Breakwater Wave Energy Absorber, Hualien, Taiwan, ROC—R.L. Wiegel

PART I

THEORETICAL AND OBSERVED WAVE CHARACTERISTICS

Taichung Harbor—H.S. Hou

CHAPTER 1

Maximum Entropy Spectral Estimation for Wind Waves

Jorge Calderón Alvarez*
Adolfo Marón Loureiro*

Some results are presented on the application of new spectral estimation techniques using AR and ARMA models, also known as Maximum Entropy Methods, to wind wave spectral analysis. The results are compared with those obtained with conventional FFT methods. The application of some mathematical methods for model order selection is included. The relation between the optimum order and different spectral parameters is investigated.

1 INTRODUCTION

In recent years new spectral analysis methods have substituted traditional ones in many fields of science and technology. Among them, one of the most useful ones is based in the use of autorregresive models (AR), moving average models (MA) and mixed autoregressive-moving average models (ARMA). Using these models for spectral analysis is directly related with what is called Maximum Entropy Methods (MEM) for spectral estimation.

Among the possible advantages of these methods found in the literature, we can note the following ones:

1- No assumption is made about how the time series behaves outside the known interval, while FFT methods assume that the series either is zero or repeats itself indefinitely outside such interval.

2- The spectral density can be easily represented by a small number of parameters (the model coefficients).

3- Good spectral estimates can be obtained from a very short sample of the time series.

4- Real time forecasting can be easily performed.

Although wind wave spectra have been lately analysed by means of the Fast Fourier Transform (FFT), some attemps of using ARMA models have come out in publications. However, to the authors knowledge, no systematic applications for wind wave analysis have been reported.

Generally speaking, we can say that ME methods try to model a non-deterministic, discrete and stationary stochastic process $\{x(t)\}$ by an ARMA model of finite orders (m,n), whose mathematical description is as follows:

$$\sum_{k=0}^{m} b_k \, x(t-k) = \sum_{k=0}^{n} a_j \, w(t-j) \quad ; \quad t = 1, 2, 3, \ldots \qquad (1)$$

* Programa de Clima Marítimo y Banco de Datos Oceanográficos. Dirección General de Puertos y Costas. Po. Castellana, no.16, 28046-Madrid, Spain.

Where

m is the order of the autorregresive part,
n is the order of the moving average part,
a_k are the autorregresive coefficients,
b_j are the moving average coefficients and
w(t) is a zero mean white noise with variance σ^2.

The energy spectral density of the process corresponding to this model is given by the following expression:

$$S(f) = \frac{\sigma^2 \cdot A(z) \cdot A(1/z)}{B(z) \cdot B(1/z)} \quad \text{in the Nyquist range } f \in [0, \frac{1}{2T}] \tag{2}$$

where:

$$A(z) = \sum_{k=0}^{m} a_k \cdot z^k \quad , \ z = \exp(-i2\pi Tf) \ ;$$

$$B(z) = \sum_{j=0}^{n} b_j \cdot z^j \quad \text{and}$$

T is the duration of the time series.

Many different algorithms have been proposed for estimating the coefficients a_k and b_k from the data, most of them dealing with pure autorregresive models. We have selected two of the most used ones for AR models and other two for ARMA models.

First the algorithm due to Burg (1975) which tries to minimize the sample white noise variance in successive steps, beginning with an AR (1) model and increasing the order by one on each step up to a previously selected order. The method, fast and stable, is based on the powerful Levinsons' recurrence.

The second is the least square method which tries to minimize the prediction error by a least square procedure. The equations are solved by an algorithm proposed by Marple (1980). The method is easier in conception than that Burg's one but the stability of the solution is not warranted.

For ARMA models we have tried first the Box-Jenkins (1970) algorithm, but difficulties in the convergence for large series as those of wave records, make it impossible its use in a rutinary basis. Nevertheless some results are presented in this paper. At the present time we are implementing a much stable algorithm known as Yule-Walker modified method (Kay-Marple, 1982), but no results are available for this paper.

One of the main difficulties that appear when applying AR and ARMA models, is the selection of the adequate order. We have chosen for this study three criteria among those found in the literature:

a) "Final Prediction Error"(FPE): Based in the fact that for a stationary process the prediction error (σ^2) should be stationary. It gives the following performance index:

$$(FPE)_m = \frac{N + (m+1)}{N - (m+1)} \sigma_m^2 \quad (3)$$

where: N is the total number of points in the time series and σ_m^2 is the white noise variance for order m.

b) "Akaike's information criteria"(AIC): Assumes that data are normally distributed and tries to find the order for which the model distribution better approximates the data distribution. The following index results:

$$(AIC)_m = - N \log (\sigma_m^2) + 2m \quad (4)$$

c) "AR Transfer function"(CAT). Tries to minimize the difference between the shape of the spectral density corresponding to an AR (m) model and the underlying AR(∞) model. The resulting index is:

$$(CAT)_m = \frac{1}{N} \sum_{j=1}^{m} \frac{N - j}{N\sigma_m^2} - \frac{N - m}{N\sigma_m^2} \quad (5)$$

These are indexes that tend to lower values as the model fits the data better. They have performed successfully in other scientific fields. We hope that comparison of these indexes with FFT results will give a good idea about their validity.

We have not tried similar criteria for ARMA models.

2 WAVE DATA

The primary objective of the study was to evaluate the behaviour of the different algorithms when applied to real wave data in a rutinary basis. Therefore we choose a set of successive wave buoy records that were representative of the different sea states that are encountered in practice.

The total number of records is 230 and were measured by a waverider buoy moored in deep water at the Bay of Biscay (North of Spain) and covering the period of February and part of March 1984. Is in this time of the year when the worst sea states are more likely ot occur. Each record consists of 2048 sea surface elevation points with a sample rate of 2 Hz measured each three hours.

3 AR ANALYSIS OF WAVE DATA

In order to do a comparative analysis between the AR estimation algorithms and the FFT estimation,we calculated with Burg and Least squares algorithms the estimation of AR models with order varying from 1 to 40 for all the set of records.We made some tests with selected wave records trying to confirm the highest order of the AR model where the main peak of the spectrum was completely developed, finally we chose the 40.

We selected seven spectral parameters in order to characterize each spectrum:

Condition A: Direction: North West
H_s = 4.75 m
T_p = 14.5 sec.
γ = 1.4

Condition B: Direction: North
H_s = 3.50 m
T_p = 14.5 sec
γ = 1.4

The first layout has been considered only under wave condition A.

As a resume, comparisons have been made in these three cases:

Test 1: First layout and Wave Condition A
Test 2: Second layout and Wave Condition A
Test 3: Second layout and Wave Condition B

The comparison carried out with both modelling systems (physical and numerical) is based on relative significant wave height.

An example of the results of the numerical modelling and the results of the physical model can be seen simultaneously in Fig. 2

In Figures 6,7,8 comparisons of the three tests can be seen. In the horizontal axis distance in meters have been represented and in the vertical axis significant wave heights relative to the entrace are shown.

5.2.- TEST 1

The lines along which comparisons have been made can be seen in fig. 3 and the comparisons along these lines are shown in fig. 6

In all the cases the agreement is seen to be excellent.

5.3.- TEST 2

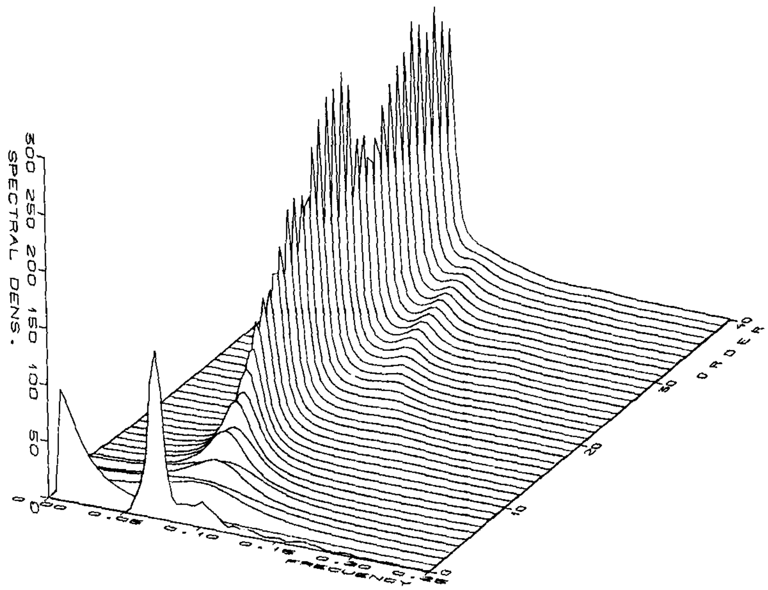
The lines along which comparisons have been made can be seen in fig. 4 and the comparisons along these lines are shown in fig. 7

In all the cases the agreement of the results of the physical and the numerical model is found to be good except for the points in which the influence of the reflection

Therefore the Maximum Entropy spectra (MES) were normalized with FFT spectra, trying to condense all the available information (40 AR models and 7 FFT estimations; raw spectrum and 3 smoothed spectra with two different windows;for each wave record).The mean and variance of normalized parameters were calculated for each possible combination. Then we could represent in graphs the spectral parameters against the order of the AR model for each FFT estimation (Figure 3).

First of all, we analysed the results obtained by both methods (Burg and Least squares),and they were very much the same,though the LS method was more unstable.Sometimes we could find records were the model did not converge and in most cases convergence is very slow compared to Burg method. Therefore the resulting conclusions for one method are valid for the other one.

The same can be said for different degrees of smoothings and smoothing windows with very similar results. Plots of normalized mean an variance can almost be superimposed.



Fiture 2. Dependence of spectral shape with order of AR model. The first spectrum is the corresponding FFT spectrum.

The normalized spectral moments tend very quickly to 1,having a small variance (Figure 3),the other 3 parameters have greater variance and need higher orders to stabilize.Around 20 for Fp and more than this order for the other parameter,Qp and Smax. Obviously,the normalized Qp and

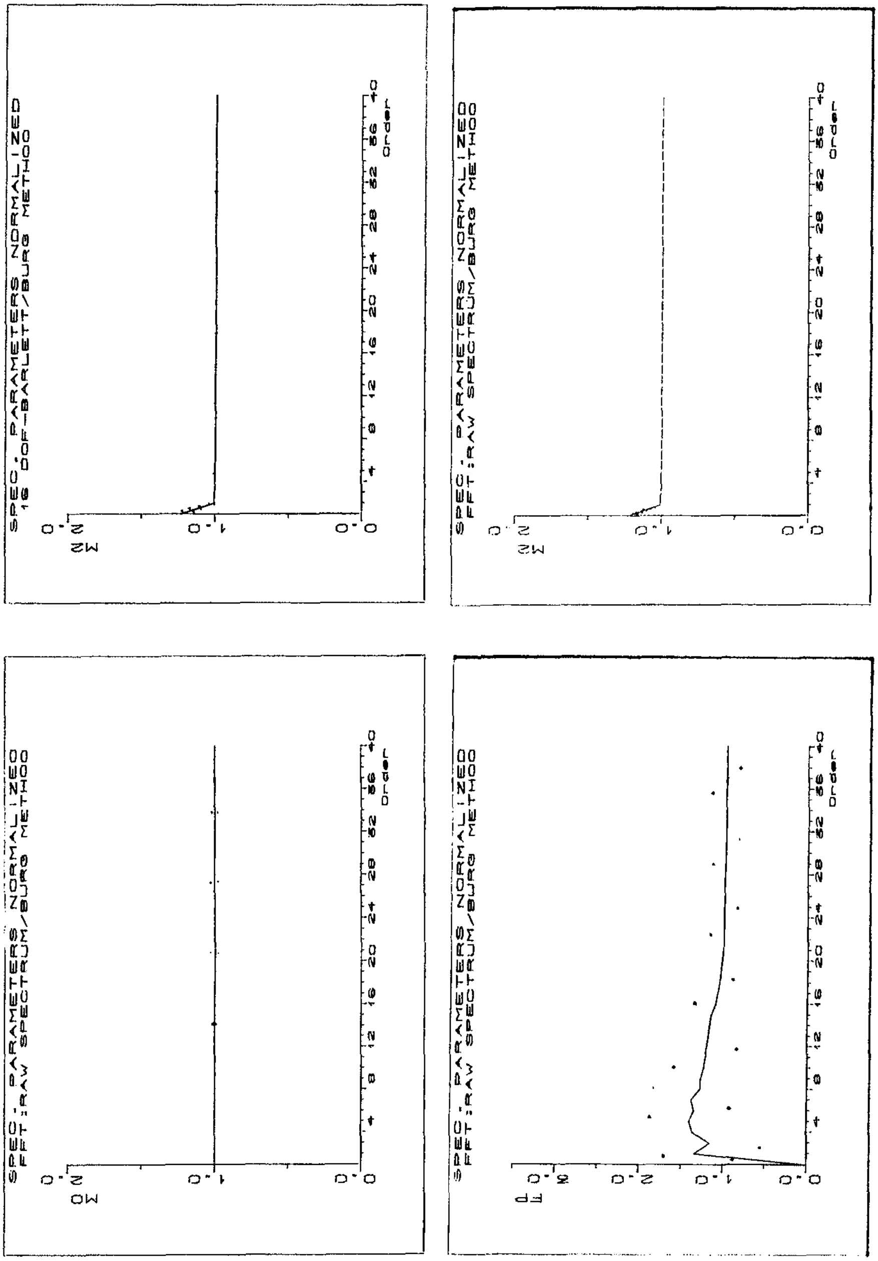


Figure 3. Examples of spectral parameters normalized.Mean (continuous - line) and standard deviation (dotted line) are represented.

Smax do not tend to 1 because their values depend largely of the degree of smoothing and the kind of used window. Though behaviour of Qp is more regular than Smax (Figure 3).

4 DETERMINATION OF THE ORDER OF AR MODELS

In order to be able to apply AR algorithms to wave data in a rutinary basis, some mathematical way of selecting the optimum order for the model should be available for its implementation in the computer. The importance of the selected order in the results is clearly illustrated in figure 2 where it can be seen how the peak frequency is converging very slowly to the true value as the order increases while it is very poorly defined for the low orders. Also some strong oscillation of the spectral value at the peak can be observed. These characteristics are common to most of the studied records.

We tried with the three criteria explained before. These criteria are based in some mathematical indexes which can be easily introduced in the programs. Theoretically, one has to evaluate such indexes for successive orders and select that order which gives the minimum value. We found actually that the indexes are monotonically decreasing when applied to waves for orders up to forty.

Therefore we choose as possible optimum order the one corresponding to the first point in the index variation curve where somo tendency to stabilization seems to happen. As a second candidate we choose the highest point were such an stabilization appears. Obviously, this method is somehow subjective and difficult to model in the computer. Actually we did the selection by hand.

Table I presents the best mean values of the normalized parameters (those nearer 1) and their corresponding standard deviations, as well as the order for which these values are obtained. The normalization is made with FFT spectra of 8 an 32 degrees of freedom. lt can be seen that different parameters need very different orders to reach an optimum.

TABLE I
RESULTS FOR BURG ALGORITHM

		M_0	M_1	M_2	M_4	F_P	S_{MAX}	Q_P
	ORDER	9	33	4	5	20	40	10
8dof	MEAN	.9999	.9998	.9978	1.0119	1.0097	1.0260	1.0442
	STD	.0025	.0164	.0064	.0195	.1600	.3180	.3120
	ORDER	36	30	3	6	20	11	9
32dof	MEAN	.9999	.9999	1.0030	.9935	.9983	1.0360	.9807
	STD	.0187	.0191	.0073	.0190	.1076	.4270	.2900

TABLE I-Results for Burg algorithm. Parameters normalized by FFT results with 8 & 32 d.o.f.(Barlett w.). Order optimun for each parameter.

In table lI the mean values of the same normalized parameters are

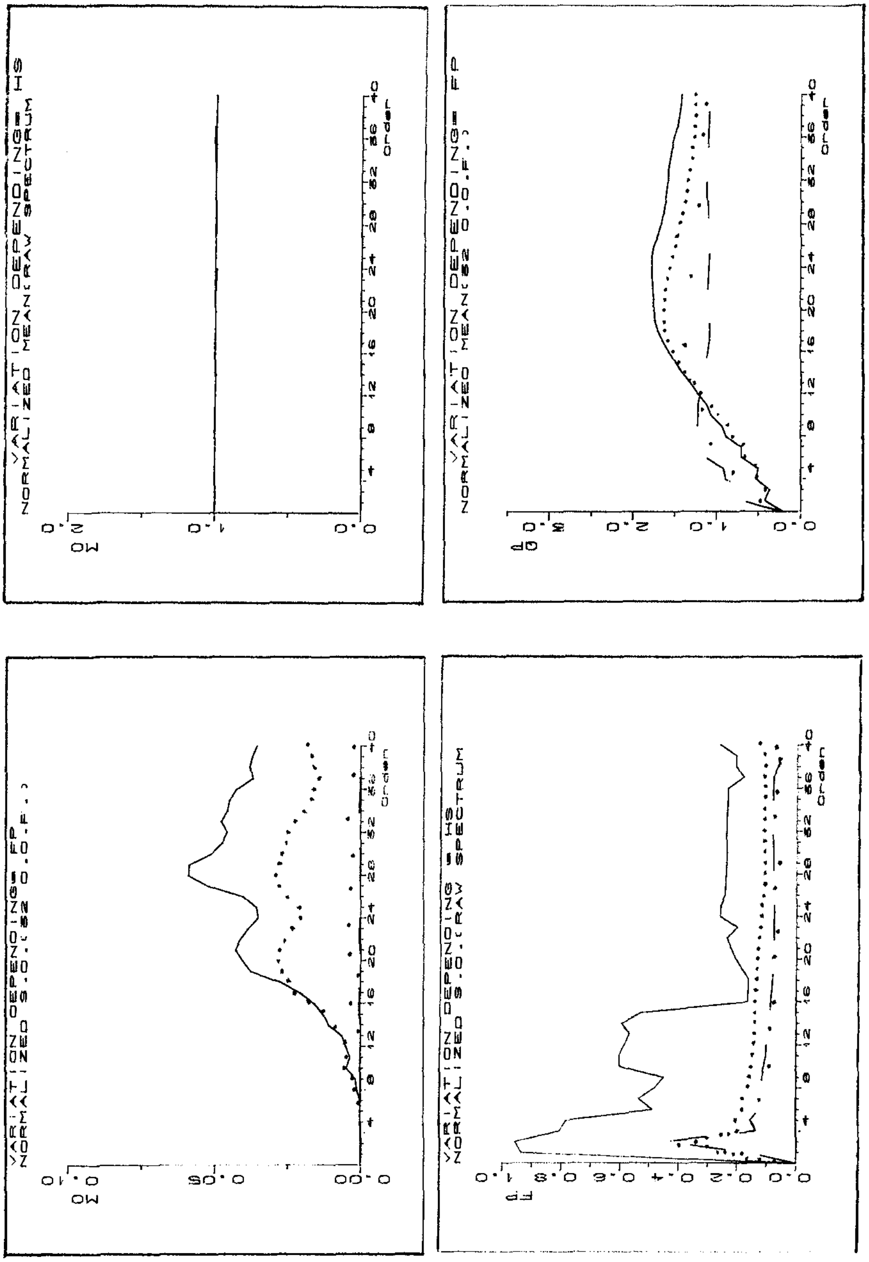


Figure 4. Variation of the order of AR model depending on different statistical parameters.

given, but in this case they correspond to the first order selected for each record by the procedure explained before when applied to the FPE criteria. Comparing both tables it can be seen that although the results obtained with the FPE orders are a little worse than the optimum ones,

TABLE II
RESULTS FOR BURG ALGORITHM
ORDER SELECTED BY FPE CRITERION

		M_0	M_1	M_2	M_4	F_P	S_{MAX}	Q_P
8dof	MEAN	.9999	.9959	.9952	.9862	1.0853	1.0177	1.3812
	STD	.0162	.0106	.0069	.0183	.2144	.4060	.3899
32dof	MEAN	.9999	.9959	.9947	.9865	1.0722	1.5653	1.4262
	STD	.0161	.0106	.0071	.0184	.1673	.5850	.4317

TABLE II-Results for Burg algorithm. Parameters normalized by FFT results with 8 & 32 dof. (Barlett w.). Order selected by FPE criterion.

the differences are not very significative except for the Qp parameter.

Of the two orders selected for each criteria we decided to deal only with the smaller one because the second one tend to be very high and no important improvement appear in the results when compared with the first one.

The three criteria gave very similar results and only a slight better performance of the FPE criteria can be mentioned.

Finally, we drew plots of the criteria selected order (mean and standard deviation) for groups of records with different values of the spectral parameters in order to see if any correlation could be found between order and spectral parameters. In general no good correlation was observed for the different parameters considered (see fig. 5 for example) and only some correlation seems to exist with the significative wave height Hs (see fig. 6). This last result is in agreement with that found by Houmb (1981) by other means.

5 VARIATION OF THE ORDER OF AR MODEL DEPENDING ON DIFFERENT STATISTICAL PARAMETERS.

We sought a possible correlation between the optimum order of the model with some spectral parameters. We first tried with significant height (Hs) and subsequently with the peak frequency (Fp).

We divided our set of data in different groups of Hs in such way that they were homogeneous, by subtracting randomly some records from those groups having more data. The five groups were divided as follows:

1.- 0. $\leq$ Hs $<$ 1.
2.- 1. $\leq$ Hs $<$ 2.
3.- 2. $\leq$ Hs $<$ 3.
4.- 3. $\leq$ Hs $<$ 4.
5.- 4. $\leq$ Hs

We repeated again the procedure described in point 3,normalizing the spectral parameters of the 40 estimated orders through Burg algorithm , via FFT estimations with 8, 16 and 32 degrees of freedom and Barlett window. In order to compare the results we represent in the same plot the curves if the mean of the five groups,and also in a separated plot the curves of the absolute value fo the standard deviation (Figure 3).

We can deduce some interesting behaviours.For the normalized spectral moments there is no distinct difference among the various groups,for both the means and the s.d., the latter ranges over a very small interval. Figure 3 shows superimposed curves corresponding to all the groups.

Concerning the Fp we found similar behaviour. The only difference found is that the first group has a turning point around the 16th. order that moves towards 1 faster than the rest of the groups,which fit into a similar shape. The s.d. for the first group is higher than for the rest of the groups but with a shape of the same sort. We cannot blame this difference to the order of the model but rather to the changing Fp, resulting from the FFT spectra for such a small Hs.

The Goda parameter behaves alike in all groups for low orders,and takes higher values for the high orders of the big groups. The s.d. is not so uniform but we did not find differences that would lead us to any conclusive result.

We tried again with the same procedure but dividing the sample in five groups according to the Fp:

1.- $0.05 <= Fp < 0.07$
2.- $0.07 <= Fp < 0.09$
3.- $0.09 <= Fp < 0.11$
4.- $0.11 <= Fp < 0.15$
5.- $0.15 <= Fp$

We uniformized the groups not taking into account the fifth group, since it is not statistically significant for having almost no sample (Figure 4).

The means of the normalized moments do not change whichever the FFT estimation used is, but s.d. decrease in the high orders for those groups having a higher Fp. For the normalized Fp those groups with high frequencies tend towards 1 (around order 18) faster than those of low frequencies (group number 2 about order 24 and group 1 more than order 30). The behaviour of s.d. is somewhat more irregular for high orders, but we found it quite similar to that described for the normalized moments.

Finally the Qp, the normalized values of groups with higher Fp tend towards 1 faster than groups with low Fp. For the s.d.the values tend to 0. starting from a particular order but its even faster as higher the group is.

We will talk about the analysis of all these results later on when we come to the conclusions.

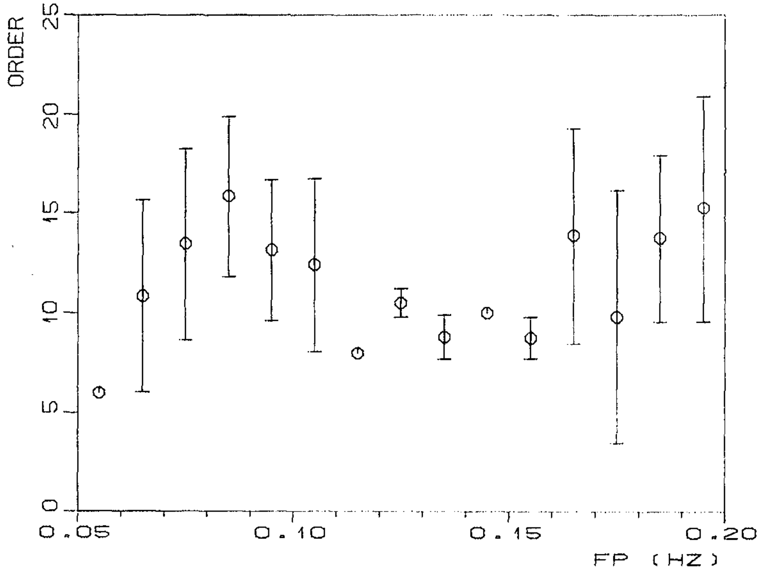


Figure 5.- Dependence of order selected by FPE criteria with peak frequency.

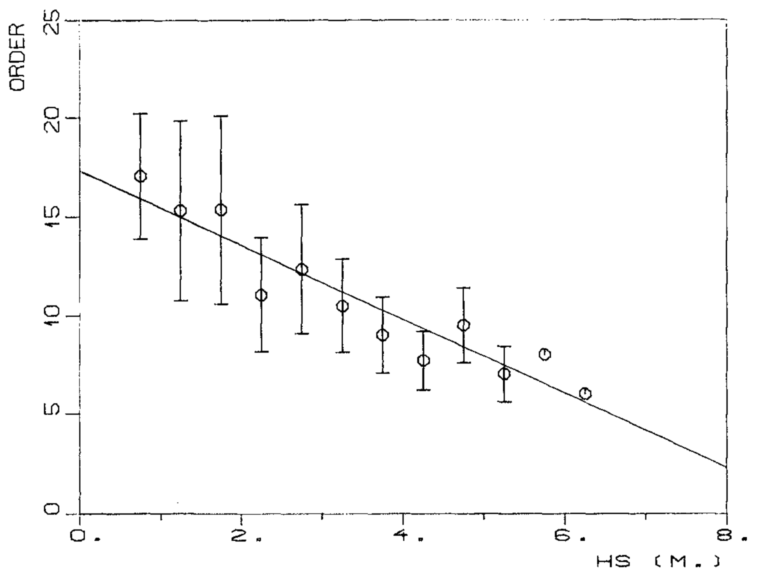


Figure 6.- Dependence of order selected by FPE criteria with significant wave height.

6 APPLICATION OF ARMA MODELS

As a first trial we selected the Box-Jenkins algorithm for ARMA model estimation, being this algorithm one of the most widely used in other fields. But we found that this algorithm performs quite poorly for the large number of data points that constitute the wave records. Computing time is prohibitive and, what is worse, the convergence to a solution is not assured giving rise to frequent running time errors.

Therefore we found it very difficult to apply this algorithm to wave data in a rutinary basis and this made it impossible to carry out a complete comparative analysis as was made for pure AR models.

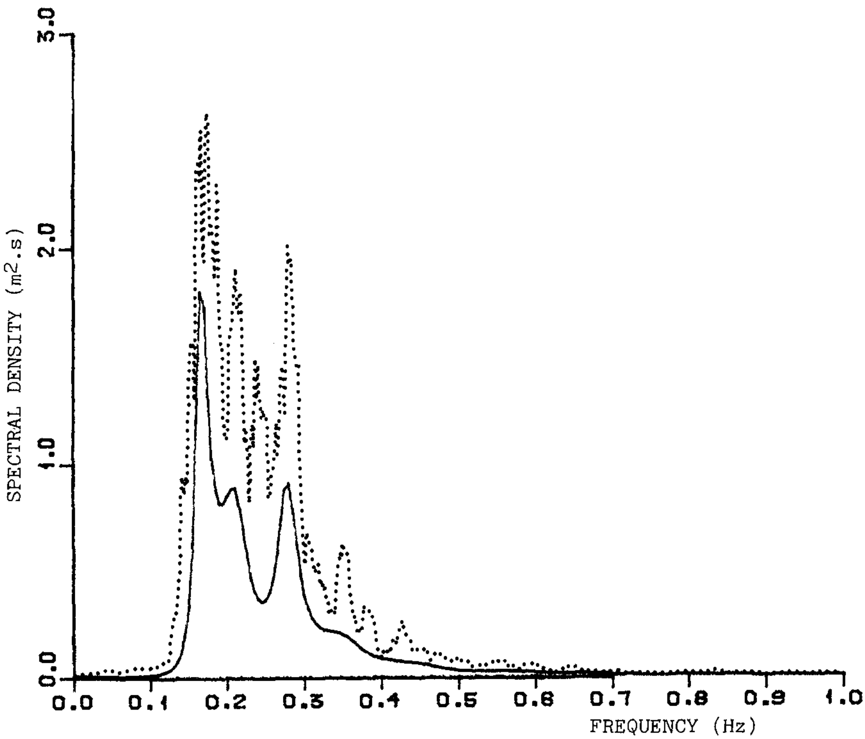


Figure 7. Example of AR spectrum obtained with an ARMA (20, 2) model (continuous line) compared with the corresponding FFT spectrum (dotted line).

We applied the algorithm to just a few cases in order to appreciate if any improvement in the spectral shape was achieved with respect to AR models. Figure 7 is one example of the results obtained using a ARMA (20,2) model to obtain the spectra of a complicated sea state. It can be seen that a very good spectrum is obtained with a very low MA order. The same was observed for most of the particular cases studied. Specially, it is very interesting to note the way in which the sharp increase in spectral density at the low frequencies is reproduced. This differs from the pure AR models which have difficulties in modelling such part of the spectra (see fig. 1).

At present, we are implementing a new ARMA model estimation algorithm based on the Yule-Walker modified algorithm that we hope will perform much better than the Box-Jenkins one, from the computational point of view.

7 APPLICATION TO SHORT SERIES

One of the main advantages claimed for the ME methods is that they are specially well suite for application to very short time series. This fact could be very helpful in the recovery of spectral information from records where an important part of the signal is lost or distorted.

In order to study if this characteristic could be of interest in the case of wave records, we divided some of the measured records in shorter series of 128 and 256 points and compare the evolution of the estimated spectra for successive short series with that of the complete series. The same was made using FFT techniques.

We observed that AR spectra for short series had a much better resolution than the corresponding FFT spectra and from this point of view the use of AR spectra is recommended. On the other hand, it was also clear that such a small number of points was a sample too short to represent adequately the underlying sea state.

8 CONCLUSIONS

Very shortly we resume the main conclusions obtained during this study:

8.1 - Systematic application of AR models:

- The different AR algorithms used give very similar results. Burg's algorithm being faster and more stable is recommended.
- Visual comparisons between AR and FFT spectra are better made using FFT spectra smoothed to 8 d.o.f. wiht a Barlett window.
- Normalized spectral moments tend to 1 for orders over 5.
- One-peaked spectra are well represented for orders over 12, while two-peaked spectra need orders around 30. Peak frequency is well represented for orders higher than 20.

8.2 - ARMA models:

- Box-Jenkins algorithm for ARMA estimation is inadequate for waves because of long computer time and instability of the solution.
- Anyway, ARMA models seem to better represent wave spectra, mainly in the low frequency range.
- Algorithms better than Box-Jenkins' one can surely be developed.

8.3 - AR order determination:

- FPE criteria seems to behave a little better than AIC and CAT, although differences are relatively small.
- Orders selected using mathematical criteria tend to be much lower than those deduced from comparison with FFT results.
- There is some negative correlation between order selected by mathematical criteria and significative wave height. Nevertheless, the evolution of spectral parameters with order seems to indicate that the correlation should be positive.

8.4 - Relation between order and spectral parameters:

- As stated before, the dependence of model order on significative wave height is not clear. It seems that higher orders are needed to correctly represent the peak frequency for higher wave heights.
- On the other side, there seems to be a high dependence between peak frequency and model order, lower peak frequencies needing higher orders.

8.5 - Application to short series:

- Spectra for short series have higher frequency resolution than the corresponding FFT spectra.
- But, spectra obtained from short series are poor representatives of the underlying sea state.

9 ACKNOWLEDGEMENTS

Dr. Juan Jose Egozcue Rubi, professor at Barcelona University, Spain, implemented the algorithms and introduce the authors to this field. This research is part of the works carried out by the "Programa de Clima Maritimo y Banco de Datos Oceanograficos" of the Spanish Authority of Harbours and Coasts.

10 REFERENCES

BURG, J.P. (1975). Maximum entropy spectral analysis. Ph. D. Thesis, Standford U., California.

EGOZCUE RUBI, J.J (1986). Temas de procesos estocasticos y analisis espectral. Pub. of Programa de Clima Maritimo, no. 15. (Spanish).

HOUMB, O.G. and T. OVERVIK (1981). Some applications of maximum entropy spectral estimation to ocean waves and linear systems response in waves. App. Ocean Res. Vol. 3, No 4.

KAY, S.M. and S.L. MARPLE Jr. (1983). Spectrum analysis, a modern perspective. Proc. IEEE, 69, 1380-1419.

MARPLE Jr., S.L. (1980). A new autorregresive spectrum analysis algorithm. IEEE Trans. Acoust. Speech, Signal Process., ASSP -28, 441-454.

Box, G.E. and G.H. JENKINS (1970). Time series analysis: forecasting and control. Holden-Day, San Francisco.