}_0^2 = a \hat{P}_0 \quad \sigma_m^2 \approx \hat{\sigma}_m^2 = a P_m$$

From numerical experience, it has been found that BFB estimated correction (17) is always one. We conclude that, for fixed BFB prediction filter  $G_m$ , the BFB estimate of  $\sigma_m^2$ , say  $P_m$ , is the optimum in the likelihood sense. This does not happen in FBLs in which the scale factor (17) can differ from unity in several orders of magnitude.

The maximum likelihood corrected prediction power  $\hat{\sigma}_m^2$  is not equal to  $\hat{P}_m$ , although those two values are usually quite close.

If a, given by (17), is different from unity and the estimated AR model is to be used in simulation studies, then there will be convenient, in case B, to take  $\hat{\sigma}_m^2$  as the prediction error power. Only when we are sure that  $\hat{P}_0$  is close to  $\sigma_0^2$  (case A.) we may choose  $P_m$  as prediction error power.

The computation of (17) is an awful task if N is a large number. Then, it is convenient ---

(see, for example, Kay, /16/) to use the following identity.

$$(19) P_m X^t T_{N-1}^{-1} X = G_m^t V_m G_m$$

where

$$(20) V_m \text{jk} = \sum_{r=1}^{N-j-k} x_{r+j} x_{r+k} \quad j, k=0, 1, \dots, m$$

The identity (19) was presented by Box- -- Jenkins, /2/ ). It is based in a factoriza-- tion of  $T_{N-1}^{-1}$  in terms of  $G_m$ .

In this fashion, the computational cost of -- the quadratic form in (17) is drastically reduced as m is lower than N and  $T_{N-1}$  has not - to be inverted. The matrix  $V_m$  is easily calcu-- lated by a simple recurrence (Dickinson /8/).

After formula (19) and introducing variance - scale factor, the logarithm of likelihood (16) appears to be proportional to

$$(21) H(a, G_m, X) = - \log |T_{N-1}| - N \log a - \frac{G_m^t V_m G_m}{a P_m}$$

where the determinant /10/

$$|T_{N-1}| = P_m^{N-m-1} \prod_{j=1}^{m-1} P_j$$

is easily calculated from  $G_m$  making use of -- step down Levinson's procedure. In the numerical examples, The function (21) has been - used to measure the sample likelihood.

## 5. NUMERICAL EXAMPLES.

An analysis has been made of a series of 35 annual means of the H component of the geo-- magnetic field recorded at Observatorio del Ebro (Tortosa, Spain) from 1943 to 1977.

The data have been smoothed and the linear - trend has been substracted. The signal is -- shown in Figure 1. Some signals related to - these data have been analyzed by Courtillot et. al./6/).

From the above mentioned series a value of -  $\hat{P}_0 = 527.35$  is obtained. The FBLs analysis -

provides stable prediction filters up to order  $m=10$ , and the use of Fougere algorithm is necessary for orders greater than 10.

Total power FBLs estimate  $\tilde{P}_0$  appear to be -- quite apart from  $\hat{P}_0$  (Table 1). In contrast, BFB  $\tilde{P}_0$  appears to differ slightly from value  $\hat{P}_0$ .

Overfitting effects are clearly detected for higher order FBLs models and it is quite --- apparent his down-biasing. Moreover, impor-- tant rounding errors appear in Fougere algo-- rithm for models greater then 20, and these results in Table 1 have to be regarded as a mere approximation.

The main features of the above described ana-- lysis are

- The estimated BFB total powers  $\tilde{P}_0$  are cente-- red around  $\hat{P}_0$  value.

- Most of the estimated FBLs total powers  $\tilde{P}_0$  are well above  $\hat{P}_0$  (except for orders greater than 26).

- Likelihood scaling of FBLs variance seems ha-- ve little importance in simulation experien-- ces, but it provides an useful comparison -- between likelihoods associated to BFB and -- FBLs models.

- When as an effect of overfitting a model is close to singularity ( $m \geq 27$ ) the likelihood scaling causes the power to get around  $\hat{P}_0$ . The shape of the analyzed signal leads to -- case B., and, in a certain sense, we can --- accept FBLs as the adecuate method of analy-- sis, and  $\hat{\sigma}_m^2$  as the estimator of the predic-- tion error power.

Figure 2 shows two simulations of the signal based on each of the BFB and FBLs models, -- using for both of them the same realization of white noise, however scaled to the predic-- tion error power ( $P_m$  in BFB,  $\hat{\sigma}_m^2$  in FBLs).

A comparison of power spectral densities of BFB and FBLs for an AR(12) can be seen at Fi-- gure 3. The expected peak at an eleven years period appears with a small shifting between 11.7 and 12.0 years, despite these periods -

are masked by low frequencies.

The annual means of the Zurich Sunspot Numbers from 1943 to 1977 were also analyzed. This signal was taken up to 1957 from Chernosky and Hagan /5/ and the following years from Solar-Geophysical Data (Central Radio - Propagation Laboratory, Boulder, Colorado), and it is shown in Figure 4.

This second signal provides us a simple example of case A., where  $\hat{P}_0$  can be taken as a good estimate of  $\sigma_0^2$ .

After we have subtracted the mean value, we obtain  $\hat{P}_0 = 2886.6$ . The results of the analysis are shown in Table 2.

The Marple FBLs algorithm does not give stable AR models from  $m = 17$  onwards, as there happened in our first example. Therefore we use again Fougere technique for those unstable models.

No significative bias from  $\hat{P}_0$  is detected in neither BFB and FBLs algorithms. The variance likelihood correction causes little modification of power and likelihood.

It is worth noting that likelihood of FBLs models is not necessarily greater than the one obtained from BFB method.

There is no problem in assuming  $\hat{P}_0$  as the variance of the signal and  $P_m$  as the variance of the prediction error in this second example.

## 6. REMARKS AND CONCLUSIONS.

In many applications we have an interest in variance or power measurements, for instance process simulations to study the maxima and minima effects.

We often get case A., where  $\hat{P}_0$  as given by (1) is a good estimator of signal power, so that we choose

$$P_m = \hat{P}_0 \prod_{j=1}^m (1-g_{jj}^2)$$

as variance of the AR model input white noise.

In case A., usually there are no important differences between BFB and FBLs analysis, as long as we will not deal with sinusoids; then, the BFB method seems suitable, being simple and numerically stable, and could be often recommended to avoid the precision problems and computing time inherent to Fougere algorithm, thought as an alternative to Marple recurrent algorithm when an unstable AR model is got.

However we can get to case B., in which  $\hat{P}_0$  is not reliable estimator. Then it would be advisable to study the power estimate  $\tilde{P}_0$  as given by both BFB and FBLs methods. For fixed and acceptable model from BFB we would get

$$E(\tilde{P}_0 | \sigma_0^2 = \hat{P}_0, G_m) = \hat{P}_0, \quad E(P_m | \sigma_0^2 = \hat{P}_0, G_m) = P_m$$

so that BFB  $\tilde{P}_0$  would not give significantly different estimate from  $\hat{P}_0$ .

However, in the FBLs method,  $\hat{P}_0$  and  $P_0$  may differ by several orders of magnitude, and the hypothesis

$$\sigma_0^2 = \hat{P}_0, \quad \sigma_m^2 = \tilde{P}_m$$

may be unrealistic.

It can be necessary a likelihood variance correction, that leads in many cases to greater likelihood and better estimate.

Burg's method cannot be corrected in this way, because the fitted autoregressive model maximizes its likelihood at value  $\hat{P}_0$ .

If it is necessary to use large order AR simulation models the least-squares method should be avoided, since it underestimate  $\sigma_m^2$ .

These large order models are better analyzed with maximum likelihood estimatory, like the one presented by Burg et.al. in 1982.

TABLE 1.

m	$\tilde{P}_0$ (BFB)	$\tilde{P}_0$ (FBL)	$\hat{G}_0^2$ (FBL)	FBL		
				H, a=1	H, a(18)	H, BFB
2	4.71 e02	8.70 e02	9.27 e02	-131	-125	-128
3	4.85 e02	9.47 e02	9.91 e02	-130	-122	-123
4	4.99 e02	1.18 e03	1.23 e03	-134	-117	-120
5	4.82 e02	1.36 e03	1.39 e03	-134	-111	-115
6	4.92 e02	1.38 e03	1.38 e03	-130	-107	-111
7	4.69 e02	2.52 e03	2.68 e03	-192	-107	-110
8	4.79 e02	2.00 e03	2.08 e03	-160	-106	-109
9	4.95 e02	2.03 e03	2.03 e03	-159	-106	-109
10	5.14 e02	2.42 e03	1.74 e03	-139	-101	-105
11	5.19 e02	6.66 e03	6.41 e03	-404	-102	-104
12	5.05 e02	2.61 e04	2.77 e04	-1768	-105	-103
13	5.17 e02	1.38 e04	1.40 e04	-882	-104	-103
14	5.25 e02	2.78 e03	2.81 e03	-188	-95	-98
15	5.23 e02	1.71 e04	1.66 e04	-1041	-94	-96
16	5.46 e02	1.29 e04	1.22 e04	-757	-94	-95
17	5.30 e02	4.43 e04	3.95 e04	-2530	-96	-95
18	5.33 e02	8.33 e04	1.01 e05	-6508	-101	-93
19	5.49 e02	3.44 e04	4.25 e04	-2733	-99	-91
20	5.55 e02	1.43 e03	1.58 e03	-114	-82	-87
21	5.77 e02	1.80 e03	1.82 e03	-125	-82	-87
22	5.88 e02	2.19 e03	2.29 e03	-149	-84	-86
23	5.97 e02	1.03 e07	6.37 e06	-4.2 e5	-102	-85
24	6.18 e02	7.18 e03	6.25 e03	-369	-76	-84
25	6.79 e02	4.86 e04	4.51 e04	-84	-29	-84
26	7.12 e02	3.18 e05	4.87 e05	-3.2 e4	-96	-83
27	6.11 e02	1.19 e00	7.17 e02	-27	-26	-73
28	5.99 e02	5.45 e-1	8.39 e02	-30	-26	-73
29	6.25 e02	1.58 e01	5.17 e02	-45	-45	-71
30	5.60 e02	9.25 e-1	4.89 e02	-64	-64	-71
31	6.34 e02	7.22 e-3	4.34 e02	-64	-63	-70
32	7.60 e02	4.04 e-1	5.71 e02	-68	-68	-70
33	1.05 e02	9.40 e-2	5.36 e02	-70	-70	-70

$$\hat{P}_0 = 5.27 e02$$

Analysis of geomagnetic field (component H), 1943-1977.

TABLE 2.

m	$\tilde{P}_0$ (BFB)	$\tilde{P}_0$ (FELS)	$\hat{\sigma}_0^2$ (FELS)	FELS		
				H, a=1	H, a (18)	H, BFB
2	2.95 e03	2.83 e03	2.77 e03	-255	-255	-255
3	3.04 e03	2.90 e03	2.76 e03	-250	-250	-250
4	3.04 e03	2.78 e03	2.65 e03	-250	-250	-250
5	2.98 e03	2.53 e03	2.48 e03	-250	-249	-249
6	2.98 e03	2.94 e03	2.85 e03	-249	-249	-249
7	2.94 e03	3.05 e03	3.00 e03	-249	-249	-249
8	2.95 e03	2.94 e03	2.91 e03	-249	-249	-248
9	3.03 e03	3.36 e03	3.29 e03	-245	-244	-244
10	3.12 e03	2.98 e03	2.92 e03	-245	-245	-244
11	3.21 e03	3.15 e03	2.97 e03	-243	-243	-243
12	3.24 e03	2.99 e03	2.80 e03	-243	-243	-242
13	3.22 e03	2.53 e03	2.40 e03	-242	-241	-241
14	2.82 e03	1.74 e03	1.98 e03	-244	-241	-241
15	2.95 e03	1.68 e03	1.84 e03	-244	-241	-240
16	3.00 e03	9.54 e02	2.07 e03	-242	-240	-240
17	2.93 e03	2.73 e03	2.94 e03	-240	-240	-235
18	2.91 e03	1.78 e03	2.23 e03	-239	-238	-235
19	2.95 e03	1.17 e03	1.62 e03	-241	-237	-234
20	2.96 e03	9.21 e02	1.32 e03	-242	-234	-231
21	2.77 e03	7.74 e02	1.33 e03	-228	-219	-228
22	2.96 e03	2.09 e02	1.21 e03	-254	-245	-226
23	2.69 e03	2.98 e04	1.78 e04	-341	-223	-226
24	2.06 e03	9.43 e-3	1.65 e03	-116	-112	-224
25	2.08 e03	1.11 e-1	3.20 e03	-142	-142	-221
26	2.18 e03	3.63 e00	5.11 e03	-192	-185	-220
27	2.36 e03	3.92 e00	4.99 e03	-200	-194	-216
28	1.87 e03	1.08 e01	8.55 e03	-228	-198	-214
29	1.97 e03	1.22 e00	2.14 e03	-189	-188	-213
30	2.02 e03	1.11 e-1	5.35 e03	-218	-210	-212
31	1.70 e03	1.77 e-4	4.90 e03	-220	-214	-211
32	1.91 e03	4.29 e-3	1.09 e04	-280	-227	-211
33	2.19 e03	3.45 e-4	3.67 e03	-217	-215	-210

$$\hat{P}_0 = 2.89 \text{ e03}$$

Analysis of Zurich Sunspot numbers, 1943-1977 .



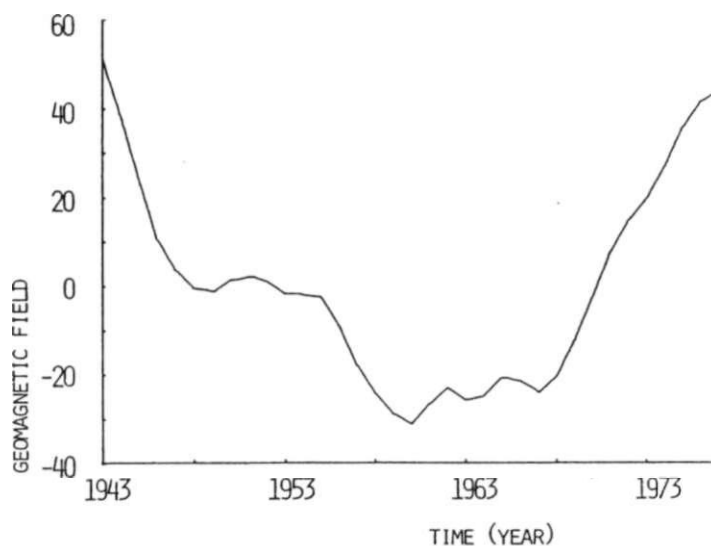


FIGURE 1 Annual means of terrestrial magnetic field (component H) recorded at Observatorio del Ebro (Tortosa, Spain) from 1943 to 1977. The signal was smoothed and linear trend subtracted.

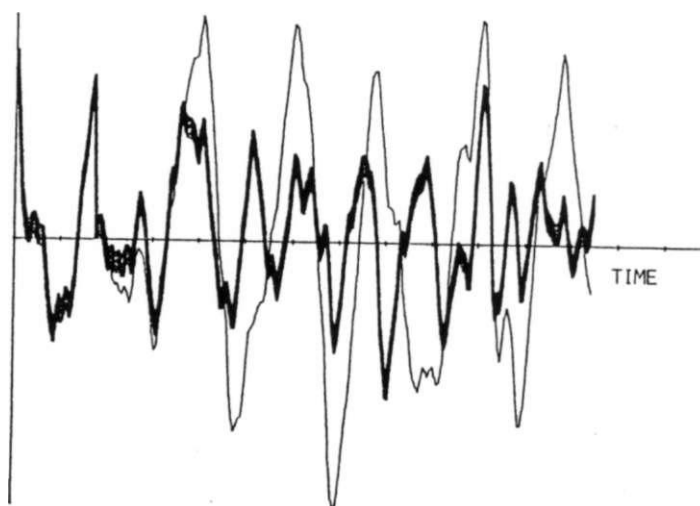


FIGURE 2. Simulation of geomagnetic field.  
Thin line: Fougere's AR(12) with  $\hat{\sigma}_m^2$   
Thick line : Burg's AR(12) with  $P_m$

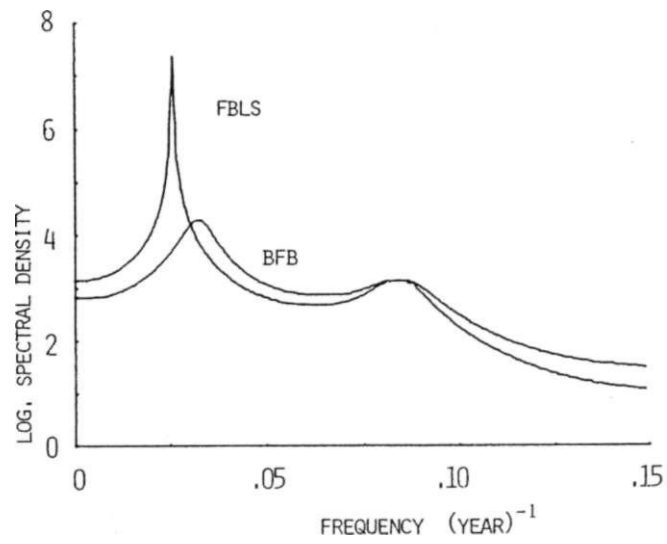


FIGURE 3 Power Spectral density of geomagnetic field estimated by BFB and FBLs with an AR(12)

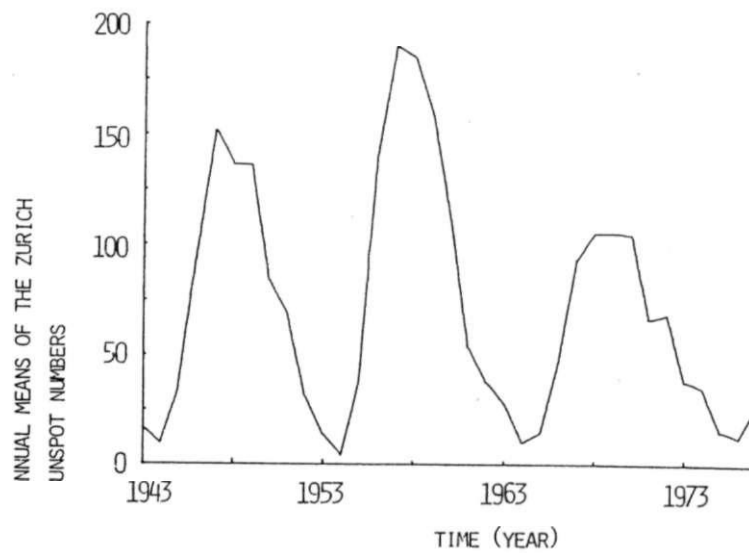


FIGURE 4 Annual means of the Zurich sunspot numbers, from 1943 to 1977

7. APPENDIX.

TOTAL POWER ESTIMATOR IN THE BFB METHOD.

Following (6) we take the prediction error -- power estimated with an AR(m) model to be

$$(A1) \hat{P}_m = \frac{1}{2(N-m)} \sum_{k=1}^{N-m} \left( e_{1m}^2(k) + e_{2m}^2(k) \right)$$

with  $e_{1m}(k)$ ,  $e_{2m}(k)$  defined as in (7).

From Levinson's procedure (5), error recurren ces are derived

$$(A2) \begin{aligned} e_{1m}(k) &= e_{1,m-1}(k+1) + g_{mm} e_{2,m-1}(k) \\ e_{2m}(k) &= e_{2,m-1}(k) + g_{mm} e_{1,m-1}(k+1) \end{aligned} \quad k = 1, 2, \dots, N-m$$

By substituting (A2) in (A1) we get

$$(A3) \hat{P}_m = \frac{1}{2(N-m)} \sum_{k=1}^{N+1-m} (1+g_{mm}^2) (e_{1,m-1}^2(k) + e_{2,m-1}^2(k)) - (e_{1,m-1}^2(1) + e_{2,m-1}^2(N-m+1)) (1+g_{mm}^2) + 4g_{mm} \sum_{k=1}^{N-m} e_{1,m-1}(k+1) e_{2,m-1}(k)$$

If we are working with BFB, the last term in (A3) can be compared to the numerator of (8) and substituted into (A3). On the other hand the denominator of (8) can be transformed, - after adding and substracting adequate terms, in a function of  $\hat{P}_m$ , so that finally we get the  $\hat{P}_m$  recurrence:

$$(A4) \hat{P}_m = \frac{N-m+1}{N-m} (1-g_{mm}^2) \hat{P}_{m-1} - \frac{1-g_{mm}^2}{2(N-m)} (e_{1,m-1}^2(1) + e_{2,m-1}^2(N-m+1))$$

This recurrence applies also to complex data, although we assume real data along this paper.

On the other hand, expression (A4) is similar to the one introduced by Andersen (1978).

From (A4) we obtain by index recurrence

$$(A5) \hat{P}_m = \frac{N \hat{P}_0}{N-m} \prod_{k=1}^m (1-g_{kk}^2) - \frac{1}{2(N-m)} \sum_{j=1}^m u_j \prod_{r=j}^m (1-g_{rr}^2)$$

where for any  $j = 1, 2, \dots, m$

$$(A6) u_j = e_{1,j-1}^2(1) + e_{2,j-1}^2(N+1-j)$$

Using (A5), we can write the total power estimator defined in (12) as

$$(A7) \tilde{P}_0 = \frac{N}{N-m} \hat{P}_0 - \frac{1}{2(N-m)} \sum_{j=1}^m u_j \prod_{r=1}^{j-1} (1-g_{rr}^2)^{-1}$$

Now, we start from hypothesis  $\sigma_0^2 = \hat{P}_0$  and -- fix the prediction filter  $G_m = (1, g_{m1}, \dots, g_{mm})^t$  or the equivalent reflection coefficients --  $g_{jj} \quad j=1, 2, \dots, m$ , while the terms  $u_j$  are still to be taken as random variables, and the --- errors  $e_{1,j-1}(1)$ ,  $e_{2,j-1}(N+1-j)$  being normal random variables of zero mean.

If the prediction filter  $G_m$  is a step in the Levinson's recurrence (6) to reach the true AR model, the variance of both errors in --- (A6) is  $P_{j-1}$ . Therefore, if this condition - applies, the random variables

$$\left[ \frac{e_{1,j-1}(1)}{P_{j-1}^{1/2}} \right]^2, \quad \left[ \frac{e_{2,j-1}(N+1-m)}{P_{j-1}^{1/2}} \right]^2$$

are distributed as a  $\chi^2$  of a single degree - of freedom. Therefore, if we take expectations in (A7) we set

$$(A8) E(\tilde{P}_0) = \hat{P}_0$$

and multiplying by  $\prod_{j=1}^m (1-g_{jj}^2)$  we also set

$$(A9) E(\hat{P}_m) = P_m$$

We recall that derivation above relies on -- two main hypothesis, namely

$$\sigma_0^2 = \hat{P}_0$$

-  $G_m$  is the true prediction filter, or  $G_m$  is an intermediate step in Levinson's recurrence to get the true model.

Therefore, if a strong difference is found from BFB  $\hat{P}_0$  to  $\hat{P}_0$ , the above hypothesis are probably false.

## 7. REFERENCES.

- /1/ ANDERSEN, N., "Comments on the Performance of Maximum Entropy Algorithms", Proc. IEEE, 66, 11, p.1581, 1978.
- /2/ BOX, G.E.P., and JENKINS, G.M., "Time series analysis, Forecasting and Control" Holden Day, San Francisco, 1970.
- /3/ BURG, J.P., "Maximum Entropy Spectral Analysis, Ph.D. Thesis Stanford Univ., California, U.S.A., 1975.
- /4/ BURG, J.P., D.G. LUENBERGER AND D.L. WENGER, "Estimation of structured Covariance Matrices", Proc. IEEE, 70, 9, 963-974, 1982.
- /5/ CHERNOSKY, E.J., AND M.P. HAGAN, "The Zurich Sunspot Numbers and its variations for 1700-1957", J. Geophys. Res., 63, 4, 775-778, 1958.
- /6/ COURTILOT, V., J.L. LE MOUËL AND N. MAYAUD, "Maximum entropy spectral analysis of geomagnetic activity index aa over 107-year interval", J. Geophys. Res., 82 2641-49, 1977.
- /7/ COURTILOT, V. AND J.L. LE MOUËL, "On the long-period variations of the Earth's magnetic field from two month to 20 years", J. Geophys. Res. 81, 2941-50, 1976.
- /8/ DICKINSON, B., "Two recursive estimates of autoregressive models based on maximum likelihood", J. Statist. Computation Simulation, 7, 85-92, 1978.
- /9/ FOUGERE, P.F., "A solution to the problem of spontaneous line splitting in maximum entropy power spectrum analysis", J. Geophys. Res., 82, 105-53, 1977.
- /10/ KAY, S.M., "Recursive maximum likelihood estimation of autoregressive processes", Trans. IEEE, ASSP-31, 1, 1983.
- /11/ MARPLE, L., "A new autoregressive spectrum analysis algorithm", IEEE Trans., ASSP-28, 4, 441-454, 1980.
- /12/ NUTTALL, A.H., "Spectral analysis of a univariate process with bad data points, via maximum entropy and linear predictive techniques" Naval Underwater System Center, technical paper 5503, march 1976.
- /13/ ULRYCH, T.J., AND R.W. CLAYTON, "Time series modelling and maximum entropy", Phys. Earth Planet. Interiors, 12, 188-200, 1976.