

VOLUME 3

Fourth International Conference on Coastal and Port Engineering in Developing Countries Proceedings of the COPEDEC IV

Rio de Janeiro, Brazil
25-29 September 1995.



AUTOREGRESSIVE MODELLING OF MEASURED SEA WAVES OFF WEST COAST OF INDIA

By

S Mandal, V Sanil Kumar & B U Nayak
Ocean Engineering Division
National Institute of Oceanography
Dona Paula, Goa 403 004, INDIA.

ABSTRACT

The spectral analysis of measured sea waves can efficiently be carried out by the Fast Fourier Transform (FFT) technique. Even though many present techniques can be used for the generation of time series on sea waves from theoretical wave spectra (Pierson-Moskowitz, JONSWAP, etc.), they may not be computationally efficient in providing accurate results for a given application. A more efficient method is therefore described here to generate time series data on sea waves and to provide spectral relationships with minimum loss of information. The random process is expressed by the autoregressive (AR) algorithm which is a polynomial function. The optimal AR polynomial is obtained by using the Akaike information criterion method. The AR parameters of the estimated spectrum are very few in number as compared to the number of values of FFT spectrum or time series waves. Hence it provides a potential advantage of storing on-line long term data on sea waves.

1 INTRODUCTION

It is often necessary that in the absence of measured time series on sea waves, theoretical power spectra (Pierson-Moskowitz, JONSWAP, etc.) are considered for various offshore applications. Even though many present techniques are used for the generation of time series on sea waves [Borgman (1969), Tucker, Challenor and Carter (1984)], they may not provide either accurate results or computationally efficient method for specific offshore applications. Presently available methods (digital time series or spectral values) are not be useful in storing random sea wave information at remote places for long periods. All these facts have led to the development of parametric modelling techniques.

Various parametric system identification techniques are applied to ocean wave dynamic systems. Spanos and Hansen (1981) used maximum entropy method which was referred by the authors as Linear Prediction Theory (LPT). This is an autoregressive (AR) algorithm for digital simulation of sea waves and can be used as an efficient alternate method. They used Pierson-Moskowitz spectrum as the target spectrum which can be obtained as the output of the AR model. Time series waves were determined as the output of the recursive digital filter to a white noise input. Although AR spectral estimates fluctuated, in an average sense the AR model approximated reliably the Pierson-Moskowitz spectrum. It was emphasized that the numerical studies were considered as preliminary, and the important aim of the study was to make potential usefulness of AR algorithm

for ocean engineering applications. Even though Fines et al (1981) described autoregressive moving average (ARMA) algorithm, they have made some statistical studies applying AR algorithm to the time series waves measured off Norwegian coast. Medina and Sanchez-Carratala (1991) established a robust AR model of JONSWAP wave spectrum, where the addition of white noise with a variance of $0.0025m_0$ (m_0 is the spectral moment) makes an acceptable fitting of the AR spectrum to JONSWAP spectrum. Witz and Mandal (1991) modified the LPT method and then applied to time series for offshore dynamic systems. Mandal et al (1992) showed that the initial accurate AR estimates will lead to estimate the better and more reliable reduced order ARMA model.

In this paper, AR algorithm is formulated based on the least square error approach and applied to time series of measured sea waves off Karwar on the west coast of India.

2 AUTOREGRESSIVE MODEL

The time series $\{y_r\}$ is said to be an autoregressive process, AR(p) of the order p and it is generated from the relationship

$$y_r = - \sum_{k=1}^p a_k y_{r-k} + \sqrt{p_w} U_r \quad (1a)$$

Here y_r is the rth sample of the discrete stochastic process and U_r is the white noise with variance p_w . Equation (1a) can be written in polynomial form as

$$y_r = \frac{\sqrt{p_w}}{a(z^{-1})} U_r \quad (1b)$$

where,

$$a(z^{-1}) = 1 + \sum_{k=1}^p a_k z^{-k} \quad (1c)$$

The AR(p) model is also called as an all-pole or maximum-entropy model (Burg, 1975).

Therefore, the transfer function of the AR(p) process can be expressed as

$$H_{AR}(z) = \frac{\sqrt{p_w}}{1 + \sum_{k=1}^p a_k z^{-k}} \quad (2)$$

The estimated power spectrum in frequency domain (ω), $S_{YY}(\omega)$ of the time series y_r can be written as

$$\begin{aligned} S_{YY}(\omega) &= |H_{AR}(e^{i\omega})|^2 \\ &= \frac{p_w}{\left| 1 + \sum_{k=1}^p a_k e^{-ik\omega} \right|^2} \end{aligned} \quad (3)$$

Where w_m is the cut-of frequency in radians.

The autocorrelation function, R of the finite set of the time series data, i.e., $\{y_0 < y_1 < \dots < y_M\}$, can be calculated from the relationships

$$R = f(M, \lambda) \sum_{t=\lambda+1}^M y_t y_{t-\lambda} \quad (4)$$

where, $f(M, \lambda) = 1/(M-\lambda)$ for the unbiased sample correlations, and $f(M, \lambda) = 1/M$ for the biased sample correlations.

Let ϵ be the error between the target spectrum, $S(w)$, and the estimate power spectrum of the AR output, $S_{YY}(w)$, is expressed as

$$\epsilon = \frac{b_0^2}{2w_b} \int_{-w_b}^{w_b} \frac{S(w)}{S_{YY}(w)} dw \quad (5)$$

Substituting $S_{YY}(w)$ [equation (3)] in above equation and minimizing ϵ , the parameters a_k are determined. Therefore, minimisation of error, ϵ , brings the following important relationship

$$\frac{d\epsilon}{da_1} = 0 \quad (6)$$

$$a_k R_{i-k} = -R_i \quad \text{for } k=1, 2, \dots, p \text{ and } i=1, 2, \dots, p \quad (7)$$

Equation (7) can be written in Toeplitz matrix form as

$$\begin{bmatrix} R_0 & R_1 & R_2 & \dots & R_{p-1} \\ R_1 & R_0 & R_1 & \dots & R_{p-2} \\ R_2 & R_1 & R_0 & \dots & R_{p-3} \\ \dots & \dots & \dots & \dots & \dots \\ R_{p-1} & R_{p-2} & R_{p-3} & \dots & R_0 \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \dots \\ a_p \end{bmatrix} = - \begin{bmatrix} R_1 \\ R_2 \\ R_3 \\ \dots \\ R_p \end{bmatrix} \quad (8)$$

These are also known as Yule-Walker equations. Using equations (8), i.e., when the parameters $a_1, a_2, a_3, \dots, a_k$ and noise variance p_w are determined, then the time series y_t can be generated digitally by using recursive equation (1). Here one has to obtain optimal spectral estimates for the values of a_k and p such that AIC between the target and estimated spectra must be minimum.

3 OPTIMAL MODEL ORDER SELECTION

The Yule-Walker method requires a minimum of p (where p is the AR model order) correlation coefficients for the solution of the well-determined set of equations of an AR process. Mandal (1992) has

demonstrated the improvement in spectral estimation accuracy with a larger number of equations using more correlation coefficients. However, there are higher computational involvement associated with a higher order model. The question is how high a model order should one select so that the solution remains computationally efficient from the application point of view and the model can represent an acceptable target. There are various methods available by which one can select the optimal model order, e.g.,

- a) Auto Correlation Matching (ACM) (Spanos, 1983)
- b) Least Squares Error (LSE) (Kay, 1988)
- c) Final Prediction Error (FPE) (Akaike, 1974)
- d) Akaike Information Criterion (AIC) (Akaike, 1976)
- e) Minimum Description Length (MDL) (Rissanen, 1983)
- f) Criterion Autoregressive Transform (CAT) (Kay, 1988)

The LSE method is commonly used owing to its simplicity, but can not be applied to any random process. If the process is purely convergent, the least squares error will decrease with increasing model order. Hence, using the LSE method, it is sometimes difficult to select the model order which is the optimal representation of the stochastic process. Other method based on the AIC overcome this problem. The AIC method is expressed as

$$\text{AIC} = M [\ln(\text{ER})] + 2p \quad (9)$$

where M is the number of time series data points and ER is the average value of the sum of the squared error of the data or target spectrum with the estimated values from 1 to M. If the autocovariance sequences of the stochastic process follow first convergence and then divergence, one should select p for the autocovariance sequences just before it diverges.

4 RESULTS AND DISCUSSIONS

The wave measurements were carried out by the Ocean Engineering Division, National Institute of Oceanography, Goa at different locations along the west coast of India at varying water depths under different sponsored projects. For the present study some monsoon wave data collected off Karwar on the west coast of India at 75 m water depth. Even though ocean dynamic systems are random in nature, the confused ocean surface elevations can be characterised by their power spectral estimation using system identification technique, namely AR model.

Figure 1 shows the location of the waverider buoy. Based on the Akaike Information Criteria (AIC) method, the optimal AR model orders are selected. Table 1 shows the selected AR model orders for various time series wave data sets (from 25 June 1989, 1820 hours to 4 July 1989, 0620 hours). Various statistical parameters (significant wave height (H_w), zero crossing period (T_w), maximum power spectral energy (E_{max}), spectral width (ϵ) and spectral peakedness (Q_p) were estimated based on optimal AR models as shown in Table 1. A comparative study of statistical parameters estimated by the FFT and AR methods are shown in Table 1 which indicates very little differences. This shows that the time series wave data set can be represented by the AR model.

A typical time series data set (dated 28 June 1989 0920 hours) is considered for the detailed analysis of autocorrelation and AIC. The auto correlation variation of the time series is shown in Figure 2a. Figure 2b shows variation of AIC. From Figures 2a,b the optimal model order is selected as 18. Therefore, the AR (18) model optimally represents the time series (dated 28 June 1989 0920 hours). The FFT power spectrum of the time series of 2048 data points is calculated using block of 256 data points and 8 averaged values. The spectral estimate comparison of the AR (18) and FFT models shows hardly any difference. Therefore, time series waves of 2048 data points can be represented by only 18 values which are the AR parameters. This shows a massive reduction of values which help to store more information of longer period in the field on-line recording unit. AR spectra of the listed time series shown in Table 1 are estimated and seem to show the establishment of AR modelling. Some typical spectral comparisons are shown in Figure 3.

The study on AR modelling of theoretical Pierson-Moskowitz wave spectra was carried out by Spanos (1983). The detailed study showed that if we use the PM spectrum for AR modelling, it does possess a zero of infinite order at $\omega=0$. Therefore, numerical difficulties exist in approximating the PM spectrum by AR algorithms which are best suited for the all-pole dominated spectrum. This problem does not arise for measured time series of sea waves.

5 CONCLUSIONS

AR algorithm is formulated based on least squares error method and optimal AR model orders are selected based on the AIC method. Based on the present study, the AR modelling of time series of measured sea waves is established, whereas AR modelling of theoretical wave (Pierson-Moskowitz) spectrum has a drawback owing to the presence of zero of infinite order at zero frequency according to Spanos (1983).

The study presented herein also shows that estimation of spectra by AR algorithm can be used for on-line data collection programme. Based on spectral estimation of many time series wave data sets, it is observed that the number of AR parameters are a few in number as compared to the number of time series data points. This method has potentially greater advantage of storing wave information of longer duration than currently possible because of the reduced size of the data sets. The wave information can be stored in the form of AR parameters and hence later on one can retrieve the power spectrum with minimum loss of information.

ACKNOWLEDGEMENTS

The authors would like to gratefully acknowledge the support of the Director, National Institute of Oceanography, Goa, India.

REFERENCES

- 1 Akaike, H. 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control, AC-19(6), December, pp 716-723.
- 2 Akaike, H. 1976. Time series analysis and control through parametric models. Proc. First Applied Time Series Symp. held at Tulsa, Oklahoma.
- 3 Borgman, L.E. 1969. Ocean wave simulation for engineering design. Journal of Waterways and Harbor Division, ASCE, 94(4), November, pp 557-593.
- 4 Burg, J. P. 1975. Maximum entropy spectral analysis. PhD thesis, Dept of Geophysics, Stanford University, CA, 1975.
- 5 Fines, S., Houmb, O.G., Mo, K. and Overtick, T. 1981. On the estimation of wind wave spectra by parametric methods. Coastal Engineering, 5, pp 97-109.
- 6 Heteren, J.V., Keijser, H. and Schaap, B. 1988. Comparison wave directional measuring systems. Applied Ocean Research, 10(3), pp 129-143.
- 7 Kay, S. M. 1988. Modern Spectral Estimation: Theory and Application. Prentice-Hall, Englewood Cliffs, N.J.
- 8 Mandal, S., Witz, J. A. and Lyons, G. J. 1992. Reduced order ARMA spectral estimation of ocean waves. Applied Ocean Research, 14(5), pp 303-312.
- 9 Mandal, S. 1992. On the identification and parametric modelling of offshore dynamic systems. PhD thesis, University of London.
- 10 Medina, J.R and Sanchez-Carratala, C.R. 1991. Robust AR representation of ocean spectra. Journal of Engineering, ASCE, 117(12), pp 2926-2930.
- 11 Rissanen, J. 1983. A universal prior for integers and estimation by minimum description length. The Annals of Statistics, 11(2), pp 416-431.
- 12 Spanos, P.D. and Hansen, J.E. 1981. Linear prediction theory for digital simulation of sea waves. Journal of Energy Resources Technology, ASME, 103, pp 243-249.
- 13 ySpanos, P.D. 1983. ARMA algorithms for ocean wave modelling. Journal of Energy Resources Technology, ASME, 105, pp 300-309.
- 14 Tucker, M.J., Challenor, P.G. and Carter, D.J.T. 1984. Numerical simulation of a random sea: a common error and its effect on wave group statistics. Applied Ocean Research, 6(2), pp 118-122.
- 15 Witz, J.A. and Mandal, S. 1991. Application of system identification to the reduction of structural monitoring data. Final Integration of Design Techniques Project Report MON5, Santa Fe laboratory for Offshore Engineering, Dept. of Mechanical Engineering, University College London, October'91.

NOMENCLATURE

a_k, p	Coefficients and model order of AR process (denominator)
b_i, q	Coefficients and model order of ARMA process (numerator)
$a(z-1)$	Polynomial of AR process
E_{max}	Maximum power spectral energy
ϵ	Spectral width
$H_{AR}(z)$	Transfer function of AR process
$H_{ARMA}(z)$	Transfer function of ARMA process
H_w	Significant wave height

- M Total number of digital data of the random process
- Q_p Spectral peakedness
- R_i Correlation coefficients of the random process
- $S(w)$ Target power spectrum
- $S_{yy}(w)$ Estimated power spectrum of AR model
- T, T_z Sampling period (s) and zero crossing period (s)
- U_x Gaussian white noise
- w Frequency in rad/sec
- z z-transform variable

Table 1 Wave statistical parameters calculated by FFT and AR methods

Day	Time	p	Hs(m)		Tz(m)		Emax(m ² /Hz)		ε		Q _p	
			FFT	AR	FFT	AR	FFT	AR	FFT	AR	FFT	AR
June 1989												
25	1820	10	2.495	2.495	5.62	5.60	6.77	6.77	0.82	0.82	1.73	1.70
25	2120	8	2.200	2.200	5.52	5.51	5.57	5.52	0.81	0.81	1.69	1.69
26	0020	7	2.192	2.192	6.56	6.55	6.67	8.08	0.76	0.77	2.06	2.48
26	0320	7	2.075	2.075	6.34	6.33	4.78	6.02	0.78	0.78	1.95	2.25
26	0920	12	2.062	2.062	6.28	6.28	4.00	3.45	0.77	0.77	1.90	1.82
26	1220	7	2.079	2.079	6.28	6.27	5.37	7.30	0.76	0.76	2.10	2.71
26	1520	32	2.580	2.576	5.07	5.04	5.97	6.67	0.80	0.80	1.36	1.36
26	1820	11	2.246	2.242	6.58	6.57	9.06	11.34	0.77	0.77	2.59	2.96
26	2120	21	2.511	2.511	6.83	6.83	9.16	8.81	0.78	0.78	2.27	2.25
27	0020	22	2.052	2.052	5.95	5.95	3.68	3.70	0.75	0.75	1.65	1.64
27	0320	13	2.428	2.428	5.44	5.42	4.58	4.28	0.79	0.79	1.41	1.39
27	0620	19	2.402	2.403	5.26	5.24	5.18	5.53	0.78	0.78	1.48	1.49
27	0920	21	2.656	2.657	5.63	5.62	7.27	6.69	0.77	0.77	1.63	1.60
27	1220	24	2.517	2.518	6.22	6.21	4.98	4.48	0.74	0.74	1.66	1.65
27	1520	12	2.436	2.436	5.81	5.79	5.37	5.04	0.79	0.79	1.56	1.48
27	1820	14	2.598	2.598	6.32	6.31	5.67	5.56	0.75	0.75	1.70	1.69
27	2120	25	2.643	2.643	6.64	6.63	6.57	6.34	0.77	0.77	1.70	1.69
28	0020	11	2.818	2.818	6.30	6.29	7.47	8.77	0.77	0.77	1.87	1.85
28	0320	14	2.587	2.588	5.28	5.26	5.27	4.13	0.78	0.79	1.40	1.32
28	0620	15	2.771	2.771	6.35	6.34	6.67	7.64	0.73	0.74	1.67	1.69
28	0920	18	2.485	2.485	6.40	6.39	5.27	5.21	0.76	0.76	1.79	1.79
28	1220	19	2.629	2.630	5.62	5.59	8.86	9.63	0.80	0.80	1.79	1.88
28	1520	17	2.452	2.453	6.33	6.33	5.08	4.85	0.75	0.75	1.82	1.80
28	1820	21	2.627	2.627	6.34	6.33	6.07	6.43	0.73	0.74	1.77	1.76
28	2120	17	2.620	2.620	6.40	6.40	5.27	5.09	0.75	0.75	1.66	1.64
29	0020	16	2.840	2.841	6.81	6.80	10.95	11.78	0.80	0.80	2.17	2.17
29	0320	14	3.249	3.250	5.90	5.89	8.66	7.01	0.78	0.78	1.57	1.47
29	0620	15	3.297	3.298	5.93	5.92	8.66	8.04	0.79	0.79	1.52	1.49
29	0920	21	3.204	3.204	6.55	6.54	9.56	9.51	0.76	0.76	1.81	1.81
29	1220	17	3.301	3.301	6.49	6.47	10.95	12.14	0.80	0.80	1.90	1.93
29	1520	14	3.494	3.494	6.78	6.77	10.95	10.37	0.74	0.74	1.88	1.88
29	1820	14	3.901	3.901	7.32	7.32	20.91	19.64	0.74	0.74	2.41	2.41
29	2120	19	3.422	3.422	7.20	7.20	13.94	12.56	0.76	0.76	2.05	2.03
30	0020	15	3.473	3.474	6.23	6.22	11.95	12.08	0.80	0.80	1.64	1.65
30	0320	27	3.276	3.277	5.79	5.77	8.96	9.81	0.82	0.82	1.47	1.50
30	0620	24	3.445	3.445	6.69	6.69	10.95	11.46	0.75	0.75	1.79	1.80
30	0920	19	3.376	3.376	6.72	6.72	11.95	11.71	0.78	0.78	1.92	1.91
30	1220	21	3.118	3.118	7.22	7.22	9.96	9.19	0.78	0.78	1.77	1.72
30	1520	18	3.458	3.458	6.61	6.60	13.94	14.03	0.81	0.81	1.94	1.93
30	1820	14	3.624	3.624	6.43	6.41	11.95	12.89	0.81	0.81	1.58	1.57
30	2120	19	3.400	3.401	7.00	6.99	9.96	9.47	0.76	0.77	1.73	1.71

Table 1 (Contd.)

Day	Time	p	Hs(m)		Tz(m)		Emax(m ² /Hz)		ε		Q _p	
			FFT	AR	FFT	AR	FFT	AR	FFT	AR	FFT	AR
July 1989												
1	0020	21	3.456	3.456	7.49	7.49	10.95	11.39	0.79	0.79	1.89	1.85
1	0320	21	3.313	3.313	7.08	7.08	10.95	9.51	0.79	0.79	1.84	1.83
1	0620	18	3.442	3.442	7.06	7.06	15.93	15.89	0.76	0.76	1.95	1.93
1	0920	19	2.936	2.936	6.77	6.76	9.66	8.68	0.78	0.78	1.79	1.74
1	1220	19	3.658	3.658	7.36	7.36	17.93	16.07	0.76	0.76	2.15	2.11
1	1520	14	3.368	3.369	6.25	6.23	10.95	12.52	0.83	0.84	1.69	1.70
1	1820	19	3.250	3.250	5.69	5.68	10.95	10.75	0.81	0.81	1.45	1.44
1	2120	18	3.354	3.354	7.34	7.34	15.93	14.84	0.78	0.78	2.14	2.12
2	0020	18	3.411	3.411	6.34	6.32	13.94	13.47	0.85	0.85	1.53	1.50
2	0320	20	3.030	3.030	7.28	7.27	10.95	10.78	0.80	0.80	1.87	1.84
2	0620	18	3.179	3.179	7.63	7.63	13.94	12.46	0.81	0.81	2.06	2.06
2	0920	20	2.861	2.862	7.01	7.00	9.16	9.38	0.77	0.77	1.85	1.85
2	1220	13	3.021	3.022	6.63	6.61	16.93	15.57	0.84	0.84	2.19	2.12
2	2120	23	2.663	2.663	7.23	7.22	9.96	10.05	0.79	0.79	1.99	2.00
3	0020	13	2.949	2.949	6.44	6.42	12.94	12.52	0.83	0.83	2.05	2.02
3	0320	23	2.802	2.803	5.96	5.93	7.37	7.77	0.83	0.83	1.67	1.71
3	0620	10	2.833	2.834	6.03	6.01	12.94	11.80	0.84	0.84	1.97	1.89
3	0920	27	3.224	3.227	5.28	5.25	11.95	12.73	0.83	0.83	1.56	1.55
3	1220	22	2.534	2.534	7.10	7.09	8.86	9.09	0.82	0.82	2.12	2.10
3	1520	19	3.021	3.001	7.51	7.48	20.91	21.23	0.79	0.79	3.00	3.16
4	0120	20	2.670	2.670	7.61	7.61	9.46	9.47	0.78	0.78	2.26	2.26
4	0320	22	2.609	2.609	7.68	7.67	8.76	9.66	0.78	0.78	2.31	2.29
4	0620	17	2.739	2.740	6.49	6.47	9.26	9.00	0.85	0.85	1.85	1.81

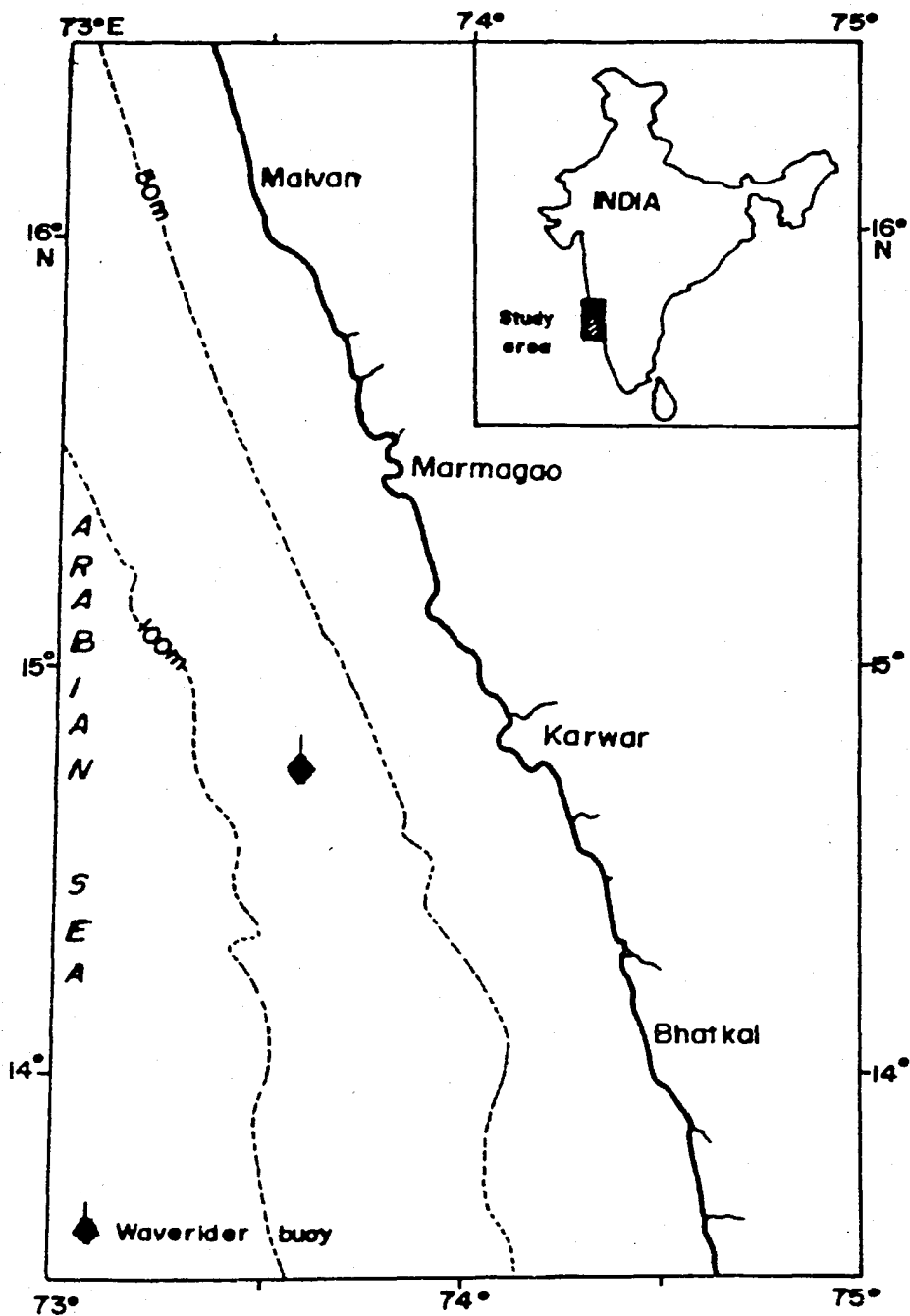


Fig. 1 Location map

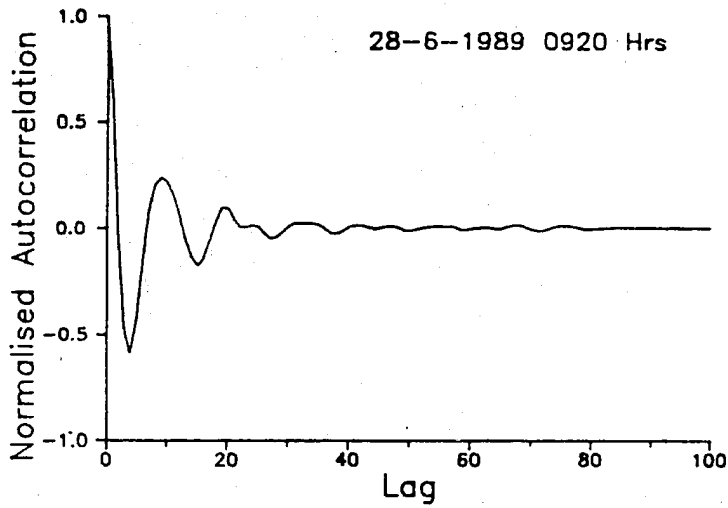


Fig. 2a Normalised autocorrelation of data set

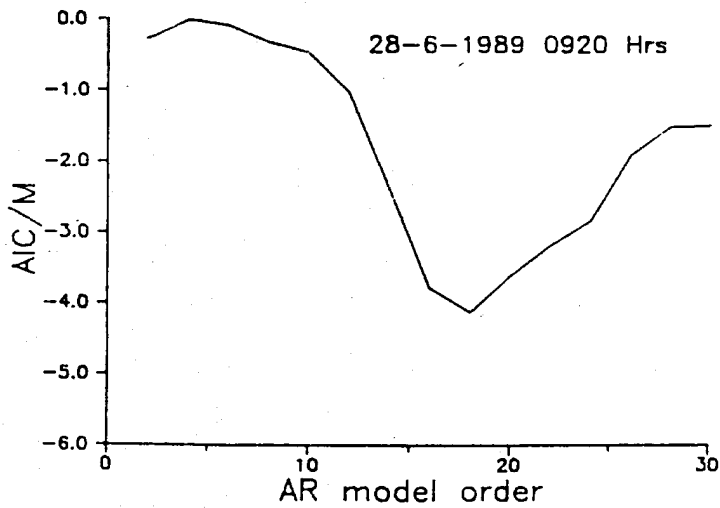


Fig. 2b Akaike Information Criterion (AIC)

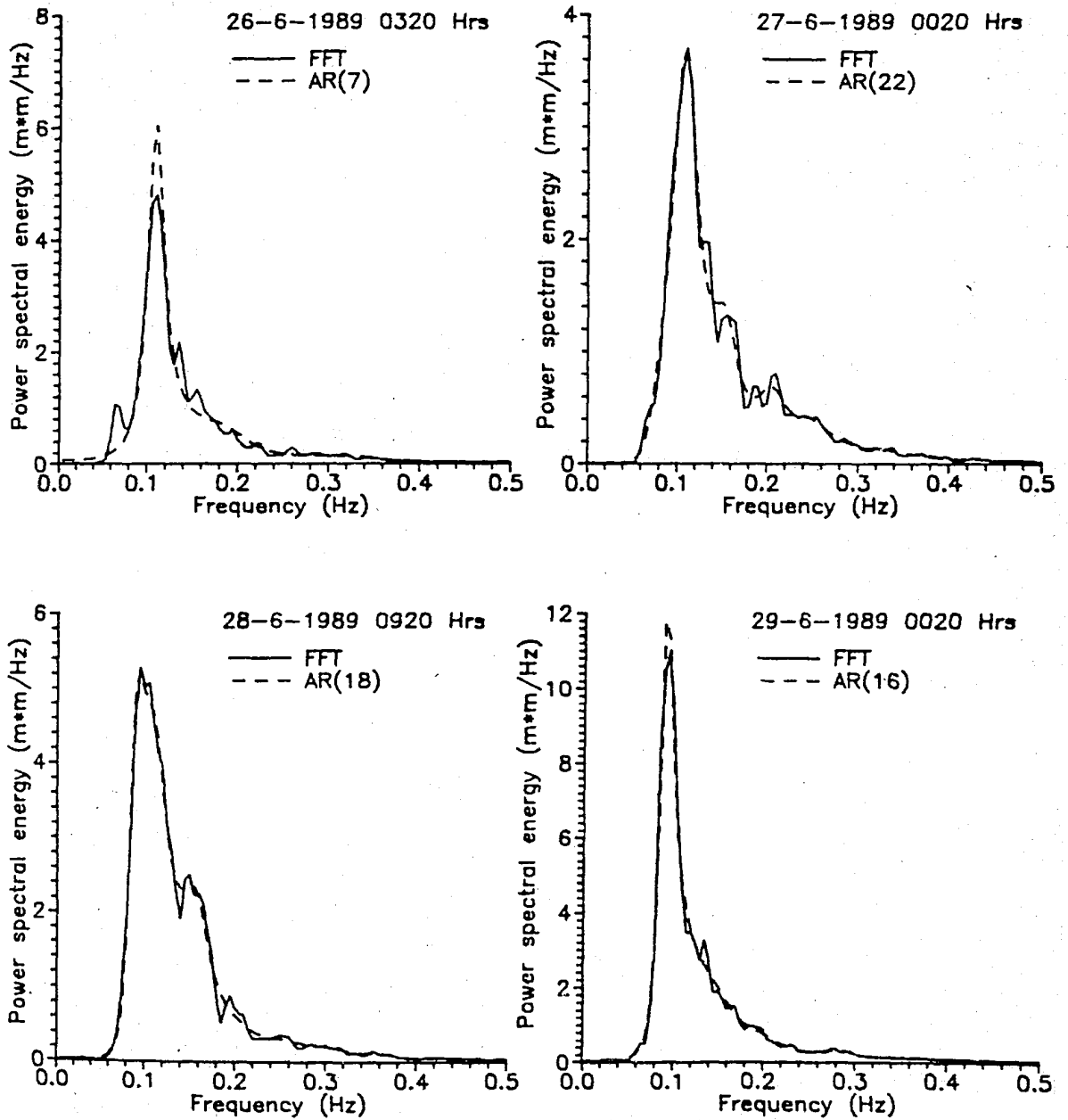


Fig. 3 Power spectral energy estimated by FFT and AR methods

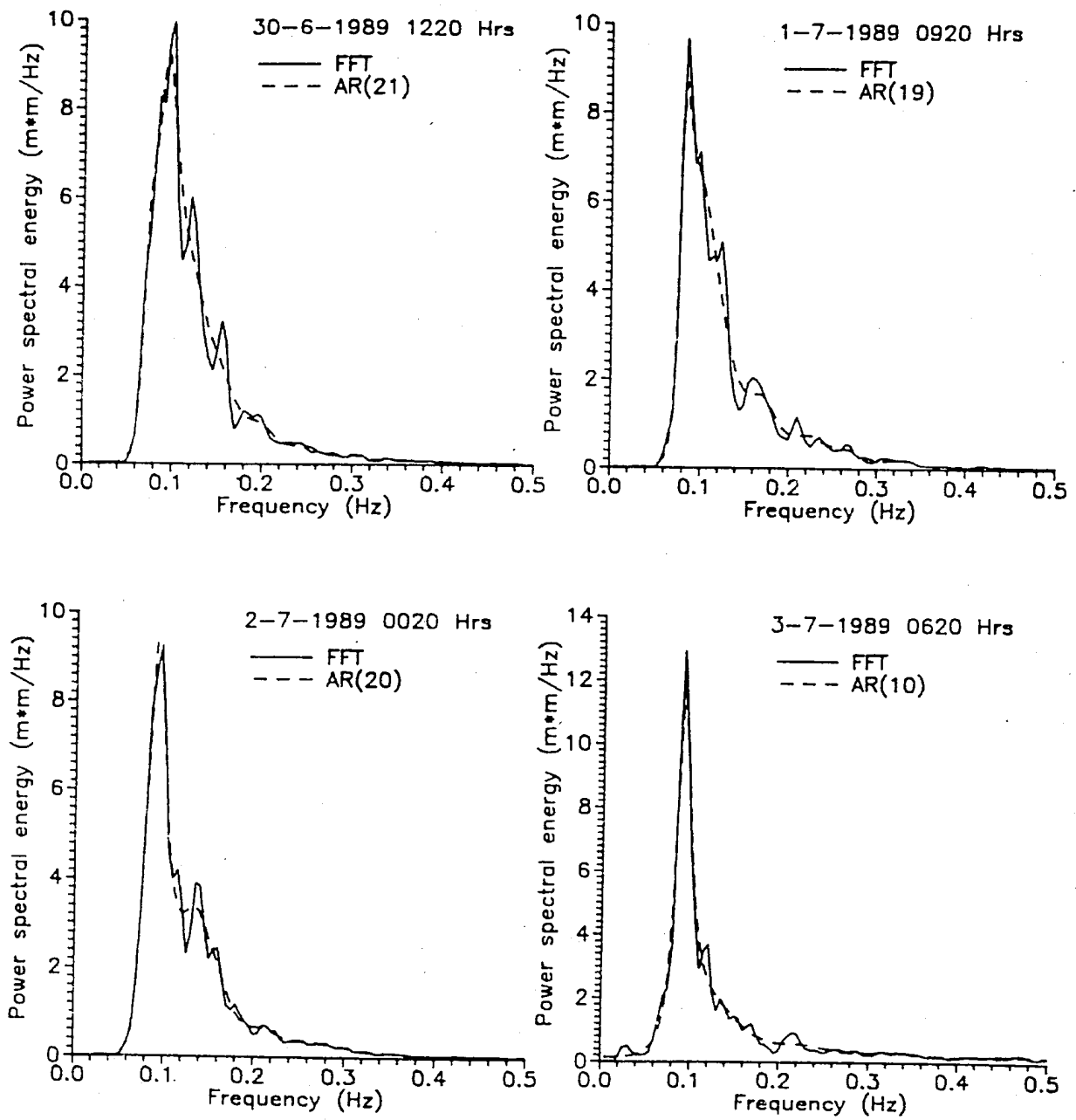


Fig. 3 Power spectral energy estimated by FFT and AR methods