ON THE IDENTIFICATION AND PARAMETRIC MODELLING OF OFFSHORE DYNAMIC SYSTEMS

A thesis submitted for the degree of Doctor of Philosophy
in the Faculty of Engineering, University of London

by

SUKOMAL MANDAL

Santa Fe Laboratory for Offshore Engineering
Department of Mechanical Engineering
University College London

August 1992

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

To

my parents

ABSTRACT

This thesis describes an investigation into the analysis methods arising from identification aspects of the theory of dynamic systems with application to full-scale offshore monitoring and marine environmental data including target spectra. Based on the input and output of the dynamic system, the *System Identification* (SI) techniques are used first to identify the model type and then to estimate the model parameters. This work also gives an understanding of how to obtain a meaningful matching between the target (power spectra or time series data sets) and SI models with minimal loss of information.

The SI techniques, namely, *Autoregressive* (AR), *Moving Average* (MA) and *Autoregressive Moving Average* (ARMA) algorithms are formulated in the frequency domain and also in the time domain.

The above models can only be economically applicable provided the model order is low in the sense that it is computationally efficient and the lower order model can most appropriately represent the offshore time series records or the target spectra. For this purpose, the orders of the above SI models are optimally selected by *Least Squares Error, Akaike Information Criterion* and *Minimum Description Length* methods.

A novel model order reduction technique is established to obtain the reduced order ARMA model. At first estimations of higher order AR coefficients are determined using modified Yule-Walker equations and then the first and second order real modes and their energies are determined. Considering only the higher energy modes, the AR part of the reduced order ARMA model is obtained. The MA part of the reduced order ARMA model is determined based on partial fraction and recursive methods. This model order reduction technique can remove the spurious noise modes which are present in the time series data. Therefore, firstly using an initial optimal AR model and then a model order reduction technique, the time series data or target spectrum can be reduced to a few parameters which are the coefficients of the reduced order ARMA model.

The above univariate SI models and model order reduction techniques are successfully applied for marine environmental and

structural monitoring data, including ocean waves, semi-submersible heave motions, monohull crane vessel motions and theoretical (Pierson-Moskowitz and JONSWAP) spectra.

Univariate SI models are developed based on the assumption that the offshore dynamic systems are stationary random processes. For nonstationary processes, such as, measurements of combined sea waves and swells, or coupled responses of offshore structures with short period and long period motions, the time series are modelled by the *Autoregressive Integrated Moving Average* algorithms.

The *multivariate autoregressive* (MAR) algorithm is developed to reduce the time series wave data sets into MAR model parameters. The MAR algorithms are described by feedback weighting coefficients matrices and the driving noise vector. These are obtained based on the estimation of the partial correlation of the time series data sets. Here the appropriate model order is selected based on auto and cross correlations and multivariate Akaike information criterion methods. These algorithms are applied to estimate MAR power spectral density spectra and then phase and coherence spectra of two time series wave data sets collected at a North Sea location. The estimation of MAR power spectral densities are compared with spectral estimates computed from a two variable fast Fourier transform, which show good agreement.

CONTENTS

GLOSSARY OF KEY SYMBOLS

Notations

a_k, b_ι	Coefficients of the ARMA process
$a(z^{-1})$	Polynomial of the ARMA process (denominator)
$b(z^{-1})$	Polynomial of the ARMA process (numerator)
$\wp$	Backshift notation
$C_{xx}(\tau)$	Auto-covariance of the random process x at lag τ
$C_{xy}(\tau)$	Cross-covariance of the random processes x and y at lag τ
$\delta_{s,t}$	Kronecker delta ($d_{s,t} = 1$ if $s = t$, else $d_{s,t} = 0$)
$\delta(\tau)$	Direc function
$\sum\varepsilon_\iota^2$	Sum of least squares error
E [.]	Expectation operator
e^b	Backward linear prediction error
e^f	Forward linear prediction error
f	Frequency in Hz.
f_N	Nyquist frequency in Hz.
g	Gravitational acceleration (9.81 m/s^2)
h_k	Impulse response function
$H_k(z)$	Transfer function of h_k in z-transform, which is also called as the Laplace transform of the impulse response
$H_{AR}(z)$	Transfer function of Autoregressive process in z-transform
$H_{MA}(z)$	Transfer function of Moving Average process in z-transform
$H_{ARMA}(z)$	Transfer function of Autoregressive Moving Average process in z-transform
M	Total number of digital data of the random process
N	Number of modified Yule-Walker equations ($N \geq 2p$), Number of correlations are selected based on number of modified Yule-Walker equations
N1	Total number of discrete values of FFT (M12) spectrum (Figure 4.1a)
$n(z^{-1})$	A polynomial

P^b	Backward residual variance matrix
P^f	Forward residual variance matrix
$\Phi(\omega)$	Magnitude squared coherence at frequency ω
$p(x)$	Probability density function of variable x
p	Autoregressive model order (denominator)
q	Moving average model order (numerator)
p_1	Number of first order mode of the initial AR polynomial
p_2	Number of second order mode of the initial AR polynomial
R_k or R_λ	Autocorrelation function with lag k or λ
ρ_x	Variance of the random variable x
$S(\omega)$	Target power spectrum
$S_1(\omega)$	Single-sided power spectrum
$S_{yy}(\omega)$	Estimated power spectrum of the SI model
$\theta(\omega)$	Phase spectrum at frequency ω
T	Sampling period
u	Input driving noise process
v	Wind speed
ω	Angular frequency in rad/sec
ω_b	Angular cut-off frequency in rad/sec
Ω_1	Energy of first order mode
Ω_2	Energy of second order mode
w_r	Gaussian white noise (a sequence of independent random variables which are normally distributed)
z	z-transform variable

Abbreviations

AIC	Akaike information criterion
ACM	Autocorrelation matching
ACS	Autocorrelation sequence
AR	Autoregressive algorithm
AR(p)	AR of order p

ARMA	Autoregressive moving average algorithm
ARMA(p,q)	ARMA where AR and MA parts have orders p and q respectively
ARIMA	Autoregressive integrated moving average algorithm
ARIMA(p,d,q)	ARIMA where AR and MA parts have orders p and q respectively with d number of differencing
BIBO	Bounded-input and bounded-output
CAT	Criterion autoregressive transfer function
cov	Covariance
FFT	Fast Fourier transform
FPE	Final prediction error
FLPR	Forward linear prediction residual technique
IAR	Inverse autoregressive method using modified Yule-Walker equations
JONSWAP	JOint North Sea WAve Project
J10	JONSWAP spectrum with wind speed 10m/s and $\gamma = 3.3$
LSE	Least squares error
MDL	Minimum description length
MA	Moving average algorithm
MA(q)	MA of order q
MYW	Modified Yule-Walker equations
MAR	Multivariate autoregressive algorithm
PM	Pierson-Moskowitz spectrum
PM10	PM spectrum with wind speed 10m/s
PM10TR	True PM spectrum with wind speed 10m/s
PM10TA	Taylor approximated PM spectrum with wind speed 10m/s
PM20TA	Taylor approximated PM spectrum with wind speed 20m/s
PM30TA	Taylor approximated PM spectrum with wind speed 30m/s
PSE	Power spectral energy
PDF	Probability density function
SISO	Single input and single output
SI	System identification
SR	Sampling rate
SVD	Singular value decomposition

Trf	Transfer function of structural motion in sea waves
var	Variance
WSS	Wide sense stationary

Notational conventions

$Q^{1/2}$	Matrix square root of positive definite matrix Q $(Q^{1/2})^T Q^{1/2} = Q$
$Q^{T/2}$	$[Q^{1/2}]^T$
$H_k(z)$	$= \sum_{n=-k}^{n=k} h(n)\ z^{-n}$
$\lvert . \rvert$	Absolute value

Superscript * denotes complex conjugate

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

1.1 General

In offshore engineering theoretical models are extensively used in the design of offshore structures. These theoretical models describe the marine environment and offshore structure, and are applied to predict the structure's service life and response to extreme conditions. Often these theoretical models require validation to ensure that the engineers have high confidence in their design. One common design verification procedure is to carry out model scale tests. Model scale tests are also often useful for providing additional design data. As a result of the assumptions made in the structural modelling process and the natural variation of environmental conditions, there is always an element of uncertainty associated with the predicted response of the final design. Owing to these factors, many offshore structures are equipped with structural monitoring systems which collect long term data on the marine environment and the structure's response. Examples of typical data collected are wave elevations, fluid velocities and accelerations, wind velocity and direction, structural accelerations and structural strains. These full scale data are useful for verification of the predicted response of the structure and for future improvements in designs. The task of establishing dynamic models from these full scale data is known as *System Identification* and it is this field that this thesis addresses. It should be noted that collecting, storing and handling vast volumes of such data in the marine environment is an expensive and complicated procedure.

Accordingly measurement programmes need to be properly planned.

The system identification method allows one to build up mathematical models of dynamic systems based on experimental data obtained from such systems. Since dynamic systems are abundant in many fields, system identification (SI) techniques (Eykhoff, 1974, Ljung, 1987, and Soderstrom and Stoica, 1988) have widespread applications. In the fields of Communications, Mechanical Engineering and Geophysical Engineering, SI techniques are used for spectral analysis (Kay, 1988, Priestley, 1981, 1988, Bendat and Piersol, 1980, 1986), adaptive filtering, fault detection, linear prediction (Karl, 1989, Franklin and Powell, 1980 and Safak, 1989) and many other purposes. In Systems and Control Engineering, SI techniques are used to obtain proper models for design of prediction algorithms, simulation, or synthesis of regulators. In business and economics, SI techniques are used to forecast in business inventories, goods production, etc. (Harvey, 1987, Nazem, 1988, Pankratz, 1983). Most of the SI modellings in business and earthquake engineering (Beck, 1979) are for nonstationary processes. SI techniques are also being successfully used in other fields such as Biology and Environmental Science to develop models for in depth scientific understanding. More recently SI techniques are also being applied in Offshore fields (Samii and Vandiver 1984, Mason and Ullmann, 1990, Broome and Pittaras, 1990, Jefferys and Goheen, 1990, Witz and Mandal, 1991, Mandal, Witz and Lyons, 1992, Worden et al, 1992). However, because of limitations in the understanding of the complex ocean phenomena, SI techniques have not been used extensively in this field compared with other fields. Proper SI modelling can not only solve mass data storage and handling problems, but also can detect structural failures and other unknown phenomena in the presence of noise in measured offshore dynamic systems.

1.2 System identification procedures

Measured data can be used to identify and establish proper models by SI techniques for solving many unknown phenomena for which many laboratory experiments could not appropriately identify or solve the problems. SI algorithms are becoming powerful tools for solving those unknown/complex phenomena in offshore fields. If the observed (offshore) time series data can be well-defined by some mathematical rules, it is said to be a true system. But in practise, comparing certain aspects of the physical systems with its mathematical description can never establish an exact or true relationship. However, mathematical rules can establish as best as possible a representation to describe the physical system, from the point of view in usefulness rather than truth. Sometimes, owing to lack of measured data, approximation of the system by SI methods could reasonably well describe the dynamic system.

The dynamic system can be defined by an Input-Output system which is shown in Figure 1.1.

For constructing a model from the data, the SI procedure, in general, can be expressed in the steps given below.

a) *Input-Output* data - The input-output data are required for a specifically designed experiment. Since collecting any offshore time series data on a long-term basis is most expensive, one must have prior knowledge (as much as possible) about what are the data to be measured and when to measure. Therefore, the *experiment design* gives some choices so that the data become maximally informative. Several choices are available in *experiment design* . These are: which signals to measure,

when to measure them, which signals to manipulate, how to manipulate them, and how to choose presampling filters and how to polish them by removing trends.

b) Define some SI models which are to be used with the data and choose the most suitable one. This is the most difficult and important step of the SI procedure. If one does not know the physical background of the system as prior knowledge or information being available , standard linear models could be used (Ljung, 1987 and Marple, 1987). If the prior information which leads to the dynamic system is known, this may be a good starting point for choosing a suitable model. Some knowledge about the nature of relationships between the measured signals can lead to appropriate selection of model structures. Fast Fourier transform (FFT) analysis of the data (Newland, 1984) also can lead to a meaningful selection of SI model order, based on the spectral shapes. This type of model set, whose parameters are viewed in order to adjust the fit to the data and do not reflect physical concepts, is called a *black box*. The black box type of models are used for adaptive prediction of the dynamic system. The model set with adjustable parameters for physical interpretation, called a *grey box*, is used for estimation of parameters of the dynamic system.

c) The assessment of the model quality depends on the performance of the model. One must identify the best model which provides the best representation or reproduction of measured data.

d) Once a particular model is chosen based on certain criteria, it is to be confirmed that the estimated model is a realistic approximation of the actual system. This is known as *validation* of the model. So one has to test whether the estimated model is good enough for its purpose. Model

validation involves various procedures to assess how the model relates to observed data and to its intended use. The simplest test is to compare the estimated power spectrum with that obtained from the Fast Fourier Transform (FFT) analysis. The second test may be to compare the output time series of the SI model with the actual output. The output time series of the optimal SI model should give a fairly good match with the actual output. Another way of validating the SI model is to estimate the residuals of the model. The residuals estimation is based on the condition that the difference between the model output and the actual output should be a white noise process. If the estimated residuals of the SI model is closer to white noise, the better the model is. One can plot the residuals and its FFT spectrum and then look to see whether they are similar to those of a white noise process. Other checks can be made by using various statistical tests. In most practical applications, we are more interested in estimating the optimal SI model parameters rather than estimating the noise or residual model.

The above procedures can be briefly described in the loop as shown in Figure 1.2.

1.3 Modelling

The models are principally categorized into parametric and nonparametric types. Some of the features of both types are briefly discussed in this section and importance of parametric modelling in ocean engineering field is highlighted. The SI models, namely, AR, MA, ARMA and ARIMA algorithms can be used for parameter estimation of univariate random processes related to the ocean engineering field.

A SI model can be defined by any mathematical representation which approximates the relation between the input and output of a dynamic system. The SI models can be classified into two principal categories.

a) *Parametric models*: These are a particular type of SI model where the essential features of the input-output relations are described in the form of parametric polynomials. These parameters must be assigned values before the model is completely specified. Prior information will assist to determine the assigned values. Various parametric model structures are available (see Ljung, 1987 and Safak, 1989). In general the parameters must be estimated from the input and output of the dynamic system. As an example, time series data of a stochastic process could be expressed in the time domain by the difference equation as

$$y_t = -\sum_{k=1}^{p} a_k y_{t-k} + \sum_{l=0}^{q} b_l w_{t-l} \tag{1.1}$$

where a_k and b_l are unknown parameters of the model to be estimated, w_t is the white noise as input and y_t is the t^{th} sample of the discrete stochastic process as output.

Assuming that the model is linear and time-invariant, equation (1.1)

can be expressed in polynomial form in terms of z-transform notation as

$$Y(z) = H(z)\, W(z) \qquad (1.2)$$

$$\text{where} \quad H(z) = \frac{B(z^{-1})}{A(z^{-1})}$$

The transfer function H(z) containing unknown parameters A(z) and B(z) is to be estimated (see equation AII-5, Appendix-II).

b) *Nonparametric models*: These types of models have unknown parts in the form of functions rather than parameters. In these types of system it is required to assume

i) Finite memory

ii) Time-invariant

iii) Linearity for simplification of the problem

For nonparametric modelling, the system can be treated as a *"black box"*, since its aim is to determine a function which relates the input to the output without recourse to any prior information about the internal structure of the system. As an example, a time-invariant linear model with a single input and single output (SISO) can be expressed by the impulse function, h(t), and its input-output relationship is

$$y(t) = \int_0^\infty h(\tau)\, x(t-\tau)\, d\tau \qquad (1.3)$$

where x(t) is the input and y(t) is the output. Equation (1.3) can be written in z-transform as

$$Y(z) = H(z)\, X(z) \qquad (1.4)$$

Here h(t) and H(z) are arbitrary transfer functions to be estimated from the

input and output of the dynamic system. Whereas the transfer functions in equation (1.1) and (1.2) are in the form of unknown parameters. Therefore different identification procedures are required for each of the above cases. More details with examples on nonparametric modelling are given in Ljung (1987), Chapter 6.

Classical methods of estimating power spectral energy (PSE) from the time series data use discrete-time Fourier transform operations of the infinite autocorrelation sequence (ACS). This relationship between the PSE and ACS can be considered as a *nonparametric* description of second order statistics of the random process. In the case of a *parametric* model, the PSE of the time series model is a function of model parameters which are to be estimated.

The major motivation for use of parametric models of stochastic processes is the apparent higher spectral resolution achievable with these models than that achievable by classical approaches. The parametric approach could give users choices in ability to fit an assumed model with few parameters. Depending on the order selection and estimation of model parameters, the model will yield the *least squares error* which could be minimized by optimal selection of model order. The SI models, namely AR, MA and ARMA algorithms are formulated (see Chapter 3) to determine the spectral estimates of the stationary marine environmental and offshore monitoring data. Similarly for nonstationary dynamic offshore processes, ARIMA algorithms are formulated for parameter estimation as described in Chapter 5.

1.4 The application of system identification to offshore engineering problems

Most of the offshore engineering problems may be approximately solved using conventional analytical and empirical methods. Many complicated offshore phenomena, owing to lack of in-depth knowledge, are analysed based on empirical methods. From the design point of view, safety factors selected sometimes seem to be higher than required, making the design more costly. Structural fatigue, structural damping and external loading on structures in offshore environment are very different from the conventional land-based structural problems for which theories are well established. Even though the offshore dynamic problems are solved generally based on linear assumptions and modified land-based structural theories, in actual fact there are many unknown phenomena happening in the offshore environment for which present theories may not be of sufficiently useful. It is for these reasons that many offshore platforms have been equipped with structural monitoring systems to observe full scale dynamic responses which eventually will yield data which can be used for improvement in designs, manufacturing and operations of offshore structures. These full scale structural monitoring data need to be properly analysed. Using the conventional approaches, these analysed data may sometimes result in misleading information owing to unknown phenomena included in the data.

The data obtained from offshore structural monitoring systems can be analysed in many ways depending on the user's interest. It can be categorised into three primary uses. The first is that the data may be used in long term statistical models where there is little theoretical basis for

establishing extreme events (Patel and Witz, 1991). The structural monitoring data also provides a cumulative loading history which is important in establishing the service life of the structure and related inspection intervals. A third approach is to analyse the structural monitoring data by applying *system identification* techniques which relate the structure's response and environmental excitation. The structural models identified from measured data using SI algorithms may be compared with the theoretical models used in the design process. This leads to the estimation of appropriate *filter* parameters.

Most of the offshore time series recordings collected are in their raw form. The time series raw data can be plotted and visually inspected for quality of the data. Any abnormal data recorded in the time series can be easily noticed. An alternative way in which one can detect the bad data is by using a residual plot (Ljung, 1987, Chapter 16.5). These data may not provide a good identification of the dynamic system. Based on the study by Witz and Mandal (1991), it is shown that the SI techniques can be applied to remove the disturbances or noise modes present in offshore time series data. Much better identification can be obtained if the data is preprocessed prior to identification. The preprocessing may involve removal of mean and erroneous large peaks, filtering, sampling interval selection and synchronization of input and output.

The dynamic characteristics of any offshore system are complicated, and can be modelled by SI techniques. Removal of the mean, simplifies solution of the problem. Sometimes there may be erroneous large peaks at several points in the time series data owing to various reasons, such as radio or electrical interference and temporary sensor failure. These errors certainly disturb the identification which may cause misleading results. Large erroneous peaks should be removed prior to identification.

Selection of sampling interval also plays an important role in the SI

parameter estimation. The sampling interval is directly related to the frequency resolution of the time series data through the cut-off frequency (Nyquist frequency, f_N). The frequencies beyond f_N are folded back and superimposed over the lower frequencies in the spectrum. This process is called *aliasing*. To avoid aliasing, data should be filtered using an anti-aliasing filter which is a low-pass filter with cut-off frequency, f_N. In general, in offshore dynamic systems, the phenomena are described up to a certain frequency which may be much smaller than f_N. If this is the case, one need not use a high sampling rate. For high sampling rates, the SI algorithm will only identify the high frequency part of the dynamic system (Ljung and Soderstorm, 1983). Sometimes one may need a high-pass filter to eliminate the very low-frequency drifts in the time series data.

The last step in the preprocessing of data is the synchronization of the input and output. If data are not recorded in a synchronous way, it can be tackled by properly selecting the time delay between input and output during the identification. However, it is often difficult to select an appropriate time delay which must be obtained by trial and error. Therefore, it is better to use synchronized data without a time delay.

1.5 Review of previous work

1.5.1 Univariate SI models and applications

Cadzow (1980) presented a method for generating an ARMA spectral estimate of wide sense stationary time series data. The method is based on a set of error equations which are dependent on the ARMA model parameters. The deviation of these error equations with respect to the ARMA model parameters leads to the ARMA spectral estimates. Even though the method developed by Cadzow gives high performance as compared to the maximum entropy method (Burg, 1975) where much higher order estimates needed to reproduce the actual spectral estimates, an improved and more efficient ARMA spectral estimate model can be formulated as described in Chapter 3.

Generally, the numerical generation of sea wave records is computed based on the superposition of several harmonic waves. Although this approach is simple, it requires a large number of harmonic components and considerable computer time. The first to apply *linear prediction theory* were Spanos and Hansen (1981). This was an autoregressive (AR) algorithm for digital simulation of sea waves and can be used as an alternative, efficient method. They used the 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. Even though 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 prime aim of attention was to the potential usefulness of AR algorithm for ocean engineering applications.

Based on studies of *linear prediction theory* applied to sea wave estimation by Spanos and Hansen (1981), ARMA algorithms were applied to similar studies by Spanos (1983). It was shown that one should carefully select the sampling interval for developing an AR approximation of a Pierson-Moskowitz spectrum. Spanos also suggested that a quite high order AR modelling should be used to ensure proper matching. This is not always true. For least squares error converging with increasing model orders, one can always select a high model order. Otherwise for proper matching one should follow model order selection methods as described in Chapter 3. Spanos assumed ARMA algorithms were of the form of single degree-of-freedom linear spring-mass-damper system. Accordingly Spanos had investigated the use of a least squares approximation of ARMA modelling of Pierson-Moskowitz spectrum. Even though the ARMA spectrum appeared very similar to the Pierson-Moskowitz spectrum and gave improved results with higher order ARMA models, the matching inequality remained.

Later on, much improved ARMA modelling of Pierson-Moskowitz spectrum was carried out by Spanos and Mignolet (1986). It was shown that the fluctuations in the AR spectrum were associated with the presence of poles of the transfer function in the vicinity of the unit circle. They had shown how the initial AR approximation could lead to efficient ARMA models of the Pierson-Moskowitz spectrum. Using the Taylor approximation of exponential terms of the Pierson-Moskowitz spectrum, the AR spectral estimates reduced sharp fluctuations but could not

eliminate them. They used two alternative procedures to obtain ARMA coefficients:

a) *Auto/cross Correlation Matching (ACM)* where an ARMA representation of the AR filter can be obtained by matching the output auto-correlations with input-output cross-correlations.

b) *Power Order Matching (POM)* the ARMA model equivalent to the AR model. Here, by equating like powers of z (where the SI models are defined by z-transform), the coefficients of the ARMA models are obtained from the known initial AR coefficients.

Mignolet and Spanos (1987) presented a unified approach in determining ARMA algorithms for simulating a random process based on the target spectrum. The ARMA algorithms were obtained by relying on a prior AR approximation of the target spectrum. AR to ARMA procedures were formulated by minimizing the frequency domain error. For determining ARMA parameters two approaches, namely, ACM and POM were studied in detail. It was shown that there are computational advantages of the POM procedure over the ACM procedure in terms of the size of the system of linear equations.

A finite-order stationary ARMA model was obtained from an infinite-order AR model by equivalence relation as described by Li, Zhu and Dickinson (1989). In practice, it is not possible to obtain the parameters of an infinite order AR model. Therefore, some approximations have to be made by using a higher order AR model. Li et al presented a comparison study for two methods of estimating ARMA(p,q) parameters p and q . One method was derived directly from the equivalence relation (Graupe, Krause and Moore, 1975). The other

one was derived by Li and Dickinson (1986, 1988) based on an iterated least square regression approach, where the ARMA parameters p and q are determined by first obtaining a p+q order AR model and then solving a set of linear equations similar to the method by Graupe et al.

Rosen and Porat (1989) presented a class of estimators based on the sample covariances for the time series data with missing observations. They proposed an algorithm which is based on nonlinear least squares fit of the sample covariances determined from the time series data to the covariances of the assumed ARMA model. While collecting time series data, the pattern of missing data can be quite arbitrary. As cases of special interest, Rosen and Porat considered two patterns: (a) random Bernoulli pattern - the data missing is of a fixed probability and misses are independent and (b) deterministic periodic pattern - missing data points are repeated periodically. The ARMA algorithm presented by them is asymptotically optimal, i.e., the error variance tends to the smallest value when the amount of data tends to infinity.

Cadzow (1982) attempted to establish the fundamental approach to the generation of rational modelling of wide sense stationary time series data. Rational modelling was carried out based on the modified Yule-Walker (MYW) equations which characterize the autocorrelation sequence of the rational time series data. By taking an overdetermined model parameter approach, a procedure for reduction in data-induced model parameters was obtained and then improvement in the modelling performance was carried out. Furthermore, adapting a singular value decomposition representation of the MYW equations to this procedure, a desired rational model order determination method was achieved. This approach yields low order high quality spectral estimates using short data

lengths.

Beex and Scharf (1981) proposed a systematic procedure to covariance sequence approximation for parametric spectrum modelling. This approach was represented by the approximation of a covariance sequence of a wide sense stationary process in a modal decomposition. For the special class of processes with modal decomposition, there is a random synthesis algorithm that may be used for time series data reduction. They used the first order mode decomposition technique.

Friedlander and Porat (1984) made some clarifications and put in proper perspective the various issues related to the MYW method. While reviewing the different versions of the MYW method, they exposed the common framework of the stochastic process leading to MYW equations and fitting a rational model to a noisy impulse response of a linear time-invariant system. They also emphasized the importance of using a combination of an overestimated order model and an overdetermined set of equations. They used a procedure for removing spurious noise modes based on the modal decomposition of the sample covariance sequence as proposed by Beex and Scharf (1981). The concept of modal energy is defined to select the signal related models and to discard the noise modes. They also reviewed the singular value decomposition method for solving MYW equations. The estimation of MA spectral parameters seems to be the more difficult part of the ARMA spectral estimation problem. Even though this paper described many techniques for evaluating MA spectral parameters, none of them guarantee the positive definiteness of the MA correlation sequence, which sometimes estimates negative spectra.

Samii and Vandiver (1984) presented a numerically efficient

technique for time domain simulation of water particle velocities and accelerations corresponding to a target wave spectrum. The ARMA algorithms were applied for estimating time series data on wave velocities and accelerations. They used theoretical Bretschneider wave spectra for ARMA spectral estimation. Based on available significant wave height and zero upcrossing period, the peak frequency and corresponding peak power spectral energy of the Bretschneider velocity spectrum can be obtained. The output of the ARMA model was processed by a series of numerical convolutions. Each convolution accounts for a horizontal or vertical shift to a different spatial location. Numerical differentiation of the vertical velocity yields acceleration at each point. Horizontal velocities and accelerations were determined by using a Hilbert transform. Simulation steps were discussed and presented in a block diagram.

Samaras, Shinozuka and Tsurui (1985) developed a technique to generate the sampling functions of a Gaussian vector process using an ARMA model. They used a two-stage least squares method to determine the coefficient matrices of the ARMA models. The numerical example showed that the sampling functions generated by the method developed by them reproduced the target correlations extremely well. This was observed between the analytical target auto and cross correlations and the corresponding correlations obtained from the generated sample functions. But they did not carry out optimal selection of model orders which is important for better estimation of ARMA representation.

Popescu and Demetriu (1990) had shown that the nonstationary analysis technique and representation of the quasi-stationary data blocks of earthquake ground motions through parametric ARMA models provide an efficient and flexible description of the observed motion by a few

parameters. The problem of nonstationary time series data was solved by segmenting the original data into different data blocks which were considered to be quasi-stationary. Using evaluation by the *Akaike information criterion*, an ARMA model was fitted for each quasi-stationary data block. The original data and predicted data obtained from ARMA models in a case study were compared by evaluating a number of statistical characteristics such as cumulative energy, cumulative root-mean-square acceleration, short-time energy, short-time autocorrelation, and short-time spectrum distributions. It was shown that there is a good acceptance match for all above statistical distributions.

In general many authors developed ARMA process based on the initial higher order AR approximation. Spanos and Mignolet (1990) introduced a new concept as an alternative approach of estimating ARMA process from the initial MA approximation, where the time series data are poles dominated. The MA parameters were obtained first relying on the maximization of an energy-like quantity, then the ARMA algorithms were derived from the initial MA approximation. This was achieved by relying on the minimization of frequency domain errors. It was observed that the initial higher order AR approximation sometimes creates a sensitiveness for proper representation of ARMA process estimation. Spanos and Mignolet emphasized that the MA to ARMA approximation technique can be used as an alternative approach of AR to ARMA modeling of the stochastic process. It was observed that the MA to ARMA of order (p, q) approximation sometimes shows very good results for the model orders $q \geq p$.

Li and Ko (1988) applied autoregressive, moving average and autoregressive moving average models to structural failure detection and

monitoring for offshore applications. Li and Ko used the Green's function of the time series and the impulse response function of a vibrating system for formulation in terms of AR and MA parameters while MA parameters were checked by the inverse function of the time series for the control of numerical convergency. The ARMA(2n,2n-1) was developed under the strategy of modelling as described by Pandit and Wu (1983). The variation of the absolute value of the modal characteristic roots could be used as the index for judgement of the degree of relative damage. The location of damage was judged by comparing of the variations of dispersions of the major modes detected by different accelerometers. Later on Li (1991) conducted a series of progressive tests in sea trials of ships to obtain the response signals of structural vibration aboard ship and process the random time series data by a linear difference stochastic modelling (ARMA) technique to study the estimation of system damping ratios of the ship hull-girders in response to random environment. For solving the dynamic equation of multi-degree systems of ship structures, one can choose reasonable values of modal damping ratios which can be determined using ARMA(2n,2n-1) model.

Mourjopoulos and Paraskevas (1991) carried out all-pole (i.e., autoregressive) and all-zero (i.e., moving average) model approximations of transfer functions with application to acoustic signals. While examining the above modelling, two main problems associated with transfer functions are apparent. These are their high arithmetic order, and their sensitivity and dependence on specific source placement. An all-pole transfer function model order was optimized using an information criterion (Akaike, 1974) method. It was shown that in a time series digital acoustic signals, the optimal order of an all-pole model increases with an increase of data points (two to the power of an integer number). It was

found that all-pole models present significant advantages over all-zero models because of their lower order and lower sensitivity to source placement variation. An all-zero model can yield an exact model of a spectrum within an approximation error. However it generates a mismatched time domain model owing to sensitivity to changes in source placement. The sensitivity of an all-zero model to source placement variations can be explained by the nature of acoustic signals. Transfer function zeros result from local cancellations of multi-path sound components which are easily disturbed by changes in source positions.

Kaplan, Jiang and Dello-Stritto (1981) described a sequential estimation technique of system identification for determining coefficients of Morrison equation which is used to describe the hydrodynamic loading on slender offshore structures. The method of system identification was applied to estimate the state-space variables and parameters in a noisy nonlinear dynamic system. The least squares estimate of state variables was obtained from minimizing the integral of weighted mean square errors. They also studied on-line filtering action and compared measured forces, velocities, etc. with estimated values in the time domain. The unknown force coefficients estimated by the SI technique were found to be reasonably constant and the estimated force time histories generally exhibited Morrison's equation model results.

Methods relevant to identification of linear and nonlinear behaviour of structures subjected to environmental loadings, such as ground motion owing to earthquakes (seismic motion), wind generated pressures and ocean wave forces were reviewed by Imai, Yun, Maruyama and Shinozuka (1989). The methods used were least squares, instrumental variables, maximum likelihood and extended Kalman filter. The dynamic

characteristics of the structural behaviour could be described by system identification models. Models commonly used in structural engineering are state-space and ARMAX (Auto-Regressive Moving Average eXogeneous) models derived from the ordinary differential equations describing equilibrium of the structures. Numerical simulation studies were carried out for identification of the aerodynamic coefficients of a suspension bridge under wind forces, drag coefficients of an offshore structure under wave forces, and displacement and stiffness ratio of a building structure subjected to seismic excitation. Numerical results showed that the instrumental variables and maximum likelihood methods provide good estimates for a linear system and the extended Kalman filtering technique gives excellent estimates for non-linear system.

Jefferys and Goheen (1990) carried out studies on parametric modelling of marine dynamic systems, such as the dynamics of a floating body in waves, and the surge radiation forces of a tension leg platform (TLP). They used two methods to estimate the transfer functions. The first one is the indirect frequency response curve fitting method which produces reliable transfer function models. This method optimizes in the square norm sense in the frequency domain. The second method involves three stages: production of input-output time-domain series by inverse Fourier transformation; identification of ARMA discrete time models by SI, and then a transformation back to continuous time by a mapping from z-plane to s-plane. Both the above methods were shown to work on frequency response data derived from a known transfer function.

Mason and Ullmann (1990) carried out an experimental study on evaluation of structural damping in a major diagonal member of an

offshore steel jacket in the fabrication yard. The response histories (displacement and acceleration) of the diagonal member in air were measured and then the system identification technique, namely, output error algorithm which is not restricted to single degree-of-freedom system was applied. It was shown that the response histories reconstructed from system parameters identified by the output error method were often visually indistinguishable from the corresponding recorded data.

An advanced statistical method was applied to analyze the wave induced forces acting on the free-to-surge vertical cylinder (Sajonia and Niedzwecki, 1990). The experimental data were used to develop an autoregressive wave force model which is capable of accounting for localized flow history effects. It was noticed that a high frequency force component which was not accounted for in the Morrison equation was quantified using the AR model. The AR model with higher order can improve the force prediction for the above experiment. Using experimental data and Morrison equation resulted in a root mean square error of 24% and multiple correlation coefficient of 0.71. The AR model reduced the root mean squares error from 24% to 9% and increased the multiple correlation coefficient from 0.71 to 0.83, and accounts for the high frequency components. The research study carried out showed significant improvements in hydrodynamic force prediction through the development of a wave force model which was expressed by an autoregressive algorithm.

Broome and Pitteras (1990a) carried out work on adaptive ship motion prediction which is based on mathematical models generated in real time by using system identification techniques. They used ARMAX algorithm with the option of obtaining ARMA and AR parameters

estimation. Using a recursive least squares algorithm, the unknown SI parameters were determined based on the input-output measurements and then the one-step-ahead predictor was estimated. To determine the appropriate model order they used a loss function which is described in the form of a least squares error variation with model orders. At first an AR(5) model was chosen as the least squares error did not decrease appreciably for model orders greater than 5. A low order ARMA(2,4) was also chosen which could represent ship motion fairly well. The best approach chosen was based on the search of ARMA models from (2,1) to (2,15) which showed that the optimal AIC was at the ARMA(2,14) model. Later on an AR(20) model was used for ship motion prediction based on the study carried out by Broome and Pitteras (1990b). The method used here could not perform well for the large amplitude ship roll motions which needed to be investigated. However prediction theory used by them can be applied for short time prediction of ship motions.

Based on studies by Spanos and Hansen (1981) and Spanos (1983), it was observed that the AR spectrum fluctuation are in the peak frequency region of the theoretical ocean wave spectrum. And even by selecting higher model orders, the AR spectral shape remains sharply deviated from theoretical ocean wave spectrum near the peak frequency region. This instability problem was studied by Medina and Sanchez-Carratala (1991) to represent a robust AR algorithm of the theoretical JONSWAP ocean wave spectrum. They also introduced a reasonable criterion for qualifying the goodness of fitting a proposed ARMA model to a target ocean wave spectrum. After establishing this criterion and considering most existing techniques to define ARMA models, a new robust AR representation could be obtained. An extra white noise was added with a variance of $0.0025m_0$ (where m_0 is the zeroth spectral moment) to the AR

model to generate a robust AR representation of the JONSWAP spectrum.

A broad perspective review on spectrum analysis of discrete time series was carried out by Kay and Marple (1981). Many new techniques were developed in the sixties, seventies and early eighties. These include the classical periodogram, classical Blackman-Tukey, autoregressive (maximum entropy), moving average, ARMA, maximum likelihood, Prony and Pisarenko methods. These were all presented in a unified framework and with common nomenclatures. All methods were tabulated for comparative studies including their type of model structures, key references, appropriate equations for computation of each spectral estimate, advantages and disadvantages.

Very recently Pires et al (1992) carried out a comprehensive study on adaptive AR modelling of measured sea waves off the Portuguese continental coast and Azores islands with water depth ranging from 40 to 100m. They demonstrated that the adaptive AR modelling performed better in high seas than in calm seas. One justification could be the existence of a certain level of noise originated in the measuring, recording and digitizing equipment. The domination of noise in calm sea waves may sometimes lead to erroneous results.

1.5.2 Determination of optimal model orders

Anderson (1963) discussed the determination of the order of dependence in a Gaussian autoregressive process explicitly as a multiple decision problem. A sequence of tests of the models was carried out starting from highest order to the lowest order. This procedure can be applied to a real problem provided one specifies the level of significance of the test for each order of the model. As it is difficult to choose the levels of significance, the essential problem of optimal model order determination remains. The loss function of the decision procedure defined by the probability of making incorrect decisions leads to a situation where the order of the true structure will always be infinite.

Akaike (1969) first introduced a criterion which is called the *Final Prediction Error* (FPE). The appropriate AR model order was selected based on the FPE method, where the average error variance is minimized. Here the average error variance is the mean of the sum of squared errors between the target and AR process. For the AR process, the FPE is defined as

$$\mathrm{FPE}(p) = \varepsilon_i \frac{M+(p+1)}{M-(p+1)} \tag{1.5}$$

where M is the total number of samples, p is the model order, and ε_i is the average of the sum of errors between observed and estimated data. The appropriate AR model order, p is selected for which the FPE is minimum. If the random process purely consist of zeros, the model orders are fairly well selected by FPE method. However offshore dynamic systems are

random in nature and mostly consist of poles and zeros. Therefore one should define criteria for model order selection for specific applications.

Akaike (1974) reviewed statistical hypothesis testing in time series analysis. It was noticed that the hypothesis testing procedure was not adequately defined as the procedure for statistical model identification. The problem of determining a finite order model structure can be solved by approximation of the true structure by the model. Based on reviewing the classical maximum likelihood estimation procedure, Akaike introduced a new estimate which is called after his name as the *Akaike Information Criterion* (AIC) estimate. It is defined as

AIC = - 2 Log [Maximum likelihood] + 2 [Number of adjusted parameters] (1.6)

For analysis of any random process, the exact definition of likelihood function is generally too complicated for practical use. Therefore some approximation is made based on the Gaussian distributed random process. Instead of maximum likelihood, the mean log-likelihood was chosen for the criterion of fit of a statistical model. From the above relation, a more applicable simplified AIC is defined by Marple (1987) as

$$AIC = M \, Ln \, [\varepsilon_i] + 2 \, p \qquad (1.7)$$

This provides a versatile procedure for statistical model identification. The ambiguities inherent in the application of conventional hypothesis testing procedure for statistical model identification resulted in very limited practical utility. Since the procedure based on the AIC estimate can be implemented without the aid of subjective judgement, many statistical identification procedures with the AIC estimate could be made

practical. It must be noted that the AIC method can not be compared with a hypothesis testing method unless the latter is specified with the required levels of significance.

Parzen (1974) described some of the important concepts and techniques which may help to provide a solution of the stationary time series problem. A comparative study between the measured signal plus noise and the ARMA representation was carried out. He reviewed prediction theory and developed criteria of closeness of AR, MA and ARMA models to the 'true' models. The central role of the infinite AR transfer function was developed and then tested with time series modelling. He also introduced a criterion for selecting model orders and it is termed as *Criterion Autoregressive Transfer* (CAT)

$$\mathrm{CAT}(p) = \frac{1}{N}\sum_{i=1}^{p}\frac{1}{\bar{\varepsilon}_i} - \frac{1}{\bar{\varepsilon}_p} \tag{1.8}$$

where $\bar{\varepsilon}_i = \frac{N}{N-i}\varepsilon_i$

This criterion will determine a finite AR model order which is optimal.

Many researchers had noticed that the order selected by AIC is too low for non-autoregressive processes. Kashyap (1980) had found that the AIC is not statistically consistent. The result was a tendency to overestimate the order as the data record length increases which was also noticed by Mourjopoulos and Paraskevas (1991). Because of inconsistency of the AIC estimation, Rissanen (1983) developed a variant information theoretical criterion to the AIC which is called *Minimum Description Length* (MDL). Rissanen introduced the MDL principle for minimization

of the number of binary digits required to represent the observed data. This method gives an optimum length relative to a class of parametrically given distributions. It permits estimation of the number of parameters in statistical models, their values and the model structure. Rissanen also described a procedure for truncating the real valued Maximum Likelihood estimates to an optimum precision for the final criterion. The MDL criterion had been shown to lead to strongly consistent estimates of the model parameters and their numbers in AR and ARMA processes.

Jones (1974) carried out a study on identification of model orders and AR spectrum estimation. While selecting model order, Jones used the AIC method. He extended the AIC method from univariate AR process to multivariate AR processes. The comparative study between AR spectrum estimation and classical spectrum estimation was shown to be consistent in respect of the model order selection. He tested the above from the analysis of the large amount of digital data from the biological [two channels of electroencephalographic (EEG) data from a human newborn] and physical sciences [meteorology, i.e., wind data at two stations]. It showed that the AIC method worked very well for model order estimation, and the multivariate AR model order estimation proposed by him showed consistently good results.

1.5.3 Multivariate SI processes

Morf and Kailath (1975) introduced some new algorithms for recursive estimation algorithms for linear systems based on the Kalman (1960) filter technique. The solution of a matrix can be obtained using square-root algorithms in the least squares sense. They first presented an instantaneous derivation of a form of the previously known covariance square-root array algorithms. Then it was shown how the assumption of constant model parameters could be used to reduce the number of variables in the array algorithm. Finally updating equations were obtained by explicitly specifying the orthogonal transformations used in the array methods.

Based on the above technique for estimation of the square-root matrix, Morf, Vieira, Lee and Kailath (1978) applied this to stationary discrete time scalar processes. The autocorrelation of a stationary discrete time scalar process can be characterised by the partial autocorrelation function which is a sequence of values less than or equal to unity. They had shown that the matrix covariance function of a multivariate stationary process could be characterized by a sequence of matrix partial correlations, which were obtained using techniques based on forward and backward prediction of the time series data sets, where singular values are less than or equal to unity in magnitude. The squares of the singular values of a matrix, R are the eigenvalues of RR^T. Morf et al. presented a procedure to estimate a sequence of matrix partial correlations directly from the multivariate data. From these estimates Morf et al. uniquely determined minimum phase multivariate AR parameters and hence a unique power spectral estimate was obtained.

Strand (1977) described a multivariate complex maximum entropy (autoregressive) method for estimating spectral parameters. He extended this to generalize the univariate Burg reflection coefficient estimation (Burg, 1975) process to multivariate complex time series. It was shown that least-squares estimation of complex matrix reflection coefficients using inverse-power weighting provides a sequence of positive definite power matrices. This yields a resulting positive definite autocovariance matrix. Based on the preliminary numerical results obtained from a monochromatic signal with noise, it was noticed that superior spectral resolution can be expected from the extended multivariate Burg processes.

Jones (1978) reviewed the univariate algorithms and discussed multivariate generalizations of autoregressive algorithms using residuals within the data span. The problems encountered in generalizing Burg's maximum entropy algorithm to multivariate time series were highlighted. Burg's algorithm did not generalize directly since the forward and backward autoregression matrices are not the same in the multivariate case, and the forward and backward one-step prediction error covariance matrices are different. Therefore this leads to different estimates of the power spectrum. This problem does not arise in case of univariate AR processes, where only a single sequence of reflection coefficients exists. This fundamental difficulty was overcome by Morf et al (1978) where normalized partial correlations were directly obtained from the multivariate data and then multivariate transfer functions and prediction error covariance matrices were determined.

Lin (1987) proposed a multivariate ARMA model for prediction of a ship's response to random waves. Each degree of freedom of ship motions and each wave measurement was considered as one of the inputs and

outputs of the ARMA filter. Using the laboratory test data, it was shown that the agreement between predicted and measured response of the ship model was very good. The ARMA filter takes into account the directional response of the ship to the ocean waves. The directional effects of waves on pitch prediction were examined by a tri-variate ARMA model (which includes two sets of wave measurements and the pitch measurement) and a bi-variate ARMA model (which includes a wave measurement and the pitch measurement). It was shown that the tri-variate ARMA model is superior to the bi-variate ARMA model for different prediction time steps. This demonstrated that the directional effects of waves are very important for ship motion prediction. The coupling effect on ship motion prediction was studied by a quad-variate ARMA model (pitch, heave and the two wave measurements). This showed further improvement in prediction of ship motions as compared to that of tri-variate ARMA modelling.

1.6 Objectives

Offshore structures are designed based on the selection and application of theoretical models representing marine extreme conditions. Owing to limitations in assumptions while modelling a structure and variation of extreme marine conditions, there is always an element of uncertainty associated with the predicted response of the final design. For these reasons, full scale real time long term data on the marine environment and the structure's response are measured enabling future improvement in design. Collecting, storing and handling vast volumes of such data in the marine environment is not a simple process. Also there may be noise or disturbances collected along with actual response of the structure or environmental data which must be correctly dealt with.

The study presented herein has investigated the analysis methods arising from identification aspects of the theory of dynamic system. Using SI methods, parametric models of dynamic systems are determined based on input and output of the stochastic processes. Full-scale offshore monitoring data or marine environmental data can be substantially reduced to a few parameters of the SI models which can also further produce the estimated spectrum of the stochastic process for immediate requirements. At the initial stage, the parametric models' orders are obtained based on auto correlation and or cross correlation, least squares error, Akaike information criterion and minimum description length methods. Further reduction of the model orders can eliminate spurious noise or disturbances. Using SI techniques, the parametric models can also be applied to simulate time series data based on the target spectrum.

Sometimes nonstationary time series offshore data which are a combination of high frequencies and low frequencies are observed. These

types of data can be modelled provided the nonstationary time series data are transformed to stationary time series data by a differencing technique prior to modelling as demonstrated in Chapter 5.

The multivariate autoregressive (MAR) modelling of ocean waves can provide not only the required power spectra but also the phase and coherence which lead to information about wave directionality. Using traditional FFT methods, one can estimate the power spectrum. However, there might be spikes due to disturbances which one may not be able to discard. Using MAR modelling the required spectra can be estimated as shown in Chapter 6.

1.7 Summary of contribution

The principal aim of this research work has been to formulate a practical approach which will allow the optimal estimates of parameters of structural models to be determined systematically from the time series recordings of ocean waves and offshore structural motions. This approach can be used for *in situ* data reduction, which can store much longer duration information of offshore dynamic systems than the present methods of data storage.

The linear SI models are chosen here partly because they are simple and easily formulated, and partly because they are a natural starting point. Because the linear time-invariant discrete models are commonly used in dynamic design, identification of these types of models is of practical importance in offshore dynamic systems. These can be used either in the response spectrum approach or through the SI models and particular offshore time series records to estimate full response histories. The main aims of this work have been to investigate (a) how well the time-invariant linear discrete models fit with target spectra of ocean waves and offshore structural motions in stationary and nonstationary cases, (b) how the model orders can be reduced to obtain acceptable reduced order ARMA models, and (c) multivariate autoregressive models, and their practical applications.

Since the use of SI techniques as applied to ocean dynamic system is an unusual approach for offshore engineers and ocean scientists, an introduction has been included covering system identification procedures, parametric and non-parametric modelling, application problems of offshore raw data in Chapter 1. For on-line estimation, one has to use the

delay-time or recursive filtering technique. The study presented herein considers off-line estimation procedure as a number of data sets or blocks to be used for parameter estimation. The parametric modelling technique is suitable for off-line estimation.

The main previous works relevant to the study presented herein are described in Chapter 1 (section 1.5). Most of this related work and theoretical development took place in the electrical, system and control engineering fields. In the offshore field, very little related work has been carried out but that which has is reported here. The first part of the review of previous work mainly concentrates on establishing the SI techniques, namely, autoregressive, moving average and ARMA algorithms which have been subsequently developed for offshore dynamic systems. The Yule-Walker equations may be used to obtain the AR model. For rational modelling, an extended version of Yule-Walker approach yields efficient ARMA models. Some of these algorithms are used in other fields. The second part describes the appropriate selection of the SI model order. This part optimizes the model structure and minimizes the computational time and hence leads to a more efficient representation of the dynamic system. Finally multivariate autoregressive modelling is reviewed.

Chapter 2 gives an overview of stochastic processes and linear systems leading to an information criterion. This chapter describes how well the random processes can be represented by the linear time-invariant discrete models and leads to the estimation of power spectral energy. Random processes are also shown to lead to the estimation of an information criterion. The basic notations for the SI modelling introduced in this chapter are used in later chapters.

In Chapter 3, the identification of the various suitable SI models for given target processes is carried out as highlighted by system identification

procedures (section 1.2). The SI models, namely, AR, MA and ARMA models are formulated. A meaningful AR power spectrum can be obtained from the target (either power spectra or time series data) based on the appropriate selection of model order. Because of some inherent properties of AR modelling, the Pierson-Moskowitz wave spectrum is shown to need modification for better results. The estimation of MA parameters is carried out from the perspective of Fourier approximation of a decomposition of the target spectrum. The ARMA modelling is carried out based on the initial AR parameters estimation. Two approaches, namely, power order matching and inverse AR methods are used for estimation of the ARMA parameters. A comparative study between those two approaches of the ARMA modelling is carried out for theoretical wave spectra (PM and JONSWAP) and measured ocean waves. These SI modellings are also applied to offshore structural motions.

Model order reduction techniques applied to the ARMA algorithm are described in Chapter 4. Firstly the initial higher model order is selected based on the AIC or MDL method and then the AR coefficients are determined using the modified Yule-Walker equations. Then the first and second order real modes are obtained from the AR polynomial. A method of calculating energy in each mode is described in Appendix-V. Each modal energy contributed for the SI model is determined and then only higher energy modes which form the AR part of the reduced order ARMA model are considered. The moving average part is calculated based on partial fraction and recursive methods. This reduced order ARMA modelling is applied to ocean waves and offshore structural motion.

Chapter 5 describes the application of SI modelling to nonstationary offshore dynamic systems. The nonstationary time series data set can be

modelled using autoregressive integrated moving average (ARIMA) algorithms which are defined and formulated depending the nature of the nonstationary process. Nonstationary generated ocean waves and offshore platform deck (Magnus) displacements are used for ARIMA modelling.

Multivariate autoregressive modelling and its application to ocean waves are presented in Chapter 6. The MAR model is formulated based on the estimation of the residual variance matrices and partial correlations of the multivariate processes. Here the appropriate model orders are selected based on auto and cross correlations and the multivariate AIC methods. These algorithms are applied to estimate the power spectral energies and their phase and coherence spectra of two time series wave data sets collected at a North Sea location.

The last two chapters give an overall discussion and conclusions based on the various SI modellings and their applications to offshore dynamic systems.

1.8 Applications

As indicated in the foregoing sections the techniques developed and adapted in this study have been applied to a number of sources of real marine environmental and offshore structural monitoring data. Some of this which has been made available is considered propriety. As a result full details of some of the sources can not be made available herein. The data has been made available from

Table 1.1 Measured time series data sets of offshore dynamic systems and sources.

Source	Location/ Offshore structures	Data set	Date
Rijkswaterstaat, Netherlands	North Sea waves from MPN platform	Waves M12 and M17	22- Jan- 1980
Indian Navy	Sea waves from West Coast of India	Waves K15	28- May- 1989
BP, UK	Magnus Platform	Deck Hor. Acceleration	22 - Oct- 1987
Santa Fe Drilling Co, UK	Semisubmersible Rig135	Heave Acceleration HA28	11- Aug- 1989
McDermott, UK	DB50 Crane vessel A	Roll	23 - Oct- 1987
Anon.	Crane vessel B	Pitch, Roll, Heave Acc.	12- Apr- 1992

CHAPTER 2

STOCHASTIC PROCESSES AND LINEAR SYSTEMS OF SPECTRAL ESTIMATION

2.1 Introduction

Power spectral estimation has generally been a traditional research area for statisticians (Anderson, 1971). Recently it has been extensively used for engineering applications (Bendat and Piersol, 1986, Newland, 1984, Witz and Mandal, 1991). Most of the statistical analyses are carried out based on restrictive assumptions about the nature of the data, that is, whether it is Gaussian distributed. The art of spectral estimation lies more on empirical relationship than on a theoretical one. Using a spectral analysis method, any signal or time series random process can be characterized in the frequency content. The strengths of the signal or time series in the frequency domain can be quantified by the power spectral energy estimation techniques.

Power spectral energy estimation has traditionally been based on Fourier transform techniques. A primary motivation for this research study in alternative methods is to improve performance with minimum loss of information while expressing the spectral estimates by a few parameters of the rational functions.

This chapter gives an overview of linear systems (Sinha, 1991, Hannan and Deistler, 1988, Oppenheim and Schafer, 1975, 1989) and

stochastic processes leading to the estimation of information criteria and power spectral energy of dynamic systems. Many of the notational conventions introduced in this chapter are used in later chapters. Stochastic processes leading to the estimation of linear system parameters and power spectral energy are presented in section-2.2. A *linear system* (discrete or continuous) can be expressed by superposition of responses. Suppose that there are two input signals to the response of a linear system, then the system can be expressed by simply the sum of the separate system responses to each individual input signal. The linear system is said to be *time-invariant* if the inputs and outputs are time independent, for example, an input x_t produces y_t and x_{t+t_0} produces y_{t+t_0} for any time shift t_0. Since the present study is restricted to discrete system (as offshore time series data are collected that way), linear time-invariant discrete processes are presented in section-2.3. Section-2.4 describes the discrete linear processes which can be analyzed by using the Fourier series transform. While formulating random processes in the form of linear systems, one has to verify or check the model quality. One such important test is the stability criterion which is described in section-2.5.

2.2 Stochastic processes leading to information criteria and PSE estimation

The concepts of probability and stochastic process theory is briefed here for formal introduction of information criteria and power spectral energy estimation. This section is segmented into four subsections dealing with probability and random variables, random processes, power spectral energy estimation, and convergence of random sequences.

2.2.1 Probability and random variables

Let X be a random variable of some experimental outcome which cannot be exactly predicted in advance. Mathematically the properties of X are quantified by a distribution function, F(x), which is the probability that the random variable X has a value less than or equal to x or Pr(X<x). F(x) is a nondecreasing function with limiting values of $F(-\infty) = 0$ and $F(\infty) = 1$. The probability density function (PDF), p(x) is expressed as

$$p(x) = \frac{dF(x)}{dx} \tag{2.1}$$

For discrete random variables, X takes one of the finite number of values $x_1, x_2, \ldots$ with corresponding probabilities $p_1, p_2, \ldots$ which must satisfy the conditions

$$p_i \geq 0 \quad \text{and} \quad \sum_i p_i = 1 \tag{2.2}$$

Then the distribution function,

$$F(a) = Pr[X < a] = \sum_{x_i < a} p_i \tag{2.3a}$$

If $a \le X < b$ then

$$Pr[a \le X < b] = F(b) - F(a) \tag{2.3b}$$

which is a piecewise constant function with a jump of height p_i at x_i (Figure 2.1).

For continuous random variable, the distribution function can be defined as

$$F(a) = \int_{-\infty}^{a} p(x)\, dx \tag{2.4}$$

and p(x) must satisfy

$$p(x) \ge 0 \quad \text{and} \quad \int_{-\infty}^{\infty} p(x)\, dx = 1 \tag{2.5a}$$

If $a \le X < b$ then

$$Pr[a \le X < b] = \int_{a}^{b} p(x)\, dx \tag{2.5b}$$

An important parameter of a random variable is its expectation or mean value, denoted by E[x] which is given by

$$E[x] = \sum_{i} x_i\, p_i \qquad \text{(discrete)} \tag{2.6a}$$

$$E[x] = \int_{-\infty}^{\infty} x\, p(x)\, dx \qquad \text{(continuous)} \tag{2.6b}$$

This is also called the first moment of x. The expectation of x squared is defined as

$$E[x^2] = \sum (x_i)^2\, p_i \qquad \text{(discrete)} \tag{2.7a}$$

$$E[x^2] = \int_{-\infty}^{\infty} x^2\, p(x)\, dx \qquad \text{(continuous)} \tag{2.7a}$$

This is also called the second moment of x.

The expectation of a function, g(x) of random variable x can be directly calculated using the PDF of x as

$$E[g(x)] = \int_{-\infty}^{\infty} g(x)\; p(x)\, dx \qquad \text{(continuous)} \tag{2.8a}$$

$$E[g(x)] = \sum g_i\, p_i \qquad \text{(discrete)} \tag{2.8b}$$

The variance, ρ of the random variable is the mean squared deviation of the random variable from its mean.

$$\text{var}\{x\} = E[x^2] - (E[x])^2 = \rho \tag{2.9}$$

The standard deviation, σ of x is

$$\sigma = \sqrt{\text{var}\{x\}} = \sqrt{\rho} \tag{2.10}$$

The covariance is defined as the statistical correlation between one random variable and another random variable. If X_1 and X_2 are any pair of finite variance random variables, the covariance of X_1 and X_2 can be expressed as

$$\begin{aligned} \text{cov}\{X_1X_2\} &= E[(X_1 - E[X_1])(X_2 - E[X_2])] \\ &= E[X_1X_2] - E[X_1]\,E[X_2] \end{aligned} \tag{2.11}$$

If X_1 and X_2 are said to be independent for a purely random case (i.e., white noise) if

$$\text{cov}\{X_1X_2\} = 0$$

The uniform distribution of a real variable x is quantified by a uniform PDF as

$$p(x) = \frac{1}{b-a} \qquad \text{for} \quad a \le x \le b \tag{2.12}$$

The Gaussian or normal distribution of a real variable x with mean, $\bar{x}$ and variance ρ is characterised by a PDF given by

$$p(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left[-\frac{(x-\bar{x})^2}{2\sigma_x}\right] \qquad \text{for} \quad -\infty < x < \infty \tag{2.13}$$

Here it is assumed that the random variable, x_i (for i = 0, 1, 2, ... M-1) is statistically independent. A class of estimators is known as *maximum likelihood estimates* [see Appendix-I] which are based on a consideration of the joint probability of M observed values as a function of the parameter to be estimated. The maximum likelihood estimate is the value of the parameter for which the probability of the observed values is

a maximum. The maximum likelihood estimates method can be used to determine the best estimate of the model parameters. Equation (2.13) can be used to estimate the information criterion of the random process for appropriate SI model order selection.

Let $x_1, x_2, \ldots x_M$ represent the results of M independent observations of a random variable with PDF p(x). If the parameter family of the function is expressed by $f(x|\theta)$ with a vector parameter θ which is to be optimized, the average log-likelihood can be expressed as

$$\frac{1}{M}\sum_{i=1}^{M} \ln f(x_i|\theta) \tag{2.14}$$

As M tends to infinity, the above average tends, with probability one, to

$$MLL = \int p(x) \ln f(x|\theta)\, dx \tag{2.15}$$

where MLL is the average or mean log-likelihood. From the efficiency point of view of the maximum likelihood estimate, it must be highly sensitive to small deviations of $f(x|\theta)$ from p(x). With some modification of equation (2.15) and using information theory, Akaike (1974, 1976) derived a final form of the information criterion, AIC as

$$AIC = -2 \ln [\text{ maximum likelihood }] + 2k \tag{2.16}$$

where k is the number of independently adjusted parameters to be selected.

2.2.2 Random processes

A discrete random process is a collection or ensemble of real or complex discrete sequence of time series observed values of any experiment. Mathematically it is just a collection $\{X_t, t \in T\}$ of observations of a random variable. Here T has the connotation of time, i.e., $\{X_t\}$ is a continuous time process if T is an interval, say, [a, b]; or (X_t) is a discrete time process if T contain only integer values.

The present work is limited to discrete time processes, i.e., digitized time series records of ocean waves, offshore structural motions, etc. Therefore, the random process theory presented here is restricted to discrete time processes. There are two ways to define the time series which can be modelled by discrete time processes:

a) Time series data which are only available in discrete form
b) Time series data which are produced by sampling continuous data

In the second case, one should be careful about the appropriate sampling rate to be chosen so that discrete time series data fairly represent the continuous data.

If T =[1, 2, ... M], then the random process is expressed as $\{X_t\} = [X_1, X_2, ... X_M]$, and its probabilistic behaviour is given by the joint distribution of the M random variables involved.

Even though an offshore dynamic system is a stochastic process with time, for calculation simplicity one can define this process as *stationary* if its distribution does not vary with absolute time, i.e., for any $[t_0, t_1, t_2, ... t_n]$, the distribution of the n vector random variables

$[X_{t_1}, X_{t_2}, \ldots\ldots X_{t_n}]$ is the same as that of $[X_{t_1+t_0}, X_{t_2+t_0}, \ldots\ldots X_{t_n+t_0}]$. This shows that the origin of time is irrelevant and the joint distribution of the random variables depends only on time interval separating them. Therefore this process has well-defined mean and covariance function. The process is said to be a *wide-sense stationary* (WSS) if its mean, a constant, and auto-correlation are independent of absolute time and the autocorrelation depends only on the relative time.

Let x_t be a stationary random process. At time index τ, the mean or expected value is defined as

$$\bar{x} = E[x_t] = E[x_{t+\tau}] \tag{2.17}$$

The auto-correlation of the random process at two different time indices t and $t+\tau$ is expressed as

$$R_{xx}(\tau) = E[x_t x_{t+\tau}] \tag{2.18}$$

In engineering applications, the term autocorrelation is normally defined as a relative quantity, called the normalized autocorrelation which lies between zero and unity. The autocovariance of the process is the auto-correlation process with mean removed,

$$\begin{aligned} C_{xx}(\tau) &= E[(x_t - \bar{x})(x^*_{t+\tau} - \bar{x}^*)] \\ &= E[x_t x^*_{t+\tau}] - E[x_t]\, E[x^*_{t+\tau}] \\ &= R_{xx}(\tau) - \bar{x}\,\bar{x}^* \end{aligned} \tag{2.19}$$

If the above process has zero mean for all t, then the auto-correlation and the auto-covariance are identical,

$$C_{xx}(\tau) = R_{xx}(\tau) \tag{2.20}$$

The cross-correlation of two random processes x_t and y_t can be defined as

$$R_{xy}(\tau) = E[x_t\, y_{t+\tau}] \tag{2.21}$$

Similarly cross-covariance is defined as

$$\begin{aligned} C_{xy}(\tau) &= E[(x_t - \bar{x})\,(y^*_{t+\tau} - \bar{y}^*)] \\ &= E[x_t y^*_{t+\tau}] - E[x_t]\,E[y^*_{t+\tau}] \\ &= R_{xy}(\tau) - \bar{x}\,\bar{y}^* \end{aligned} \tag{2.22}$$

If two random processes are uncorrelated then

$$C_{xy}(\tau) = 0 \qquad \text{for all } \tau \tag{2.23}$$

Some useful properties are

$$\left.\begin{aligned} &R_{xx}(0) \geq R_{xx}(\tau) \\ &R_{xx}(-\tau) = R^*_{xx}(\tau) \\ &R_{xx}(0)\,R_{yy}(0) \geq \left|R_{xy}(\tau)\right|^2 \\ &R_{xy}(-\tau) = R^*_{yx}(\tau) \end{aligned}\right\} \tag{2.24}$$

The above properties are valid for all integers τ. From these properties one can verify that autocorrelation must be a maximum at $\tau = 0$.

Based on equation (2.17), the random process can be rewritten after removing the mean ($\bar{x}$) from the original process as

$$\tilde{x}_t = x_t - \bar{x} \tag{2.25}$$

Let a linear system be expressed as

$$\tilde{x}_t = \sum_{k=1}^{p} a_k \tilde{x}_{t-k} \tag{2.26}$$

To fit this model [equation (2.26)] of order p [where k=1, 2, ...p], one can start with least squares method which needs the mean square of the residuals (Res_p) as

$$Res_p = \frac{1}{M}\sum_{t=1}^{M}(\tilde{x}_t - \sum_{k=1}^{p} a_k \tilde{x}_{t-k})^2 \tag{2.27}$$

Res_p is to be minimized with respect to unknown parameters $\{a_1, a_2, ..., a_p\}$ assuming that $\tilde{x}_t = 0$ for $t \leq 0$. To determine these unknown parameters, one has to use *Yule-Walker* equations, where equations (2.17) to (2.23) are needed. This relationship can be written in Toeplitz matrix, TA, as

$$[TA].\{a_k\} = \{C_k\} \tag{2.28}$$

where

$$TA = \begin{bmatrix} C_{xx}(0) & C_{xx}(1) & \cdots & C_{xx}(p-1) \\ C_{xx}(1) & C_{xx}(0) & \cdots & C_{xx}(p-2) \\ C_{xx}(2) & C_{xx}(1) & \cdots & C_{xx}(p-3) \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ C_{xx}(p-1) & C_{xx}(p-2) & \cdots & C_{xx}(0) \end{bmatrix}$$

and

$$\{C_k\} = \begin{Bmatrix} C_{xx}(1) \\ C_{xx}(2) \\ C_{xx}(3) \\ \vdots \\ \vdots \\ C_{xx}(p) \end{Bmatrix}$$

Based on residual estimates [equation (2.27)] and Yule-Walker equations, model order selection methods (AIC and MDL) are established as described in the later chapters.

2.2.3 Power spectral energy estimation using FFT

Fourier series techniques play an important role for the analysis of WSS processes, which leads to the so called power spectral theory of stationary processes.

The z-transform of the auto correlation and cross correlation sequences which are determined from time series x and y are defined as

$$S_{xx}(z) = \sum_{\tau=-\infty}^{\infty} R_{xx}(\tau)\, z^{-\tau}$$
$$S_{xy}(z) = \sum_{\tau=-\infty}^{\infty} R_{xy}(\tau)\, z^{-\tau} \qquad (2.29)$$

This leads to the definition of power spectral energy. The above equations are expressed in frequency scale

$$S_{xx}(f) = T \sum_{\tau=-\infty}^{\infty} R_{xx}(\tau) \exp(-j2\pi f\tau T) \qquad (2.30)$$

$$S_{xy}(f) = T \sum_{\tau=-\infty}^{\infty} R_{xy}(\tau) \exp(-j2\pi f\tau T) \qquad (2.31)$$

Here S(f) is a density function which represents the distribution of power with frequency, f and j is the square root of -1. The Fourier transform of the auto-correlation sequence is often referred to as the *Weiner-Khintchine* theorem. Owing to the properties of correlations, PSE must be a real and positive. For the autocorrelation to be strictly real valued,

$$R_{xx}(-\tau) = R_{xx}(\tau) \qquad (2.32)$$

Then the PSE can be expressed as

$$S_{xx}(f) = 2T \sum_{\tau=0}^{\infty} R_{xx}(\tau) \cos(2\pi f\tau T) \qquad (2.33)$$

Here

$$S_{xx}(-f) = S_{xx}(f) \tag{2.34}$$

which means that it is a symmetric function.

Therefore, the time series can be expressed by its autocorrelations which can be used to determine power spectral energy. This relationship is also shown in Figure 2.2.

The discrete white noise process with zero mean can be defined in the form of autocorrelation function as

$$R_{xx}(\tau) = \rho_x \, \delta(\tau) \tag{2.35}$$

where

$$\begin{aligned} \delta(\tau) &= 1 \quad \text{for } \tau = 0 \\ &= 0 \quad \text{for } \tau \neq 0 \end{aligned} \tag{2.36}$$

$\delta(\tau)$ is the discrete impulse function. This says that a white noise process is uncorreleted with all time lags except at $\tau = 0$. Therefore the PSE of the white noise process becomes

$$S_{xx}(f) = 2T\rho_x \tag{2.37}$$

which is a constant for all frequencies. The reason for the name *white noise* is by analogy with white light which has an approximately flat frequency spectrum.

2.2.4 Convergence of stochastic processes

While investigating random processes, one may wish to ask many questions such as whether a given process is stationary, whether parameter estimates converge to their true values with increasing data points, and so on. To answer the above questions, one has to study the convergence of sequences of random processes.

Let $\{X_k\} = X_1, X_2, \ldots$ be a non-random sequence of real numbers. Then one can say that $\{X_k\}$ converges to X, i.e.

$$X_k \to X \quad \text{as} \quad k \to \infty$$

$$\lim_{k \to \infty} X_k = X \tag{2.38}$$

The process $\{X_k\}$ converges to some random variable X if and only if $\{X_k\}$ is a *Cauchy sequence*, i.e, $|X_n - X_m| \to 0$ as n,m $\to \infty$. This means that for any small value $\varepsilon > 0$ there exists n(e) such that $|X_n - X_m| < \varepsilon$ for all n,m $\geq$ n(e). Here the definition of a Cauchy sequence means only to the elements of the sequence themselves and it does not consider any possible limit points.

Some properties of the convergence of random processes are highlighted in Appendix-I. For further details on convergence of random processes, one can refer to texts by Pollard (1984), and Hannan and Deistler (1988).

2.3 Linear time-invariant discrete processes

Linear time-invariant discrete processes form the most important class of random processes such as time series of ocean waves, response of offshore structures etc. Here some basic concepts are described below which will be instrumental in development of the SI models. The linear time-invariant discrete process can be expressed as

$$y_t = \sum_{k=-\infty}^{\infty} h_k x_{t-k} \qquad \text{for} \quad t = 0, 1, 2, \ldots \tag{2.39}$$

where x_t is the input, y_t is the output and h_k is the impulse response or weighting function. The above process is said to be time-invariant if its response to a certain input does not depend on absolute time. And the above process is also linear in the sense that the output response to a linear combination of inputs is the same linear combination of the output responses of the individual inputs.

The input-output relationship [equation (2.39)] with the transfer function can be expressed as

$$\begin{aligned} R_{xy}(\tau) &= \sum_{k=-\infty}^{\infty} h(k) R_{xx}(\tau - k) \\ R_{yx}(\tau) &= \sum_{k=-\infty}^{\infty} h^*(-k) R_{xx}(\tau - k) \\ R_{yy}(\tau) &= \sum_{m=-\infty}^{\infty} h(\tau - m) \sum_{k=-\infty}^{\infty} h^*(-k) R_{xx}(m - k) \end{aligned} \tag{2.40}$$

Denoting the z-transform of the transfer function , h(k) as

$$H(z) = \sum_{k=-\infty}^{\infty} h(k) z^{-k} \tag{2.41}$$

the power spectral energies are obtained as

$$\begin{aligned} S_{xy}(z) &= H(z)\, S_{xx}(z) \\ S_{yx}(z) &= H^{*}(1/z^{*})\, S_{xx}(z) \\ S_{yy}(z) &= H(z)\, H^{*}(1/z^{*})\, S_{xx}(z) \end{aligned} \tag{2.42}$$

If h(k) is real then

$$H^{*}(1/z^{*}) = H(1/z) \tag{2.43}$$

Using the concept of z-transform, equation (2.39) can be written as

$$Y(z) = H(z)\, X(z) \tag{2.44}$$

where X(z) and Y(z) are the z transforms of the variables x and y.

The formation of equation (2.44) is described in Appendix-II. The transfer function, H(z) describes a complete characterization of the process. Here x_t is assumed to be a stationary random sequence with autocorrelations, $R_{XX}(\tau)$ and power spectral energy, $S_{XX}(f)$ related as given by

$$R_{xx}(\tau) = E[x_t\, x_{t+\tau}^{*}] = \int_{-\infty}^{\infty} S(\omega) e^{-j\omega\tau T} d\omega \tag{2.45}$$

and

$$S(\omega) = \frac{1}{2\omega_b} \sum_{\tau=-\infty}^{\infty} R_{xx}(\tau)\, e^{j\omega\tau T} \tag{2.46}$$

ω_b is the cut-off frequency which satisfies Nyquist relation as

$$T = \frac{\pi}{\omega_b} \tag{2.47}$$

The output system y_t is also a stationary random sequence whose power spectral energy is

$$\begin{aligned} S_{yy}(\omega) &= H^*(e^{j\omega\tau T})\, S_{xx}(\omega)\, H^*(e^{j\omega\tau T}) \\ &= S_{xx}(\omega) \left| H(e^{j\omega\tau T}) \right|^2 \end{aligned} \tag{2.48}$$

Equations (2.46) and (2.48) are periodic functions of period, $2\pi/\omega$.

As a simple case, assume that the input process x_t is a band-limited discrete white noise. The auto-correlation and power spectral energy of the input white noise process can be described as

$$R_{xx}(\tau) = 2\omega_b\, I\, \delta(\tau) \tag{2.49}$$

and

$$S_{xx}(\omega) = I \tag{2.50}$$

where I and $\delta(\tau)$ are identity matrix and discrete impulse function respectively. $\delta(\tau)$ is also known as the *Kronecker delta*. Therefore an approximate transfer function, $H(e^{j\tau\omega T})$ can be determined by several procedures based on the input and output of any dynamic system of random processes. For example, an important causal discrete linear time invariant process can be described by a constant coefficient linear difference equation in which input (white noise), w_t and output, y_t are related as

$$y_r = -\sum_{k=1}^{p} a_k y_{r-k} + \sum_{\iota=0}^{q} b_\iota w_{r-\iota} \tag{2.51}$$

Here the estimators $a_1, a_2, \ldots a_p$ and $b_0, b_1, \ldots b_q$ characterise the linear process.

The z-transform of the above equation can be written as

$$Y(z)\left[1+\sum_{k=1}^{p} a_k z^{-k}\right] = W(z)\left[\sum_{\iota=0}^{q} b_\iota z^{-\iota}\right] \tag{2.52}$$

Therefore the transfer function H(z) can expressed as

$$H(z) = \frac{Y(z)}{W(z)} = \frac{\sum_{\iota=0}^{q} b_\iota z^{-\iota}}{1+\sum_{k=1}^{p} a_k z^{-k}} \tag{2.53}$$

Here both the polynomials can be factored into its roots (A_k and B_ι) and can be expressed by

$$H(z) = \frac{b_o \prod_{i=1}^{q}(1 - B_i z^{-1})}{\prod_{k=1}^{p}(1 - A_k z^{-1})} \tag{2.54}$$

where b_0 is a scaling factor which can be determined from the random process. The upper roots $B_1, B_2, \ldots B_q$ are called the *zeros* of H(z) and the lower roots $A_1, A_2, \ldots A_p$ are called the *poles* of H(z). For a stable minimum phase linear process, all the poles and zeros of the transfer function H(z) lies inside the unit z-plane circle. i.e.,

$$|A_k| < 1 \quad \text{and} \quad |B_i| < 1 \tag{2.55}$$

The definition of the stability for linear systems is given in section 2.5. If any of the poles lies outside the unit circle in the z-plane, the system can be said to be anticausally stable.

2.4 Discrete-time Fourier series transform

The existence of the Fourier transform and its inverse for a given function x(t) can be determined by one sufficient condition

$$\int_{-\infty}^{\infty} |x(t)| dt < \infty$$

One can refer to Kay (1988) and Marple(1987) for details of less restrictive sufficient conditions for the existence of the Fourier transform.

Based on the usual definition of the Discrete Fourier Transform (DFT), X(f) of the discrete data samples, x_t (total M points) can be expressed as

$$X(f) = \sum_{t=0}^{M-1} x_t \, e^{-j2\pi ft/M} \tag{2.56}$$

Inverse of DFT is

$$x_t = \frac{1}{M} \sum_{f=0}^{M-1} X(f) \, e^{j2\pi ft/M} \tag{2.57}$$

Here the sampling period (T) is not taken care for proper units of spectral estimates. This can be modified for the discrete-time Fourier series as given below.

$$X(f) = T \sum_{t=0}^{M-1} x_t \, e^{-j2\pi ft/M} \quad \text{for} \quad 0 \le t \le (M-1) \tag{2.58}$$

$$x_t = \frac{1}{MT}\sum_{f=0}^{M-1} X(f)\, e^{j2\pi ft/M} \qquad \text{for} \qquad 0 \le f \le (M-1) \tag{2.59}$$

Therefore, power spectral energy of the discrete-time Fourier series is written as

$$S(f) = |X(f)|^2 = T^2 \left| \sum_{t=0}^{M-1} x_t\, e^{-j2\pi ft/M} \right|^2 \tag{2.60}$$

2.5 Stability of linear systems

A linear system is said to be *stable* if the output is bounded for all bounded inputs. The responses of the system will be bounded if and only if the roots of the AR polynomial, the denominator of the transfer function are less than one. In other words, a linear time-invariant discrete system is said to be stable in the bounded input and bounded output (BIBO) sense if all the roots of the AR polynomial have magnitudes less than unity. Otherwise the system is unstable.

If we define the transfer function in the form of poles of the random process, then the system is said to be stable and causal provided that the poles lie within the unit circle of the z-plane. This is also known as the minimum-phase filter. If the system is anticausal and stable, the poles have to lie outside the unit circle.

To explain these two cases, consider a general AR process. A causal transfer function

$$H_{AR}(z) = 1/A(z) \quad = \quad \left[1 + \sum_{k=1}^{p} a_k z^{-k}\right]^{-1} \tag{2.61}$$

will lead to the parametric model

$$x_t \quad = \quad -\sum_{k=1}^{p} a_k x_{t-k} \quad + \quad w_t \tag{2.62}$$

Similarly an anticausal transfer function

$$H^*_{AR}(z^*) = 1/A^*(1/z^*) \quad = \quad \left[1 + \sum_{k=1}^{p} a^*_k z^{-k}\right]^{-1} \tag{2.63}$$

yields the parametric model

$$x_t \quad = \quad -\sum_{k=1}^{p} a^*_k x_{t+k} \quad + \quad w_t \tag{2.64}$$

Using equation (2.64) x_t can be generated based on the future values x_{t+k}. The causal stable model given by equation (2.62) is said to be the *forward prediction model,* and the anticausal one given by equation (2.64) is called as the *backward prediction model.* The forward and backward prediction models are used later in multivariate AR parameter estimation. For the MA model stability, one can choose the transfer function as the minimum-phase filter which will guarantee a stable and causal inverse filter. For more about the stability of the linear systems one can refer texts by Sinha (1991), and Soderstrom and Stoica (1989).

CHAPTER 3

PARAMETRIC MODELLING OF UNIVARIATE RANDOM PROCESSES

3.1 Identification of parametric models

While identifying a parametric model, one has to look into the random process to see whether it should be identified by off-line or on-line techniques. On-line identification is needed if the purpose is to track parameters slowly varying in time. Whereas off-line identification is used batchwise where all recorded data is processed simultaneously. Even for on-line identification of the process with unknown dynamic properties, one should first use off-line identification in order to validate the model.

For identifying a particular model which is most suitable for the random process under consideration one should follow the scheme as highlighted in Figure 1.2; i.e., an experiment has to be designed (select inputs, outputs, sample interval, total number of sample, etc.); a model set and model structure has to be chosen (choose linear/nonlinear characteristics of model, model order, parametrization, etc.); an identification criterion has to be selected (prediction error methods or correlation methods) and a procedure for validating the chosen model has to be devised.

When selecting an identification method the purpose of the

identification should be clear to users, since it may express both the type of model which is required and what accuracy is sought. A crude model can be adequate for the purpose of rough estimation, while high accuracy is needed for better representation of the processes such as marine environments and structural motions, where theoretical models need to be verified.

The reliability of the optimal estimates of SI parameters depends on how accurately the SI model represents the 'true' system. Theoretical dynamic systems can be best described by the optimal estimates of SI parameters (Ljung, 1987). However, in practical cases with the presence of noise, one has to draw a line between the fitness of acceptability and the 'true' system, i.e., a particular SI model can give a best fit which is very close to the 'true' system, but may not exactly describe the 'true' system. Therefore, an optimal estimate of the SI model can describe an approximation to the physical processes occurring in the real system.

Based on previous works on random processes, the present study considers AR, MA and ARMA models of SI methods which are to be applied to the target (marine environmental and structural monitoring time series data). Since the stochastic processes in an ocean state have either poles or zeros or combination of both, the above three models can fairly represent the above process. In general, the above SI algorithms are described in Spanos (1983), Marple (1987) and Kay (1988). Here the above models are formulated and applied to stochastic data of the ocean state.

3.2 Description of system identification models

It is sometimes too difficult to select an appropriate SI model for spectral estimates of any random process. Based on studies carried out by Spanos (1986), Lin (1987), Spanos and Mignolet (1987), Witz and Mandal (1991), Mandal, Witz and Lyons (1992), and others, SI algorithms are applied to a random process in ocean dynamic states. The study presented herein considers autoregressive, moving average and autoregressive moving average algorithms which are formulated leading to power spectral estimates as described below.

3.2.1 Autoregressive (AR) model

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

$$y_r = -\sum_{k=1}^{p} a_k y_{r-k} + b_0 w_r \tag{3.1a}$$

Here y_r is the r^{th} sample of the discrete stochastic process and w_r is Gaussian white noise. Equation (3.1a) can be written in polynomial form as

$$y_r = \left[\frac{b_0}{a(z^{-1})}\right] . w_r \tag{3.1b}$$

where

$$a(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3} + \ldots\ldots + a_p z^{-p}$$
$$= 1 + \sum_{k=1}^{p} a_k z^{-k}$$

The AR(p) model is also called as an *all-pole model* and is illustrated in Figure 3.1.

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

$$H_{AR}(z) = \frac{b_0}{1 + \sum_{k=1}^{p} a_k z^{-k}} \tag{3.2}$$

and whose input is a white noise process w_r. The scaling constant b_0 can be obtained by using the equation

$$b_0^2 = \frac{1}{2\omega_b}\left[R_0 + \sum_{k=1}^{p} a_k R_k\right] \tag{3.3}$$

The sampling period, T is generally defined by the cut-off frequency, ω_b, through the Nyquist relationship

$$T = \frac{\pi}{\omega_b} \tag{3.4}$$

Therefore, the estimated power spectrum $S_{yy}(\omega)$ of y_r can be written as

$$S_{yy}(\omega) = \left|H_{AR}(e^{i\omega T})\right|^2 = \frac{b_0^2}{\left|1+\sum_{k=1}^{P} a_k e^{-ik\omega T}\right|^2} \tag{3.5}$$

The autocorrelation function R_λ of the target/observed power spectrum is defined as

$$R_\lambda = \int_{\frac{\pi}{T}}^{\frac{\pi}{T}} S(\omega).\cos(\lambda\omega T).d\omega$$
$$\lambda = 0,\quad 1,\quad 2,\quad 3..... \tag{3.6a}$$

where, $S(\omega)$ is the target/observed spectrum. In the absence of a target spectrum, if the available information is a finite set of time series data, i.e. $\{ 0 < y_t < M\}$, then the true correlation coefficients R_λ can be calculated from the relationships

$$R_\lambda = f(M,\lambda) \cdot \sum_{t=\lambda+1}^{M} y_t . y_{t-\lambda} \tag{3.6b}$$

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

The error, ε between the target spectrum, $S(\omega)$, and the estimated power spectrum of the AR output, $S_{yy}(\omega)$, is expressed as

$$\varepsilon = \frac{b_0^2}{2\omega_b} \cdot \int_{-\omega_b}^{\omega_b} \frac{S(\omega)}{S_{yy}(\omega)} d\omega \tag{3.7}$$

Substituting $S_{yy}(\omega)$ [equation (3.5)] into the above equation gives

$$\varepsilon = \frac{1}{2\omega_b}\int_{-\omega_b}^{\omega_b} S(\omega)\left|1+\sum_{k=1}^{p} a_k e^{-ik\omega T}\right|^2 d\omega \qquad (3.8)$$

By minimizing ε, the parameters a_k are determined. Therefore, minimisation of error, ε brings the following important relationship

$$\frac{d\varepsilon}{da_i} = \frac{1}{\omega_b}\left[R_i + \sum_{k=1}^{p} a_k R_{|i-k|}\right] = 0$$

or

$$\sum^{p} a_k R_{|i-k|} = -R_i \qquad (3.9)$$

Equation (3.9) can be written in *Toeplitz matrix* form as given below

$$\begin{bmatrix} R_0 & R_1 & R_2 & \cdots & R_{p-1} \\ R_1 & R_0 & R_1 & \cdots & R_{p-2} \\ R_2 & R_1 & R_0 & \cdots & R_{p-3} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ R_{p-1} & R_{p-2} & R_{p-3} & \cdots & R_0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \cdots \\ a_p \end{bmatrix} = - \begin{bmatrix} R_1 \\ R_2 \\ R_3 \\ \cdots \\ R_p \end{bmatrix} \qquad (3.10)$$

These are also known as *Yule-Walker* equations. Using equations (3.3) and (3.10), i.e.; when the parameters a_1, a_2, a_3, ... a_p and b_0 are determined, then the time series y_r can be generated digitally by using the recursive equation (3.1). Here one has to obtain optimal spectral estimates for the values of a_k and p such that the AIC or MDL between the target and estimated spectra must be a minimum.

3.2.2 Moving average (MA) model

The time series $\{y_r\}$ is said to be an moving average process of the order q, MA(q) if it is generated from the numerical scheme

$$y_r = \sum_{\iota=-q}^{q} b_\iota w_{r-\iota} \tag{3.11}$$

Here y_r is the r^{th} sample of the discrete stochastic process and w_r is the Gaussian white noise. The MA(q) model is also called as an *all-zero model* and is shown in Figure 3.2.

In general, any weakly stationary process can be expressed as the output of an infinite (in practical applications a finite value of q is chosen) order digital filter whose transfer function is defined as

$$H_{MA}(z) = \sum_{\iota=-q}^{q} b_\iota z^{-\iota} \tag{3.12}$$

and whose input is a white noise process. As one is looking for the best fit of the target spectrum, $S(\omega)$, the MA coefficients b_ι are calculated by making use of the Fourier coefficients of the square roots of the target spectrum as follows

$$b_\iota = \frac{T}{\pi}\int_0^{\frac{\pi}{T}} Q(\omega)\cos(\iota T\omega)\, d\omega$$

$$\text{where} \quad S(\omega) = Q(\omega).Q^*(\omega) \tag{3.13}$$

If $S(\omega)$ is real then $Q(\omega) = \sqrt{S(\omega)}$. Therefore, the estimated power spectrum $S_{yy}(\omega)$ of the MA process can be written as

$$S_{yy}(\omega) = H_{MA}(e^{i\omega T}).H^{*}_{MA}(e^{i\omega T}) \tag{3.14}$$

If the available information is in the form of time series, then the MA coefficients, b_ι can be estimated from either of the following relations:

(a) Equating power spectra estimated by the FFT [equation (2.29)] and equation (3.14) gives

$$R_{xx}(k) = \begin{cases} \sum_{\iota=k}^{\iota=q} b^{*}_{\iota} b_{\iota-k} & \text{for } k = 0,1,2,\ldots\ldots,q \\ R^{*}_{xx}(-k) & \text{for } k = -q,-(q-1),\ldots\ldots,-1 \end{cases} \tag{3.15}$$

The above equations are valid for the general case. The present study estimates the SI coefficients in real form. Therefore, equations (3.15) can be solved to obtain the MA coefficients. However for higher order MA models, estimation of the MA coefficients using equations (3.15) becomes more complicated. In this case an alternative method can be used as described in (b).

(b) By equating the AR(p) process with the MA(q) process, [equation (3.1a) with equation (3.11)] a_k is expressed as the impulse response of $1/b(z^{-1})$. This can can be written as

$$H_{MA}(z) = b(z^{-1}) = \frac{1}{a(z^{-1})} \tag{3.16}$$

Therefore,

$$a(z^{-1}) = \frac{1}{b(z^{-1})} \tag{3.17}$$

Hence $H_{MA}(z)$ can be obtained from the AR polynomial, $a(z^{-1})$ [see example 2 of Appendix II]. Now if the impulse response of $1/b(z^{-1})$ decays to zero for a lag greater than p, then the AR(p) process is a good approximation to the MA(q) process.

Equation (3.11) can be used for simulating a time series by the weighted average of 2q+1 white noise deviates which moves in time. The advantage of this model is that one does not require feedback values of time series. However in the case of equation (3.1) the AR process requires not only the feedback mechanism of the time series but also the initial y_{r-1} values.

3.2.3 Autoregressive moving average (ARMA) model

Let the time series $\{y_r\}$ be an Autoregressive Moving Average (ARMA) process of the orders (p, q) with $p \geq q$. It is defined by the difference equations

$$y_r = -\sum_{k=1}^{p} a_k y_{r-k} + \sum_{\iota=0}^{q} b_\iota w_{r-\iota} \tag{3.18a}$$

Here y_r is the r^{th} sample of the discrete stochastic process and w_r is the Gaussian white noise process. The above equation (3.18a) can be expressed

in polynomial notation as

$$y_r = \left[\frac{b(z^{-1})}{a(z^{-1})}\right] . w_r \qquad (3.18b)$$

where

$$a(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3} + \ldots\ldots + a_p z^{-p}$$
$$= 1 + \sum_{k=1}^{p} a_k z^{-k}$$

and

$$b(z^{-1}) = b_0 + b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3} + \ldots\ldots + b_q z^{-q}$$
$$= \sum_{\iota=0}^{q} b_\iota z^{-\iota}$$

The ARMA(p,q) model is also known as a *pole-zero model* and is illustrated in Figure 3.3.

Now the ARMA(p,q) process can also be expressed as the output of a digital filter whose transfer function in terms of z-transform is defined as

$$H_{ARMA}(z) = \frac{\sum_{\iota=0}^{q} b_\iota z^{-\iota}}{1 + \sum_{k=1}^{p} a_k z^{-k}} \qquad (3.19)$$

and whose input is a discrete white noise process w_r.

The estimated power spectrum $S_{yy}(\omega)$ is written as

$$S_{yy}(\omega) = \left|H_{ARMA}(e^{i\omega T})\right|^2 \qquad (3.20)$$

There are various procedures available to obtain the unknown coefficients a_k and b_ι [see Graupe et al (1975), Gersch and Yonemoto (1977), Friedlander (1983)]. Based on applications [Samii and Vandiver (1984), Samras et al (1985), and others] of random processes in marine problems, two alternative procedures to determine a_k and b_ι are described below.

3.2.3a Power order matching (POM)

This technique is to match the power of AR and ARMA algorithms. Equating the transfer functions of AR(m) [replacing a_k and b_0 by $\hat{a}_k$ and $\hat{b}_0$ in equation (3.1)] and ARMA(p,q) models given by equations (3.2) and (3.19) the following relationship can be expressed

$$\hat{b}_0\left(1+\sum_{k=1}^{p} a_k z^{-k}\right) = \left(\sum_{\iota=0}^{q} b_\iota z^{-\iota}\right).\left(1+\sum_{\iota=1}^{m} \hat{a}_\iota z^{-\iota}\right) \tag{3.21}$$

where, it is assumed that $p+q \le m$ and $p \ge q$.

Now equating the same power of z in the above relation, the following equations are obtained to determine the unknown coefficients a_k and b_ι.

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_q \end{bmatrix} = \frac{1}{\hat{b}_0}\cdot\begin{bmatrix} \hat{a}_1 & 0 & 0 & \dots & 0 \\ \hat{a}_2 & \hat{a}_1 & 0 & \dots & 0 \\ \hat{a}_3 & \hat{a}_2 & \hat{a}_1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \hat{a}_q & \hat{a}_{q-1} & \hat{a}_{q-2} & \dots & \hat{a}_1 \end{bmatrix}\begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_{q-1} \end{bmatrix} + \frac{1}{\hat{b}_0}\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_q \end{bmatrix} \tag{3.22}$$

$$\begin{bmatrix} a_{q+1} \\ a_{q+2} \\ a_{q+3} \\ \vdots \\ a_p \end{bmatrix} = \frac{1}{\hat{b}_0} \cdot \begin{bmatrix} \hat{a}_{q+1} & \hat{a}_q & \hat{a}_{q-1} & \dots & \hat{a}_1 \\ \hat{a}_{q+2} & \hat{a}_{q+1} & \hat{a}_q & \dots & \hat{a}_2 \\ \hat{a}_{q+3} & \hat{a}_{q+2} & \hat{a}_{q+1} & \dots & \hat{a}_3 \\ \vdots & \vdots & \vdots & \vdots & \\ \hat{a}_p & \hat{a}_{p-1} & \hat{a}_{p-2} & \dots & \hat{a}_{p-q} \end{bmatrix} \cdot \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_q \end{bmatrix} \tag{3.23}$$

and

$$\begin{bmatrix} \hat{a}_{p+1} & \hat{a}_p & \hat{a}_{p-1} & \dots & \hat{a}_{p+1-q} \\ \hat{a}_{p+2} & \hat{a}_{p+1} & \hat{a}_p & \dots & \hat{a}_{p+2-q} \\ \hat{a}_{p+3} & \hat{a}_{p+2} & \hat{a}_{p+1} & \dots & \hat{a}_{p+3-q} \\ \vdots & \vdots & \vdots & & \vdots \\ \hat{a}_{p+q} & \hat{a}_{p+q-1} & \hat{a}_{p+q-2} & \dots & \hat{a}_p \end{bmatrix} \cdot \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_q \end{bmatrix} = 0 \tag{3.24a}$$

or,

$$\begin{bmatrix} \hat{a}_p & \hat{a}_{p-1} & \dots & \hat{a}_{p+1-q} \\ \hat{a}_{p+1} & \hat{a}_p & \dots & \hat{a}_{p+2-q} \\ \hat{a}_{p+2} & \hat{a}_{p+1} & \dots & \hat{a}_{p+3-q} \\ \vdots & \vdots & & \vdots \\ \hat{a}_{p+q-1} & \hat{a}_{p+q-2} & \dots & \hat{a}_p \end{bmatrix} \cdot \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_q \end{bmatrix} = -b_0 \cdot \begin{bmatrix} \hat{a}_{p+1} \\ \hat{a}_{p+2} \\ \hat{a}_{p+3} \\ \vdots \\ \hat{a}_{p+q} \end{bmatrix} \tag{3.24b}$$

$$\text{with} \qquad b_0 = \hat{b}_0 \tag{3.25}$$

First, $b_1, b_2, b_3 \ldots b_q$ are determined from equation (3.24b); thereafter a_1, a_2, a_3, ... a_k are calculated from equations (3.22) and (3.23).

3.2.3b Inverse AR filter using MYW equations

Multiplying both sides of the equation (3.18a) by $y_{t-q-\iota}$ and taking expected values, the correlation coefficients of the process y_t is

$$R_\lambda = E\{y_\iota y_{\iota-\lambda}\}$$

and

$$R_\lambda = R_{-\lambda}$$

Then we get

$$R_{q+\iota} + a_1 R_{q+\iota-1} + a_2 R_{q+\iota-2} + \ldots\ldots + a_p R_{q+\iota-p} = 0 \quad \text{for} \quad 1 \le \iota \le N - q \tag{3.26}$$

Equations (3.26) are often called the *Modified Yule-Walker* **(MYW)** equations. The simple ordinary Yule-Walker equations can be obtained with q equal to zero. If the observed or target spectrum, $S(\omega)$, is given, then R_λ can be determined from equation (3.6a) or equation (3.6b). Now the AR part of the ARMA algorithm is obtained by solving the above MYW equations. There are various methods available to estimate the MA part of the ARMA algorithm. An efficient computational method is described below.

Based on the modal decomposition method and combining all partial fractions to a common denominator, the power spectrum of the ARMA model can be described as

$$S_{yy}(z) = \sum_{i=-\infty}^{\infty} R_i z^{-i}$$

$$= \frac{n(z)}{a(z)} + R_0 + \frac{n(z^{-1})}{a(z^{-1})} \tag{3.27}$$

where, the causal part of the autocorrelation sequence is

$$\frac{n(z^{-1})}{a(z^{-1})} = \sum_{i=1}^{\infty} R_i z^{-i} \tag{3.28}$$

and

$$n(z^{-1}) = n_1 z^{-1} + n_2 z^{-2} + \ \ldots \ + n_p z^{-p} \tag{3.29}$$

Therefore $n(z^{-1})/a(z^{-1})$ is a linear system whose impulse response is the one-sided covariance sequence.

The ARMA spectrum is defined as

$$S_{yy}(z) = \frac{b(z).b(z^{-1})}{a(z).a(z^{-1})} \tag{3.30}$$

Comparing equations (3.27) and (3.30), the MA part of the ARMA spectrum can be expressed using numerators as

$$b(z).b(z^{-1}) = n(z).a(z^{-1}) + R_0.a(z).a(z^{-1}) + n(z^{-1}).a(z) \tag{3.31}$$

We define

$$\frac{1}{a(z)} = \sum_{i=0}^{\infty} h_i z^{-i} \tag{3.32}$$

Where h_i can be determined by a recursive filter technique. Here $n(z^{-1})$ can be directly estimated from R_i using the approximate

relationship in the time-domain

$$\begin{bmatrix} h_0 & 0 & 0 & \dots & 0 \\ h_1 & h_0 & 0 & \dots & 0 \\ h_2 & h_1 & h_0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ h_{p-1} & h_{p-2} & h_{p-3} & \dots & h_0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ h_{N-1} & h_{N-2} & h_{N-3} & \dots & h_{N-p} \end{bmatrix} \cdot \begin{bmatrix} n_1 \\ n_2 \\ n_3 \\ \vdots \\ n_p \\ \vdots \\ n_N \end{bmatrix} = \begin{bmatrix} R_1 \\ R_2 \\ R_3 \\ \vdots \\ R_p \\ \vdots \\ R_N \end{bmatrix} \tag{3.33}$$

Equation (3.33) is an overdetermined system. This system will not have a solution unless the R_i are associated with an ARMA process of order (p, p) or lower. Assuming this not to be the case, one can determine n_i in the least squares sense using the singular value decomposition (SVD) technique. Once n_i are determined, the MA part of the ARMA model is calculated using equation (3.31). Hence the final ARMA spectrum can be obtained from coefficients of $a(z^{-1})$, $n(z^{-1})$ and R_0. By factorisation of equation (3.31), it easy to obtain $b(z^{-1})$. Hence the ARMA algorithms are determined. The quality of the ARMA spectral estimate described in the above two cases will depend on how accurately the initial higher order AR parameters are selected.

3.3 Univariate optimal model order selection

The modified Yule-Walker method requires a maximum of p+q (where p and q are the ARMA model orders) correlation coefficients for the solution of the well-determined set of equations of an ARMA process. Friedlander (1983) has demonstrated the improvement in spectral estimation accuracy with a larger number of equations and more correlation coefficients. However, there are high computational costs associated with an overdetermined set of equations and higher order models. The question is how high a model order should one select so that the solution remains computationally feasible 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, i.e.,

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) (Marple, 1987)

The LSE method is commonly used owing to its simplicity, but can not be applied to all random processes. If a 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 SI model order which is the optimal representation of the stochastic process. Other methods based on the AIC or MDL overcome this problem.

The optimal model order of the initial AR process can be selected by the AIC and MDL which are expressed as

$$AIC = M\,[\ln(ER)] + 2p \tag{3.34a}$$

$$MDL = M\,[\ln(ER)] + p\,\ln(M) \tag{3.34b}$$

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 estimated values from 1 to M. Another way of estimating ER is based on the estimation of variance of the initial higher order AR process using the forward linear prediction residual (FLPR) technique. The ARMA model is derived from the initial higher order AR model. The higher order AR model needs a finite order which can be determined by the FLPR technique. The FLPR (based on the AR process) is expressed as

$$e^f(n) = x(n) + \sum_{k=1}^{p} a(k).x(n-k) \tag{3.35}$$

where x(n) represents the discrete data and a(k) are the AR coefficients. This expression is identical to the AR algorithm except for the term on the left hand side of the equation which is not a driving white noise process. The FLPR from a finite discrete data set may, or may not, be a white noise process. While fitting an AR model to a finite data set, here we assume that the FLPR is a white noise process, so that it allows one to equate the AR parameters to the linear prediction coefficients. The sum of the squares of the FLPR is estimated as

$$ER = \frac{1}{M-p}\sum_{n}\left|e^f(n)\right|^2 \tag{3.36}$$

This ER is used in equations (3.34) to find the value of p that minimizes the AIC or MDL.

An approximate expression for the AIC as given in Friedlander and Porat (1984) is

$$AIC = N \log\{(ER_p)/(N-p)\} + 2p. \tag{3.37}$$

where, ER_p is the sum of the squared error from p+1 to N ($N \geq 2p$). Here the MYW method is used for solving an overdetermined set of equations. Let the MYW be written as

$$R_{q+\tau} + a_1 R_{q+\tau-1} + a_2 R_{q+\tau-2} + \ldots\ldots + a_p R_{q+\tau-p} = e_{q+\tau} \quad \text{for} \quad 1 \leq \tau \leq (N-q) \tag{3.38}$$

then ER_p is estimated as the sum of the squares of $e_{q+\tau}$. Here one has to find the value of p that minimizes AIC. If the autocovariance sequences of the stochastic process follow the convergence rule, then for higher values of p, the estimate will improve. However, if the autocovariance sequences first converge and then diverge one should select p for the autocovariance sequences just before divergence.

3.4 Applications of AR, MA and ARMA modelling

Sometimes a particular SI model becomes ill-conditioned owing to reasons such as not taking into account poles and zeros, optimal model orders and the number of Yule-Walker equations. Here the theoretical (Pierson-Moskowitz and JONSWAP) wave spectra are considered for AR, MA and ARMA algorithms of SI methods. Measured time series waves and offshore structural motions are then considered for the above algorithms. Subsequently a reduced model order technique is used for further reduction of data in the form of reduced ARMA model coefficients as described in Chapter 4.

3.4.1 Ocean waves

Ocean waves represent a random process. In many occasions and locations we may not have measurements of the wave process, and so theoretical estimates are used for design purposes. These may be obtained from spectral formulation such as the Pierson-Moskowitz or JONSWAP ocean wave spectra. Here these theoretical wave spectra are used to establish system identification algorithms in the form of rational modellings. Since most of the theoretical spectra represent random processes which can be modelled by SI algorithms, it is simple to generate time series ocean waves using equations (3.1), (3.11) and (3.18).

3.4.1a Theoretical Pierson-Moskowitz power spectrum

The first systematic and reliable way of establishing an ocean wave spectrum was carried out by Pierson and Moskowitz (1964) and is widely accepted for the waves of a fully developed sea. The Pierson-Moskowitz spectrum is defined as

$$S(\omega) = \frac{A}{\omega^5} \exp\left(-\frac{B}{\omega^4}\right) \tag{3.39}$$

where, $A = 0.0082\, g^2$

$B = 0.74\, (g/v)^4$

v = wind speed in m/sec at a height of 19.5m above mean water level

g = gravitational acceleration in m/sec^2

$\omega = 2\pi f$

f = frequency in Hz

Spanos and Hansen (1981) have shown that the Pierson-Moskowitz spectrum could be approximated reliably by the AR process in an average sense, but the AR spectrum exhibits sharp fluctuations. These fluctuations could not be removed either by increasing the filter's orders or by selecting a larger number of equations in the Toeplitz matrix [equations (3.10)]. While examining equation (3.39), it clearly shows that the Pierson-Moskowitz spectrum does possess a zero of infinite order at the frequency, $\omega=0$. Therefore, numerical difficulties exist in approximating the Pierson-Moskowitz spectrum by AR algorithms which are best suited for all pole-dominated spectra. It is shown that an eight term Taylor expansion of the exponential term of equation (3.39) expresses an excellent approximation to the Pierson-Moskowitz spectrum for a wide range of wind velocities

most commonly used by engineers. The Taylor expansion of equation (3.39) is expressed as

$$S(\omega) = \frac{A.\omega^{27}}{\omega^{32} + B\omega^{28} + \frac{B^2}{2!}\omega^{24} + \frac{B^3}{3!}\omega^{20} + \frac{B^4}{4!}\omega^{16} + \frac{B^5}{5!}\omega^{12} + \frac{B^6}{6!}\omega^{8} + \frac{B^7}{7!}\omega^{4} + \frac{B^8}{8!}} \tag{3.40}$$

This equation has been successfully applied (Spanos, 1983) to represent a reliable AR process of the Pierson-Moskowitz spectrum. Here it reduces the order of the zero of the Pierson-Moskowitz spectrum at $\omega=0$ from infinity to only 27. Figures 3.4 shows comparison between a Taylor approximated Pierson-Moskowitz spectrum and a true Pierson-Moskowitz spectrum for various wind speeds. It is observed that there are very small variations at the low frequency region. This is owing to the approximation of the true Pierson-Moskowitz spectrum. Otherwise, there are hardly any differences. The Taylor approximated Pierson-Moskowitz spectrum is an excellent fit with the MA(29) (the model order is chosen based on the autocorrelation of the process, Figure 3.5) spectrum as shown in Figure 3.6 as compared to the true Pierson-Moskowitz spectrum (Figure 3.7) where some differences are observed in the peak frequency region. This is owing to the presence of the zero of infinite order in the true Pierson-Moskowitz spectrum. Figures 3.8 shows the comparison between the Pierson-Moskowitz spectrum and ARMA(20,20) spectrum. Improvement of matching occurs owing to consideration of more time steps of the autocorrelation of the Pierson-Moskowitz spectrum which is a convergent process.

3.4.1b Theoretical JONSWAP power spectrum

The JONSWAP (JOint North Sea WAve Project) spectrum is established based on the North Sea wave measurements and analysis carried out by Hasselmann et al (1973). This spectrum is defined as

$$S(\omega) = \alpha . g^2 . \omega^{-5} . \exp\{-(\omega_p/\omega)^4\} . \gamma^{\Gamma}$$
$$\Gamma = \exp\{-(\omega - \omega_p)^2/(2\sigma^2\omega_p^2)\} \qquad (3.41)$$

The same Pierson-Moskowitz spectrum with $\alpha = 0.0082$, $\omega_p = 0.74\ (g/v)^4$, $\sigma_a = 0.07$, $\sigma_b = 0.09$ and $\gamma = 3.3$ is compared with MA(29), ARMA(25,25) spectra (Figures 3.9 and 3.10) which show a very good fit.

3.4.1c Measured time series of ocean waves

A long term wave measurement programme was carried out by the Indian Navy off the West Coast of India at 16m water depth as shown in Figure 3.11. A set of 2048 measured wave data points (K15) at 2Hz sampling rate are used for wave power spectral estimation using AR, MA and ARMA algorithms. The order of the model is determined based on an autocorrelation process (Figure 3.12) and the LSE (Figure 3.13) method. The ARMA(p,q) algorithms are determined based on the power of order matching method. Equating transfer functions of initial AR(m) and ARMA(p,q) algorithms and considering the like powers of z [see equations (3.21) to (3.25)], the coefficients, b_l and a_k are determined. Based on the LSE and autocorrelation methods, the model orders are

selected. Here the selected model orders are AR(40), MA(30) and ARMA(20,20). A comparison study on both IAR and POM methods (Figure 3.14) show that there are some differences. The advantage of POM method over IAR method is that knowing the initial AR process the ARMA coefficients are obtained taking into account all energies contributed from initial AR estimates, whereas the approximate technique based on the MYW equations is applied to the IAR method. Comparative results of LSE of ARMA spectra between POM and IAR techniques for PM, JONSWAP and K15 waves are tabulated (Table 3.1), which also confirm that POM method estimates less LSE as compared to LSE by the IAR method. Both of the above ARMA spectral estimation techniques can be applied without much loss of information. Figure 3.15 shows good agreement between the SI models' spectra and the FFT spectrum. It is observed from the autocorrelation and LSE plots that this power spectrum is a pure convergent process. Therefore, by increasing the model orders, it yields improved results.

3.4.2 Semisubmersible (Santa Fe Rig135) heave motion

Santa Fe Rig135 is representative of semisubmersibles used for hydrocarbon exploration and production in the North Sea (Figure 3.16). Heave acceleration of semisubmersible (Santa Fe Rig135) collected by the IDAS load and motion monitoring package developed by the Department of Mechanical Engineering, University College London is considered herein. A typical 2048 digital heave acceleration data set (SF28) collected at 2.5 Hz interval (Figure 3.17a) is considered to estimate the power spectrum

using SI algorithms. Figure 3.17b shows the normalised autocorrelation of the initial AR algorithm for the above data. This shows that the acceleration data are not purely convergent. Based on the autocorrelation process (Figure 3.17b), it is observed that the model order can be optimised at 98. The AR(98) model shows good matching with the FFT spectrum (Figure 3.18). Here ARMA(50,50) and ARMA(49,49) spectral estimates show very little difference (Figures 3.19 and 3.20). The above two ARMA spectral estimates can be used for further reduction of ARMA model orders. The normalized variation of MA parameters (Figure 3.21) of the above data shows very low optimal model order as compared to AR and ARMA processes. Figures 3.22 and 3.23 show the comparison between MA spectra with 32 and 34 coefficients and the FFT heave acceleration spectrum which fits very well. This type of data in which mostly zeros are present in the SI algorithm may require a higher order AR or ARMA model. Zeros of any random process can best be fitted by a MA algorithm.

3.4.3 Monohull crane vessel motions

Two vessels, namely, Vessel-A and Vessel-B are considered here. The former is the DB50, the latter is not disclosed owing to reasons of confidentiality, but is also a monohull crane vessel.

3.4.3a Crane Vessel-A

As described above for the Rig135 heave motions, similar observations are noticed for the McDermott's DB50 crane vessel roll

motions. A record of 2048 data points collected at a rate of 4 Hz is considered for this study. Based on the autocorrelation process (Figure 3.25), the initial higher AR model order is estimated which is shown to be more than 200. The AR or ARMA algorithm can not be properly applied for such a high model order owing to ill conditioning of Toeplitz matrix values. However for such a case a very low order AR model can generally be fitted to the measured spectrum. Starting from the first order of the autoregressive process the AR(3) model seems to provide a very close fit to the above process as shown in Figure 3.26. The above data can easily be fitted with the MA algorithm which is shown in Figures 3.27 and 3.28. Here the 2048 time series data points can be represented by only 30 parameters which are the coefficients of the MA(30) model.

3.4.3b Crane Vessel-B

A later version of the offshore monitoring unit, IDAS (see section-3.4.2) was installed on monohull crane vessel-B in the North Sea in April 1992 and recorded the vessel's motions at different headings and sea conditions. Here one set of such a time series of 1024 data points in head seas is analysed. The sampling rate of the time series was 5 Hz. The measured data consist of pitch, roll, strap down vertical (heave) acceleration, longitudinal acceleration, transverse acceleration and sea wave elevation. Power spectra calculated using the FFT technique show that these are narrow banded spectra for which the autocorrelation lags are very high up to which the process is convergent. Since these time series are narrow banded, a low order AR or ARMA model can be used to

represent the measured spectra. The appropriate AR or ARMA model orders are selected based on the best estimate to represent the power spectra.

First the sea wave data is modelled using autocorrelation, LSE and AIC techniques. It shows that even though the actual model order seems to need to be very high the best fit between the measured wave and estimated AR spectra shows that model order of 24 seems to adequately represent the wave process. While examining the LSE there is not much variation between the AR(20) and AR(24) models (Figures 3.29a,b,c) which are considered for estimation of vessel's transfer function.

The heave acceleration, pitch and roll motions are analysed with SI modellings. The appropriate SI models' parameters are estimated from the vessel motions. The SI models spectra are compared with the spectra determined by FFT. Pitch motion can be represented by the AR(16) and ARMA(12,1) models. The spectra estimated by AR(16) and ARMA(12,1) models are compared with the spectra determined by FFT as shown in Figures 3.30a,b. Roll motion can be expressed by AR(29) model and its spectral comparison with FFT spectrum is shown in Figure 3.31. Similarly heave motion can be represented by the AR(13), ARMA(4,3) and ARMA(2,2) models and their spectra are compared with FFT spectrum (Figure 3.32a,b,c).

The transfer functions (Briggs and Vandiver, 1982) estimated by SI modellings for each type of vessel motions in random sea waves are compared with measured transfer function determined by the FFT technique which shows good consistency with the SI derived transfer functions. Figures 3.33a,b show the comparison between the pitch transfer functions estimated by SI modelling and FFT technique. Similarly

Figures 3.34a,b show the comparison between the roll transfer functions estimated by the SI modellings and FFT technique. These show good consistency between the estimated transfer functions by the SI modelling and FFT technique. The estimated transfer functions by SI modellings can also be used to represent heave motion as shown in Figures 3.35a,b. This study of transfer function estimation further confirms that the estimated lower order SI models can be used to represent vessel motions with significant reduction in data storage requirements.

CHAPTER 4

REDUCED ORDER ARMA MODELLING OF UNIVARIATE RANDOM PROCESSES

4.1 Introduction

Selecting a suitable model for any stochastic process is a complicated task. Spanos and Mignolet (1986) and Lin (1987) have investigated the application of SI models to random processes associated with the ocean environment. The most commonly method of model order reduction technique in Control and System Engineering (Glover, 1984) is derived based on state-space modelling. This chapter presents a different method for the reduction of the model order based on the initial optimal AR estimates of the ocean dynamic systems and then the formation of a reduced order ARMA algorithm.

4.2 Initial autoregressive parameters estimation using MYW equations

Let the time series $\{y_t\}$ be an ARMA process of the orders (p, q) with $p \geq q$ defined by the difference equations

$$y_t = -\sum_{k=1}^{p} a_k y_{t-k} + \sum_{\iota=0}^{q} b_\iota w_{t-\iota} \qquad (4.1)$$

where a_k and b_ι are the coefficients of the ARMA process. Here y_t is the t^{th} sample of the discrete stochastic process and w_t is the Gaussian white noise. Multiplying both sides of the equations (4.1) by $y_{t-q-\iota}$ and taking expected values, the MYW equations can be obtained as described in equations (3.26) [Chapter 3] and can be rewritten as

$$R_{q+\iota} + a_1 R_{q+\iota-1} + a_2 R_{q+\iota-2} + \ldots\ldots + a_p R_{q+\iota-p} = 0 \qquad \text{for} \quad 1 \le \iota \le N - q \qquad (4.2)$$

with the original Yule-Walker equations obtained by equating q to zero. In the case of spectral estimation, we have to find the true correlation coefficients, R_i and the order (p, q) of the ARMA process. The correlation values can be determined from equations (3.6).

For the ARMA process, initial higher order p coefficients (a_1, a_2, a_3 ...a_p) are obtained by solving equations (4.2).

4.3 Higher energy modes selection to form reduced order AR parameters

After determining the initial AR coefficients a_1, a_2, a_3, ... a_p from equations (4.2), it is required to determine the roots of the polynomial,

$a(z^{-1})$, which contribute most of the energy to the stochastic process. These roots of $a(z^{-1})$ containing higher energies should match closely to represent the original model or time series data without significant loss of information. In other words, the true SI model fits the required random process along with some noise components present in the system. Here we try to eliminate the noise modes thus reducing the order of the SI model.

Some of the roots of the polynomial, $a(z^{-1})$, may be real but most of the roots will be complex. Since the complex roots are conjugate, those pairs can be multiplied to form second order real modes. Therefore, the polynomial, $a(z^{-1})$ can be factored into first and second order real modes,

$$a(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2} + \ \ldots\ldots\ + a_p z^{-p}$$

or

$$a(z^{-1}) = \prod_{i=1}^{p_1}[1 + d_i z^{-1}].\prod_{i=1}^{p_2}[1 + e_{1i} z^{-1} + e_{2i} z^{-2}] \qquad (4.3)$$

where $p_1 + 2p_2 = p$ and d_i, e_{1i}, e_{2i} are coefficients.

And p_1 is the number of first order modes and p_2 is the number of second order modes.

The truncated single sided power spectrum can be expressed in a partial fraction form. The total power spectrum can split into

$$S_T(z) = S_1(z^{-1}) + R(0) + S_2(z) \qquad (4.4)$$

where, R(0) is the autocovariance with zero lag and S_1, S_2, are the single sided causal and anti-causal power spectra respectively,

$$S_1(z^{-1}) = \sum_{i=1}^{N} R_i z^{-i} = \left[\sum_{j=1}^{P_1} \frac{D_j z^{-1}}{1 + d_j z^{-1}} + \sum_{j=1}^{P_2} \frac{E_{1j} z^{-1} + E_{2j} z^{-2}}{1 + e_{1j} z^{-1} + e_{2j} z^{-2}} \right] * \pi(z) \tag{4.5}$$

Here $\pi(z)$ is the transform of a rectangular window (1, N) and * denotes the convolution operator. D_j, E_{1j} and E_{2j} can be estimated in the least squares sense.

We define

$$\frac{1}{1 + d_j z^{-1}} = \sum_{i=0}^{\infty} f_{ij} z^{-i} \tag{4.6}$$

$$\frac{1}{1 + e_{1j} z^{-1} + e_{2j} z^{-2}} = \sum_{i=0}^{\infty} g_{ij} z^{-i} \tag{4.7}$$

Here f_{ij} and g_{ij} are computed using recursion filter techniques as described in Appendix-II (examples 3 and 4). Equation (4.5) can be rewritten in the time-domain as

$$[F \; G] \cdot \begin{bmatrix} D \\ E \end{bmatrix} \approx \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_N \end{bmatrix} \tag{4.8}$$

where

$$F = \begin{bmatrix} f_{01} & f_{02} & \cdots & f_{0\,p_1} \\ f_{11} & f_{12} & \cdots & f_{1\,p_1} \\ \vdots & \vdots & & \vdots \\ f_{N-1\,1} & f_{N-1\,1} & \cdots & f_{N-1\,p_1} \end{bmatrix}$$

and

$$G = \begin{bmatrix} g_{01} & 0 & \cdots & g_{0\,p_2} & 0 \\ g_{11} & g_{01} & \cdots & g_{1\,p_2} & g_{0\,p_2} \\ g_{21} & g_{11} & \cdots & g_{2\,p_2} & g_{1\,p_2} \\ \vdots & \vdots & & \vdots & \vdots \\ g_{N-1\,1} & f_{N-21} & \cdots & g_{N-1\,p_2} & g_{N-2\,p_2} \end{bmatrix}$$

Here $R_1, R_2, R_3, R_4 \ldots R_N$ are the estimates of auto-covariances, determined from the observed/generated time-series data, or from the target spectrum. The f_{ij} and g_{ij} are determined from equations (4.6) and (4.7). Therefore, D (i.e., D_j) and E (i.e., E_{1j} and E_{2j}) can be determined from equation (4.8) which can be rewritten as

$$\begin{bmatrix} D \\ E \end{bmatrix} = \left([F \;\; G]^T [F \;\; G] \right)^{-1} [F \;\; G]^T \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_N \end{bmatrix} \qquad (4.9)$$

In order to select the higher energy modes we have to determine the energy associated with each mode. Appendix-IV describes the calculation procedure for determining the energies of the first and second order modes of a dynamic system (Porat, 1990). The energy of the first and second order modes can be obtained by replacing

$$\left.\begin{aligned} a_1 &= d_j \\ b_0 &= 0 \\ b_1 &= D_j \end{aligned}\right\} \quad \& \quad \left\{\begin{aligned} a_1 &= e_{1j} \\ a_2 &= e_{2j} \\ b_0 &= 0 \\ b_1 &= E_{1j} \\ b_2 &= E_{2j} \end{aligned}\right. \qquad (4.10)$$

in equations (AV-7 and (AV-8) of Appendix-V respectively.

The energy of a first order mode, Ω_{1j}, is given by

$$\Omega_{1j} = \frac{D_j^2}{1-d_j^2} \tag{4.11}$$

and the energy of a second order mode, Ω_{2j}, is given by

$$\Omega_{2j} = \frac{(1+e_{2j})(E_{1j}^2+E_{2j}^2)-2e_{1j}E_{1j}E_{2j}}{(1-e_{2j})(1+e_{2j}+e_{1j})(1+e_{2j}-e_{1j})} \tag{4.12}$$

The number of modes to be chosen depends on how much prior knowledge we have of the dynamic system. It is required that only higher energy modes are retained which in combination represent more or less the original system. If the reduced order of the model is unknown, it is essential to fix some limit and retain those modes whose energies are above this limit. Finally, the selected modes are multiplied to yield the final reduced order (p_3) AR part of the model.

4.4. Estimation of reduced order MA parameters

Once the AR part of the reduced ARMA model is obtained, the MA part can be calculated based on the modal decomposition and partial fraction methods. The power spectrum of the reduced order ARMA model, S_T, can be described from equation (4.4) as

$$\begin{aligned} S_T(z) &= \sum_{i=-\infty}^{\infty} R_i z^{-i} = \frac{n(z)}{a(z)} + R_0 + \frac{n(z^{-1})}{a(z^{-1})} \\ &= \frac{n(z).a(z^{-1}) + R_0.a(z).a(z^{-1}) + n(z^{-1}).a(z)}{a(z).a(z^{-1})} \end{aligned} \tag{4.13}$$

where the causal part of the autocorrelation sequence is

$$\frac{n(z^{-1})}{a(z^{-1})} = \sum_{i=1}^{\infty} R_i z^{-i} \tag{4.14}$$

$$n(z^{-1}) = n_1 z^{-1} + n_2 z^{-2} + \ \ldots \ + n_{p_3} z^{-p_3} \tag{4.15}$$

Defining the reduced order ARMA spectrum as $\{b(z).b(z)\}/\{a(z).a(z)\}$ and comparing with equation (4.13) the MA part of the ARMA spectrum can be expressed as

$$b(z).b(z^{-1}) = n(z).a(z^{-1}) + R_0.a(z).a(z^{-1}) + n(z^{-1}).a(z) \tag{4.16}$$

We can define

$$\frac{1}{a(z)} = \sum_{i=0}^{\infty} h_i z^{-i} \tag{4.17}$$

where h_i can be determined by recursive filter techniques (Appendix-II, example 2). Here $n(z^{-1})$ can be directly estimated from R_i using the approximate relationship in the time domain [from equations (4.14) and (4.17)]

$$\sum_{k=1}^{p} h_{i-k} n_k = R_i \qquad i = 1, 2, \ldots\ldots N$$
$$\text{and} \quad h_{i-k} = 0 \quad \text{for } i < k \tag{4.18}$$

Equations (4.18) form an overdetermined system. This system will not have a solution unless R_i are associated with an ARMA process of order (p, q) or lower. Assuming that this not is the case, n_i are determined in the least square sense. Rewriting equation (4.18)

$$H_{ij} \,.\, n_i = R_i$$

or,

$$n_i = \left[H_{ij}^{*} \,.\, H_{ij}\right]^{-1} .H_{ij}^{*} .R_i \tag{4.19}$$

In order to achieve an efficient computational algorithm, one can take N=p, then equation (4.18) will provide a well determined solution. In other words, in simplifying overdetermined system the following equations can be applied for determining n_i.

$$\begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_p \end{bmatrix} = \begin{bmatrix} R_1 & 0 & \ldots & 0 \\ R_2 & R_1 & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ R_p & R_{p-1} & \ldots & R_1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_{p-1} \end{bmatrix} \tag{4.20}$$

Once n_i are determined the MA part of the reduced order ARMA model is calculated using equation (4.16).

4.5 Final reduced order ARMA model

The final reduced order ARMA spectrum [equation (4.13)] can be obtained from the reduced order coefficients of $a(z^{-1})$ [from section-4.3], $n(z^{-1})$ [from section-4.4, equations (4.17) and (4.18)] and R_0 [section-4.2]. By factorisation of equation (4.16), it is straight forward to obtain $b(z^{-1})$. Thus the reduced order ARMA algorithm is determined and is expressed by the difference equations

$$y_t = -\sum_{k=1}^{p_3} a_k y_{t-k} + \sum_{t=0}^{q_3} b_t w_{t-t} \qquad (4.21)$$

where the time series $\{y_t\}$ is the t^{th} sample of a reduced order ARMA process of the orders (p_3, q_3) with $p_3 \geq q_3$, and a_k and b_t are the coefficients of the reduced order ARMA process. Here w_t is the Gaussian white noise. Now the ARMA(p_3, q_3) process can also be expressed as the output of a digital filter whose transfer function, $H_{ARMA}(z)$, is defined in terms of the z-transform as

$$H_{ARMA}(z^{-1}) = \frac{b(z^{-1})}{a(z^{-1})} \qquad (4.22)$$

where $b(z^{-1}) = b_0 + b_1 z^{-1} + b_2 z^{-2} + \ldots\ldots b_{q_3} z^{-q_3}$

$a(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2} + \ldots\ldots a_{p_3} z^{-p_3}$

and whose input is a discrete white noise process, w_t.

The estimated power spectrum, $S_{yy}(\omega)$, using $z = e^{-i\omega T}$, is written as

$$S_{yy}(\omega) = \left| H_{ARMA}(e^{i\omega T}) \right|^2 \qquad (4.23)$$

where ω is the radian frequency and T is the sampling period.

4.6 Applications of reduced order ARMA modelling

The model order reduction techniques presented in this chapter have been successfully applied to the estimation of reduced order ARMA parameters of ocean waves (theoretical spectra and measured time series) and measured offshore structural motions.

4.6.1 Ocean waves

4.6.1a Theoretical (Pierson-Moskowitz and JONSWAP) power spectra

Consider first the Pierson-Moskowitz spectrum associated with a wind speed of 10 m/s (denoted here as PM10). Figure 3.5 shows the normalised autocorrelation for this wave spectrum. Figure 4.1 presents the AIC against increasing AR model order. Consideration of Figures 3.5 and 4.1 indicates that an AR model of order 40 (AR(40)) is the optimum representation of this wave spectrum. The coefficients of the AR part of the ARMA model are calculated from the AR(40) model using the MYW method previously described. The resulting polynomial, $a(z^{-1})$ is then factorized into first and second order real modes. These modes are shown in Table 4.1. Here modes 4 to 7 are first order real modes and the remaining modes have complex conjugates. Hence pairs of complex conjugate modes are multiplied to form second order real modes. The energy in each mode is calculated using equations (4.11) and (4.12). Table 4.1 also shows the energy associated with each mode for the coefficients of the AR part of the ARMA model representation of PM10.

The number of modes to be retained for the reduced order ARMA model depends on the required level of statistical accuracy. Single peak spectra generally associated with ocean waves can be represented by a few modes which contribute the higher energies. Based on prior knowledge, some energy limit can be fixed and only those modes whose energies are above that limit need be considered. If the reduced model order is unknown, one has to predetermine the energy limit. Table 1 shows that mode 11 has the highest energy level. The energy levels presented in Table 1 are normalized by the energy associated with mode 11. Modes 9 and 10 also have significant energy levels. If only modes 10 and 11 are considered then PM10's AR(40) model reduces to an ARMA(4,4) model. The MA part of the ARMA model is determined based on partial fraction and recursive methods previously described. Figure 4.2a presents a comparison between the reduced order ARMA(4,4) spectrum and the target Pierson-Moskowitz spectrum which shows reasonably good agreement except in the low frequency region. As we are searching for an improved fit, the next higher energy mode is added to the AR part of the reduced ARMA model. Now modes 11, 10 and 9 are considered and the resultant ARMA(6,6) spectrum is determined. The ARMA(6,6) spectrum shows some improvement compared with the ARMA(4,4) spectrum as shown in Figure 4.2a. From the above study, it is observed that consideration of all modes yields the best results but at the expense of added complexity. For instance the ARMA(20,20) spectrum presented in Figure 4.2b shows a very good fit with the target Pierson-Moskowitz spectrum as it takes account of all the modes. This indicates that the initial AR(40) model was properly selected. Now suppose we take an AR(32) model and look for model order reduction. The resultant

ARMA(16,16) spectrum which takes account of all modes present does show some differences with the target Pierson-Moskowitz spectrum. This is owing to the fact that the AR(32) autocorrelation function can not take into account all component energies present in the Pierson-Moskowitz spectrum.

Figure 4.3a and 4.3b present the results of the above model order reduction technique applied to the JONSWAP spectrum. Figure 4.3a shows reasonable agreement between a low order ARMA(2,2) spectrum reduced from an initial optimal AR(60) model and the target JONSWAP spectrum. Increasing the number of modes so that the reduced ARMA(2,2) model increases to an ARMA(6,6) model results in an improved comparison between the reduced model spectrum and the target JONSWAP spectrum (see Figure 4.3b).

4.6.1b Measured time series of ocean waves

The wave measurements were recorded at a location in the southern North Sea where the water depth was 17.3m as shown in Figure 4.4. The wave gauge sensors were placed in the MPN platform as shown in Figures 4.5a,b and time series waves were recorded.

For the present study a time series wave data (M12) is used here for the estimation of the reduced order ARMA parameters. This study considers 4096 data points collected at a sampling frequency of 4Hz. Figure 4.6 shows the normalized autocorrelation of the M12 time series wave data. Figures 4.7a and 4.7b present the variation of AIC and MDL with increasing model order. In Figure 4.7a the AIC and MDL are determined

using the error variance between the target spectrum and estimated spectrum. However sometimes the AIC method may not determine the optimal model order. Using the forward linear prediction residual (FLPR) technique the AIC variation does not show the optimal order (Figure 4.7b). It may show a very high value which does not seem to be the appropriate model order. The MDL method determines the optimal model order using the FLPR technique. The advantage of the MDL method over the AIC method is that the values of p ln(M) increase with M faster than 2p. Hence the MDL method is said to be statistically consistent (Rissanen, 1983). This example shows that it is not necessary that the AIC method can always determine the optimal model order. As the ARMA parameters are derived from the initial higher order AR model which represent the time series wave data, the estimates of the ARMA model parameters require a maximum of 2p parameters of the initial AR model. From Figures 4.6 and 4.7, it is shown that an initial higher order AR model of order 44 (AR(44)) is the optimum representation of the above time series data. The coefficients of the AR part of the ARMA(22,22) are calculated from the AR(44) model using the MYW method. The resulting AR polynomial, $a(z^{-1})$, has the order as 22. While factorizing $a(z^{-1})$ to form the real modes, it is observed that all modes are complex conjugates. Therefore pairs of complex conjugate modes are multiplied to form second order real modes which are shown in Table 4.2. The energy in each mode is calculated as described in Appendix-V. Table 4.2 also shows the energy associated with each second order real mode for the coefficients of the AR part of the ARMA model which represents the time series wave data.

In the following the model reduction technique is applied to reduce the initial ARMA coefficients without significant loss of information. As described in sections 4.2 to 4.5, the number of modes to be retained for the final reduced order ARMA model depends on the required level of statistical accuracy. Generally a good time series wave data set can be represented by a few modes which contribute the higher energies. Based on prior knowledge, some energy limit can be fixed and only those modes whose energies are above that limit need be considered. If the reduced model order is unknown, one has to predetermine the energy limit. Table 4.2 shows that mode 9 has the highest energy level. The energy levels presented in Table 4.2 are normalized by the energy associated with mode 9. Modes 10 and 11 also have some significant energy levels. If only mode 9 is considered then M12's AR(44) model reduces to an ARMA(2,2) model. The MA part of the ARMA model is determined based on partial fraction and recursive methods previously described. Figure 4.8a presents a comparison between the reduced order ARMA(2,2) spectrum and the target M12 ocean wave spectrum which shows reasonably good agreement. As we are searching for an improved fit, the next higher energy modes are added to the AR part of the reduced ARMA model. Now modes 9, 10 and 11 are considered and the resultant ARMA(6,6) spectrum is determined. Figure 4.8b presents the ARMA(6,6) spectrum which shows some improvement compared with the ARMA(2,2) spectrum.

Another time series wave data set (K15) collected off the West coast of India (Figure 3.11) is also investigated here. These time series wave data, each of 2048 points, were collected at 2 samples per second. Figures 4.9a and 4.9b show the results of the reduced order ARMA spectra of K15.

Here it is seen that the ARMA(2,2) spectrum of K15 wave data provides reasonable agreement with the measured wave spectrum. Figure 4.9b shows that the ARMA(4,4) spectrum fits better than that of the ARMA(2,2) spectrum.

4.6.2 Semisubmersible (Santa Fe Rig135) heave motion

The drilling semi-submersible Santa Fe Rig135 heave motions digitized data are used here. A set of 2048 data points (sampling rate 2.5Hz) is considered. Based on the autocorrelation of this data set (Figure 3.17), it is observed that the model order will be optimally selected around 98 or 100. The ARMA(49,49) and ARMA(50,50) power spectra were compared with the measured heave acceleration spectrum (Figures 3.19 and 3.20). and then model order reduction techniques are applied. The reduced ARMA(2,2) spectra determined from ARMA(49,49) and ARMA(50,50) models show fairly good agreement with measured acceleration spectrum which are shown in Figures 4.10 and 4.11. Evidently 2048 heave data samples can be expressed in only a few parameters which are the coefficients of ARMA(2,2) model.

It is observed from the above results that the relatively low order ARMA spectrum provides reasonable agreement with the target spectra or measured time series data. This has a significant practical advantage in that a relatively long period record/target spectrum can be characterised by a low order ARMA model which is described by only a few parameters.

CHAPTER 5

ARIMA MODELLING OF UNIVARIATE NONSTATIONARY PROCESSES.

5.1 Introduction

Dynamic offshore systems are, strictly speaking, nonstationary. For application purposes, when a random process follows a normal (Gaussian) distribution, a dynamic offshore system can be viewed as a wide-sense stationary process. Sometimes long period waves (swells) approach the coast along with wind generated sea waves of relatively short period. These combined time series wave data are strictly nonstationary. Similarly measurements of offshore structural motions of long periods with short period waves yield nonstationary dynamic responses, and vice versa. These type of time series data can not be modelled using AR, MA and ARMA algorithms as described in previous chapters. Such types of nonstationary time series data can be more precisely modelled using modified SI algorithms (Box and Jenkins, 1976, Jenkins, 1979, Pankratz, 1983 and Priestley, 1988), namely, *autoregressive integrated moving average* (ARIMA).

The univariate autoregressive moving average algorithms have some advantages over other traditional methods such as:

a) The autoregressive moving average model is a family of models, not just a single model, that is, it can express any of the AR, MA and

ARMA algorithms.

b) The concepts of these models are derived from classical probability theory and mathematical statistics.

c) These models can determine optimal univariate prediction with smaller least squares error.

5.2 Definition of ARIMA model

A nonstationary time series data is subjected to some form of data transformation and differencing d times to yield a stationary time series data, z_t. If there is no differencing applied (d = 0) an ARMA model of z_t enables a wide class of stationary processes. If the differencing is greater than zero (d = 1, 2, 3, ...), the appropriately transformed and differenced time series can be modelled by the *autoregressive integrated moving average* (ARIMA) algorithms and it is denoted as ARIMA(p,d,q), where p is the number of autoregressive parameters, d is the degree of differencing, q is the number of moving average parameters.

The differencing d is a simple operation which involves determining successive changes in the values of a data series. Let the original nonstationary time series data be y_t. Then using the first differencing of y_t, define a new variable z_t as

$$z_t = y_t - y_{t-1} \qquad \text{for} \quad t = 2, 3, 4, \ldots\ldots n \tag{5.1}$$

Here the time series, z_t is called the first differences of the time series, y_t. If the first differences do not seem to show a stationary series, z_t can be redefined as the second differences (d = 2).

Differences of the first differences can be expressed as

$$\begin{aligned} z_t &= (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) \\ &= y_t - 2y_{t-1} + y_{t-2} \qquad \text{for} \quad t = 3, 4, \ldots\ldots n \end{aligned} \tag{5.2}$$

Similarly one can continue the differencing until the new time series shows a stationary process.

5.3 Backshift notation

Using the above definition we are forcing nonstationary time series into stationary time series and then estimating ARIMA parameters. However we are interested in predicting the original time series. This is not a problem as we can regain the original time series using a backshift operator, $\wp$.

The backshift operator, $\wp$, is such that the time series data can be expressed as

$$\left.\begin{aligned} \wp\, y_t &= y_{t-1} \\ \wp^2\, y_t &= y_{t-2} \\ \wp^3\, y_t &= y_{t-3} \\ &\vdots \\ \wp^d\, y_t &= y_{t-d} \end{aligned}\right\} \tag{5.3}$$

Here the backshift operator does not have any meaningful value or constant. To make sense of the backshift operator, equation (5.3) states that $\wp$ shifts time subscripts. When $\wp$ is used in an algebraic expression it must be multiplied by some other variable such as y_t. Therefore the backshift operator, $\wp$, is meaningful because it changes the time subscript on the variable by which it is multiplied as shown in equation (5.3). Multiplying a constant by any number of $\wp$'s does not alter the constant. As an example, let C be a constant

$$\left.\begin{aligned} &\wp\, C = C \\ &\wp^2\, C = C \\ &\wp^3\, C = C \\ &\quad\vdots \\ &\wp^d\, C = C \end{aligned}\right\} \qquad (5.4)$$

Therefore the first differenced time series, z_t of the original time series, y_t can be expressed by using the *differencing operator* $(1-\wp)$ as

$$\begin{aligned} z_t &= y_t - y_{t-1} \\ &= y_t - \wp\, y_t \\ &= (1 - \wp) y_t \end{aligned} \qquad (5.5)$$

As $\wp$ is not a number, $(1-\wp)$ has no numerical value; it is denoted as an operator. In equation (5.5), $(1-\wp)$ is multiplied by an original time series variable to express the first differences of that variable.

The second differenced time series data can be expressed as

$$\begin{aligned} z_t &= y_t - 2y_{t-1} + y_{t-2} \\ &= y_t - 2\wp\, y_t + \wp^2 y_t \\ &= (1 - 2\wp + \wp^2) y_t \\ &= (1 - \wp)^2 y_t \end{aligned} \tag{5.6}$$

Similarly d differenced time series data can be expressed as

$$z_t = (1 - \wp)^d\, y_t \tag{5.7}$$

This shows that the differenced stationary time series, z_t is linked to the original nonstationary time series, y_t by the operator $(1-\wp)^d$.

5.4 Formation of ARIMA algorithm

Based on the definition of the ARIMA processes and backshift notation, steps may be followed for formation of the ARIMA algorithm as given below:

a) Select nonstationary variable y_t and subtract its mean

$$\bar{y}_t = y_t - \bar{y}$$

where $\bar{y}$ is the mean value of the nonstationary time series data, y_t.

b) Multiply $\bar{y}_t$ by the differenced operator $(1-\wp)^d$ to form a stationary time series, z_t.

c) Multiply z_t by the autoregressive operator which can be expressed as

$$(1 + a_1\wp + a_2\wp^2 + a_3\wp^3 + \ldots\ldots + a_p\wp^p)$$

where a_i (i=1, 2, 3, ... p) are the autoregressive coefficients up to order p.

d) Multiply white noise w_t by the moving average operator which can be written as

$$(b_0 + b_1\wp + b_2\wp^2 + b_3\wp^3 + \ldots\ldots + b_q\wp^q)$$

where b_i (i=0, 1, 2, 3, ... q) are the moving average coefficients up to order q.

e) Equating the results from steps (c) and (d) yields the ARIMA algorithm

$$\begin{aligned}&(1 + a_1\wp + a_2\wp^2 + a_3\wp^3 + \ldots\ldots + a_p\wp^p)\{(1-\wp)^d\,(y_t - \bar{y})\}\\ &= (b_0 + b_1\wp + b_2\wp^2 + b_3\wp^3 + \ldots\ldots + b_q\wp^q)w_t\end{aligned} \qquad (5.8)$$

Equation (5.8) can be written as ARIMA(p,d,q) algorithm which is defined in section 5.2. The ARIMA(p,d,q) is a general form of MA, AR and ARMA in integrated form. So that the ARIMA (p,d,q) can be expressed in the form of either the MA as ARIMA(0,d,q), AR as ARIMA(p,d,0) or a combination of both. Once the differencing is carried out to form a stationary time series, the ARIMA models are expressed as similar to the univariate ARMA models as described in Chapter-3, except for the additional procedure to determine the nonstationary process. This additional procedure is described in the following.

5.5 Procedure for integration to regain nonstationary processes

Equation (5.7) shows that the z values are the differences of the y values, and the y values are sums of the z values. Therefore y values can be obtained by *integrating* (summing) the z values. So the ARMA(p,q) model for the stationary time series z_t is an integrated ARIMA(p,d,q) model for nonstationary time series, y_t. Even though we can build up the ARIMA model for the stationary time series, z_t, our main aim is to get back the original nonstationary time series . This can be achieved as follows.

Suppose we have the stationary time series z_t for differencing, d=1, then equation (5.5) can be rewritten as

$$\begin{aligned} y_t &= (1-\wp)^{-1} z_t \\ &= [1 + \wp + \wp^2 + \wp^3 + \ldots\ldots]z_t \\ &= z_t + \wp z_t + \wp^2 z_t + \wp^3 z_t + \ldots\ldots \\ &= z_t + z_{t-1} + z_{t-2} + z_{t-3} \ldots\ldots \\ &= \sum_{i=-\infty}^{t} z_i \end{aligned} \tag{5.9}$$

Suppose we have a differencing of d=2, we have to integrate twice to obtain y_t

$$\begin{aligned} y_t &= (1-\wp)^{-2} z_t \\ &= (1-\wp)^{-1}[(1-\wp)^{-1} z_t] \end{aligned} \tag{5.10}$$

Let the second term (in square brackets) in equation (5.10) be defined as x_t. Equation (5.10) can then be rewritten as

$$y_t = (1 - \wp)^{-1} x_t \tag{5.11}$$

where

$$x_t = (1 - \wp)^{-1} z_t \tag{5.12}$$

Therefore once x_t is obtained from equation (5.12) in the same way as shown in equation (5.9), substituting x_t in equation (5.11) determines the original nonstationary time series, y_t. In this way we can continue to integrate d number of times to obtain the original time series.

5.6 Selection of number of differencing to form stationary processes

The number of differencing, d, can be selected based on the following procedures:

a) Check the time series data visually. This mostly gives a hint to the approximate degree of differencing required. Using nonstationary time series data, initially use first differencing (d=1) and then examine the first differencing time series. If it still shows nonstationarity, use d=2 and check the second differencing time series. This way one can check the trend of time series and its mean value which tends towards zero.

b) Check the estimated autocorrelations of the original time series and the differenced time series. Variation of autocorrelation values with time lag can also sometimes give the proper selection of d values for which

differenced time series form a stationary process. The estimated autocorrelation decays with increasing lag for nonstationary process does not seem to follow like stationary cases (Bendat and Piersol, 1986).

c) Also one can examine the initial estimated AR polynomial about stable stationary process. Whether poles of AR model are less than unity in magnitude as described in section 2.3.

5.7 Applications

5.7.1 Generated nonstationary ocean waves

Wind generated waves are often called sea waves having relatively short periods compared with swell or other long period waves. When waves of short periods and long periods are measured in combination, the behaviour of such time series no longer follow stationarity conditions. If we divide the nonstationary time series into segments we find that their probability distributions are not the same. As an example, a measured time series wave data set of short periods (North Sea waves: M12) is considered here and then combined with $\mathrm{Sin}(\frac{\pi}{180}t)$ to form a nonstationary wave process. This combined process is shown in Figure 5.1. Since a basic requirement of SI modelling is to have stationary processes, visual examination of the time series data (Figure 5.1) is clearly shown to be a nonstationary process. Now it is required to find the kind of transformation that is likely to convert the nonstationary process to a

stationary process. Using the term differencing as described in section-5.6, one can convert the nonstationary process to a stationary process by using d number of differencings. Starting with d=1, the first differenced time series is estimated as shown in Figure 5.2. Examining both Figures 5.1 and 5.2, one can easily notice that the trend of long periods has disappeared in Figure 5.2 and seems to form a stationary process with constant mean.

Another way of checking the differenced time series to be stationary is to examine the estimated autocovariances for both the time series (Figures 5.1 and 5.2) as shown in Figure 5.3. This clearly shows that the differenced time series autocorrelations decay along the fixed (x-axis) line, whereas the autocorrelations of the nonstationary time series do not show clear convergence which is the basis for identification in parametric modelling of random processes.

Figure 5.4 shows the AIC variation for the autoregressive process of the first differenced time series. This shows that there is a gradual decrease in AIC values up to the model order of 32 and then there is very slow variation. It shows that the AIC(44) is minimum at the model order 44. Therefore estimations of AR(32) and AR(44) power spectra are compared with the power spectrum calculated by FFT which are shown in Figures 5.5 and 5.6. This shows that there is little difference except in the peak frequency region. Based on this study the AR(32) model can be accepted as a good model of the first differenced time series data. However based on the minimum criterion, the AR(44) model is chosen for use of the model order reduction technique. Using the model order reduction technique as described in Chapter 4, the AR(44) model is reduced to an ARMA(4,4) model by discarding noise modes. A comparative study between the FFT

power spectrum and the reduced order ARMA power spectrum shows (Figure 5.7) that the ARMA(4,4) model can fairly represent first-differenced time series data (Figure 5.2). In nonstationary terms, this model can be expressed as an ARIMA(4,1,4) algorithm.

The model stability can be verified by examining the eigenvalues of the polynomials, to determine if the eigenvalues are less than unity in magnitude. The best way to represent it is to plot poles as shown in Figures 5.8a,b. Here all poles are less than unity in magnitude. Hence these SI models are stable.

To check whether one can regain the original nonstationary time series data (Figure 5.1) from first differenced time series data (Figure 5.2) one has to integrate. Using the procedure given in section-5.5 and starting nonstationary value we get indistinguishable nonstationary time series data (Figure 5.1). Since the integrated time series data is indistinguishable from the original nonstationary time series data this plot is not shown here.

Examining the poles and zeros (Figures 5.9a, b) of the ARIMA(4,1,4) and ARIMA(2,1,2) models one notices that one complex pair of poles and zeros of ARIMA(4,1,4) are very close to the same value. If we cancel these poles and zeros then the ARIMA(2,1,2) model is formed as shown in Figure 5.10. However the ARIMA(2,1,2) spectrum shows a high PSE in the peak frequency region compared with the ARIMA(4,1,4) spectrum. Since the cancelled poles and zeros contribute some energy a better spectral estimate of ARIMA(4,1,4) algorithm is observed as shown in Figure 5.7.

5.7.2 Nonstationary displacements of jacket platform deck (Magnus)

The response of giant Magnus platform (Figure 5.11) deck (located in the northern North Sea) was measured in the form of accelerations. These time series data have been double integrated to yield displacements. The platform deck acceleration appears to be a stationary process, whereas the resultant displacements are nonstationary. This can be observed from Figure 5.12. Before SI modelling, the nonstationary time series data need to be transformed to stationary time series data. Considering first differencing, (d=1), the first differenced time series data are determined based on the procedure described in section 5.6. This is shown in Figure 5.13. This also shows that some nonstationarity remains. Differencing of the first differenced time series data, (d=2), the second differenced time series data are estimated and shown in Figure 5.14. This appears to form a stationary process. By integrating (summing) twice, the time series data of Figure 5.14 can be converted to the original nonstationary data of Figure 5.12.

Because the displacements of the platform deck form a sharp peaked power spectrum, it has a very high autocorrelation lag for which the AIC is minimum. A similar case was noticed as described in section-3.4. Such types of time series data can be modelled with minimal AR or ARMA parameters. Starting with AR(2), the estimated power spectrum is compared with the power spectrum determined by the FFT method. This result shows a good match. The AR(2) model can be expressed as a nonstationary process of form as an ARIMA(2,2,0) model as shown in Figure 5.15. Similarly increasing the AR model order, the ARIMA(3,2,0) and ARIMA(5,2,0) model parameters are estimated and their spectra are

compared with spectrum determined by FFT as shown in Figures 5.16 and 5.17. The ARIMA(5,2,0) model seems to represent well the second differenced time series data of platform deck displacements. While examining poles of the above models as plotted and shown in Figures 5.18a-c, it is observed that one pole is very close to zero in magnitude in the ARIMA(3,2,0) model. If we do not consider this pole, then the ARIMA(2,2,0) model is formed. This can be verified by comparing their spectra. Because of the additional pole has a small value present in the ARIMA(3,2,0) model, the PSE contributed by this pole estimates some energy which can be observed in the peak frequency region as compared to the ARIMA(2,2,0) spectrum (Figures 5.15 and 5.16).

CHAPTER 6

MULTIVARIATE AUTOREGRESSIVE MODELLING

6.1. Introduction

In the analysis of recorded time series of ocean wave processes we often require the power spectrum which may be obtained from a single point measurement such as local wave elevation. To obtain further information about the wave process such as directionality we need to apply multivariate spectral analysis techniques to a number of simultaneously measured wave records to determine the cross spectrum. The traditional approach to spectral analysis of ocean waves is to use methods based on the Fourier Transfer such as Welch's method (Oppenheim and Schafer, 1975). An alternative approach is to use modern spectral estimation based on univariate and multivariate autoregressive algorithms.

Univariate autoregressive algorithms have been widely applied in many fields, including ocean engineering, especially for the modelling of ocean waves (Spanos and Hansen, 1981 and Spanos, 1983). For univariate AR models, we only require autospectra of the time series data. There are many practical problems of interest where vector processes are involved. For example phase, coherence and directional spectra of ocean waves involve vector processes. In vector processes important information is present in the cross spectra rather than in the auto spectra alone. Here the desired information can be extracted using multivariate

autoregressive (MAR) algorithms. MAR algorithms can estimate not only the power spectral densities, but also phase and coherence spectra of time series wave data sets. From these MAR algorithms one can estimate wave directionality from three simultaneous wave time histories, provided that the distances between measurement locations are less than half the wave length.

This chapter describes MAR modelling for the estimation of multivariate AR matrix coefficients and prediction error covariance matrices which finally yield the power spectral estimates. The present study considers bi-variate random processes. The MAR model order is selected based on the auto and cross correlation methods and also Akaike information criterion method.

6.2. The multivariate AR model

The unidirectional ocean wave process can be described by the univariate autoregressive algorithm (Spanos, 1983 and Spanos and Mignolet, 1986). MAR processes are developed from the algorithms that have been developed for univariate autoregressive processes. The MAR model is developed to estimate power spectral densities, and then coherence and phase spectra.

The MAR process is defined as the vector recursion

$$y_t = -\sum_{k=1}^{p} A_k y_{t-k} + u_t \tag{6.1}$$

where, y_t is an m-variate vector of time variables,

$$y_t = \begin{bmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \\ \vdots \\ y_{m,t} \end{bmatrix} \tag{6.2}$$

y_{t-k} is the vector of time variables with k time steps elapsed.

A_k are the $m \times m$ autoregressive parameter matrices,

$$A_k = \begin{bmatrix} A_{11,k} & A_{12,k} & A_{13,k} & \dots A_{1m,k} \\ A_{21,k} & A_{22,k} & A_{23,k} & \dots A_{2m,k} \\ A_{31,k} & A_{32,k} & A_{33,k} & \dots A_{3m,k} \\ \vdots & \vdots & \vdots & \dots \vdots \\ A_{m1,k} & A_{m2,k} & A_{m3,k} & \dots A_{mm,k} \end{bmatrix}, \qquad k = 1, 2, 3 \dots\dots p \tag{6.3}$$

And u_t is an *mx1* vector representing the input driving noise process,

$$u_t = \begin{bmatrix} u_{1,t} \\ u_{2,t} \\ u_{3,t} \\ \vdots \\ u_{m,t} \end{bmatrix} \tag{6.4}$$

The z-transform of equation (2.1) can be written as

$$Y(z) = A^{-1}(z) \cdot U(z) \tag{6.5}$$

$$\text{where,} \quad A(z) = I + \sum_{k=1}^{k=p} A(k) \cdot z^{-k}$$

If the multivariate noise processes are white noise with constant covariance matrix, P_W, then the MAR PSE function can be written in z-transform

$$\begin{aligned} P_{MAR}(z) &= A^{-1}(z) \cdot P_w^f \cdot A^{-T}(z^{-1}) \\ &= B^{-1}(z) \cdot P_w^b \cdot B^{-T}(z^{-1}) \end{aligned} \tag{6.6}$$

where A(z) and B(z) are forward and backward transfer functions, and P_w^f and P_w^b are the respective prediction error variances of the time series data sets. The transfer functions and covariances of equation (6.6) are estimated based on the estimation of residual variance matrices and partial correlations as described by Morf et al (1978).

6.3. Multivariate AR algorithms

A stationary *m*-variate AR process may be described by the *forward* filter process at time index, n, as

$$\begin{aligned} e_p^f(n) &= y(n) + \sum_{k=1}^{p} A(k)\, y(n-k) \\ &= \underline{a}_p\, \underline{y}_p(n) \end{aligned} \tag{6.7}$$

where $e_p^f(n)$ is the forward prediction error at n, $\underline{a}_p$ is the block row vector of MAR matrix coefficients given by

$$\underline{a}_p = [I \quad A_p(1) \quad A_p(2) \quad \ldots \quad A_p(p)] \tag{6.8}$$

I is an $m \times m$ identity matrix and $\underline{y}_p(n)$ is the block column vector of multivariate data given by

$$\underline{y}_p(n) = \begin{bmatrix} y(n) \\ y(n-1) \\ \vdots \\ y(n-p) \end{bmatrix} \tag{6.9}$$

Similarly MAR processes can also be expressed by the *backward* filter process

$$\begin{aligned} e_p^b(n-1) &= y(n-p-1) + \sum_{k=1}^{p} B(k)\, y(n-p-1-k) \\ &= \underline{b}_p\, \underline{y}_p(n-1) \end{aligned} \tag{6.10}$$

where

$$\underline{b}_p = [B_p(p) \quad B_p(p-1) \quad \ldots B_p(1) \quad I] \tag{6.11}$$

$$\underline{y}_p(n-1) = \begin{bmatrix} y(n-p-1) \\ y(n-p-2) \\ \vdots \\ y(n-1) \end{bmatrix} \tag{6.12}$$

and $e_p^b(n-1)$ is the backward predictor error.

Initialization of forward and backward prediction error is

$$e_0^f(n) = e_0^b(n) = y(n) \tag{6.13}$$

The residual variance matrices, P_p^f , P_p^b , and the covariance matrix, P_p^{fb} are written as

$$\left.\begin{aligned} P_p^f &= \frac{1}{N-p}\sum_{n=p+1}^{N} e_p^f(n)\cdot e_p^{f\,T}(n) \\ P_p^b &= \frac{1}{N-p}\sum_{n=p+1}^{N} e_p^b(n-1)\cdot e_p^{b\,T}(n-1) \\ P_p^{fb} &= \frac{1}{N-p}\sum_{n=p+1}^{N} e_p^f(n)\cdot e_p^{b\,T}(n-1) \end{aligned}\right\} \tag{6.14}$$

Initialization of residual variance matrices is

$$P_0^f = P_0^b = \frac{1}{N}\sum_{n=1}^{N} y(n).y^T(n) \tag{6.15}$$

Therefore the normalized partial correlation as proposed by Morf et al (1978) can be expressed as

$$\begin{aligned} \rho_{p+1} &= (P_p^f)^{-1/2}\cdot P_p^{fb}\cdot (P_p^b)^{-T/2} \\ &= (P_p^{f\,1/2})^{-1}\cdot P_p^{fb}\cdot (P_p^{b\,1/2})^{-T} \end{aligned} \tag{6.16}$$

The logic for the square-root matrix, $P^{1/2}$, is that for any positive definite matrix, P, satisfies

$$P = P^{1/2}\,P^{T/2} \tag{6.17}$$

where $P^{T/2}$ equals $(P^{1/2})^T$. The superscript $^{1/2}$ denotes the lower triangular matrix. Similarly $P^{-T/2}$ is equal to $(P^{1/2})^{-T}$. The reflection coefficient, $a_p(p)$ provides a unique parameterization of the univariate AR process. Similarly the multivariate normalized partial correlation ρ_p provides a unique parameterization sequence for the multivariate algorithm. Therefore the new forward and backward reflection coefficients (A and B respectively) are predicted based on the partial correlation functions as

$$\begin{aligned} A_{p+1}(p+1) &= -(P_p^{f\,1/2})\,\rho_{p+1}\,(P_p^{b1/2})^{-1} \\ B_{p+1}(p+1) &= -(P_p^{b\,1/2})\,\rho_{p+1}^T\,(P_p^{f1/2})^{-1} \end{aligned} \tag{6.18}$$

Now the order for the prediction error covariance matrices is updated as

$$\begin{aligned} P_{p+1}^f &= [I - A_{p+1}(p+1)\,B_{p+1}(p+1)]\,P_p^f \\ P_{p+1}^b &= [I - B_{p+1}(p+1)\,A_{p+1}(p+1)]\,P_p^b \end{aligned} \tag{6.19}$$

Also multivariate linear prediction errors between orders p and p+1 can be established as

$$\begin{aligned} e_{p+1}^f(n) &= e_p^f(n) + A_{p+1}(p+1)\,e_p^b(n-1) \\ e_{p+1}^b(n) &= e_p^b(n-1) + B_{p+1}(p+1)\,e_p^f(n) \end{aligned} \tag{6.20}$$

Here once the residual variance matrices and their normalized partial correlations are determined, reflection coefficients and error covariances of MAR processes can be estimated using equations (6.18) and (6.19). Reflection coefficients are also known as transfer functions.

6.4 Bi-variate random processes

For bi-variate autoregressive processes, equation (6.1) can be rewritten as

$$\begin{aligned} y_{1,t} &= -\sum_{k=1}^{p} a_{11,k}\, y_{1,t-k} - \sum_{k=1}^{p} a_{12,k}\, y_{2,t-k} + u_{1,t} \\ y_{2,t} &= -\sum_{k=1}^{p} a_{21,k}\, y_{1,t-k} - \sum_{k=1}^{p} a_{22,k}\, y_{2,t-k} + u_{2,t} \end{aligned} \tag{6.21}$$

where $y_{1,t}$ and $y_{2,t}$ are the two time series data sets, $a_{i,j}$ are bi-variate autoregressive coefficients, $u_{1,t}$ and $u_{2,t}$ are white noise and p is the model order. The model order, p, is optimized by using the auto and cross-correlation processes and the AIC of the two time series data sets. The main purpose of optimizing the MAR model is to fit the model to the time series data as closely as possible. Optimization of MAR model order is essential if real time time series data is to be stored in the form of MAR parameters. This ensures that the maximum amount of information may be retrieved from the MAR parameters.

The auto and cross correlations, $r_{y_i y_j}(\lambda)$, of bi-variate processes [equations (4.1)] are expressed as

$$\begin{aligned} r_{y_1 y_1}(\lambda) &= E[y_1(t+\lambda) \cdot y_1(t)] \\ r_{y_2 y_2}(\lambda) &= E[y_2(t+\lambda) \cdot y_2(t)] \\ r_{y_1 y_2}(\lambda) &= E[y_1(t+\lambda) \cdot y_2(t)] \end{aligned} \tag{6.22}$$

where E[.] denotes the expected value.

Power spectral densities, $P_{y_i y_j}(f)$ can be described by the discrete Fourier transform of the correlation functions [equations (6.22)]

$$\begin{aligned} P_{y_1 y_1}(f) &= T \sum_{\lambda=-\infty}^{\infty} r_{y_1 y_1}(\lambda) \cdot e^{-j2\pi f \lambda T} \\ P_{y_2 y_2}(f) &= T \sum_{\lambda=-\infty}^{\infty} r_{y_2 y_2}(\lambda) \cdot e^{-j2\pi f \lambda T} \\ P_{y_1 y_2}(f) &= T \sum_{\lambda=-\infty}^{\infty} r_{y_1 y_2}(\lambda) \cdot e^{-j2\pi f \lambda T} \end{aligned} \tag{6.23}$$

where T is the sampling period and f is the cyclic frequency in Hz.

Cross spectral densities are complex and related by

$$P_{y_1 y_2}(f) = P^*_{y_2 y_1}(-f) \tag{6.24}$$

where the asterisk superscript denotes the complex conjugate.

It is evident that the cross spectrum must be less than or equal to the geometric mean of the time series y_1 and y_2 process spectra, i.e.,

$$\left|P_{y_1 y_2}\right| \leq \sqrt{P_{y_1 y_1} \cdot P_{y_2 y_2}} \tag{6.25}$$

Therefore, the 2x2 power spectral densities matrix must have a positive determinant for all frequencies. The coherence function, $\phi_{y_1 y_2}(f)$, is defined in terms of complex dimensionless form as

$$\phi_{y_1 y_2}(f) = \frac{P_{y_1 y_2}(f)}{\sqrt{P_{y_1 y_1}(f) \cdot P_{y_2 y_2}(f)}} \tag{6.26}$$

The phase spectrum, $\theta(f)$, between two time series data can be expressed as

$$\theta(f) = \tan^{-1}\left[\frac{Imag(\phi_{y_1y_2})}{Real(\phi_{y_1y_2})}\right] \tag{6.27}$$

The magnitude squared coherence, $\Phi(f)$, is defined as

$$\Phi(f) = \frac{\left|P_{y_1y_2}(f)\right|^2}{P_{y_1y_1}(f) \cdot P_{y_2y_2}(f)} \tag{6.28}$$

Since the cross spectrum is less than or equal to the mean of the product of the autospectra [equation (6.25)], $\Phi(f)$ lies between the limits of zero and unity which correspond to no and perfect coherence respectively.

6.5 Multivariate model order selection

For the univariate case, the least squares error and Akaike information criterion methods are well known and easily applied to determine the optimal model order. However, for multivariate cases, it is sometimes difficult to determine the optimal model order mainly owing to the cross correlation processes. To obtain a reliable estimate of MAR coefficients, one has to minimize the model order in order to ensure that the Toeplitz matrix does not become ill conditioned. In addition, the computational time for MAR spectral estimates is much higher than that for univariate cases and so it is desirable to select the minimum model order which can represent approximately the MAR processes of the time series data sets.

In general, exact values of the covariance function are not known and need to be estimated from the time series data y_t . One method for estimating the lagged (λ) covariances, $R(\lambda)$ is

$$R(\lambda) = \frac{1}{M-\lambda} \sum_{t=1}^{M-\lambda} y_t \, y_{t+\lambda} \tag{6.29}$$

This is an unbiased correlation for a total of M data points. The model order can be selected from the auto and cross correlations of the time series data sets up to a lag, λ, equal to the model order p, up to which the process is convergent. Alternately, one can use a multivariate version of the AIC (Jones, 1978) which can be expressed as

$$AIC(p) = M \ln\left|P_p^f\right| + 2m^2p \tag{6.30}$$

Here the model order, p, is selected for which AIC(p) is a minimum. Equation (6.30) can be used as a guideline for selecting the approximate model order. Both of the methods described above are used. In the first pass a simple method for selecting model order can be used by examining auto and cross correlations of the data sets, the best combination (from auto and cross correlations) of minimum order, p, up to which the process is convergent may be selected. Then using equation (6.30) the selected optimal model order can be confirmed.

6.6 Application of multivariate AR algorithms for ocean wave modelling

The MAR algorithms presented in this chapter based on the forward and backward prediction error techniques have been successfully applied to the estimation of power spectra, phase and coherence of two measured time series wave data sets.

The wave measurements were recorded at a location in the southern North Sea where the water depth was 17.3m. The two wave gauge sensors (denoted M12 and M17 respectively) were horizontally separated by a distance of 17.66m (Figure 4.5b). Each wave record set consisted of 4096 points sampled at 4Hz.

The approximate model order for the M12 and M17 wave data sets is determined both by the auto and cross correlation method and the AIC method, which are shown in Figures 6.1a and 6.1b. From Figure 6.1a, it is difficult to select the optimal model order, but it can be assumed that it could be between 40 and 60, up to which the processes are convergent. This difficulty is overcome when one considers the AIC which is presented in Figure 6.1b. It is evident from the Figure 6.1b that the AIC changes very little for model orders greater than 36. It can also be seen that the AIC is a minimum at the model order 44. Hence the optimal model order is selected as 44. Sometimes, if the processes are purely convergent, it is difficult to select the approximate model order by auto and cross correlation. In this case the AIC method is often the best method to select the approximate model order, because multivariate AIC takes into account of least squares error minimization of the multivariate processes.

Figures 6.2 and 6.3 show a comparison between the power spectral estimates using the MAR algorithm presented in this paper and the two

variable FFT technique for the two time series wave data sets. Here it is observed that the MAR power spectral estimates are reasonably close to those determined by the two variable FFT for frequencies where relatively high energy levels are present. Figure 6.4 presents a comparison of the cross spectral energy densities obtained from the above two methods. There is good agreement which is encouraging since the cross spectral density contains the important information needed for estimating the coherence and phase. Once the optimal model order of the MAR model is obtained based on the convergent processes, the phase and coherence of the time series data are determined using equations (6.27) and (6.28) which are shown in Figures 6.5 and 6.6. Figure 6.5 shows the comparison between coherence spectral estimates using MAR algorithms and the two variable FFT. The coherence spectrum also shows good agreement for frequencies associated with relatively high energy levels. There are some differences in the estimates of the coherence and phase spectra compared with the results obtained from the two variable FFT. This is mainly owing to the approximation of the time series data sets in the form of MAR parameters. For purely convergent processes, the optimal model order will sometimes be very high. In this case it will be more complicated to estimate the MAR parameters owing to the ill conditioning of the Toeplitz matrices.

CHAPTER 7
DISCUSSION

7.1 Univariate AR, MA and ARMA modelling

This work presented provides an insight into stochastic processes (ocean waves and offshore vessel and structural motions) by the various SI models. Here three SI algorithms, namely AR, MA and ARMA are successfully applied not only for generating time series data which are compatible with a target (Pierson-Moskowitz, JONSWAP or measured) spectrum of ocean waves, but also for optimization of SI algorithms for the reduction of the vast volume of time series measured data.

The approximation of the theoretical Pierson-Moskowitz spectrum has demonstrated that the system identification process of the Taylor approximated Pierson-Moskowitz spectrum can be the basis for developing efficient ARMA models. The presence of a zero of infinite order of the true Pierson-Moskowitz spectrum at the frequency, $\omega=0$, can be a source of ill-conditioning of the system of linear equations associated with SI models. The AR algorithm represents the current value of the time series as a linear combination of its past values and of white noise. It is based on the recursive filter (refer to Appendix-II). This process also ensures that the autocorrelation functions of the target and AR processes match at a number of time steps (lag) by which the AR model order is determined. As long as the autocorrelation function follows the convergence rule, the model order selection will be proper. If the dynamic

system consists of a convergence and divergence process, then one must be careful while selecting the model order. One should select the optimal order of the model that covers only convergence of the autocorrelation function. The main feature of the AR process is that autocorrelation values can be extrapolated for a time lag greater than the duration of the available record using equation (3.9). The MA process can generate time series data based on the linear combination of a white noise process which is weighted by suitable constants. For determining these constants several criteria can be adopted. Here these constants are associated with the Fourier coefficients of the square root of the target spectrum. The ARMA process represents the current value of the time-series data as a combination of its past values and the values of white noise. Here two approaches, namely, power of order matching (POM) and inverse AR (IAR) filter using MYW equations can be used for determining the ARMA coefficients. The model order selections of the above SI models are optimized based on the LSE and AIC methods. The results of a comparison study on LSE of both the IAR and POM methods of estimating ARMA spectra are shown in Table 3.1. It is observed there that the LSE of the ARMA-POM spectra are less than that of the ARMA-IAR spectra predicted from PM (PM10), JONSWAP (J10), and measured (K15) spectra. This also confirms the efficiency of the ARMA-POM method over the IAR method. Both of the above ARMA spectral estimation methods can be used without much loss of information.

As the model order increases the number of equations in (3.10) increases and the system of Yule-Walker equations becomes ill-conditioned for theoretical ocean wave spectral estimation. Hence the accuracy of the AR coefficients deteriorates. For the above study to be reliable, the numerical integration for obtaining autocorrelation of theoretical spectra should be performed accurately.

Even though ocean dynamic systems are random in nature, the confused ocean surface elevations can be characterised by their power spectral estimations. Similarly offshore structural motions can be expressed by their responses. Sometimes for structural motions the correlation lags up to the stage in which the process is convergent is of very high order owing to its sharp peaked nature. Hence the Toeplitz matrix can become ill-conditioned. For such a high correlation lagged dynamic system the SI model order needs to be properly selected. The lower order autoregressive model can generally be used to describe the sharp peaked dynamic system. The monohull crane vessel A and B motions show examples of such cases. Using autocorrelation, AIC, or MDL methods the selected model order of the crane vessel motion yields a very high value for which the Toeplitz matrix in equation (3.10) becomes ill-conditioned. However, the lower order AR model [AR(3)] can fairly represent the above dynamic system as shown in Figure 3.26.

The study presented herein has shown that the ocean wave process can be appropriately modelled using autocorrelation, AIC, or MDL methods and AR, MA, and ARMA algorithms. For structural motions, one should make first ones examinations at autocorrelations lags up to which the process is convergent. If it is too high, one can start from a lower order and then estimate the proper AR or ARMA model which yields a better representation of the structural motions.

7.2 Reduced order ARMA modelling

The reduced order ARMA modelling procedure presented in Chapter 4 may be summarised. First the initial autoregressive parameters of the ARMA model are obtained based on the optimal model selection methods and MYW equations for target spectra or measured time series wave data. Then a model order reduction technique is applied for further reduction of data in the form of reduced ARMA model coefficients.

Autoregressive, moving average and ARMA algorithms have been applied by other authors for the identification and modelling of target sea wave spectra and the time series random processes [Spanos (1983), Spanos and Mignolet (1987, 1990), Marple (1987), Kay (1988), Rosen and Porat (1989), and Lin (1987)]. These SI algorithms are optimized by selecting the model orders based on autocorrelation matching, least squares error, or the AIC method. Sometimes a particular SI model becomes ill-conditioned owing to a number of factors such as not taking into account all poles and zeros, optimal model orders, and the number of Yule-Walker equations. In some cases depending on the observed time series data, there may be inadequate correlation. Highly uncorrelated data can be modelled with very high model orders, and the model orders selected by the AIC may not always be an improved representation of the dynamic system. However, the proposed method for the reduction of the SI model orders described in Chapter 4, based on initial higher order model selection, gives improved statistics of the data while eliminating spurious disturbances and noise. In many situations, it is practically impossible to obtain undisturbed time series data. It is also difficult to remove disturbances from observed time series data. In the statistical sense and

with prior knowledge of the nature of the response of the system, these disturbances can be removed by the model order reduction technique.

This part of the study emphasizes the application of system identification techniques to the further reduction of marine environmental data. Since the system identification method allows one to build up mathematical models of dynamic systems based on experimental data from the systems, one would also like to establish the theoretical spectra in the form of rational modellings. In this part first theoretical PM and JONSWAP spectra are used. Even though selection of the initial model order is high, this can be reduced to a few parameters which are the reduced order ARMA coefficients. This shows that the model order reduction technique described here is also valid for the development of theoretical spectra. Here the model order reduction technique may not be practically useful for the theoretical spectra which are defined by few values such as wind speed, significant wave height etc., it can be used to verify the power spectral estimatation. This leads to the further establishing of SI models of measured ocean waves. The reduction time series data first involves a smoothing operation by a correlation technique. Here M data points are reduced to N modified Yule-Walker equations where N corresponds to the number of correlation coefficients. Equations (3.6a) and (3.6b) estimate the correlations, where the number of correlations is set by equations (4.2). Then the SI algorithms are applied to reduce the N correlation coefficients to 2p AR coefficients by the optimal model order selection method. These AR coefficients are further reduced to a final reduced order ARMA model by only considering modes with significant energy. Initial excessive smoothing (N is too small) will cause loss of resolution in the resulting spectral estimate. Very little or no smoothing (M=N) should be initially applied to signals with spectra

containing sharp peaks, whereas considerable smoothing (M>>N) can be applied to signals with broadband spectra.

The method of estimating ocean wave spectra by the reduced order ARMA algorithm represents an alternative to other methods such as those based on the fast Fourier transform (FFT) or summation of cosines. The FFT or summation of cosines analysis will include the contributions from noise or the other erroneous disturbances present in the measured time series data. The reduced order ARMA algorithm presented in Chapter 4 offers the potential to eliminate the noise modes. Also relatively few parameters which are the reduced order ARMA coefficients are needed to describe the spectral estimate compared to that based on FFT methods. This has significant advantages in terms of the data storage requirements for a long term structural monitoring system.

7.3 ARIMA modelling of nonstationary processes

The SI models described in Chapters 3 and 4 applied only to stationary processes, that is, mean, variance and autocorrelations of any random process are constant through time. The time series can be nonstationary in many ways. A time series can have a nonstationary mean, variance, autocorrelations or combination of these statistical parameters. The most common nonstationary processes can be those with some steady trend in the mean of the time series. Such types of time series can simply be transformed to a stationary time series by using the differencing technique. Other types of time series may be such that their trend in variability is more complex in nature, and transformation of such time series to stationary process may not be achieved so easily. In these cases, some suitable algebraic transformation techniques need to be developed.

In offshore dynamic systems nonstationary processes (which are a combination of short periods and long periods) are sometimes measured. SI modelling of this type of time series data can be possible through transformation of nonstationary process to a stationary process by the differencing technique. Complete removal of nonstationarity may not always be easy. From the practical application point of view an approximate transformed stationary time series is likely to be adequate for estimation of the ARIMA parameters.

The application of ARIMA modelling to nonstationary time series offshore data is examined in section 5.7. The offshore nonstationary time series are, in general, homogeneous, i.e., different segments of each time series seem to behave similarly to the rest of the time series if we allow for changes of each segment's trend which may be level and/or sloped. In

other words, if we can eliminate the slope and level of different segments of a nonstationary process, we form a seemingly stationary process. Such types of time series can be simply transformed to a stationary process by one or two time differencings as described in section 5.6.

After differencing to form a stationary process, the ARIMA(p,d,q) model is constructed. This SI modelling procedure is similar to that described in Chapter 3 and 4 except for the differencing term. Here model reduction techniques can also be applied to estimate the reduced order ARIMA model.

One should avoid unnecessary differencing which may create artificial patterns in the time series data and reduce estimation accuracy. The number of differencings can be chosen based on (a) visual inspection of time series data, (b) checking autocorrelations of original and differenced time series data, and (c) possibly examining the AR parameters to determine if they follow stationarity conditions.

To recover the original nonstationary time series the differenced time series needs to be integrated which involves summing successive values in a differenced time series.

7.4 Multivariate autoregressive modelling

Multivariate autoregressive algorithms are applied for estimating power spectral densities and then coherence and phase spectra of time series wave data sets. Before proceeding with any identification of time series data, one should first look into the data sets. In most cases, digital time series data recordings are in raw form and the records include some noise and extraneous disturbances. Even though system identification techniques can be applied to remove the disturbances as described by Witz and Mandal (1991), much better identification can be obtained if the data is preprocessed prior to identification. The preprocessing may involve the following:

1. Removal of erroneous large peaks
2. Removal of mean values
3. Band-pass filtering
4. Sampling interval selection

Once the data sets are preprocessed, they can be used for estimation of the MAR parameters. First of all an appropriate model order which can represent the MAR processes needs to be selected. There are some difficulties associated with selecting optimal model orders using auto and cross correlations of time series data sets as shown in Figure 6.1a. The AIC method overcomes this problem. The AIC method appropriately selects the minimum model order at which the AIC(p) is minimum. In general, parametric MAR models have high dimensionality. For high model orders, simultaneous estimation of so many parameters using the MAR algorithm will involve very high computational times compared with those for univariate AR algorithms. In addition, higher order matrix

coefficients can sometimes give misleading results owing to the ill conditioning of the higher order Toeplitz matrix. Many researchers have extended Burg's (1975) algorithms to the multivariate case. In doing so, two sequences of reflection coefficients are determined by forward and backward filter techniques. Briggs and Vandiver (1982) used only the MAR forward filter technique to estimate the transfer function of an offshore platform's response to environmental excitation. Using these two techniques it is straight forward to estimate power spectral densities separately. Power spectral density estimates from the two techniques are different for the same data sets. Here we have applied a partial correlation technique which has been applied to solve the estimates of the MAR matrix coefficients and predicted error covariance matrices using both forward and backward filter techniques. This is made possible by determination of square-root matrix and normalized partial correlation as described by equations (3.11) and (3.10) respectively.

For univariate AR models, we only require the autospectrum of the time series data. There are many practical problems of interest where vector processes are involved. For example; phase, coherence, and directional spectra of ocean waves involve vector processes. In vector processes important information is present in the cross spectra rather than in the auto spectra alone. Here the desired information can be extracted using MAR algorithms. MAR algorithms can estimate not only the power spectral densities, but also phase and coherence spectra of time series wave data sets. From these MAR algorithms one can estimate wave directionality from three simultaneous wave time histories, provided that the distances between measurement locations are less than half the wave length.

Using MAR modelling, one can reduce the time series data into a number of parameters which are the MAR coefficients. As an example, two time series wave data of 4096x2 points are shown to have been reduced to 132 values which are the MAR coefficients. Future development of the reduced order multivariate SI modelling will further reduce multivariate SI coefficients with minimum loss of information.

7.5 Future work

The study presented herein provides an insight into the stochastic processes such as ocean waves and offshore structural motions by the various SI models. Here AR, MA and ARMA algorithms have been successfully applied for estimation of theoretical and measured spectra of ocean waves, platform deck motion, semisubmersible and monohull crane vessels' motions. ARIMA algorithms have been applied to nonstationary structural motions. Some time series wave data have also been used for optimal selection of the above SI algorithms. However there are many limitations when model order reduction is applied. The quality of final reduced order ARMA spectral estimation purely depends on selection of the initial higher order model. If the initial higher order model is not properly selected, the final ARMA spectral estimation can cause misleading results. Moreover multipeak spectra lead to the selection of very high orders of model. So an initial higher orders model must be properly selected. Therefore, it is necessary to use many observed time series structural monitoring data and marine environmental data to verify the proposed SI models and modify the SI models accordingly. Another way of verifying the performance of SI models is to compare the target spectrum and estimated spectrum (or original data and predicted data) obtained from SI models by evaluating a number of statistical characteristics such as cumulative energy and cumulative root-mean-square distributions. This is worthy of further study.

One should properly select the sampling rate of the time series offshore data. A higher sampling rate will increase correlation lags and hence the optimal model order will increase. However a very low

sampling rate may describe acceptable spectral estimates but it may poorly represent the time series. Detailed study of such cases may be undertaken to confirm the effect of SI modelling on the sampling rate of the time series.

Based on analysis of measured monohull crane vessel motions in random sea waves the transfer function estimation by SI algorithms and FFT shows good consistency. However for better estimation of the transfer functions, one may measure and analyse different high sea state structural motions. This study will confirm the validity of estimation of the transfer functions by SI algorithms which can be used as an alternative method to the FFT technique.

Based on the study as reported here, it is clear that there is an immense opportunity for use of SI methods applied to time series in offshore monitoring and environmental data and further study is likely to be needed in order to obtain the most suitable properly selected optimal models for specific problems. Depending on the problems one needs to resolve, various SI algorithms can be formulated and used for wide sense stationary random processes. As an example, to determine the modes and the intensity of vibration of a member of the offshore structure likely to fail, SI modelling can be applied.

Similarly modified SI algorithms can be applied depending on the nature of nonstationarity of the offshore random processes. In this thesis generated nonstationary wave and real platform deck motions are analysed using ARIMA algorithms. Here the process is nonstationary in its mean. There may be nonstationarity in its variance or other form for which modified SI modellings need to be developed.

The model reduction technique as described in Chapter 4 for univariate offshore dynamic systems can be extended to multivariate offshore dynamic systems, so that the reduced order rational SI model could effectively represent the multivariate processes with minimum loss of information. This will give a more complete reduction technique for application to offshore dynamic systems.

The multivariate autoregressive algorithms described in Chapter 6 are applied to two time series data sets. These algorithms could be extended to estimate wave directionality from three simultaneous time series wave data sets collected within the distances between measurement locations are less than half the wave length.

The study presented herein has been limited to off-line SI modelling. This could also be extended for on-line SI modelling by modifying SI algorithms and using proper feedback mechanisms.

CHAPTER 8
CONCLUSIONS

The study presented herein has emphasized the application of system identification techniques to marine environmental and structural monitoring data. The SI algorithms, namely AR, MA, and ARMA have been discussed and applied not only for generating time series data which are compatible with a target (PM or JONSWAP) spectrum of ocean waves, but also for optimization of SI algorithms for the reduction of the vast volume of time series measured data (ocean waves, offshore structural motions, etc.). Further reduction of the SI models' orders can eliminate noise disturbances which are present in the recorded time series data.

The AR algorithm is formulated based on the recursive filter technique (Appendix II). This process also ensures determination of the model order which is selected as the number of time steps of lag of the autocorrelation function up to which the process is convergent. The main feature of the AR process is that autocorrelation values can be extrapolated for lags greater than the duration of available record using equation (3.9). The MA algorithm is formulated based on the linear combination of a white noise process which is weighted by suitable constants associated with the Fourier coefficients of the square root of the target spectrum. The MA process can also be generated from the AR process using relationship as described in section 3.2.2. The ARMA algorithms are formulated as the combination of AR and MA processes.

Here two approaches, namely, power of order matching (POM) and inverse AR (IAR) methods are used for determining the ARMA coefficients as described in section 3.2.3. The model order selection of the above SI models are carried out based on the LSE, AIC, or MDL method.

The approximation of the theoretical PM spectrum has demonstrated that the system identification process of Taylor approximated PM spectrum up to an order of eight can be the basis for developing efficient ARMA models. The presence of a zero of infinite order of the true PM spectrum at the frequency, $\omega=0$, can be a source of ill-conditioning of the system of linear equations associated with SI models. The above SI algorithms are successfully applied to ocean waves (theoretical PM and JONSWAP spectra and measured time series) and offshore structural dynamic systems. The time series measured data should be preprocessed (removal of erroneous peaks and mean values, selection of proper sampling interval, etc.) prior to SI modelling which will lead to a better estimation.

The correlation lag up to which the process is convergent is very high in some cases of offshore structural motions. In these types of processes one should be careful about proper selection of model order. It is observed that the lower order AR or ARMA models can generally be used to express such types of time series. These lower order SI models' transfer functions of offshore structural motions in random sea waves show consistency with that estimated by FFT.

Based on the selection of the initial higher order SI model and the MYW method, the model order reduction technique is developed. Firstly an estimation of a higher order AR is carried out and then the model

order reduction technique (based on calculation of the energy of the mode) is applied to obtain the final reduced order ARMA model. This reduction technique can remove the spurious noise modes which are present in the time series data. Using the above SI techniques, for example, 2048 digital wave data are reduced to a few values which are the reduced order parameters of an ARMA model as shown in Figure 4.9a. Therefore, once a particular SI model is validated for any stochastic process, this model can be used for subsequent reduction of time series data.

Sometimes nonstationary offshore structural motions or ocean processes may be observed. Such types of time series normally can not be modelled using SI algorithms which are generally used for wide sense stationary random processes. One has to visually observe the time series in respect of its nature of nonstationarity, i.e., levels and trends of nonstationarity. The study presented herein considered the cases where the observed nonstationary time series having nonstationary mean are modelled using differencing techniques as described in Chapter 5. To regain the nonstationary process, one has to integrate by a number of times equal to the number of differencings already undertaken for SI modellings. There may be some cases where the time series having nonstationary variance can be modelled by first taking natural logarithms and then differencing techniques [Pankratz (1983), Chapter 7]. The main purpose of forcing the nonstationary time series to stationary time series is to model the process by formulating ARIMA algorithms. Once the system seems to form a stationary process one can use SI algorithms as described in Chapters 3 and 4. The generated nonstationary ocean waves and measured nonstationary platform deck motions (displacements) are

successfully modelled using ARIMA algorithms as described in Chapter 5.

The multivariate autoregressive modelling will not only represent the random processes, it can also give information about their effects or influences on each other based on coherence and phase spectral estimates. The MAR algorithms are formulated using the forward and backward prediction error techniques, feedback weighting coefficients matrices and the driving noise vector which are obtained based on the estimation of the partial correlation of the time series data sets. Here the appropriate model order is selected based on convergence processes of the time series data sets, i.e., by auto and cross correlations and multivariate Akaike information criterion methods. These algorithms are applied to estimate MAR power spectral density spectra and then phase and coherence spectra of two time series wave data sets collected at a North Sea location. The estimation of MAR power spectral densities are compared with spectral estimates computed from a two variable fast Fourier transform, which show very good agreement. Also the estimated coherence shows good agreement with coherence determined by FFT in the peak frequency region. This MAR modelling can be extended to further development of the reduced order multivariate SI modellings which will further reduce the SI parameters with minimum loss of information.

ACKNOWLEDGEMENTS

I would first like to thank National Institute of Oceanography (CSIR, India), Goa and the Ministry of Welfare (Government of India) for providing an opportunity and financial support to do my PhD research work.

I would like to express my sincere gratitude to my supervisor, Dr G J Lyons for his valuable guidance, advice, help and support throughout this work. I am grateful to Dr J A Witz for his keen interest, frequent discussion and valuable suggestion during this research study. I am also thankful to Professor M H Patel for his valuable advice, encouragement and support.

I wish to thank my friends, Mr H I Park, Mr Z Tan, Dr J Fang, Mr A Watkins, Mr O Pearson at the Santa Fe Laboratory for Offshore Engineering and Dr G X Wu at the London Centre for Marine Technology at University College London for their friendly help and encouragement; to Miss S J Collins for her kind secretarial assistance.

I am thankful to M/s (a) Rijkswaterstaat, the Netherlands, (b) Indian Navy, (c) Santa Fe Drilling Co., UK, (d) McDermott Inc, (e) British Petroleum, UK and (f) others *Anon.* for kindly providing measured time series data on waves, semisubmersible motions, platform deck acceleration, monohull crane vessels motions. This research study would have been incomplete without these real time offshore time series data made available.

I would like to thank Dr B U Nayak and Ocean Engineering colleagues of National Institute of Oceanography, Goa, India for their encouragement. My particular thanks are due to my parents, brothers and sister, Mr Padmalava and his family for their constant inspiration, encouragement and concern; to my wife, Swapna and daughter, Sagarika for their sympathetic interest and support.

REFERENCES

1 Akaike, H. Fitting autoregressive models for prediction. *Ann. Inst. Statist. Math.*, vol. 21, **1969**, pp 243-247.

2 Akaike, H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control,* Vol-AC-19, No-6, December, **1974**, pp 716-723.

3 Akaike, H. Time series analysis and control through parametric models. *Proc. First Applied Time series Symp.* held at Tulsa, Oklahama, May 14-15, **1976**.

4 Anderson, T. W. Determination of the order of dependence in normally distributed time series. in *Time Series Analysis.*, (Ed. M. Rosenblatt), New York, Wiley, **1963**, pp 425-446.

5 Anderson, T. W. *The Statistical Analysis of Time Series*. Wiley, New York, ISBN 0-471-02900-9, **1971**.

6 Beck, J. L. Determining models of structures from earthquake records. *Ph.D. thesis*, California Institute of Technology, Pasadena, California, **1979**.

7 Beex, A. A. and Scharf, L. Covariance sequence approximation for parametric spectrum modeling. *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol-ASSP-29, No-5, **1981**, pp 1042-1052.

8 Bendat, J. S. and Piersol, A. G. *Random Data,* Wiley-Interscience Publication, N.Y., **1986**.

9 Bendat, J. S. and Piersol, A. G. *Engineering Applications of Correlation and Spectral Analysis*. Wiley-Interscience Publication, N.Y., ISBN 0-471-05887-4, **1980**.

10 Box, G. E. P. and Jenkins, G. M. *Time Series Analysis: Forecasting and Control.* Revised Edn., Holden-Day Inc., California, ISBN 0-8162-1104-3, **1976**.

11 Briggs, J. and Vandiver, J. K. Multichannel Maximum Entropy Method of Spectral Analysis Applied to Offshore Platforms. *Proc. Offshore Technology Conference,* Houston, Texas, OTC 4286, **1982,** pp 647-659.

12 Broome, D. R. and Pittaras, A. The time prediction of ship motions at sea. *Proc. Offshore Technology Conference,* Houston, Texas, OTC 6222, **1990a,** pp 199-206.

13 Broome, D. R. and Pittaras, A. Ship motion prediction. Proc. 9th Intl. Conf. *Offshore Mechanics and Arctic Engineering (OMAE),.* Houston, Texas, **1990b,** pp 303-311.

14 Burg, J. P. Maximum entropy spectral analysis. Ph.D. thesis, Dept. of Geophysics, Standford Univ., Standford, CA., **1975.**

15 Cadzow, J. A. High performance spectral estimation - A new ARMA method. *IEEE Transactions on Acoustics, Speech and Signal Processing,* Vol- ASSP-28, No-5, October, **1980,** pp 524-529.

16 Cadzow, J. A. Spectral estimation: an overdetermined rational model equation approach. *Proc. IEEE,* **1982,** pp 907-939.

17 Eykhoff, P. *System Identification: Parameter and State Estimation.* John Wiley & Sons, London, ISBN 0-471-24980-7, **1974.**

18 Findley, D. F. *Applied Time Series Analysis,* Academic Press, London, ISBN 0-12-257250-5, **1978.**

19 Franklin, G. F. and Powell, J. D. *Digital Control of Dynamic System,* Addison-Wesley Publ. Comp. Inc., Reading, ISBN 0-201-02891-3, **1980.**

20 Friedlander, B. The asymptotic performance of the modified Yule-Walker estimator. *Proc. 2nd ASSP Workshop on Spectral Estimation,* Tampa, Fla., November, **1983,** pp 22-26.

21 Friedlander, B. and Porat, B. Modified Yule-Walker method of ARMA spectral estimation. *IEEE Transactions,* Vol-AES-20, No-2, **1984,** pp 158-173.

22 Gersch, W. and Yonemoto, J. Synthesis of multivariate random vibration system: A two-stage least square AR-MA model approach. *Journal of Sound and Vibration,* Vol. 52, No-4, **1977,** pp 553-565.

23 Glover, K. All optimal Hankel-norm approximations of linear multivariable systems and their L-infinity error bounds. *International Journal of Control*, Vol-39, **1984**, pp 1115-1193.

24 Gnedenko, B. V. *The theory of probability* (Translated from the Russian by B. D. Seckler), Chelsea Publishing Co, N.Y., ISBN x-28-015810-5, **1967**.

25 Graupe, D. Krause, D. J. and Moore, J. B. Identification of autoregressive moving average parameters of time series. *IEEE Transactions on Automatic Control*, Vol-AC-20, February, **1975**, pp 104-106.

26 Hannan, E. J. and Deistler, M. *The Statistical Theory of Linear Systems*. John Wiley & Sons, New York, ISBN 0-471-80777-x, **1988**.

27 Harvey, A. C. *The Econometric Analysis of Time Series*. 2nd Edn., Phillip Allan, Deddington, UK, ISBN 0-86003-085-7, **1987**.

28 Hasselmann, K. et. al. Measurements of wind-wave growth and swell decay during the JOint North Sea WAve Project (JONSWAP). Report-12, A8(Suppl..), **1973**, Deutsches Hydro. Inst., Hamberg, Germany.

29 Imai, H., Yun, C. B., Maruyama, O. and Shinozuka, M. Fundamentals of system identification in structural dynamics. *Probabilistic Engineering Mechanics*, Computational Mechanics Publications, Vol-4, No-4, **1989**, pp 162-173.

30 Jefferys, E. R. and Goheen, K. R. Novel modelling techniques for problems in marine dynamics. Proc. 9th Intl. Conf. on *Offshore Mechanics and Arctic Engineering*,, **1990**, pp 231-238.

31 Jenkins, G. M. *Practical Experiences with Modelling and Forcasting Time Series*. GJP Publication, Channel Islands, U.K., ISBN 0-9506-423-0-4, **1979**.

32 Jones, R. H. Identification and autoregressive spectrum estimation. *IEEE Transactions on Automatic Control*, Vol-AC-19, No-6, December, **1974**, pp 139-162.

33 Jones, R. H. Multivariate autoregression estimation using residuals. *Applied Time Series Analysis*, Academic Press, N. Y. (Ed. D. F. Findley), **1978**, pp 139-162.

34 Kalman, R. L. A new approach to linear filtering and prediction problems, *Transactions ASME (J. Basic Eng.)*, Vol-82D, March, **1960**, pp 34-45.

35 Kaplan, P., Jiang, C. W. and Dello Stritto, F. J. Determination of offshore structure Morrison equation force coefficients via systen identification techniques. *International Symposium on Hydrodynamics in Ocean Engineering*, Norwegian Institute of Technology, **1981**.

36 Karl, J. H. *An Introduction to Digital Signal Processing*, Academic Press, London, ISBN 0-12-398420-3, **1989**.

37 Kashyap, R. L. Inconsistency of the AIC rule for estimating the order of autoregressive models. *IEEE Transactions on Automatic Control*, Vol-AC-25, October, **1980**, pp 996-998.

38 Kay, S. M. A New ARMA Spectral Estimator", *IEEE Transactions* Vol-ASSP-28, No-5, **1980**, pp 585-588.

39 Kay, S. M. *Modern Spectral Estimation: Theory and Application*, Prentice-Hall, Englewood Cliffs N.J. 07632, ISBN 0-13-598582-x-025, **1988**.

40 Kay, S. M. and Marple, S. L. Jr. Spectrum analysis - A modern perspective. *Proc. IEEE*, Vol-69, November, **1981**, pp 1380-1419.

41 Li, C. S. Estimation of system damping ratios of ship hull-girders in response to random environment. *Marine Structures*, Vol-4, **1991**, pp 1-15.

42 Li, S. and Dickinson, B. W. An efficient method to compute consistent estimates of the AR parameters of an ARMA model. *IEEE Transactionson Automatic Control*, Vol-AC-31, March, **1986**, pp 275-278.

43 Li, S. and Dickinson, B. W. Application of the lattice filter to robust estimation of AR and ARMA models. *IEEE Transactions on*

Acoustics, Speech and Signal Processing, , Vol-ASSP-36, April, **1988**, pp 502-512.

44 Li, C. S. and Ko, W. J. On the application of the time series in structural failure detection and monitoring for offshore applications. *Transactions on the Royal Institution of Naval Architects (RINA),* Vol-130, **1988**, pp 315-327.

45 Li, S., Zhu, Y. and Dickinson, B. W. A comparison of two linear methods of estimating the parameters of ARMA models. *IEEE Transactions on Automatic Control,* Vol-AC-34, No-8, August, **1989**, pp 915-917.

46 Lin, N.K. Real time estimation of ship motion using ARMA filtering techniques. *Proc. Offshore Mechanics and Arctic Engineering* **1987**, pp 325-329.

47 Ljung, L. *System Identification: Theory for user,* Prentice-Hall, Englewood Cliffs, New Jersey 07632., ISBN 0-13-881640-9-025, **1987**.

48 Ljung, L. and Soderstrom, T. *Theory and Practice of Recursive Identification,* MIT Press, Cambridge, Massachussete, ISBN 0-262-12095-x, **1983**.

49 Mandal, S., Lyons, G. J. and Witz, J. A. Multivariate autoregressive algorithm for ocean wave modelling, Proc. 2nd *International Offshore and Polar Engineering Conference,* ISOPE-92, San Francisco, USA, Vol-III, **1992**, pp 77-84.

50 Mandal, S., Witz, J. A. and Lyons, G. J. Reduced order ARMA spectral estimation of ocean waves, *Journal of Applied Ocean Research,* (in press), **1992**.

51 Marple, S. L. *Digital Spectral Analysis with Applications,* Prentice-Hall, Englewood Cliffs N.J. 07632, ISBN 0-13-214149-3-025, **1987**.

52 Mason, A. B. and Ullmann, R. R. Experimental evaluation of damping in a steel tubular member of an offshore platform jacket in air. *Journal of Offshore Mechanics and Arctic Engineering,* ASME, Vol-112, May, **1990**, pp 143-150.

53 Medina, J. R. and Sanchez-Carratala, C. R. Robust AR representation of ocean spectra. *Journal of Engineering Mechanics*, ASCE, Vol-117, No-12, December, **1991**, pp 2926-2930.

54 Mignolet, M. P. and Spanos, P. D. Recursive simulation of stationary multivariate random processes - part I, *Journal of Applied Mechanics, ASME*, Vol-54, September, **1987**, pp 674-680.

55 Morf, M. and Kailath, T. Square-root algorithms for least-squares estimation. *IEEE Transactions on Automatic Control*, Vol-AC-20, No-4, August, **1975**, pp 487-497.

56 Morf, M., Vieira, A., Lee, D. T., and Kailath, T. Recursive multichannel maximum entropy spectral estimation. *IEEE Transactions Geoscience Electronics*, Vol-GE-16, No-2, April, **1978**, pp 85-94.

57 Mourjopoulos, J. and Paraskevas, M. A. Pole and zero modeling of room transfer functions. *Journal of Sound and Vibration*, Vol-146, No-2, **1991**, pp 281-302.

58 Nazem, S. M. *Applied Time Series Analysis for Business and Economic Forecasting*. Dekker, New York, ISBN 0-8247-7913-4, **1988**.

59 Newland, D. E. *An Introduction to Random Vibrations and Spectral Analysis*, Langman Group Ltd, Essex, ISBN 0-582-30530-6, **1984**.

60 Oppenheim, A. V. and Schafer, R. W. *Digital Signal Processing*, Prentice-Hall International Inc, ISBN 0-13-214107-8-01, **1975**.

61 Oppenheim, A.V. and Schafer, R.W. *Discrete-Time Signal Processing*, Prentice-Hall, Englewood Cliffs, N.J. 07632, ISBN 0-13-216292-x, **1989**.

62 Pandit, S. M. and Wu, S-M. *Time Series and System Analysis with Applications*. John Wiley and Sons, N.Y., ISBN 0-471-86886-8, **1983**.

63 Pankratz, A. *Forecasting with Univariate Box-Jenkins Models: Concepts and Cases*. John Wiley & Sons, N.Y., ISBN 0-471-09023-9, **1983**.

64 Parzen, E. Some recent advances in time series modeling. *IEEE Transactions on Automatic Control*, Vol-AC-19, No-6, December, **1974**, pp 723-730.

65 Patel, M. H. and Witz, J. A. *Compliant Offshore Structures*, Butterworth-Heinemann, Oxford, ISBN 0-7506-1070-0, **1991**.

66 Pierson, W. J. and Moskowitz, L. A proposed spectrum form for fully developed wind sea based on the similarity theory of S. A. Kitaigorodskii, *Journal of Geophysical Res.*, Vol-69, No-24, **1964**, pp 5181-5190.

67 Pires, H. O., Justino, P. A. P. and Pontes, M. T. Autoregressive representation of random sea waves. Proc. 2nd *International Offshore and Polar Engineering Conference*, ISOPE-92, San Francisco, USA, Vol-III, **1992**, pp 114-118.

68 Pollard, D. *Convergence of Stochastic Processes*. Springer-Verlag, New York, ISBN 0-387-90990-7, **1984**.

69 Popescu, T. D. and Demetriu, S. Analysis and simulation of earthquake ground motions using ARMA models, *Automica*, International Federation of Automatic Control, Pergamon Press plc, Vol-26, No-4, **1990**, pp 721-737.

70 Porat, B. Israel Institute of Technology, Israel. - Personal communication. June, **1990**.

71 Press, W. H., Flannery, B. P., Teukolsky, S. A. and Vetterling, W. T. *Numerical Recipes- The art of Scientific Computing* (FORTRAN Version). Cambridge Univ. Press, Cambridge, **1989**.

72 Priestley, M. B. *Spectral Analysis and Time Series*, Academic Press, London, ISBN 0-12-564922-3, **1981**.

73 Priestley, M. B. *Non-Linear and Non-Stationary Time Series Analysis*, Academic Press, London, ISBN 0-12-564910-x, **1988**.

74 Rissanen, J. A universal prior for intergers and estimation by minimum description length. *The Annals of Statistics*, Vol-11, No-2, **1983**, pp 416-431.

75 Rosen, Y. and Porat, B. Optimal ARMA parameter estimation based on the sample covariances for data with missing observations. *IEEE Transactions on Information Theory*, Vol-35, No-2, March, **1989**, pp 342-349.

76 Safak, E. Adaptive modeling, identification, and control of dynamic structural systems II: Applications. *Journal of Engineering Mechanics*, ASCE, Vol-115, No-11, November, **1989**, pp 2406-2426.

77 Sajonia, C. B. and Niedzwecki, J. M. Wave force prediction using an autoregressive model. *Ocean Engineering*, Pergamon Press plc, Vol-17, No-5, **1990**, pp 463-480.

78 Samaras, E. F., Shinozuka, M. and Tsurui, A. ARMA representation of random vector process. *Journal of Engineering Mechanics*, ASCE, Vol-111, No-3, March, **1985**, pp 449-461.

79 Samii, K. and Vandiver, J. K. A numerically efficient technique for the simulation of random wave forces on offshore structures. *Proc. of the 16th annual Ocean Technology Conference* in Houston, Tex., OTC-4811, May, **1984**, pp 301-305.

80 Sinha, N. K. *Linear Systems*. John Wiley & Sons Inc., ISBN 0-471-62341-5, **1991**.

81 Soderstrom, T. and Stoica, P. *System Identification*, Prentice-Hall International (UK) Ltd, London, ISBN 0-13-881236-5, **1989**.

82 Spanos, P-T. D. and Hansen, J. E. Linear prediction theory for digital simulation of sea waves. *Journal of Energy Resources Technology, ASME*, Vol-103, September, **1981**, pp 243-249.

83 Spanos, P-T. D. ARMA algorithms for ocean wave modeling. *Journal of Energy Resources Technology, ASME*, Vol-105, September, **1983**, pp 300-309.

84 Spanos, P-T. D. and Mignolet, M. P. Z - transform modeling of PM wave spectrum. *Journal of Engineering Mechanics, ASCE*, Vol-112, No-8, August, **1986**, pp 745-759.

85 Spanos, P. D. and Mignolet, M. P. Recursive simulation of stationary multivariate random processes - part II, *Journal of Applied Mechanics, ASME*, Vol-54, September, **1987**, pp 681-687.

86 Spanos, P-T. D. and Mignolet, M. P. Simulation of stationary random process: Two-stage MA to ARMA approach, *Journal of Engineering Mechanics*, ASCE, Vol-116, No-3, March, **1990**, pp 620-641.

87 Stoer, J. and Bulirsch, R. *Introduction to Numerical Analysis.* Springer-Verglag, New York, ISBN LC79-014629, **1980**, Sections 5.5-5.9.

88 Strand, O. N. Multichannel complex maximum entropy (autoregressive) spectral analysis. *IEEE Transactions on Automatic Control*, Vol-AC-22, No-4, August, **1977**, pp 634-640.

89 Walker, G. On periodicity in series of related terms. *Proceedings of the Royal Society*, Vol- A13, **1931**, pp 518-532.

90 Witz, J. A. and Mandal, S. Application of system identification to the reduction of structural monitoring data. Final *Integration of Design Techniques Project* Report MON5 ('89-'91), Santa Fe Laboratory for Offshore Engineering, Dept. of Mechanical Engineering, UCL, London, October, **1991**.

91 Witz, J. A., Mandal, S. and Lyons, G. J. Ocean wave spectral estimation by reduced order ARMA algorithms. Proc. 11th Intl. Conf. on *Offshore Mechanics and Arctic Engineering*, OMAE-92, Calgary, Canada, Vol-I-A,**1992**, pp 17-24.

92 Worden, K., Stansby, P.K., Tomlinson, G.R. and Billings, S.A. On wave force analysis using system identification. Proc. 6th International Conf. on Behaviour of Offshore Structures, **BOSS-92**, London, Vol-1, **1992**, pp 415-430.

93 Yule, G. U. On a method of investigating periodicities in disturbed series with special reference to Wolfer's sunspot number. *Philosophical Transactions*, Vol-A226, **1927**, pp 267-298.

APPENDIX I

Maximum likelihood estimation and properties of convergence of random processes

The method of maximum likelihood is widely used in estimation theory. The basic concept of maximum likelihood method is simple and can be described as follows.

Let **e**, [$\mathbf{e} = \mathbf{y} - x\theta$] be a discrete white random process (error) with zero mean and variance σ^2, and $p(\mathbf{e},\theta)$ be its probability density function, where θ is some unknown parameter to be determined. Consider a set of M independent samples, $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_M$. Now we have to obtain the best estimate of these samples for the parameter θ. The estimate θ is selected in such a way that $\mathbf{e}_i$ are most likely to occur during measurements. The best estimate θ can be obtained by formulating a likelihood function and then taking into derivative with respect to θ.

Assuming the PDF of $\mathbf{e}_i$ is Gaussian distributed, $p(\mathbf{e},\theta)$ can be expressed as

$$p(e_i,\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} Exp\left(-\frac{e_i^2}{2\sigma^2}\right) \qquad \text{(AI-1)}$$

The likelihood function $L(\theta)$ is expressed as the joint probability density function of $\mathbf{e}_i$ and it can be written as

$$
\begin{aligned}
LnL(\theta) &= Ln\prod_{i=1}^{M} p(e_i,\theta) \\
&= Ln\left\{\frac{1}{(2\pi\sigma^2)^{M/2}}Exp(-\frac{1}{2\sigma^2}e^T e)\right\} \\
&= -\frac{1}{2\sigma^2}(y-x\theta)^T(y-x\theta) - \frac{M}{2}Ln(2\pi\sigma^2) \\
&= -\frac{1}{2\sigma^2}(y-x\theta)^T(y-x\theta) + C
\end{aligned}
\qquad \text{(AI-2)}
$$

where C is a constant. Taking the derivative with respect to θ of the above equation

$$
\frac{\partial}{\partial\theta}(y-x\theta)^T(y-x\theta)\Big|_{\theta=\hat{\theta}} = 0 \qquad \text{(AI-3)}
$$

This solution of estimating θ is identical to the least-squares solution. This shows that the ML estimator for Gaussian distributed errors is equivalent to the LSE estimator.

Some properties of convergence concepts for random procrsses are given below. Let x_t be an indexed sequence of random variables and x^* be a random variable, then

a) $x_t \to x^*$ as $t \to \infty$ with probability unity

$$
P(x_t \to x^*, t \to \infty) = 1 \qquad \text{(AI-4a)}
$$

b) $x_t \to x^*$ in mean square

$$
E[x_t - x^*]^2 \to 0 \quad \text{as} \quad t \to \infty \qquad \text{(AI-4b)}
$$

c) $x_t \rightarrow x^*$ in probability

For every small value $\varepsilon > 0$

$$P(|x_t - x^*| > \varepsilon) \rightarrow 0 \quad \text{as} \quad t \rightarrow \infty \tag{AI-4c}$$

d) $x_t \rightarrow x^*$ in the distribution

If the probability density functions of x and x are defined as $p_{x_t}(x)$ and $p_{x^*}(x)$ respectively , then

$$p_{x_t}(x) \rightarrow p_{x^*}(x) \tag{AI-4d}$$

The stationary random process x(n) is said to be ergodic with respect to its first and second order moments if

$$\frac{1}{M}\sum_{n=1}^{M} x(n) \rightarrow E[x(n)] \tag{AI-5a}$$

$$\frac{1}{M}\sum_{n=1}^{M} x(n+\tau)x(n) \rightarrow E[x(n+\tau)x(n)] \tag{AI-5b}$$

with probability unity as $M \rightarrow \infty$. The above two terms are also used as the sample mean and covariance of the data respectively.

If the random discrete time stationary process x(n) shows $\frac{1}{M}\sum_{n=1}^{M} x(n) \rightarrow E[x(n)]$, then the covariances of the data $R_\lambda \rightarrow 0$ as $|\lambda| \rightarrow \infty$ with probability unity as $M \rightarrow \infty$.

Proof of the above properties are given in Gnedenko (1967).

APPENDIX II

Some useful examples of recursive filters

The digital filters are applied to discrete time series by convolving input (x_i) time series with the impulse response or weighting function (f_i) of the filter. The output time series (y_p) can be obtained as

$$y_p = \sum_{i=0}^{N} f_i x_{p-i} \qquad \text{(AII-1)}$$

The convolution operation can also be represented by the product of the z-transform of the input and filter time series. Equation (AII-1) can also be expressed in the form of z-transform as

$$Y(z) = F(z).X(z) \qquad \text{(AII-2)}$$

Here X(z) is the input to the filter and Y(z) is the output. The variable z represents the operation of delaying the signal sample by one sample interval. In the form of Laplace transform, z can be expressed in frequency variable, ω

$$z = e^{-\omega T}$$

where T is the sample interval in seconds.

Some digital filters can be expressed as rational functions of z, i.e., a ratio of two polynomials in z. As an example, the digital filter can be written as

$$F(z) = \frac{B(z)}{A(z)} = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2} + \ldots + b_n z^{-n}}{1 + a_1 z^{-1} + a_2 z^{-2} + \ldots + a_n z^{-n}} \qquad \text{(AII-3)}$$

Substituting F(z) in equation (AII-2), we get

$$Y(z) + z \cdot Y(z).[a_1 + a_2 z^{-1} + \ldots\ldots] = [b_0 + b_1 z^{-1} + \ldots\ldots].X(z)$$

Therefore,

$$Y(z) = [b_0 + b_1 z^{-1} + \ldots\ldots].X(z) - z.[a_1 + a_2 z^{-1} + \ldots\ldots].Y(z) \qquad \text{(AII-4)}$$

or,

$$Y(z) = B(z).X(z) - [A(z) - 1].Y(z) \qquad \text{(AII-5)}$$

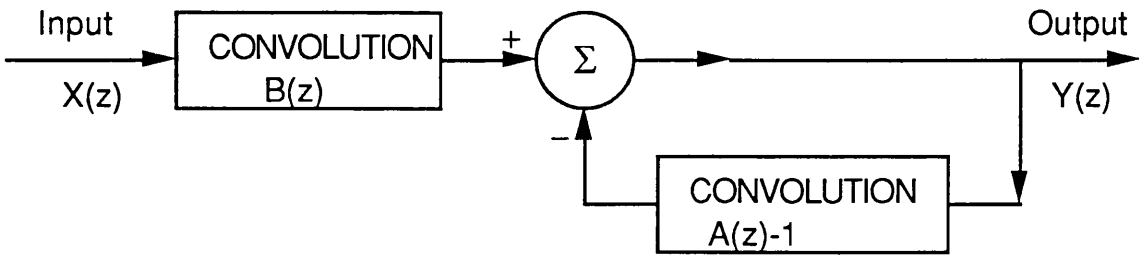

Figure AII.1 Block diagram of equation AII-5

The above figure shows a feedback mechanism of the dynamic system and may be realized in a digital computer with a feedback or

recursive equation. The recursive equation for above figure can generally be expressed as

$$y_n = [b_0 x_n + b_1 x_{n-1} + \ldots\ldots] - [a_1 y_{n-1} + a_2 y_{n-2} + \ldots\ldots]$$

or,

$$y_n = \sum_{i=0}^{n} b_i x_{n-i} - \sum_{j=1}^{n} a_j y_{n-j} \qquad \text{(AII-6)}$$

Where, the output time series is computed in the sequential order $y_1, y_2, y_3, y_4 \cdots\cdots y_n$. Here the recursive filters are physically realizable in the sense that they can not respond to an input before it occurs. Hence we must assume $y_n = 0$ for $n<0$, that is, there is no input before x_0.

Example-1

To prove equation (AII-6), let us consider a particular example of the recursive filter

$$F(z) = \frac{1.0}{1.0 - 0.5\, z^{-1}} \qquad \text{(AII-7)}$$

The above filter could be applied by expanding in a polynomial and using convolution. By using simple division

$$1.0 - 0.5\, z^{-1}) 1.0 \qquad (1 + 0.5z^{-1} + 0.25\, z^{-2} + 0.125z^{-3} + \ldots\ldots$$

$$\underline{1.0 - 0.5z^{-1}}$$

$$0.5z^{-1}$$

$$\underline{0.5z^{-1} - 0.25\, z^{-2}}$$

$$0.25\, z^{-2}$$

$$\therefore \quad F(z) = 1 + 0.5z^{-1} + 0.25z^{-2} + 0.125z^{-3} + \ldots\ldots \tag{AII-8}$$

This filter could be applied by convolving with a sufficient number of terms of the series. This filter could also be applied by using equation (AII-6), where

$a_1 = -0.5$

$b_0 = 1.0$

Therefore, equation (AII-6) becomes

$$y_n = x_n + 0.5\, y_{n-1} \tag{AII-9}$$

Now we can prove that equation (AII-9) is equivalent to equation (AII-8) by using unit impulse response of the recursive equation (i.e., input time series (x_n) as 1.0, 0.0, 0.0, . . .).

So $x_0 = 1.0$

$x_n = 0.0$ for n=0

The output of equation (AII-9)

$$y_0 = x_0 + 0.5(y_{-1}) = 1.0 + 0.5(0.0) = 1.0$$

$$y_1 = x_1 + 0.5\,(y_0) = 0.0 + 0.5(1.0) = 0.5$$

$$y_2 = x_2 + 0.5\,(y_1) = 0.0 + 0.5(0.5) = 0.25$$

$$y_3 = x_1 + 0.25\,(y_2) = 0.0 + 0.5(0.25) = 0.125 \tag{AII-10}$$

Therefore we could prove that the unit impulse response of the recursion equation is identical to the coefficients of the polynomial expansion as given by equation (AII-9). Hence the recursion equation performs the function as the convolution of the input with the filter weighting function.

Example-2

Consider an inverse of a polynomial, $a(z^{-1})$ of the time series

$$F(z) = \frac{1}{a(z^{-1})} = \sum_{n=0}^{\infty} y_n . z^{-n} \qquad \text{(AII-11)}$$

where,

$$a(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3} +$$

Equation (AII-6) can be applied as

$$y_n = x_n - [a_1 y_{n-1} + a_2 y_{n-2} + a_3 y_{n-3} + \ldots] \qquad \text{(AII-12)}$$

Using the unit impulse response or weighting function of the filter, which is identical to the convolution of the operator of the filters, the coefficients $y_0\ y_1\ y_2\ \ldots$ can be calculated from equation (AII-12).

Here $x_0 = 1.0$

$x_n = 0.0$ for $n = 0$

The output of the equation (AII-12)

$$
\begin{aligned}
y_0 &= 1.0 - [0.0] = 1.0 \\
y_1 &= 0.0 - [a_1 y_0] \\
y_2 &= 0.0 - [a_1 y_1 + a_2 y_0] \\
y_3 &= 0.0 - [a_1 y_2 + a_2 y_1 + a_3 y_0] \\
y_4 &= 0.0 - [a_1 y_3 + a_2 y_2 + a_3 y_1 + a_4 y_0]
\end{aligned}
\qquad \text{(AII-13)}
$$

Example-3

Consider a filter

$$
F(z) = \frac{1}{1 + d_1 . z^{-1}} = \sum_{n=0}^{\infty} y_n . z^{-n} \qquad \text{(AII-14)}
$$

From equation (AII-6), the output time series can be written as

$$
y_n = x_n - d_1 y_{n-1} \qquad \text{(AII-15)}
$$

Therefore, coefficients of the output are

$$
\begin{aligned}
y_0 &= 1.0 \\
y_1 &= 0.0 - d_1 y_0 \\
y_2 &= 0.0 - d_1 y_1 \\
y_3 &= 0.0 - d_1 y_2 \\
y_4 &= 0.0 - d_1 y_3
\end{aligned}
\qquad \text{(AII-16)}
$$

Example-4

Consider another type of filter

$$F(z) = \frac{1}{1 + e_1.z^{-1} + e_2.z^{-2}} = \sum_{n=0}^{\infty} y_n.z^{-n}$$

(AII-17)

From equation (AII-6), the coefficients of output time series can be expressed as

$$y_n = x_n - e_1 y_{n-1} - e_2 y_{n-2}$$

(AII-18)

Therefore, the coefficients of output are

$$\begin{aligned} y_0 &= 1.0 \\ y_1 &= 0.0 - e_1 y_0 \\ y_2 &= 0.0 - e_1 y_1 - e_2 y_0 \\ y_3 &= 0.0 - e_1 y_2 - e_2 y_1 \end{aligned}$$

(AII-19)

APPENDIX III

Modified Yule-Walker equations

An approach for the determination of successive values based on past observed values was introduced by Yule (1927). He carried out an investigation into the estimation of periodicities in disturbed series with special reference to successive annual sunspot numbers. Yule made an introductory study on generating random data based on superposing harmonic fluctuations and disturbances. Yule developed the regression approach for estimating successive values. If $y_1, y_2, y_3, \ldots\ldots y_n$ represent the series resulting from a dynamic system then successive values of that dynamic system in the absence of disturbances, having a causal relationship, can be expressed in a regression equation as

$$y_n = -[a_1 y_{n-1} + a_2 y_{n-2} + \ldots\ldots + a_p y_{n-p}] \qquad \text{(AIII-1)}$$

The above equation with disturbances is obtained by adding a term, W_n to the right-hand side.

If equations (AIII-1) hold true and provided that n is large, Walker (1931) showed that a similar equation holds approximately between the successive values of the correlation coefficients, R_t, of the terms of y_i separated by time lag t,

$$R_t = -\left[a_1 R_{t-1} + a_2 R_{t-2} + \ldots\ldots + a_p R_{t-p}\right] \qquad \text{(AIII-2)}$$

The variation of correlation coefficients, R_t, obtained from equations (AIII-2) is much smoother than that of the time series data. Based on equations (AIII-2) various relationships are found between the amplitude of corresponding terms in the Fourier series and those of the correlation coefficients. Equations (AIII-2) are known as the *Yule-Walker* equations.

Let y_t be a discrete-time zero-mean Gaussian ARMA process of order (p,q), with $p \geq q$, expressed by the difference equation

$$y_t = -\sum_{i=1}^{p} a_i y_{t-i} + \sum_{i=0}^{q} b_i w_{t-i} \qquad \text{(AIII-3)}$$

Where a_i and b_i are the coefficients of the ARMA(p,q) process, and w_t is a white Gaussian noise process.

Now multiply both sides of the equation (AIII-3) by $y_{t-q-\iota}$ and taking expected values, the correlation coefficients of the process y_t are given by

$$R_i = E\{y_t y_{t-i}\}$$

and

$$R_i = R_{-i}$$

Then we get

$$R_{q+\iota} + a_1 R_{q+\iota-1} + a_2 R_{q+\iota-2} + \ldots\ldots + a_p R_{q+\iota-p} = 0 \quad \text{for} \quad 1 \leq \iota \leq p \qquad \text{(AIII-4)}$$

Equations (AIII-4) are often called **Modified Yule-Walker (MYW)** equations. The simple ordinary Yule-Walker equations can be obtained with q = 0. In the case of spectral estimation, we have to find out the true correlation coefficients, R_i, and the order (p,q) of the ARMA process. If the observed or target spectrum is given, then R_λ can be determined from the relationships

$$R_\lambda = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} S(\omega)\cos(\lambda\omega T)\,d\omega \qquad \text{(AIII-5)}$$

where, $\lambda = 0, 1, 2, 3, \ldots$ and,

$$T = \frac{\pi}{\omega_b}$$

ω_b is designated as the cut-off frequency.

In the absence of a target spectrum, if the available information is a finite set of time series data, such as for $\{0 < t < N\}$, then the true correlation coefficients, R_λ, can be calculated from the relationships

$$R_\lambda = \frac{1}{N-\lambda} \sum_{t=\lambda+1}^{N} y_t \cdot y_{t-\lambda} \qquad \text{(AIII-46)}$$

APPENDIX IV

Calculation of the roots of a polynomial

To determine the roots of a polynomial in a straightforward way, one can use *Laguerre's method* which is briefly described below.

Let the polynomial is defined by its roots as

$$f_n(x) = (x - x_1)(x - x_2) \ldots \ldots (x - x_n) \qquad \text{(AIV-1)}$$

Taking natural logs in both side of equation (AIV-1)

$$F_x = \ln|f_n(x)| = \ln|x - x_1| + \ln|x - x_2| \ldots \ldots \ln|x - x_n| \qquad \text{(AIV-2)}$$

Then taking the first derivative with respect to x

$$\frac{dF_x}{dx} = \frac{1}{x - x_1} + \frac{1}{x - x_2} \ldots \ldots \frac{1}{x - x_n} = G \qquad \text{(AIV-3)}$$

and the second derivative with respect to x

$$\frac{d^2F_x}{dx^2} = \frac{1}{(x - x_1)^2} + \frac{1}{(x - x_2)^2} \ldots \ldots \frac{1}{(x - x_n)^2} = H \qquad \text{(AIV-4)}$$

the Laguerre formulaes make a drastic set of assumptions. They consider that the first root x_1 is located at some distance **a** from the current x_i and all other roots are located at a distance **b**

$$a = x - x_1$$
$$b = x - x_i, \qquad i = 2, 3, \ldots\ldots n \tag{AIV-5}$$

Using the notations **a** and **b** as given by equation (AIV-5), equations (AIV-3) and (AIV-4) can be writen as

$$\frac{1}{a} + \frac{n-1}{b} = G \tag{AIV-6}$$

$$\frac{1}{a^2} + \frac{n-1}{b^2} = H \tag{AIV-7}$$

Using equations (AIV-6) and (AIV-7), the solution of **a** is obtained as

$$a = \frac{n}{G \pm \sqrt{(n-1)(nH - G^2)}} \tag{AIV-8}$$

Here the sign of the denominator should be taken care. If the factor inside the square root is negative, **a** can be complex. The method of obtaining the first root x_1 is determined by iteration. First using a trial value of x, **a** is obtained using equation (AIV-8) and then (x-**a**) becomes the next trial value. This will continue untill the value **a** is sufficiently small. In this way subsequent roots are determined.

For more details one can refer to Stoer and Bulirsch (1980), and Press et al (1989).

APPENDIX V

Calculation of energies of first and second order modes of the dynamic system.

Let α_t be a pure AR process of the order n of the dynamic system,

$$\alpha_t = -a_1\alpha_{t-1} - a_2\alpha_{t-2} - \ldots\ldots - a_n\alpha_{t-n} + w_t \qquad \text{(AV-1)}$$

Where, w_t is the white noise with unit variance.

Let $\{r_0^\alpha, r_1^\alpha, r_2^\alpha, \ldots\ldots r_n^\alpha\}$ be the first n+1 covariances of $\{\alpha_t\}$.

Using the Yule-Walker equations for the order n

$$\begin{bmatrix} r_0^\alpha & r_1^\alpha & r_2^\alpha & \ldots\ldots & r_n^\alpha \\ r_1^\alpha & r_0^\alpha & r_1^\alpha & \ldots\ldots & r_{n-1}^\alpha \\ r_2^\alpha & r_1^\alpha & r_0^\alpha & \ldots\ldots & r_{n-2}^\alpha \\ \vdots & \vdots & \vdots & & \vdots \\ r_n^\alpha & r_{n-1}^\alpha & r_{n-2}^\alpha & \ldots\ldots & r_0^\alpha \end{bmatrix} \cdot \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \text{(AV-2)}$$

Let $n = 1$

$$\begin{bmatrix} r_0^\alpha & r_1^\alpha \\ r_1^\alpha & r_0^\alpha \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

or,

$$\begin{bmatrix} 1 & a_1 \\ a_1 & 1 \end{bmatrix} \begin{bmatrix} r_0^\alpha \\ r_1^\alpha \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\therefore \quad r_0^\alpha = \frac{1}{1-a_1^2}$$

$$r_1^\alpha = \frac{-a_1}{1-a_1^2} \qquad \text{(AV-3)}$$

Let $n = 2$

$$\begin{bmatrix} r_0^\alpha & r_1^\alpha & r_2^\alpha \\ r_1^\alpha & r_0^\alpha & r_1^\alpha \\ r_2^\alpha & r_1^\alpha & r_0^\alpha \end{bmatrix} \cdot \begin{bmatrix} 1 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

or,

$$\begin{bmatrix} 1 & a_1 & a_2 \\ a_1 & 1+a_2 & 0 \\ a_2 & a_1 & 1 \end{bmatrix} \cdot \begin{bmatrix} r_0^\alpha \\ r_1^\alpha \\ r_2^\alpha \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Now consider

$$\begin{aligned} \Delta &= (1+a_2) + a_1^2 a_2 - a_1^2 - a_2^2(1+a_2) \\ &= (1+a_2) - a_2^2(1+a_2) + a_1^2 a_2 - a_1^2 \\ &= (1+a_2)(1-a_2^2) - a_1^2(1-a_2) \\ &= (1+a_2)^2(1-a_2) - a_1^2(1-a_2) \\ &= (1-a_2)\left[(1+a_2)^2 - a_1^2\right] \\ &= (1-a_2)(1+a_2+a_1)(1+a_2-a_1) \end{aligned} \qquad \text{(AV-4)}$$

$$\therefore \quad \begin{bmatrix} r_0^\alpha \\ r_1^\alpha \\ r_2^\alpha \end{bmatrix} = \frac{1}{\Delta} \begin{bmatrix} 1+a_2 & -a_1(1-a_2) & -a_1(1+a_2) \\ -a_1 & 1-a_2^2 & a_1 a_2 \\ a_1^2 - a_2(1+a_2) & a_1 a_2 - a_1 & 1+a_2-a_1^2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$= \frac{1}{\Delta} \begin{bmatrix} 1+a_2 \\ -a_1 \\ a_1^2 - a_2(1+a_2) \end{bmatrix} \qquad \text{(AV-5)}$$

Let us consider b_t is an ARMA process of order n of the dynamic system i.e,

$$\beta_t = -a_1\beta_{t-1} - a_2\beta_{t-2} - a_3\beta_{t-3} \ldots - a_n\beta_{t-n} + b_0 w_t + b_1 w_{t-1} \ldots + b_n w_{t-n}$$

(AV-6)

Therefore, the variance of b_t is given by

$$r_0^{\beta} = [b_0 \quad b_1 \quad b_2 \ldots b_n] \begin{bmatrix} r_0^{\alpha} & r_1^{\alpha} & r_2^{\alpha} & \ldots\ldots & r_n^{\alpha} \\ r_1^{\alpha} & r_0^{\alpha} & r_1^{\alpha} & \ldots\ldots & r_{n-1}^{\alpha} \\ r_2^{\alpha} & r_1^{\alpha} & r_0^{\alpha} & \ldots\ldots & r_{n-2}^{\alpha} \\ \vdots & \vdots & \vdots & & \vdots \\ r_n^{\alpha} & r_{n-1}^{\alpha} & r_{n-2}^{\alpha} & \ldots\ldots & r_0^{\alpha} \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Let n = 1, i.e, for first order case,

$$r_0^{\beta} = [b_0 \quad b_1] \cdot \begin{bmatrix} 1 & -a_1 \\ -a_1 & 1 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} \cdot \frac{1}{1-a_1^2}$$

$$= \frac{b_0^2 - 2a_1 b_0 + b_1^2}{1-a_1^2}$$

(AV-7)

Let n = 2, i.e, for second order case

$$r_0^{\beta} = \frac{1}{\Delta}[b_0 \quad b_1 \quad b_2] \cdot \begin{bmatrix} 1+a_2 & -a_1 & a_1^2 - a_2(1+a_2) \\ -a_1 & 1+a_2 & -a_1 \\ a_1^2 - a_2(1+a_2) & -a_1 & 1+a_2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix}$$

(AV-8)

Here, equations (AV-7) and (AV-8) are the energies of first and second order modes of the dynamic system respectively.

Table 3.1 Comparison of LSE of ARMA power spectra between IAR and POM techniques.

	Model Order	ARMA-IAR	ARMA-POM
K15	(20, 20)	0.003346	0.002750
J10	(20, 20)	0.007545	0.001064
J10	(30, 30)	0.001358	0.001028
PM10	(20, 20)	0.000833	0.000780

Table 4.1 Modes and energy levels for PM10

Mode Number	Mode	Normalized Energy
1	$1 + 1.5453z^{-1} + 1.4740\ z^{-2}$	<1.0E-06
2	$1 + 1.4370z^{-1} + 1.0976z^{-2}$	<1.0E-06
3	$1 -. 0.0562z^{-1} + 1.1598z^{-2}$	<1.0E-06
4	$1 + 1.0018z^{-1}$	<1.0E-06
5	$1 + 0.8981z^{-1}$	<1.0E-06
6	$1 + 0.7920z^{-1}$	<1.0E-06
7	$1 + 0.4639z^{-1}$	0.000001
8	$1 - 1.4135z^{-1} + 0.7104z^{-2}$	0.06136
9	$1 - 1.3012z^{-1} + 0.6184z^{-2}$	0.52210
10	$1 - 1.1280z^{-1} + 0.5088z^{-2}$	0.97570
11	$1 - 0.8990z^{-1} + 0.3543z^{-2}$	1.00000
12	$1 - 0.5970z^{-1} + 0.1627z^{-2}$	0.01831

Table 4.2 Modes and energy levels for M12

Mode Number	Mode	Normalized Energy
1	$1 + 2.0146z^{-1} + 1.0470\,z^{-2}$	<1E-06
2	$1 + 1.7136z^{-1} + 0.9728z^{-2}$	<1E-06
3	$1 + 1.3518z^{-1} + 0.9540z^{-2}$	<1E-06
4	$1 + 0.9331z^{-1} + 1.0519z^{-2}$	<1E-06
5	$1 - 0.1283z^{-1} + 1.2933z^{-2}$	<1E-06
6	$1 - 1.4063z^{-1} + 1.5560z^{-2}$	<1E-06
7	$1 - 0.4225z^{-1} + 0.6942z^{-2}$	<1E-06
8	$1 - 2.5952z^{-1} + 1.7261z^{-2}$	<1E-06
9	$1 - 1.8786z^{-1} + 0.9311z^{-2}$	1.0000
10	$1 - 1.7675z^{-1} + 0.8763z^{-2}$	0.1482
11	$1 - 1.0679z^{-1} + 0.5032z^{-2}$	0.0002

Table 4.3 Modes and energy levels for K15

Mode Number	Mode	Normalized Energy
1	$1 - 0.0569z^{-1} + 3.9920z^{-2}$	<1E-06
2	$1 + 1.8125z^{-1} + 1.0776z^{-2}$	<1E-06
3	$1 + 1.2114z^{-1} + 0.9331z^{-2}$	0.00004
4	$1 + 0.9234z^{-1}$	0.00002
5	$1 + 0.7494z^{-1}$	0.00005
6	$1 - 0.3274z^{-1} + 0.9844z^{-2}$	0.00624
7	$1 - 1.4870z^{-1} + 1.4601z^{-2}$	<1E-06
8	$1 - 0.4412z^{-1} + 0.7595z^{-2}$	0.00501
9	$1 - 1.8268z^{-1} + 0.9248z^{-2}$	0.00232
10	$1 - 1.6260z^{-1} + 1.0357z^{-2}$	<1E-06
11	$1 - 1.3200z^{-1} + 0.7819z^{-2}$	1.00000

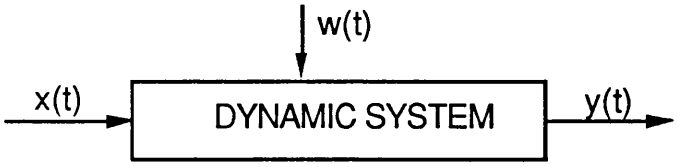

Figure 1.1 Dynamic system with input x(t), output y(t), and disturbance w(t).

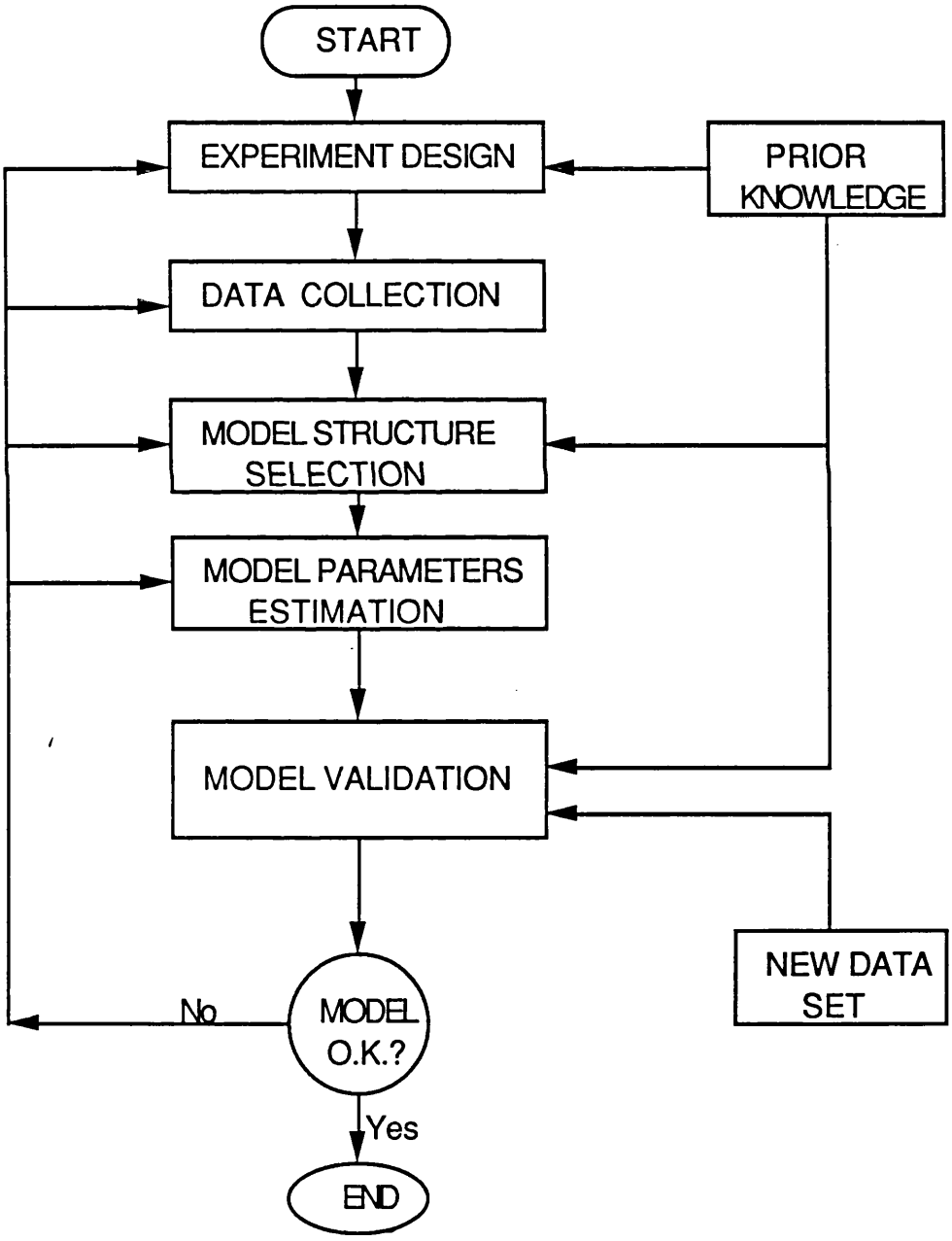

Figure 1.2 System identification loop

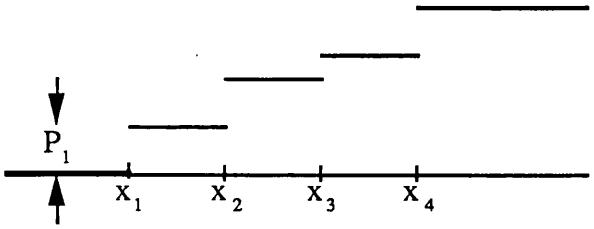

Figure 2.1 A piecewise-constant function

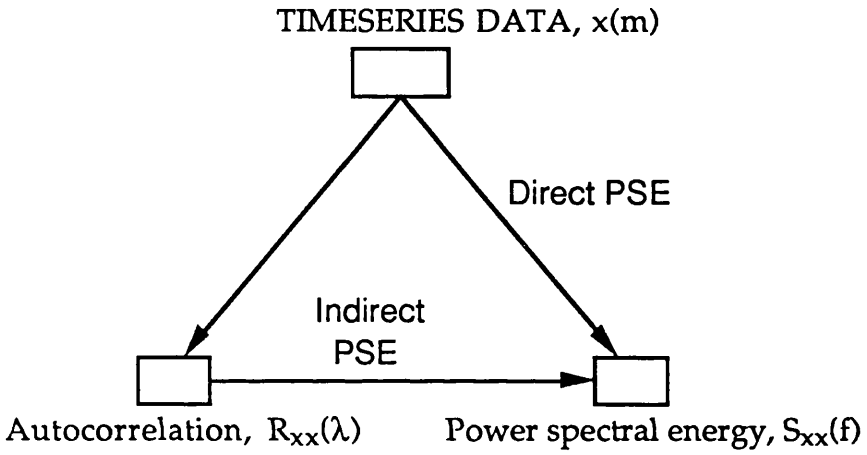

Figure 2.2 Relationships of the time series, autocorrelation and power spectral energy (PSE).

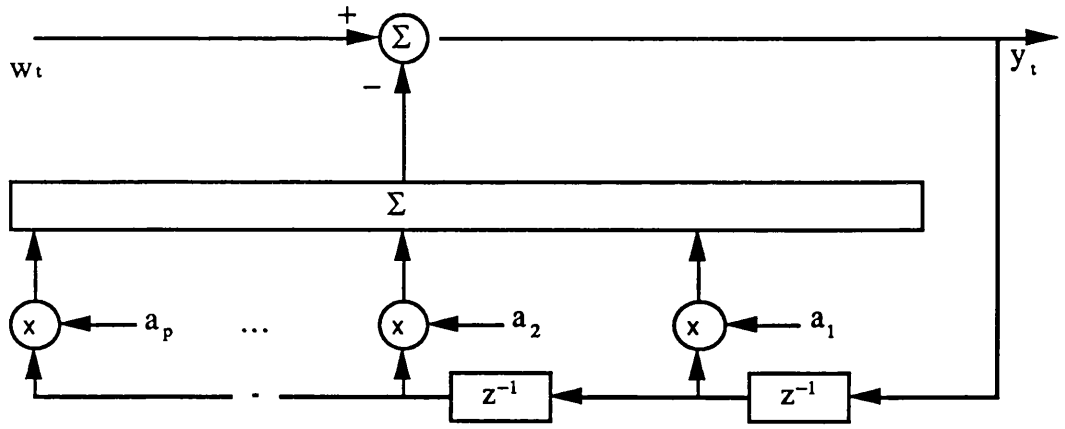

Figure 3.1 Autoregressive, AR(p) model

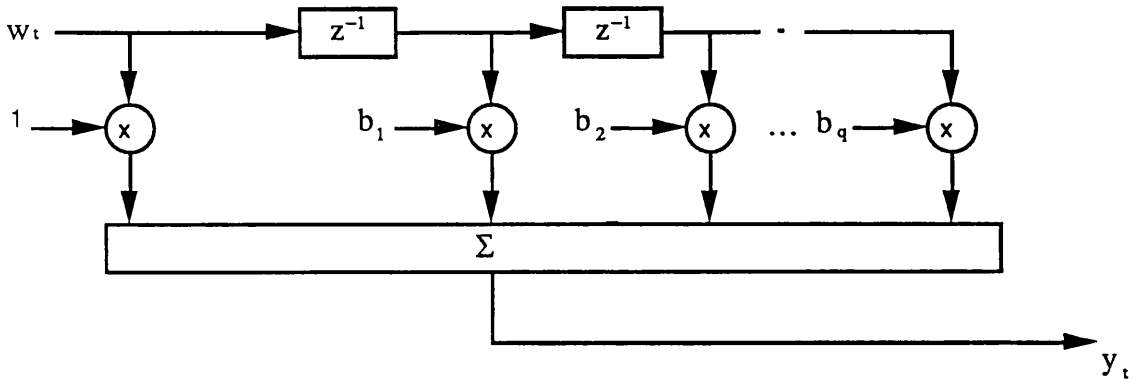

Figure 3.2 Moving average, MA(q) model

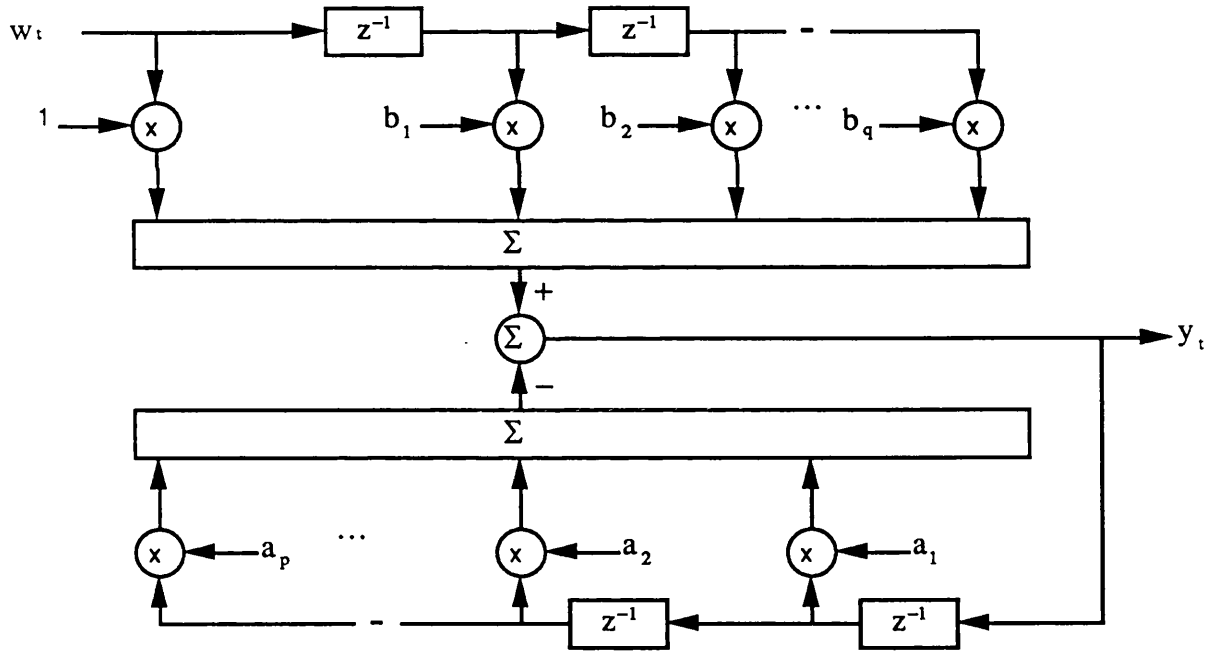

Figure 3.3 Autoregressive moving average, ARMA(p,q) model

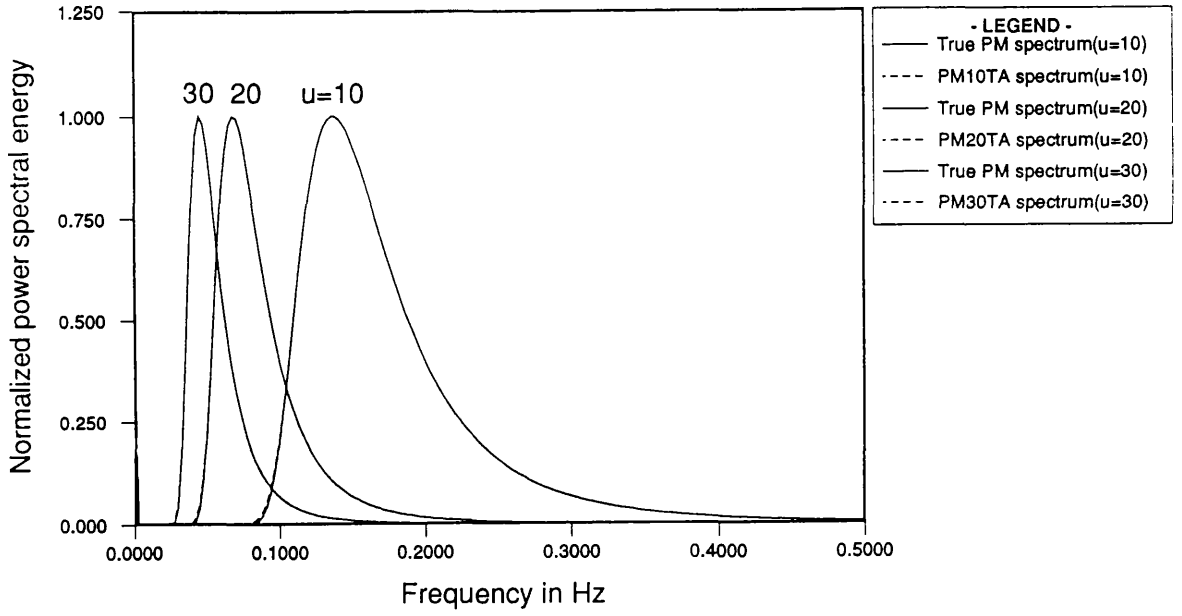

Figure 3.4 Normalized Pierson-Moskowitz (PM) spectra for wind speed, u=10, 20, 30 m/s.

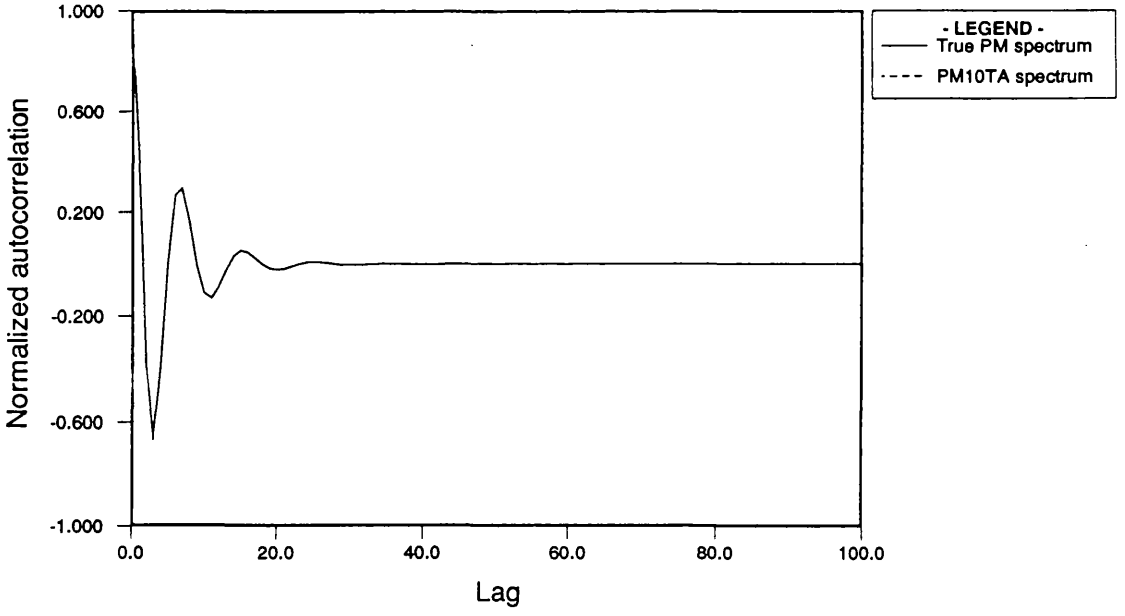

Figure 3.5 Normalized autocorrelation functions of PM10 (PM spectrum with u=10 m/s).

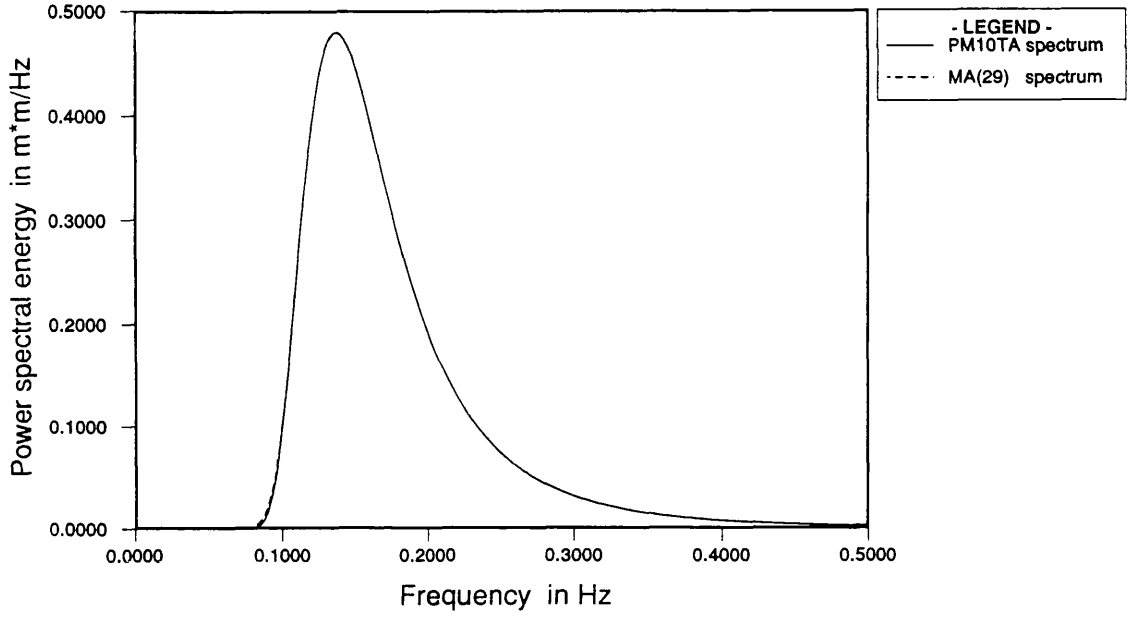

Figure 3.6 Comparison between PM10TA spectrum and MA(29) spectrum.

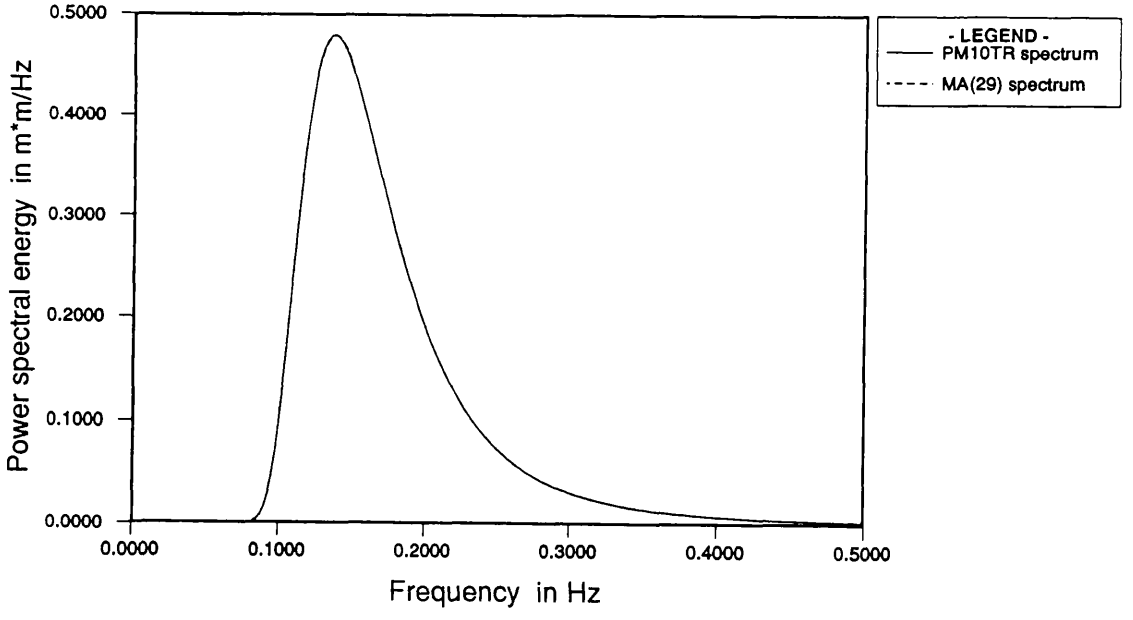

Figure 3.7 Comparison between PM10TR spectrum and MA(29) spectrum.

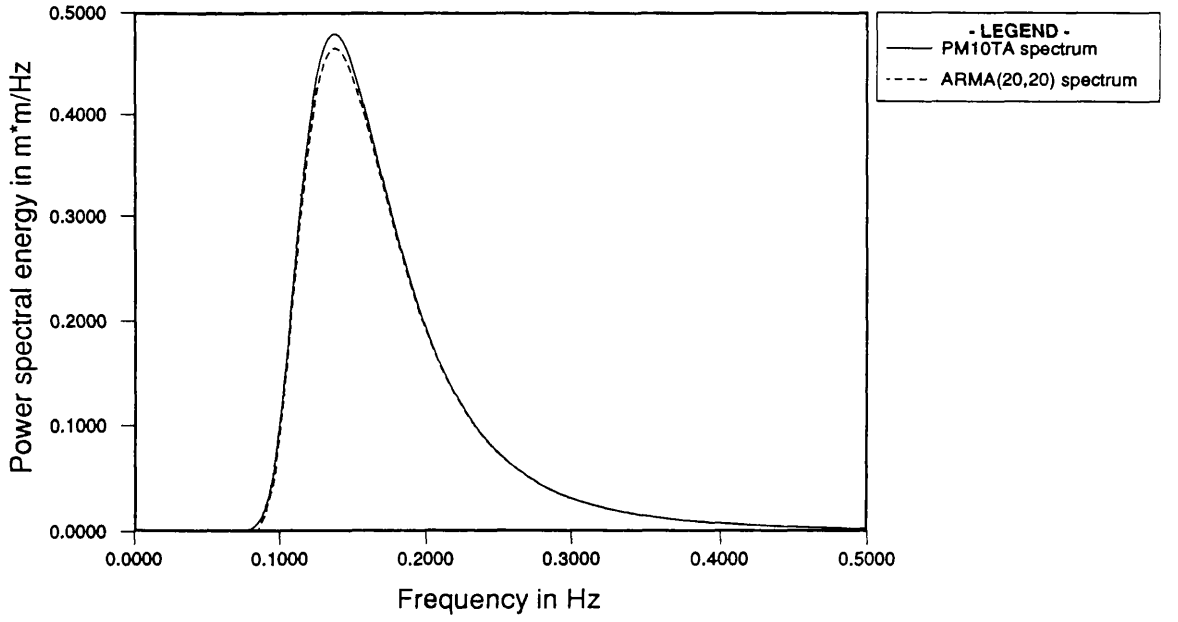

Figure 3.8 Comparison between PM10TA spectrum and ARMA(20,20) spectrum.

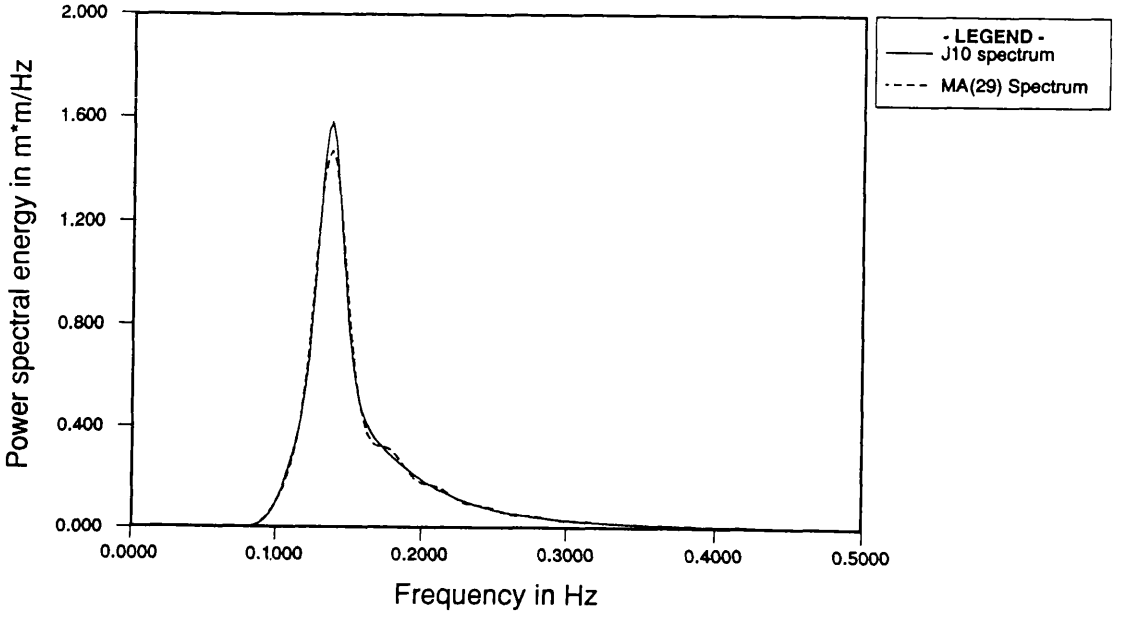

Figure 3.9 Comparison between JONSWAP (u=10m/s, γ=3.3) spectrum and MA(29) spectrum.

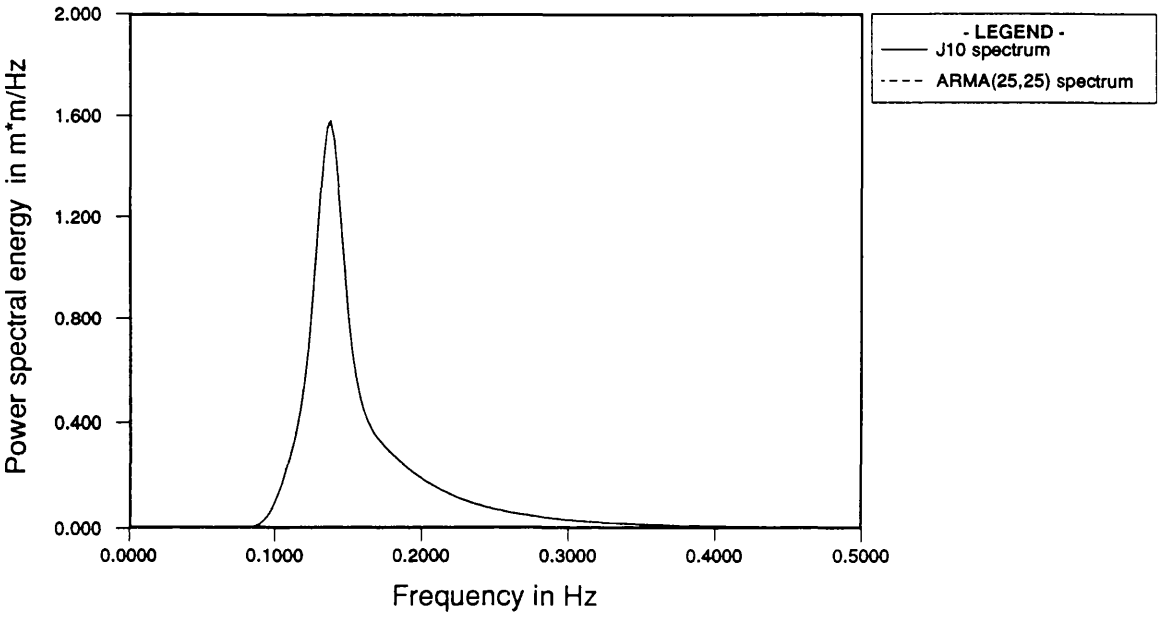

Figure3.10 Comparison between JONSWAP (u=10m/s, γ=3.3) spectrum and ARMA(25,25) spectrum.

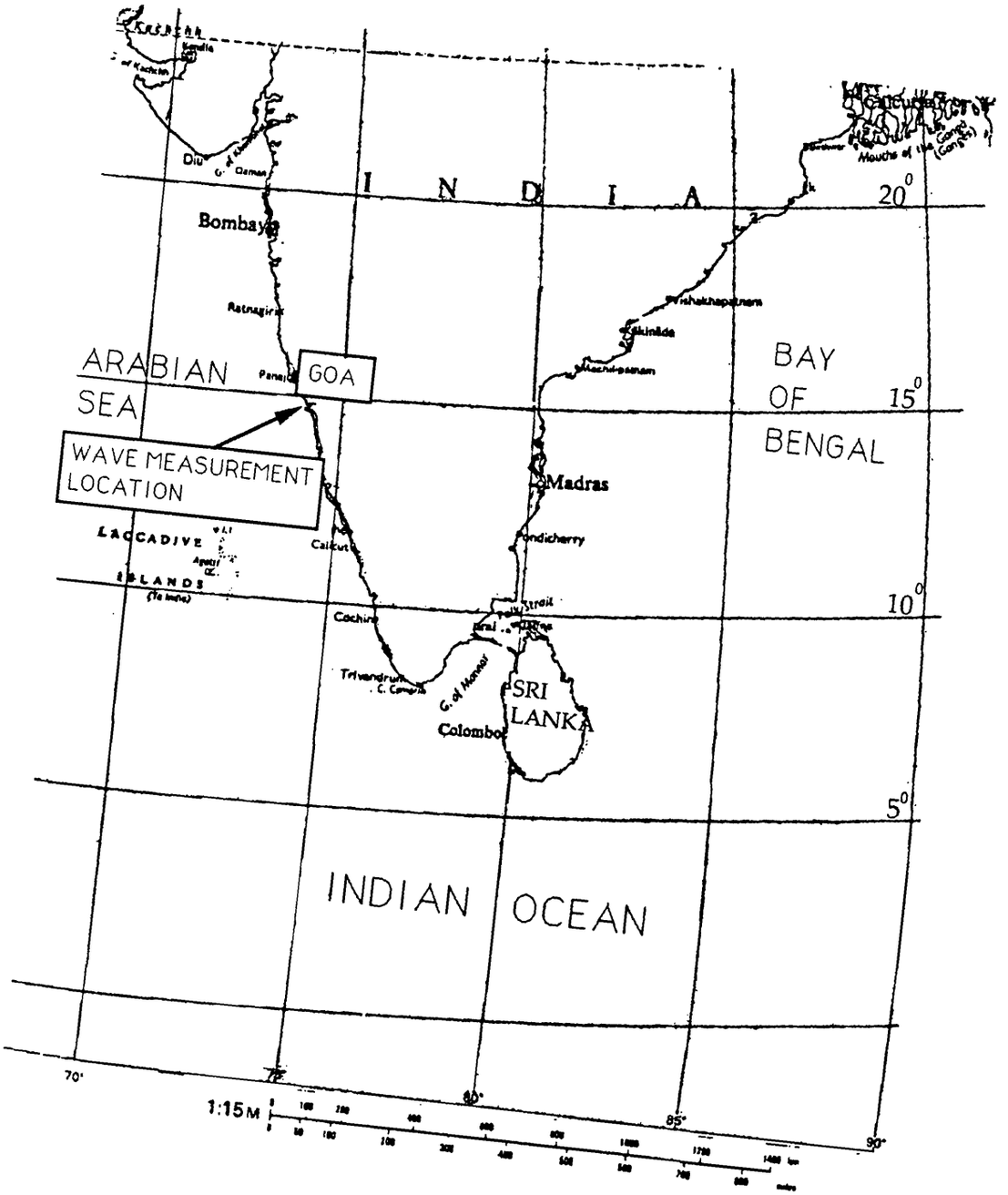

Figure 3.11 **Wave measurement location off West Coast of India.**

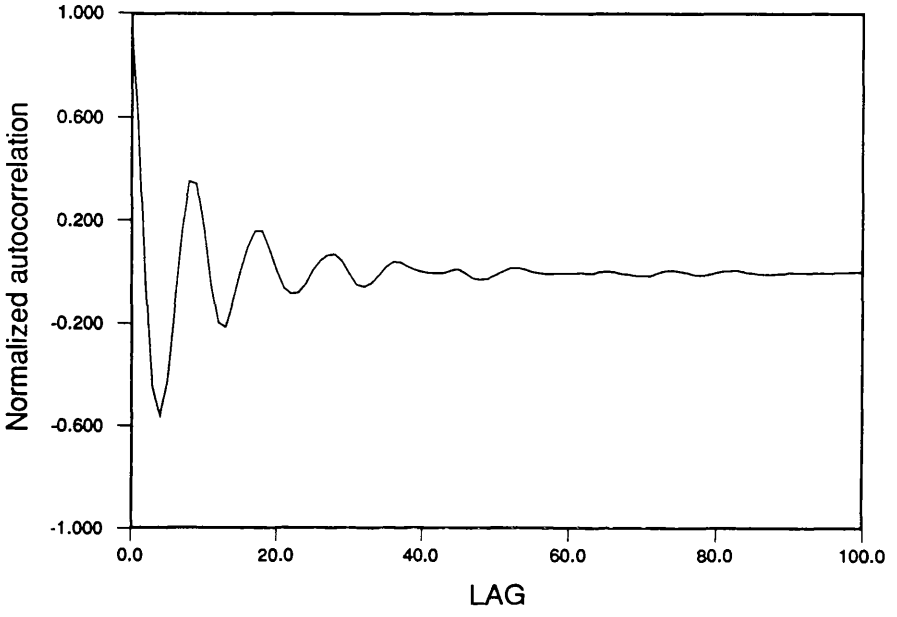

Figure 3.12 Normalized autocorrelation function of the measured time series sea waves: K15.

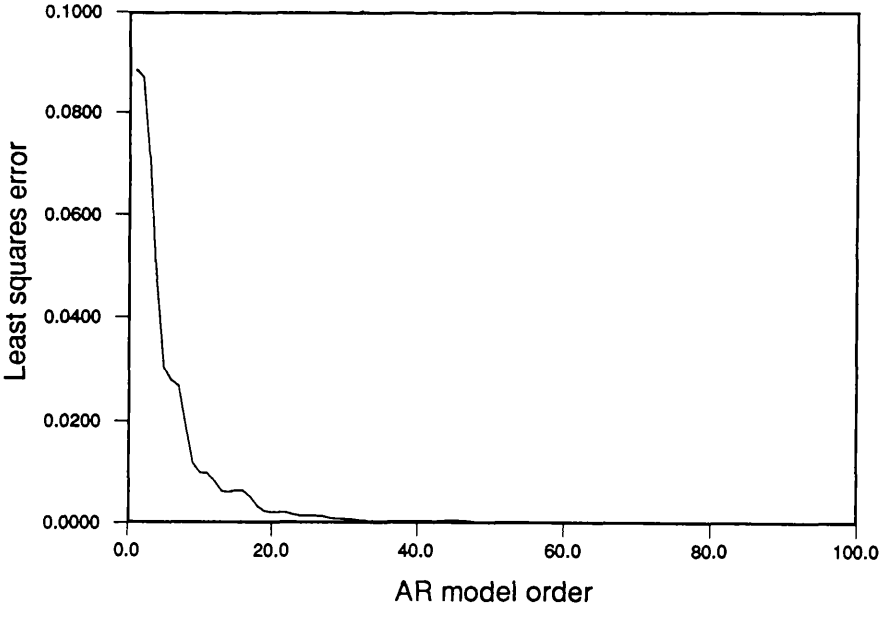

Figure 3.13 Least squares error variation for K15.

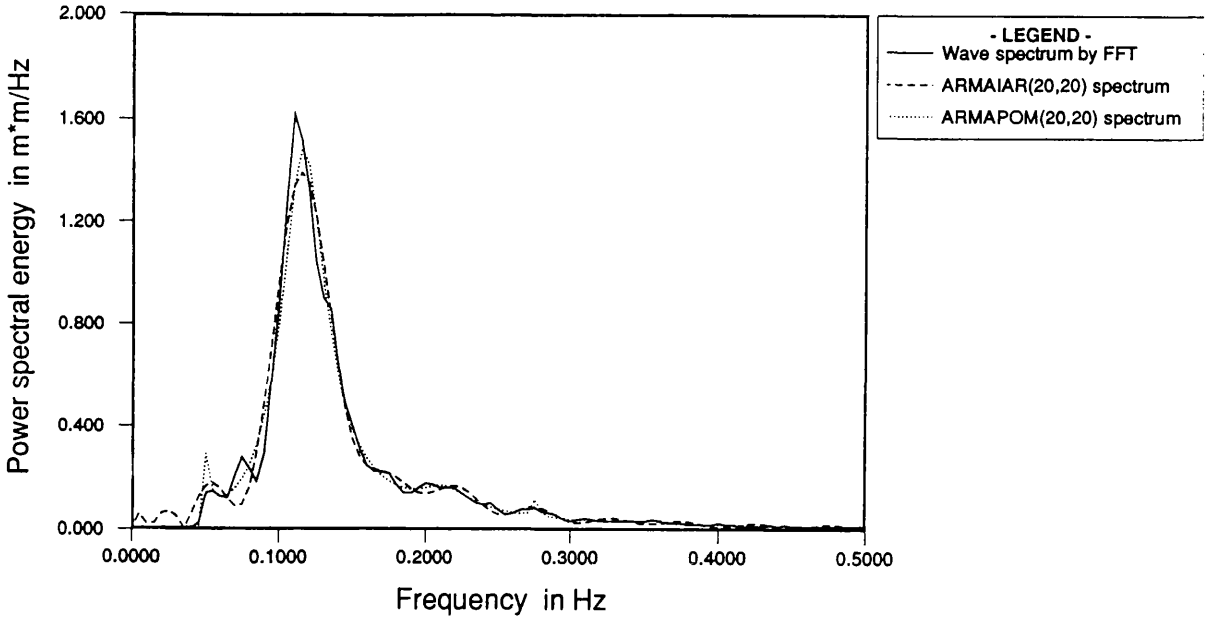

Figure 3.14 Comparison between FFT wave (K15) spectrum and ARMA(20,20) spectra estimated from AR(40) using IAR and POM methods.

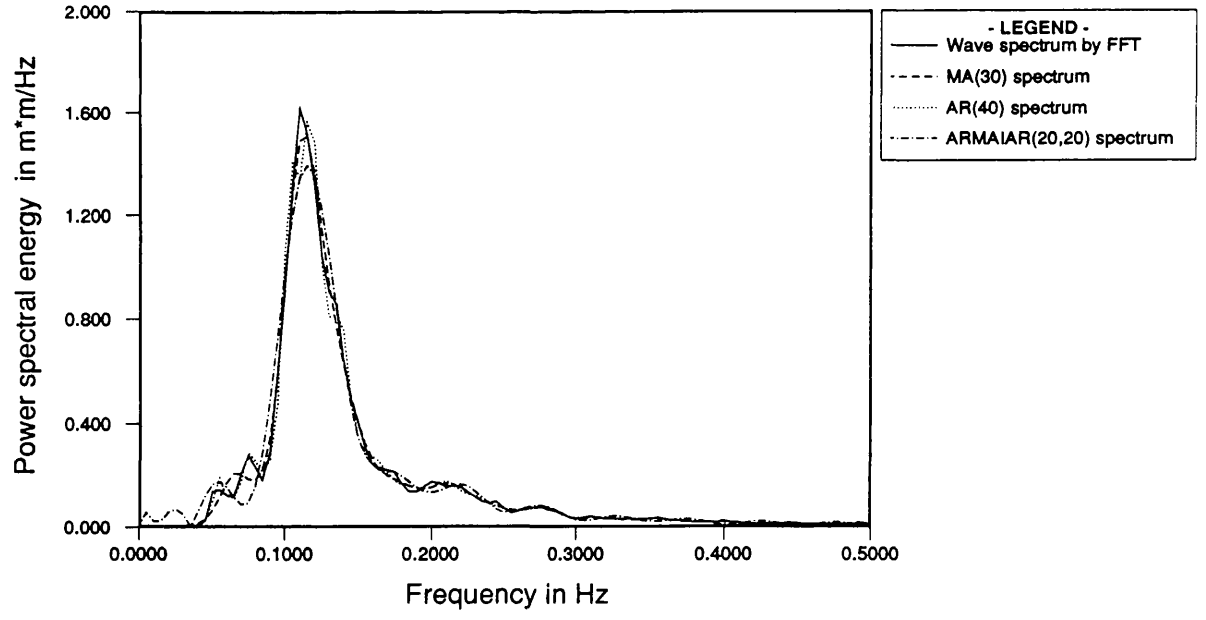

Figure 3.15 Comparison between FFT wave (K15) spectrum and MA(30), AR(40) and ARMAIAR(20,20) spectra.

Figure 3.16 Semisubmersible Santa Fe Rig135 in North Sea.

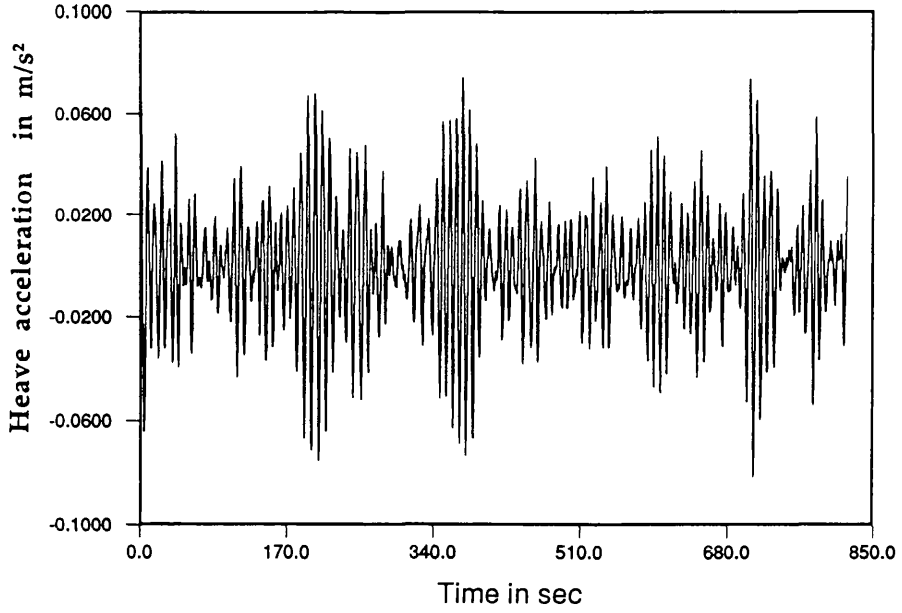

Figure 3.17a Time series heave acceleration (SR=2.5Hz) of the semisubmersible Rig135 (SF28).

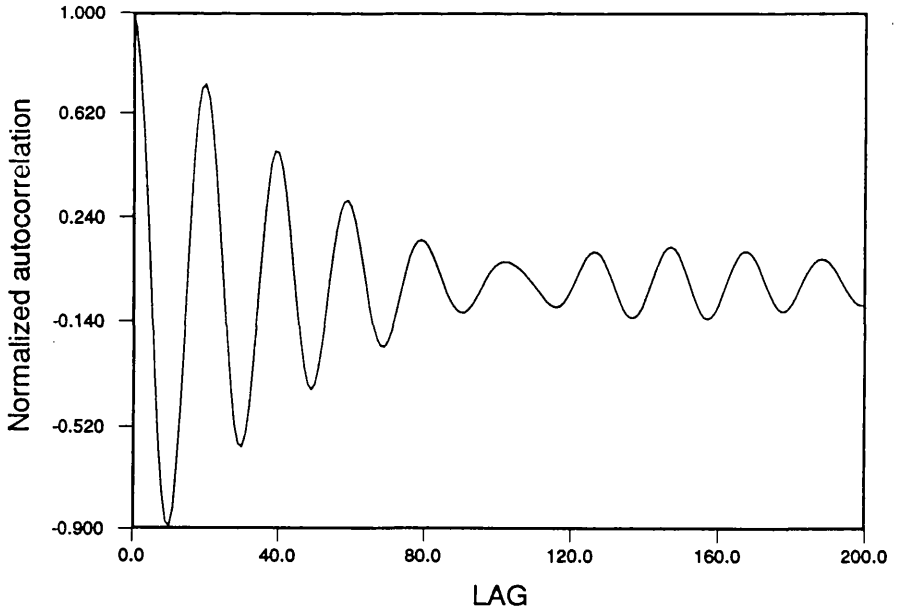

Figure 3.17b Normalized autocorrelation function of the measured time series semisubmersible heave acceleration (SF28).

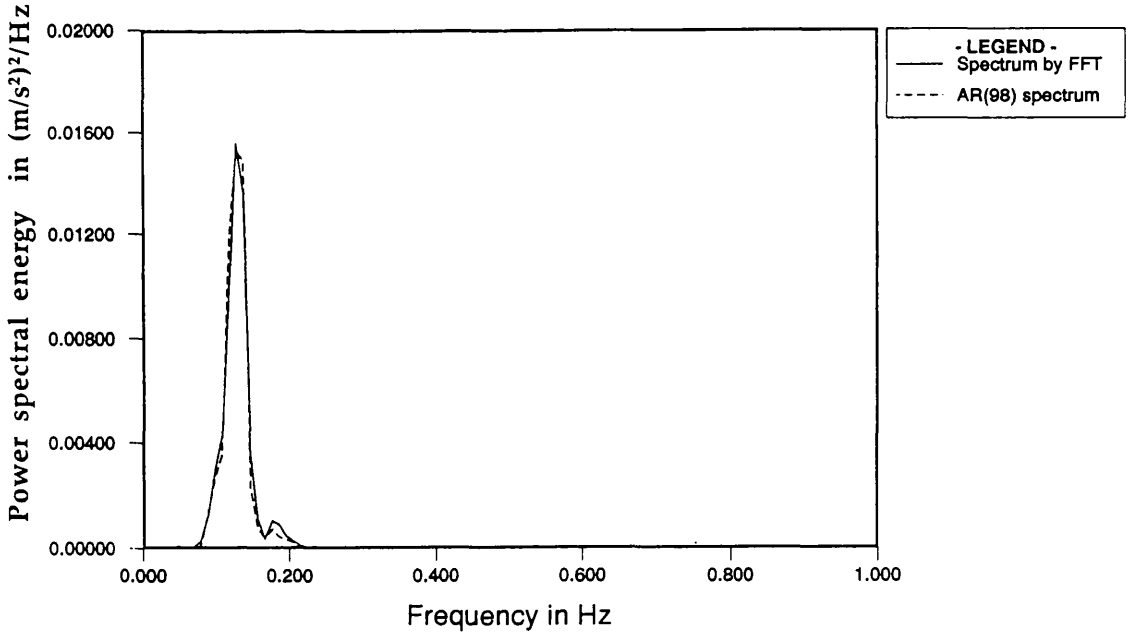

Figure 3.18 Semisubmersible heave acceleration (SF28) spectra calculated by FFT and AR(98) model.

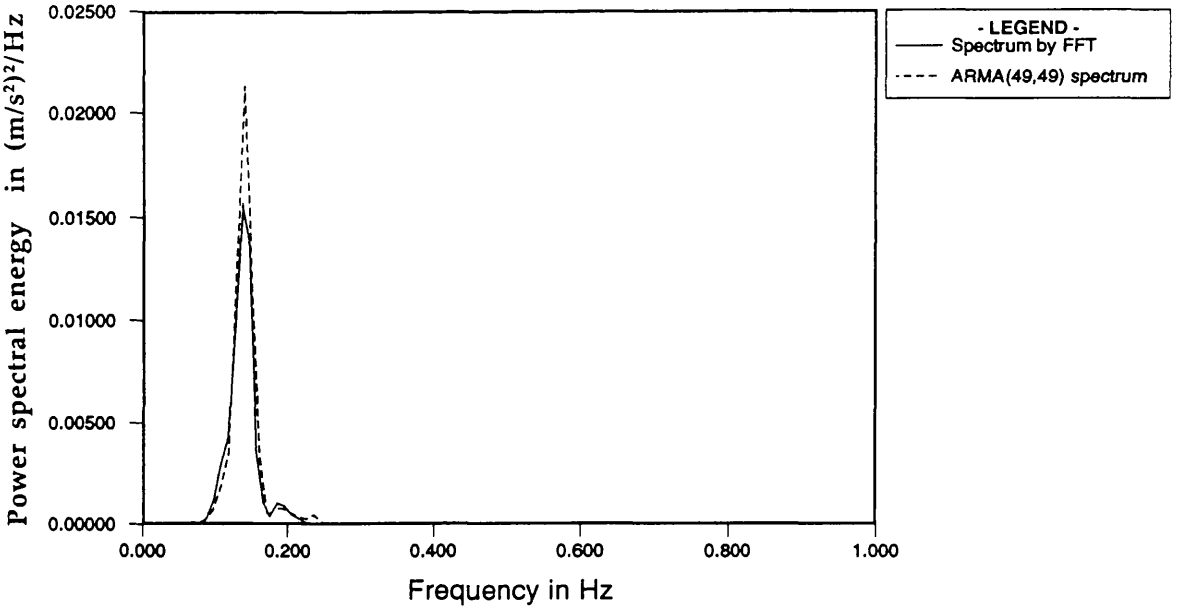

Figure 3.19 Semisubmersible heave acceleration (SF28) spectra calculated by FFT and ARMA(49,49) model.

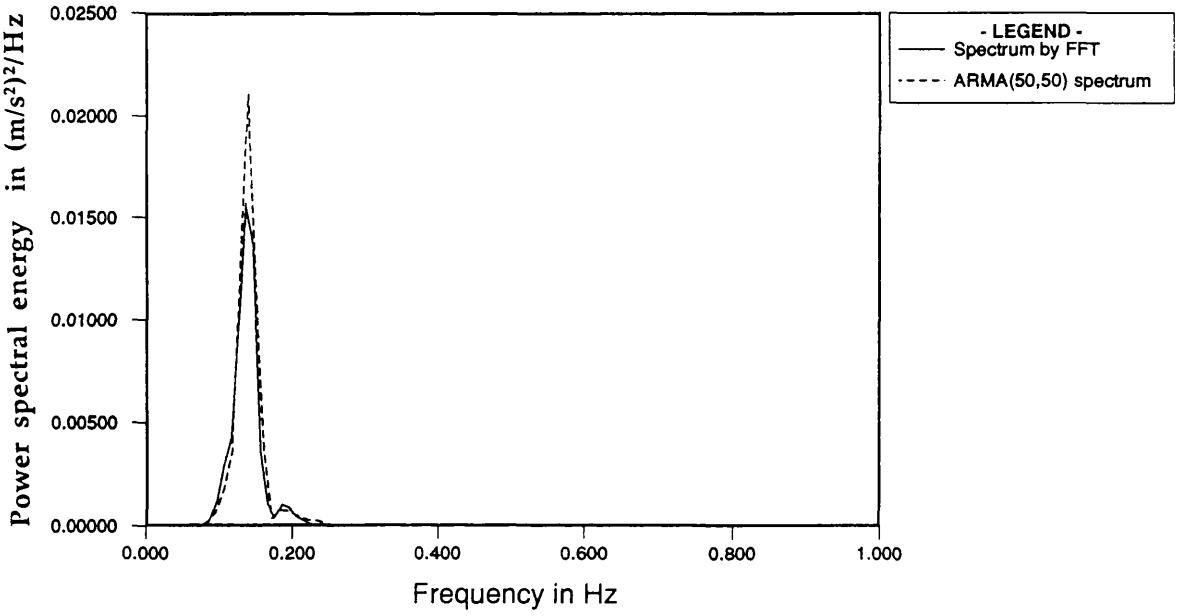

Figure 3.20 Semisubmersible heave acceleration (SF28) spectra calculated by FFT and ARMA(50,50) model.

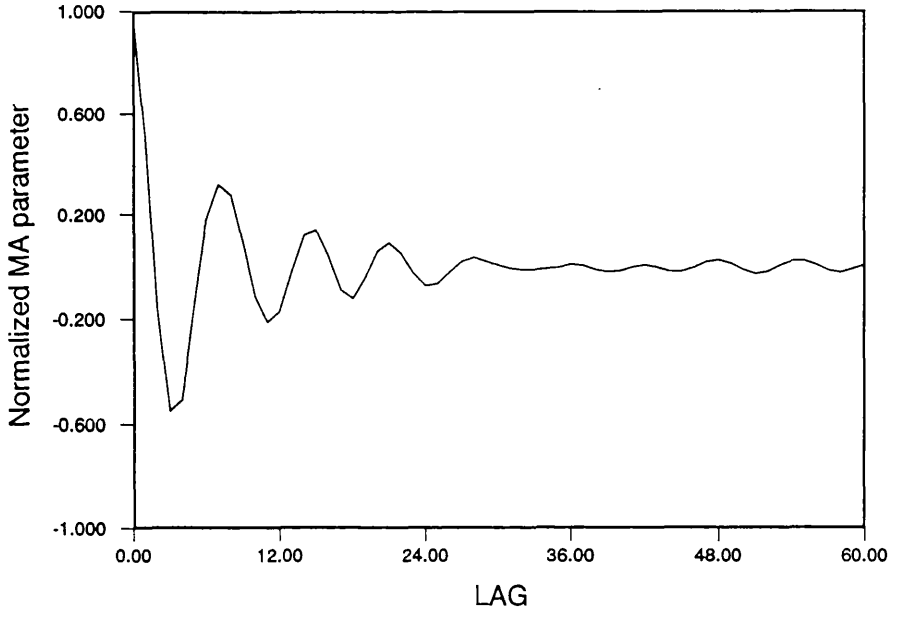

Figure 3.21 Normalized variation of the MA coefficients [equation (3.13)] of heave acceleration (SF28).

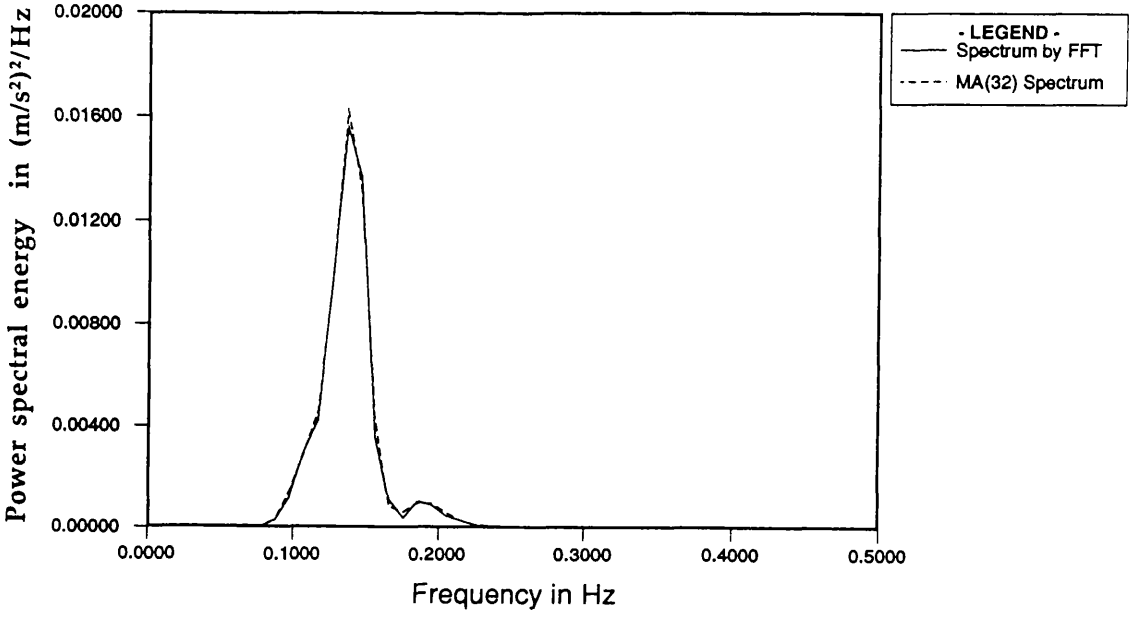

Figure 3.22 Heave acceleration (SF28) spectra calculated by FFT and MA(32) model.

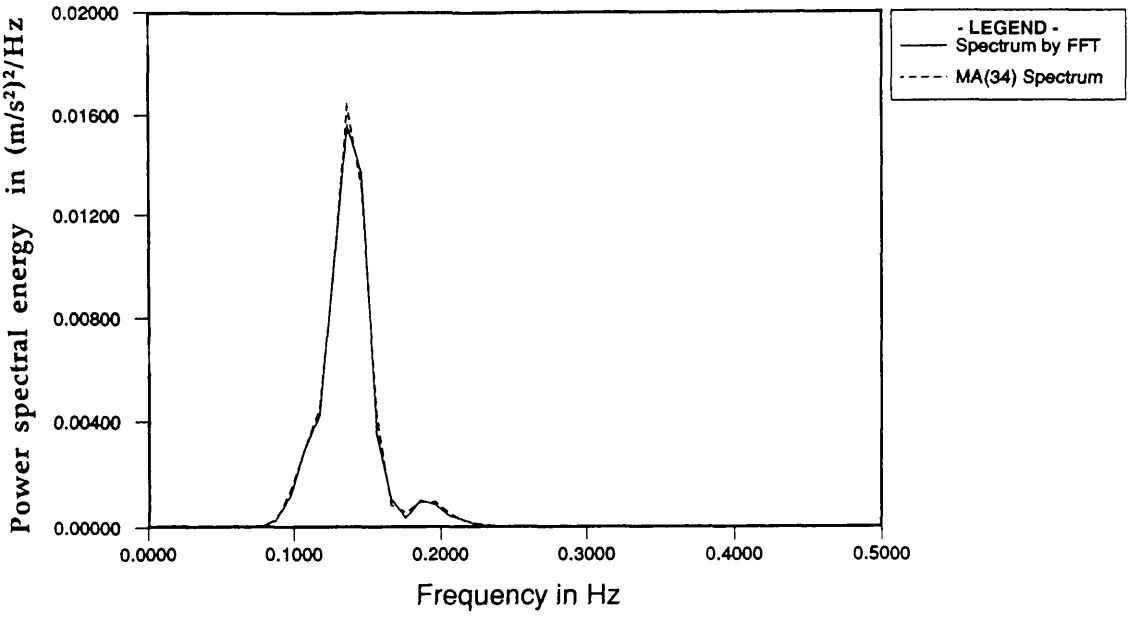

Figure 3.23 Heave acceleration (SF28) spectra calculated by FFT and MA(34) model.

Figure 3.24 Monohull crane vessel-A (McDermott, DB50) in North Sea.

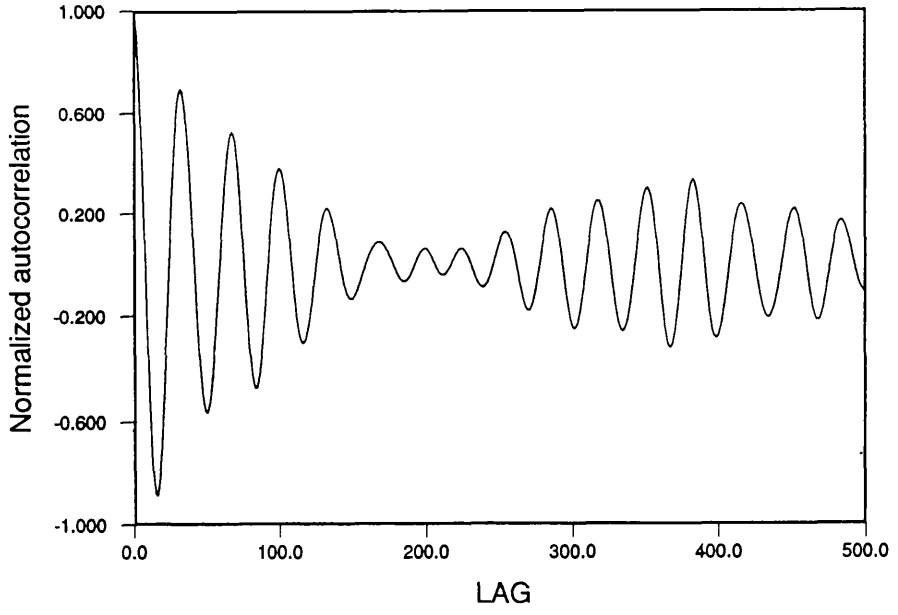

Figure 3.25 Normalized autocorrelation of monohull crane vessel-A roll.

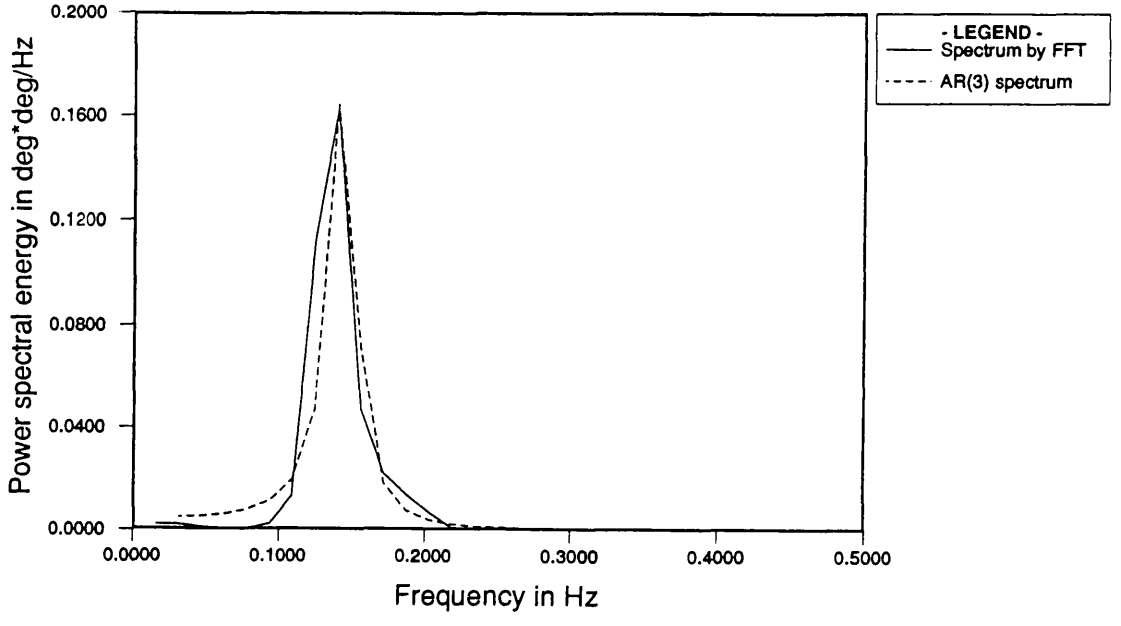

Figure 3.26 Comparison between FFT monohull crane vessel-A roll spectrum and AR(3) spectrum.

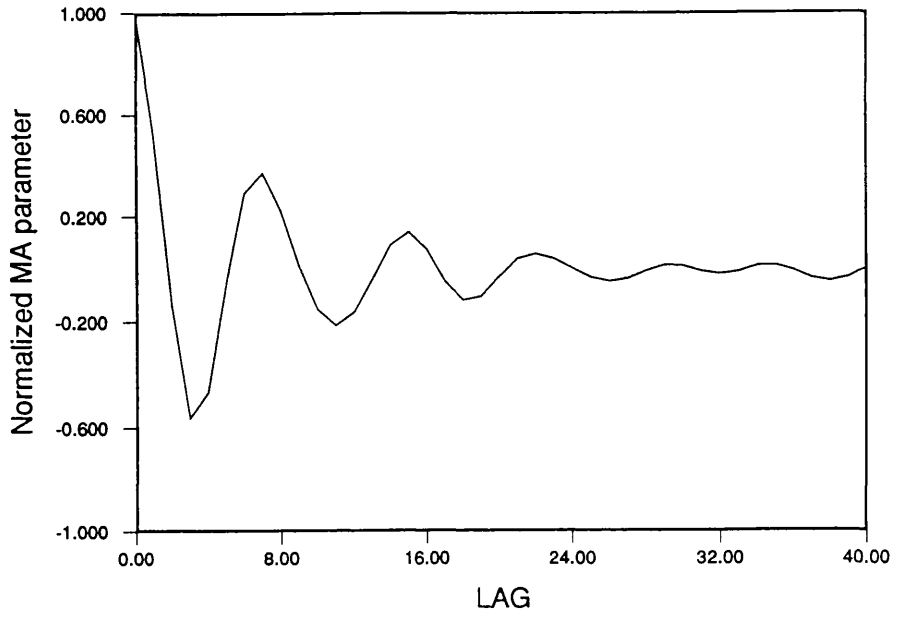

Figure 3.27 Normalized variation of the MA coefficients [equation (3.13)] of monohull crane vessel-A roll.

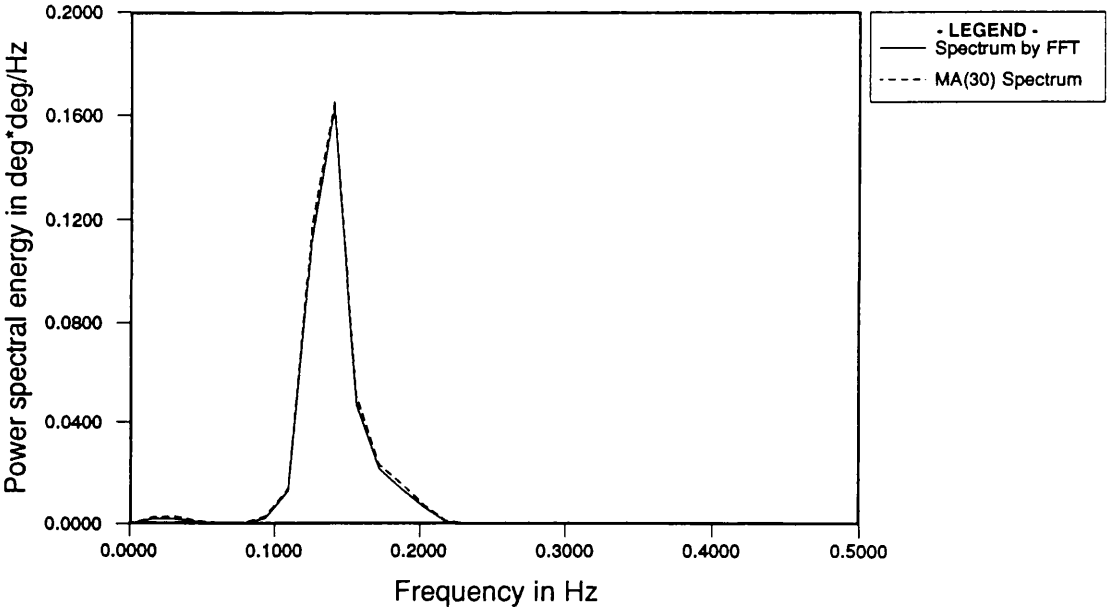

Figure 3.28 Comparison between FFT monohull crane vessel-A roll spectrum and MA(30) spectrum.

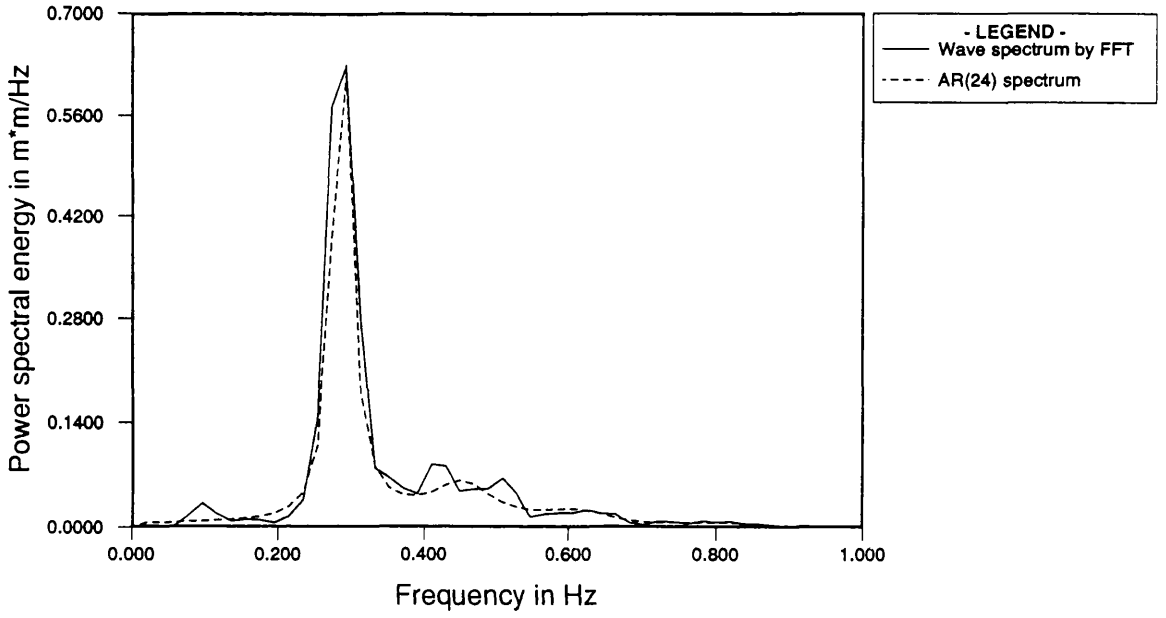

Figure 3.29a Comparison between FFT wave (near crane vessel-B) spectrum and AR(24) spectrum.

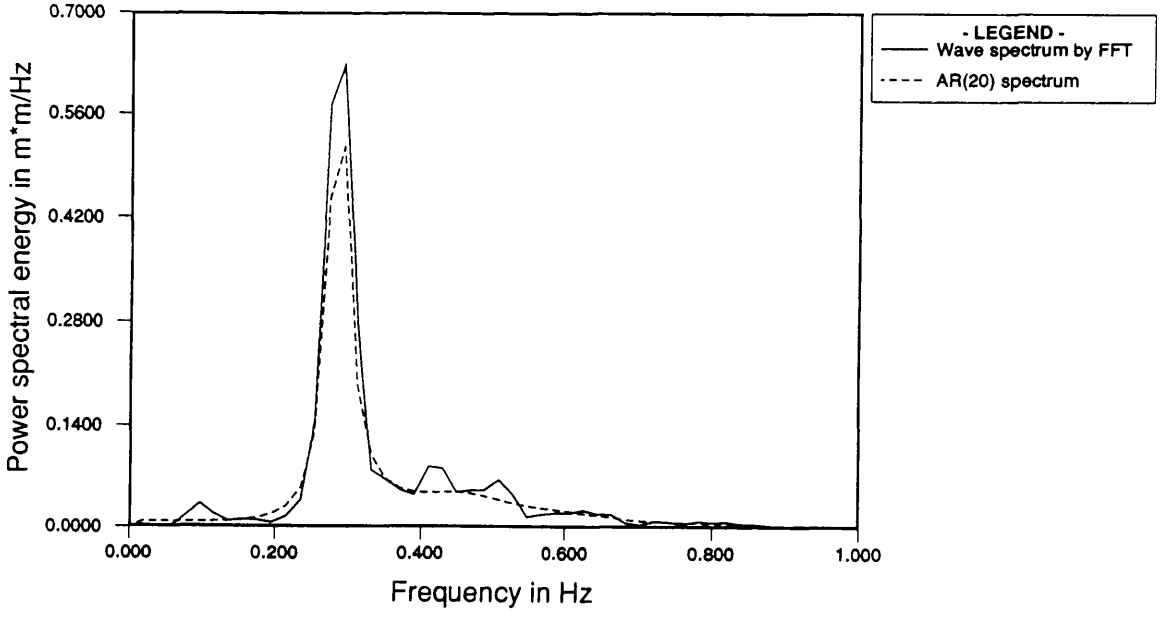

Figure 3.29b Comparison between FFT wave (near crane vessel-B) spectrum and AR(20) spectrum.

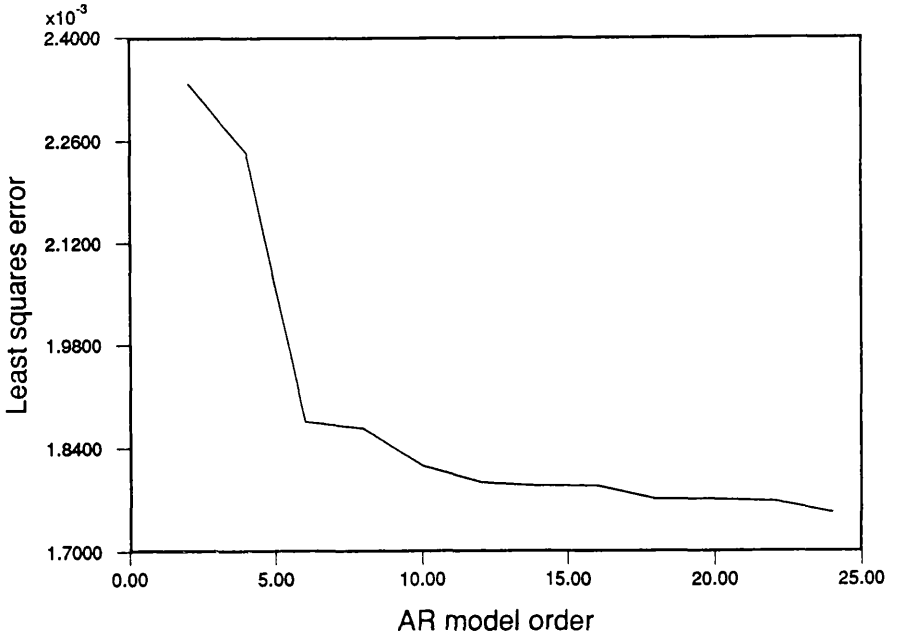

Figure 3.29c Least squares error variation for measured waves near crane vessel-B.

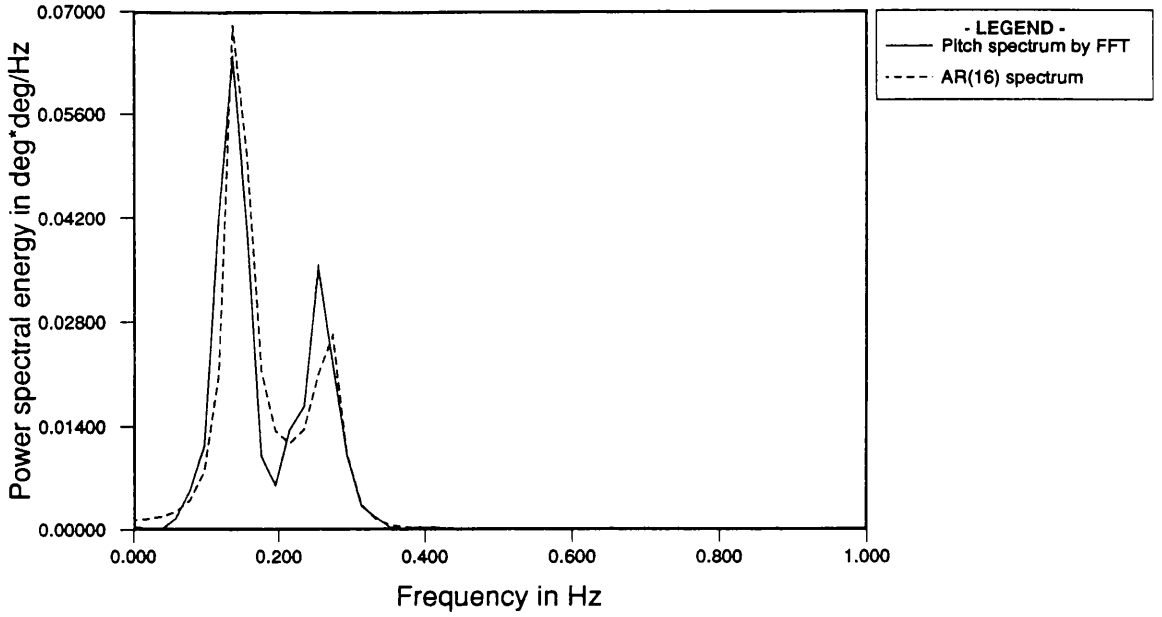

Figure 3.30a Pitch (vessel-B) spectral estimates by FFT and AR(16) model.

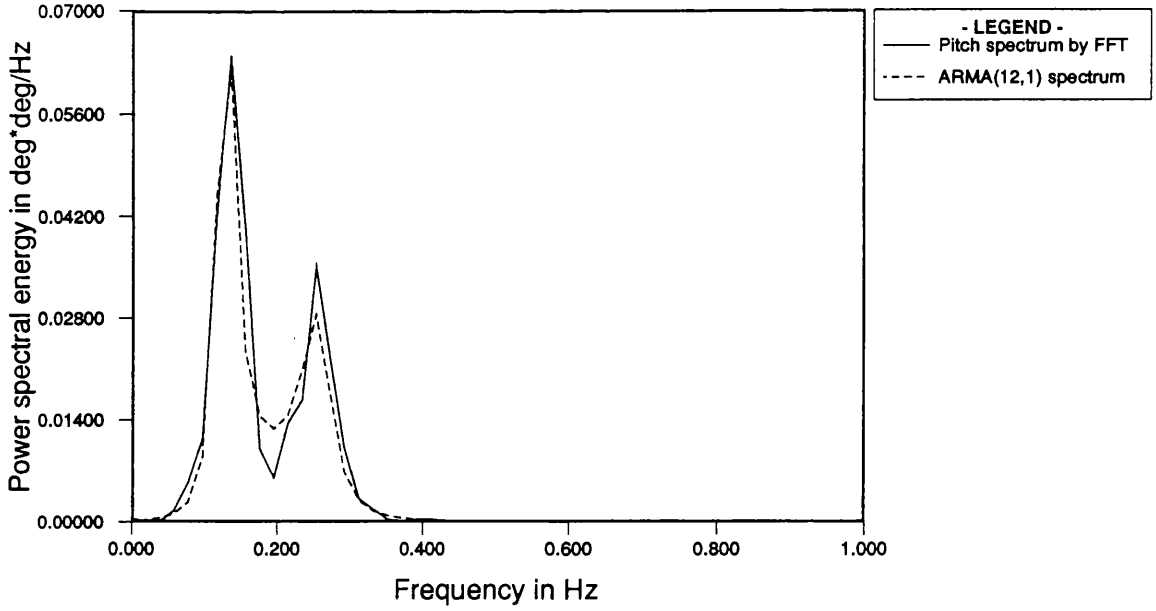

Figure 3.30b Pitch (vessel-B) spectral estimates by FFT and ARMA(12,1) model.

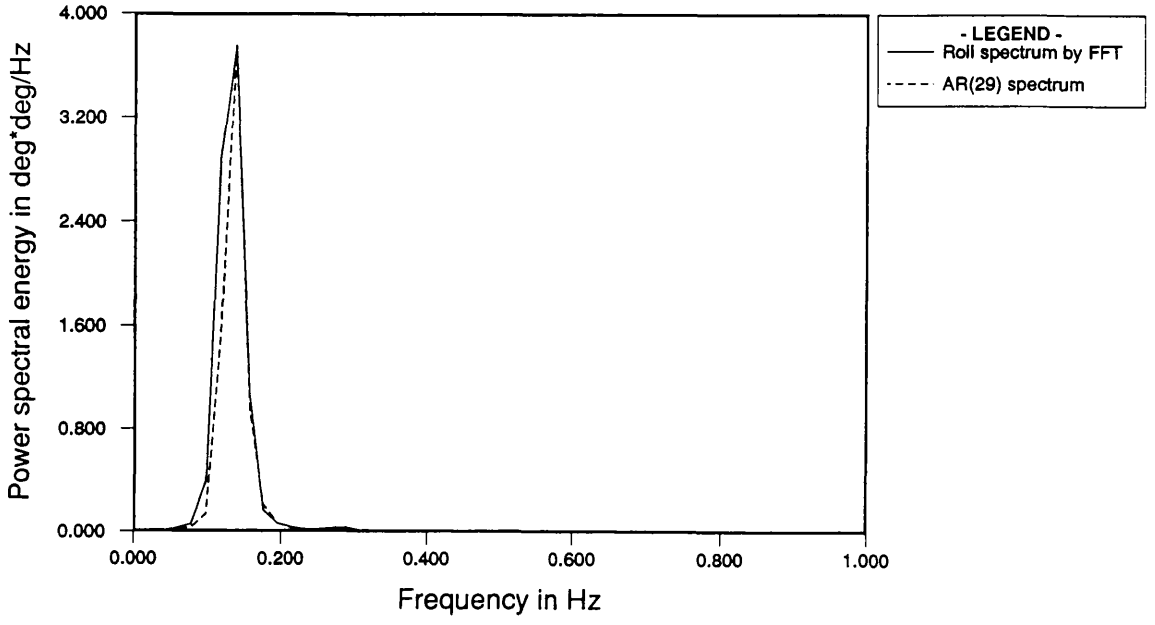

Figure 3.31 Roll (vessel-B) spectral estimates by FFT and AR(29) model.

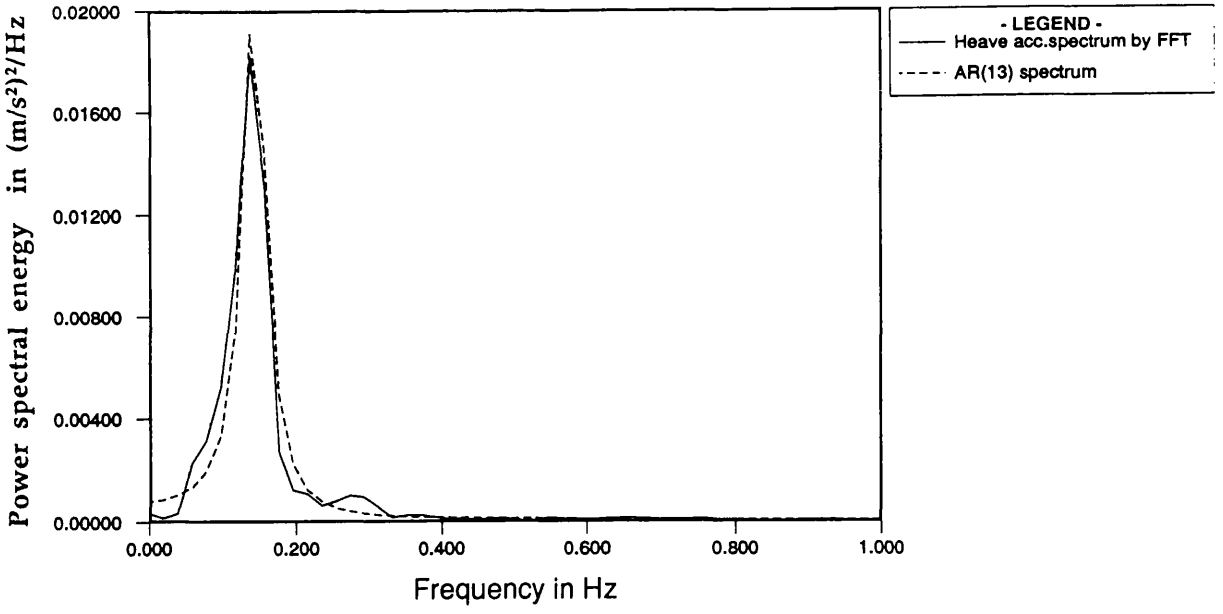

Figure 3.32a Heave acceleration (vessel-B) spectral estimates by FFT and AR(13) model.

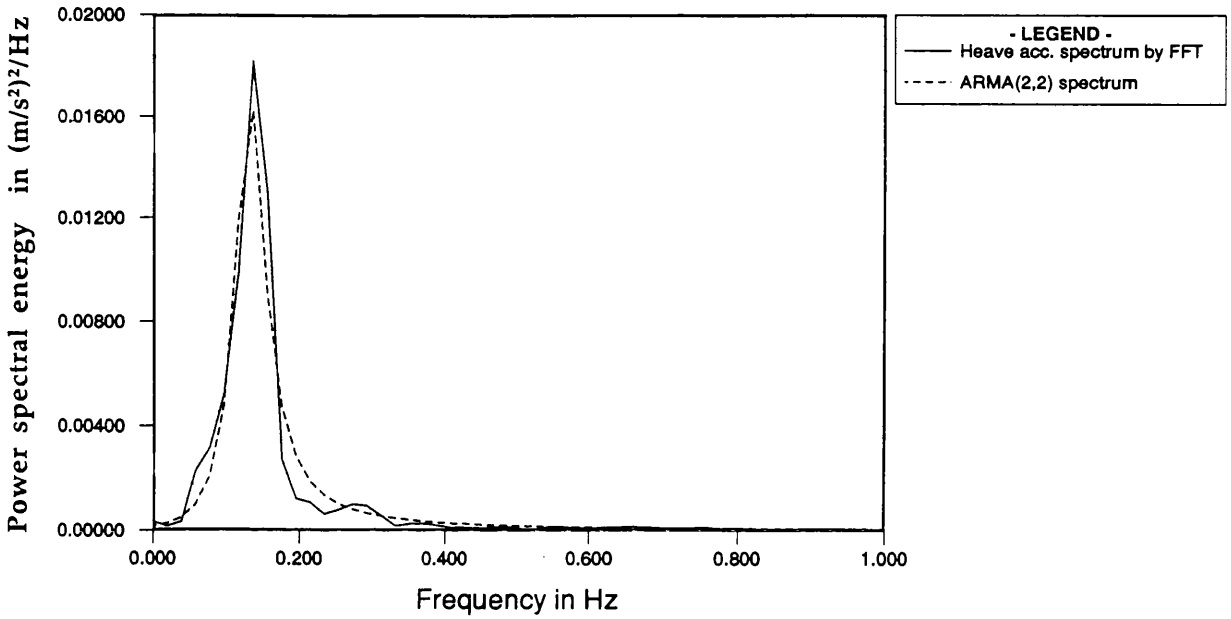

Figure 3.32b Heave acceleration (vessel-B) spectral estimates by FFT and ARMA(2,2) model.

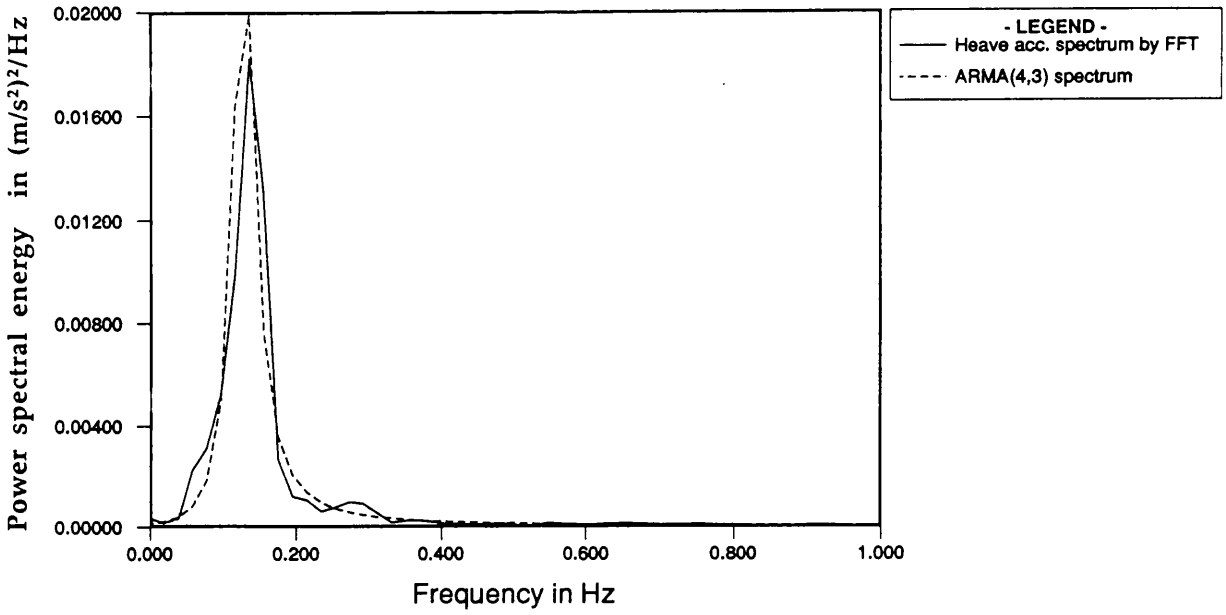

Figure 3.32c Heave acceleration (vessel-B) spectral estimates by FFT and ARMA(4,3) model.

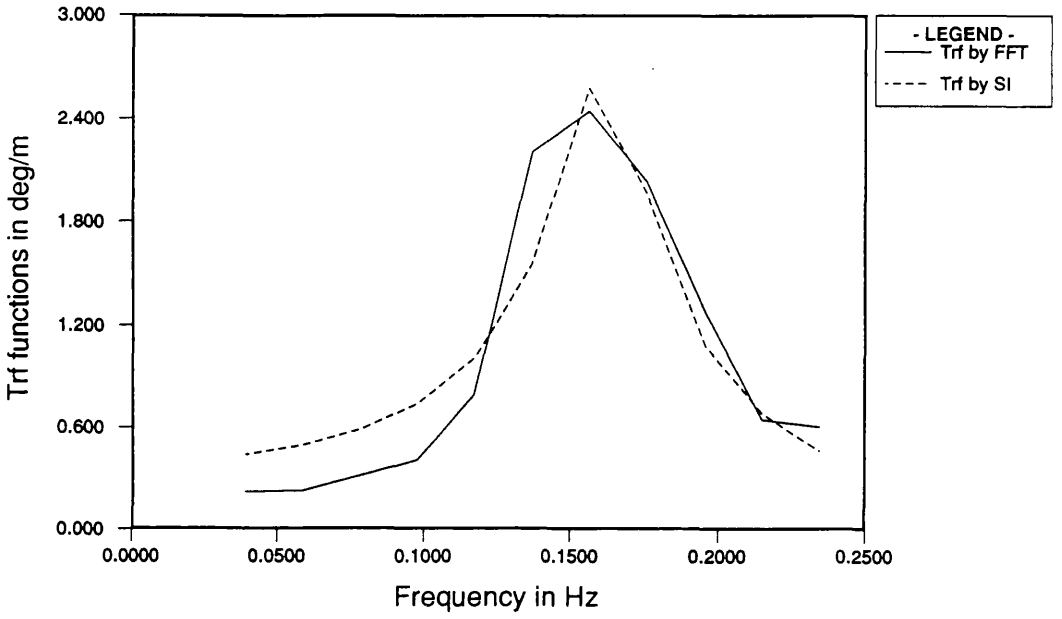

Figure 3.33a Pitch (vessel-B) transfer functions estimated by FFTand SI models with 36 parameters [from Figures 3.30a and 3.29b].

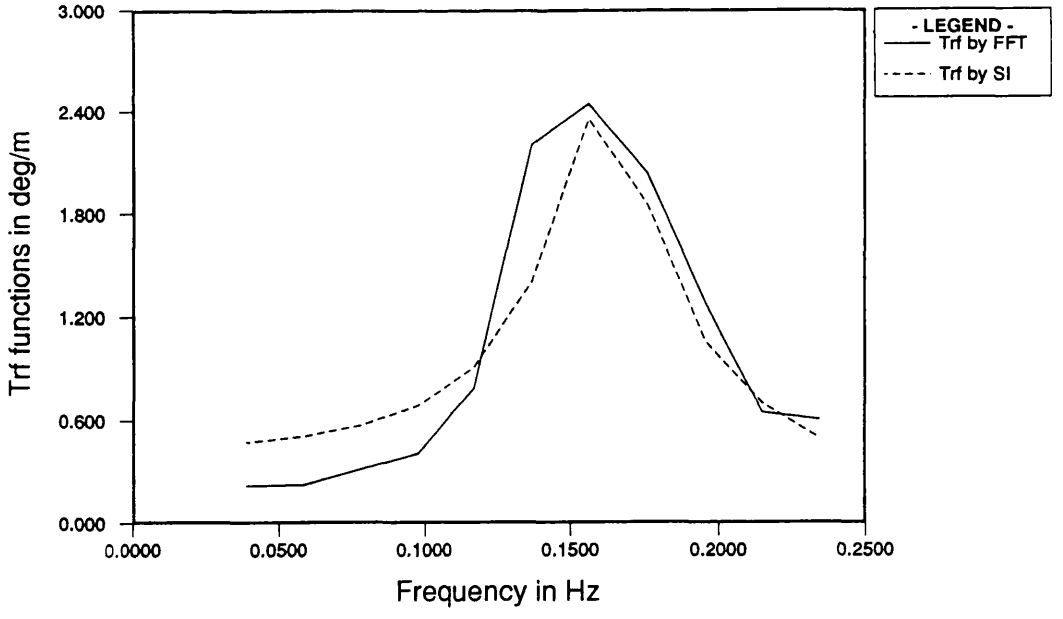

Figure 3.33b Pitch (vessel-B) transfer functions estimated by FFT and SI models with 40 parameters [from Figures 3.30a and 3.29a].

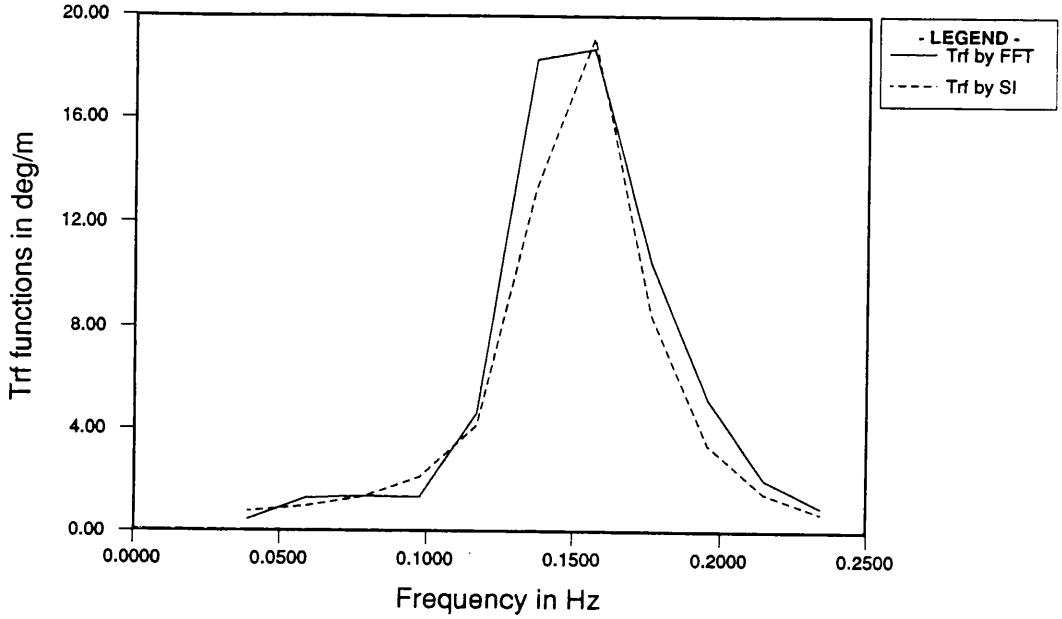

Figure 3.34a Roll (vessel-B) transfer functions estimated by FFT and SI models with 49 parameters [from Figures 3.31 and 3.29b].

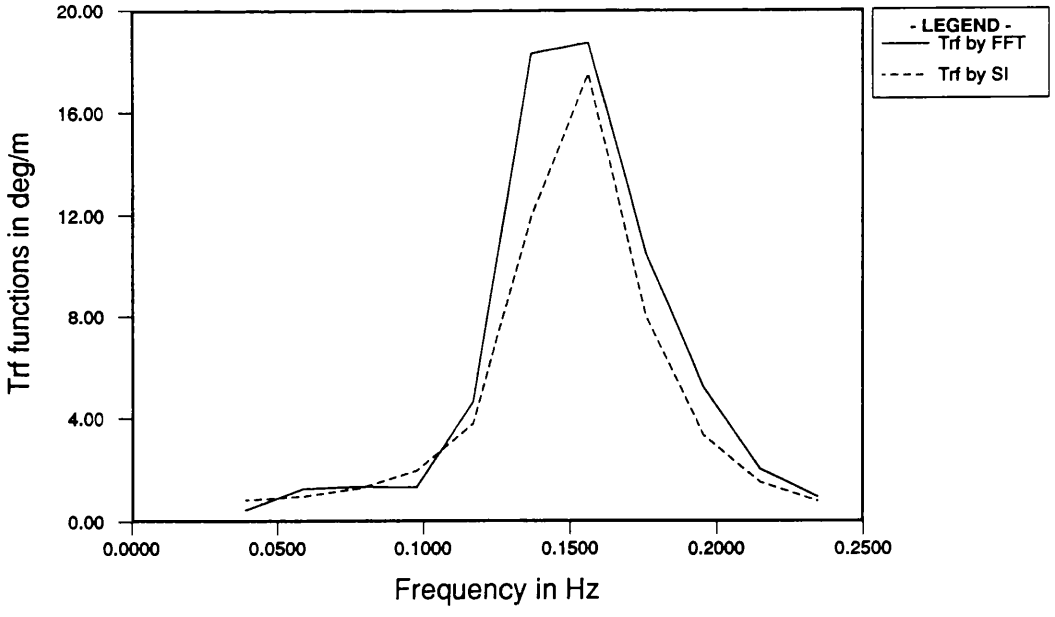

Figure 3.34b Roll (vessel-B) transfer functions estimated by FFT and SI models with 53 parameters [from Figures 3.31 and 3.29a].

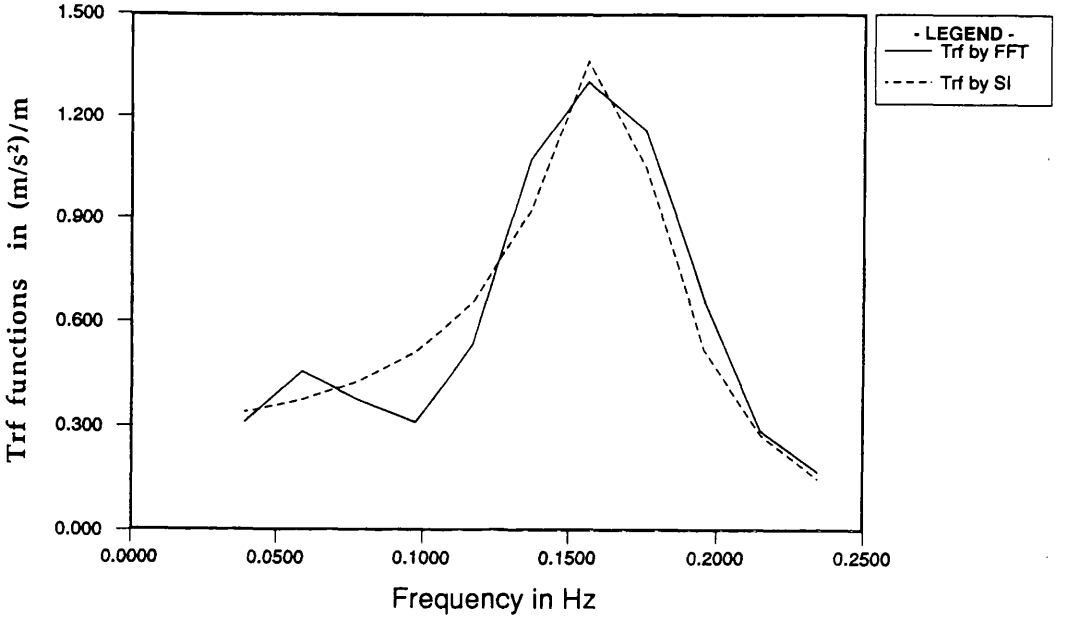

Figure 3.35a Heave acceleration (vessel-B) transfer functions estimated by FFT and SI models with 33 parameters [from Figures 3.32a and 3.29b].

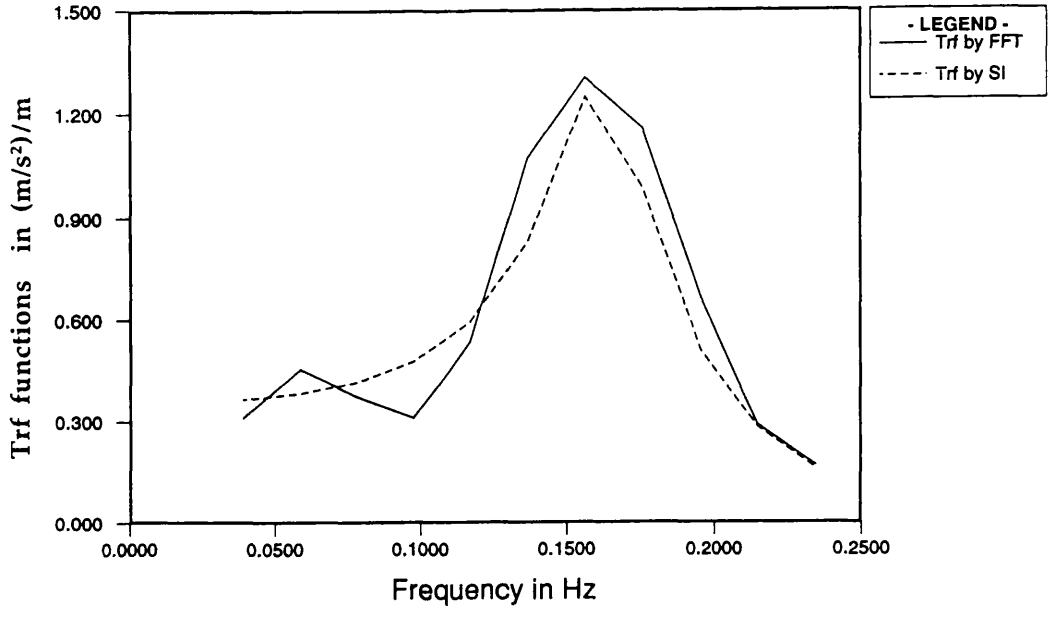

Figure 3.35b Heave acceleration (vessel-B) transfer functions estimated by FFT and SI models with 37 parameters [from Figures 3.32a and 3.29a].

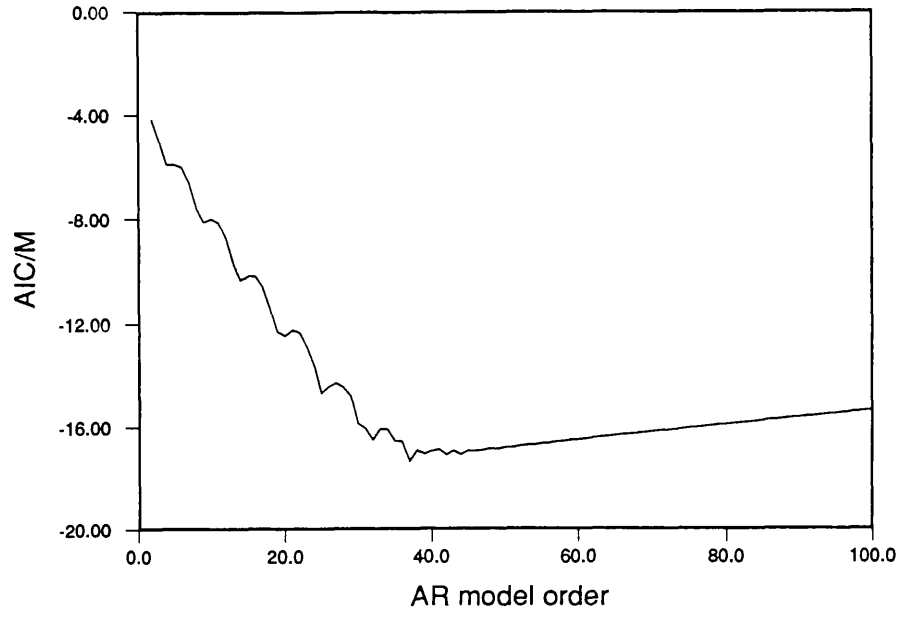

Figure 4.1 Variation of Akaike information criterion (AIC) for PM10 spectrum.

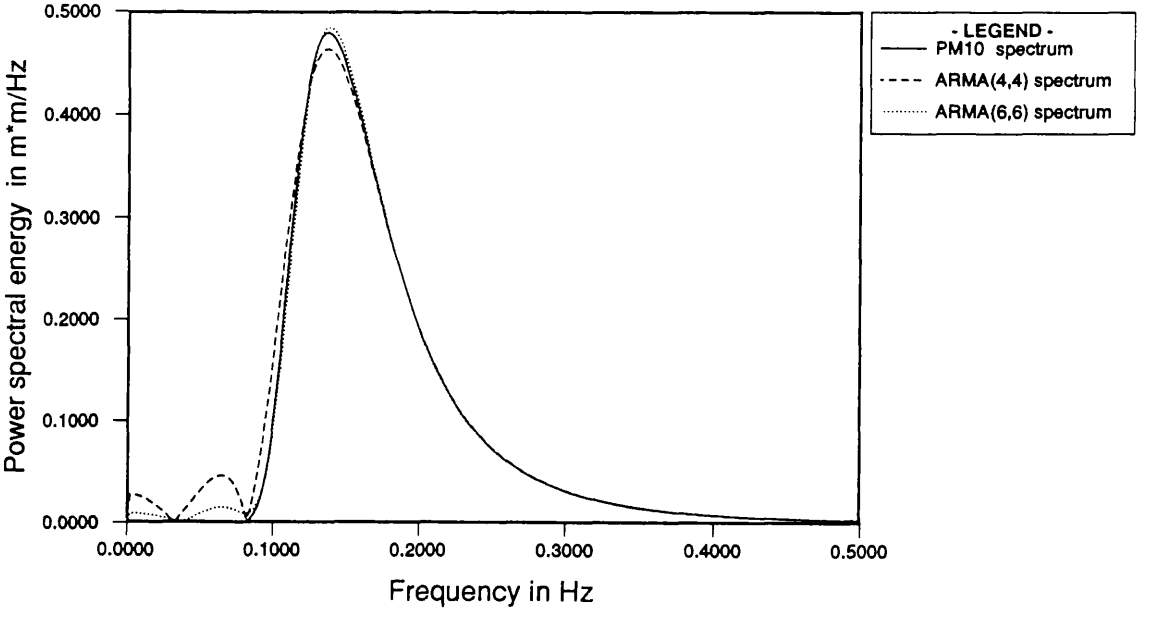

Figure 4.2a Comparison between PM10 spectrum and ARMA(4,4) and ARMA(6,6) spectra estimated from AR(40) using MYW equations and reduction method.

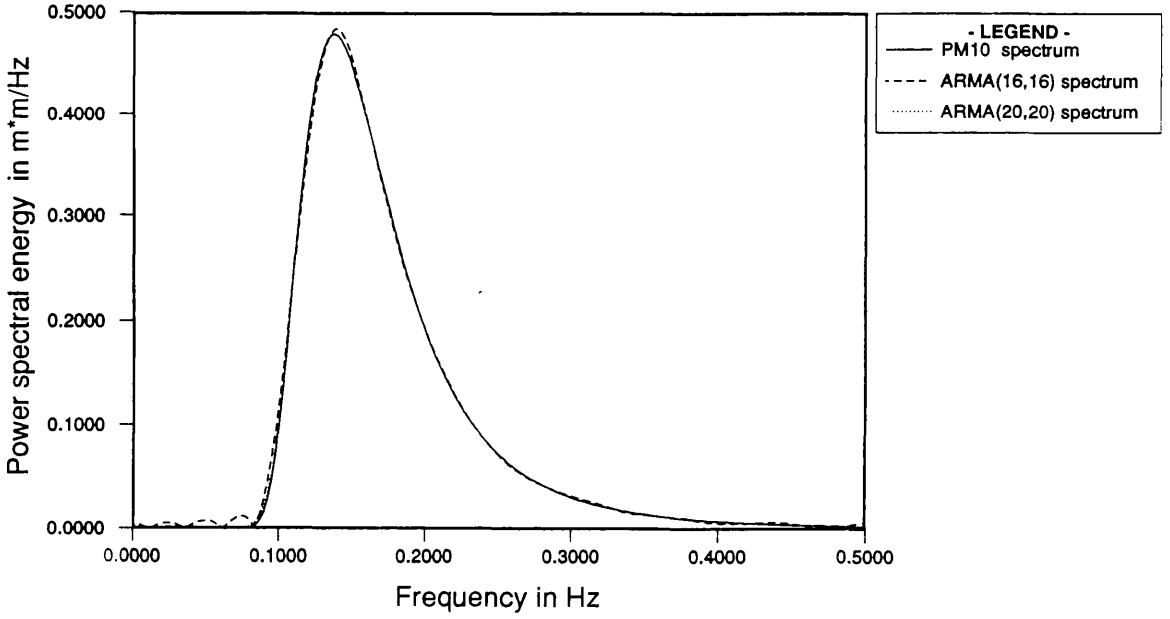

Figure 4.2b Comparison between PM10 spectrum and ARMA(16,16) and ARMA(20,20) spectra using POM method.

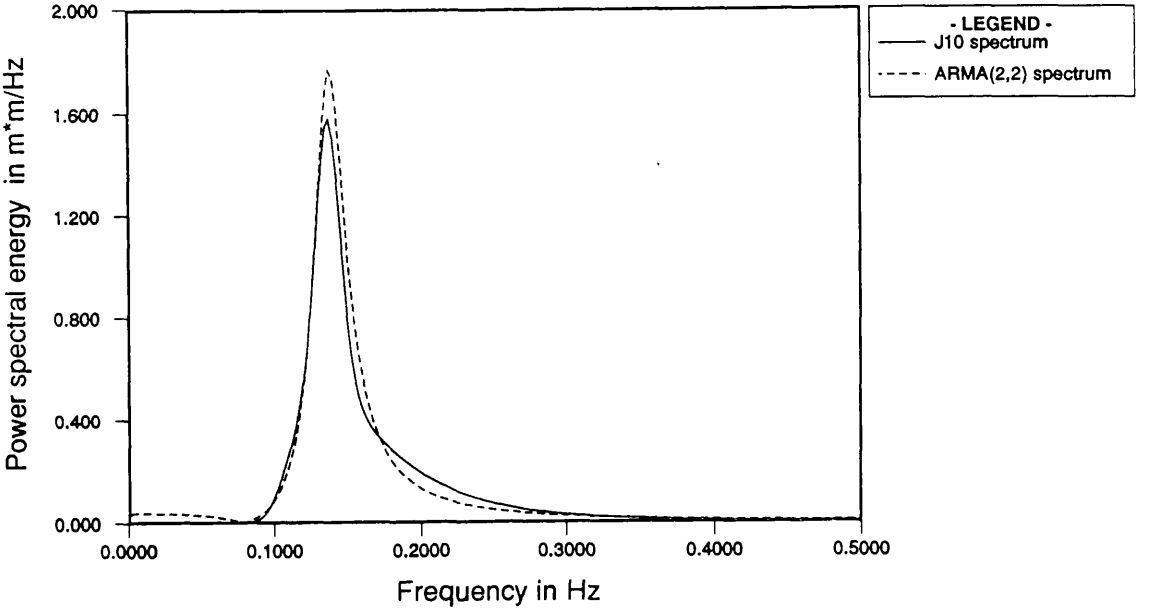

Figure 4.3a Comparison between JONSWAP (u=10m/s, γ=3.3) spectrum and ARMA(2,2) spectrum estimated from AR(60) using MYW equations and reduction method.

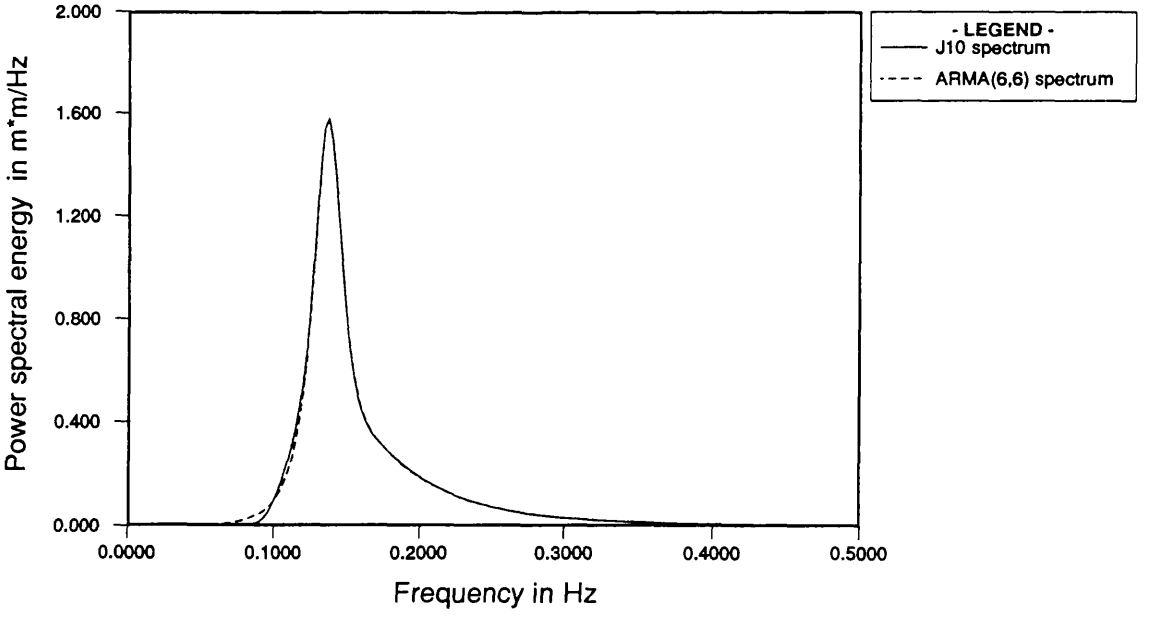

Figure 4.3b Comparison between JONSWAP (u=10m/s, γ=3.3) spectrum and ARMA(6,6) spectrum estimated from AR(60) using MYW equations and reduction method.

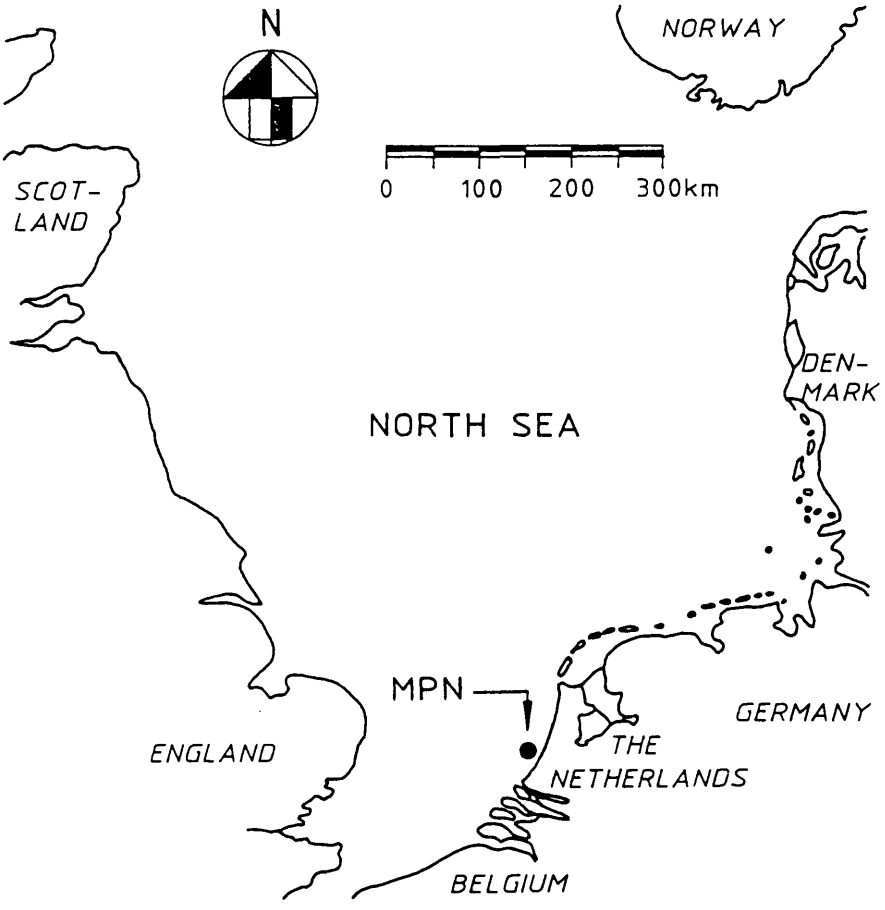

Figure 4.4 MPN platform location in North Sea.

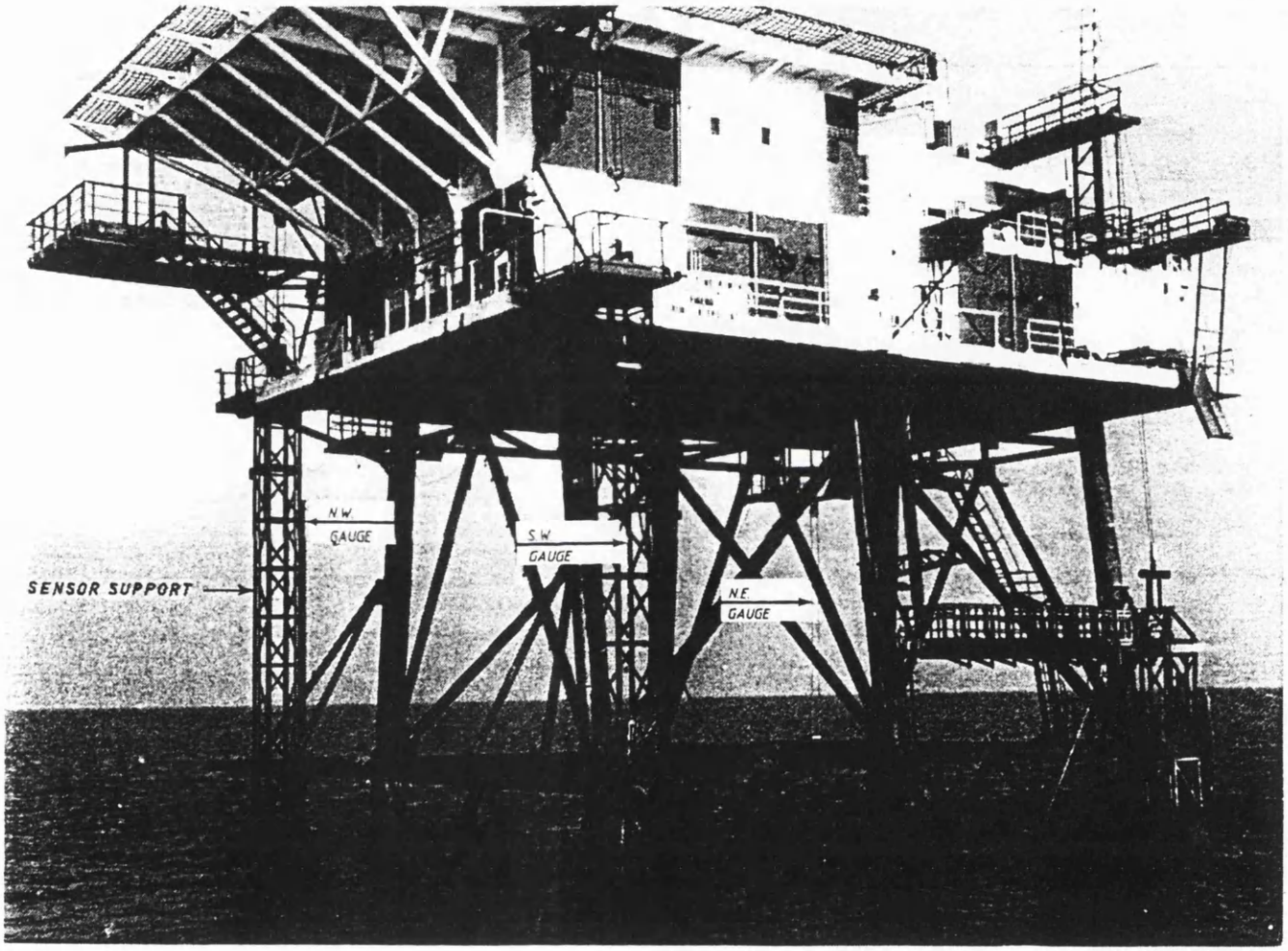

Figure 4.5a MPN platform with location of wave gauge sensors.

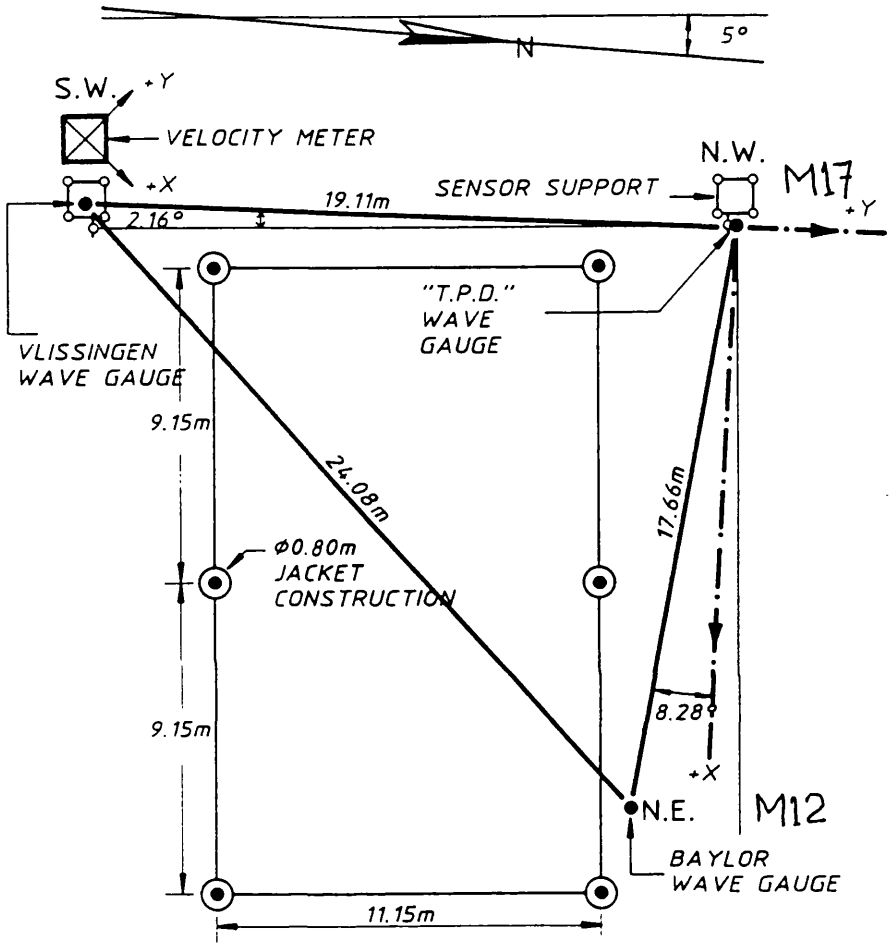

Figure 4.5b Wave gauge sensors positions (M12 and M17).

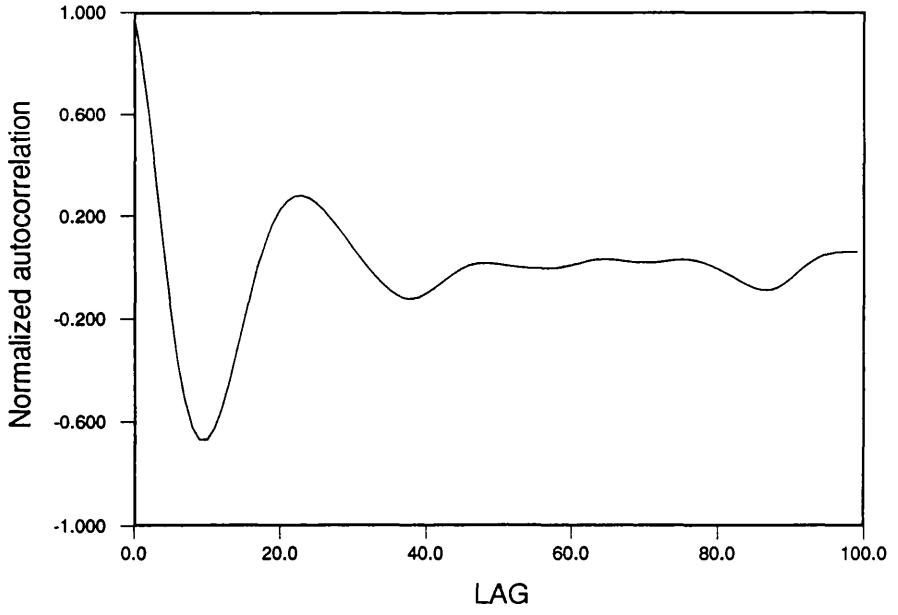

Figure 4.6 Normalized autocorrelation function of the North Sea time series wave data: M12.

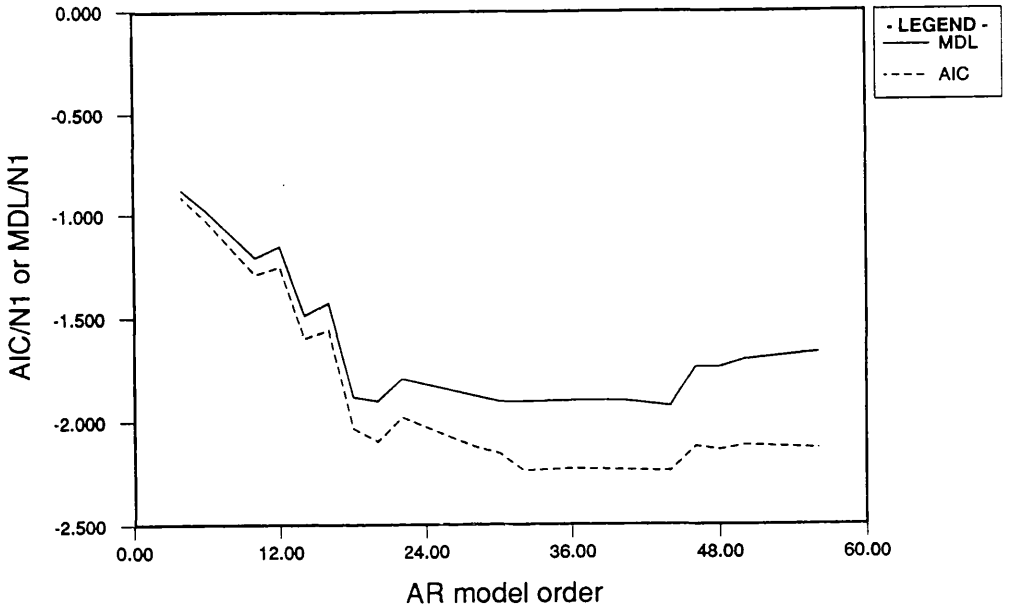

Figure 4.7a Variation of AIC and MDL (ER calculated as average square error between target spectrum and estimated spectrum) for M12.

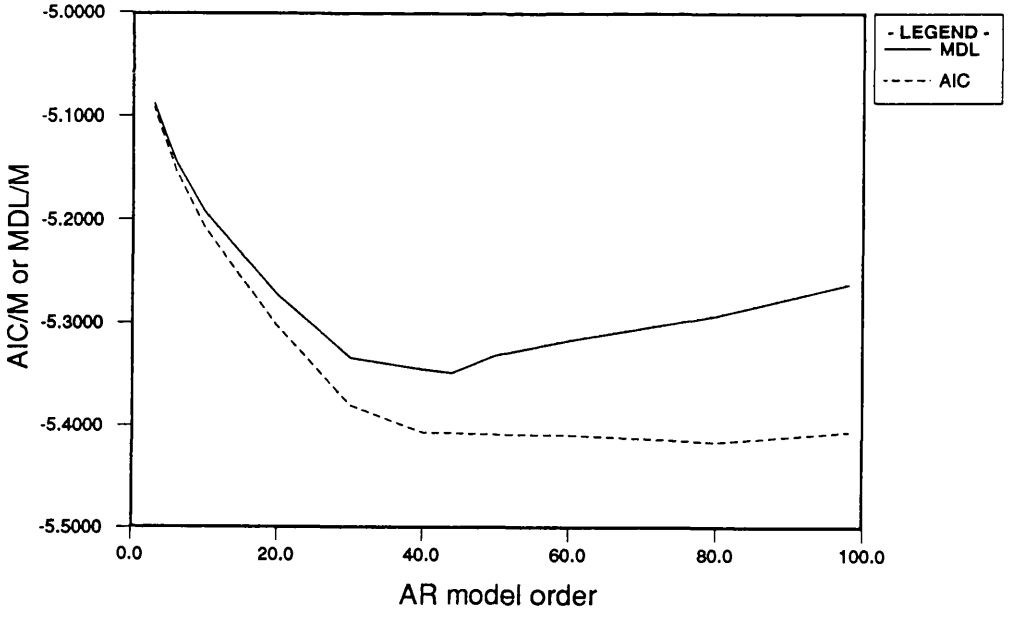

Figure 4.7b Variation of AIC and MDL (ER calculated using forward linear prediction residual technique) for M12.

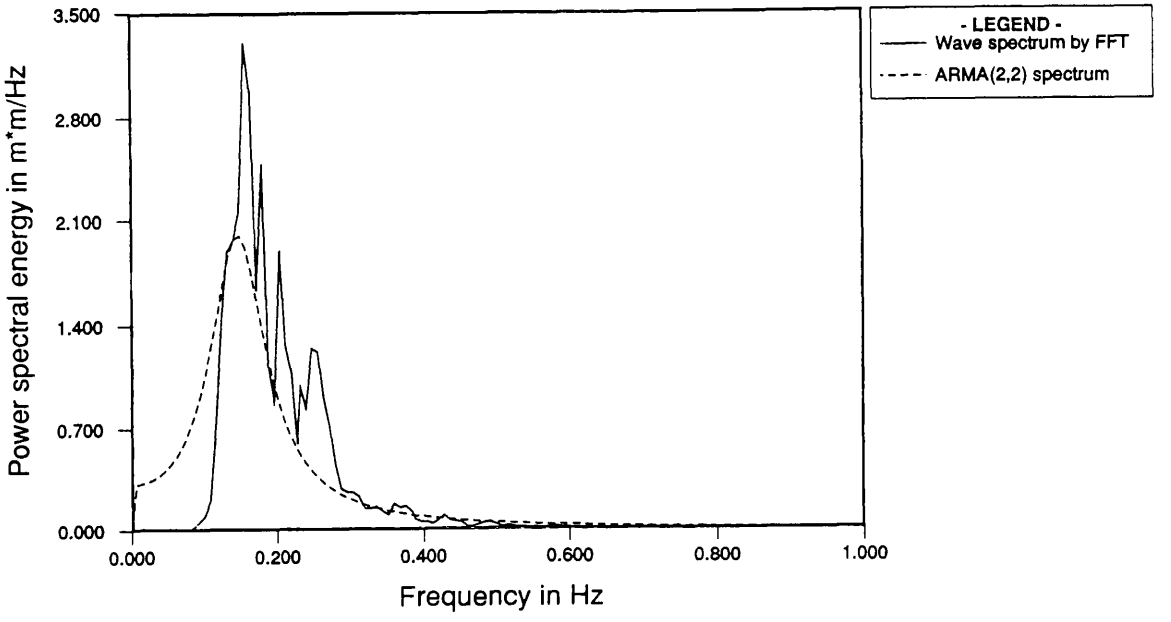

Figure 4.8a FFT wave (M12) spectrum and ARMA(2,2) spectrum estimated from AR(44) using MYW equations and reduction method.

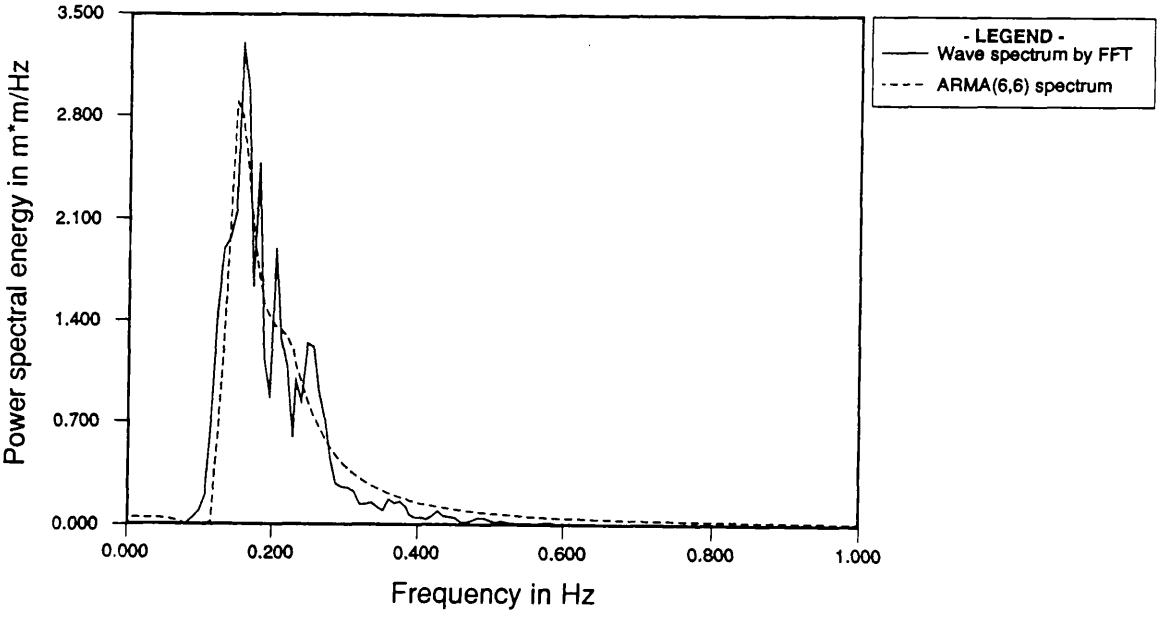

Figure 4.8b FFT wave (M12) spectrum and ARMA(6,6) spectrum estimated from AR(44) using MYW equations and reduction method.

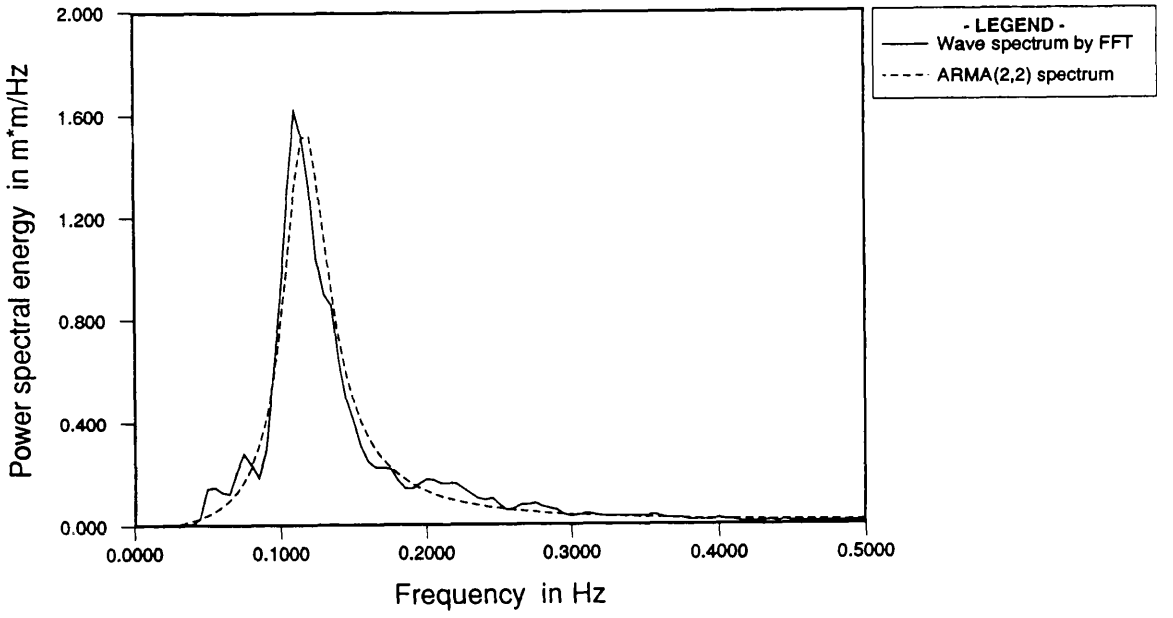

Figure 4.9a FFT wave (K15) spectrum and ARMA(2,2) spectrum estimated from AR(40) using MYW equations and reduction method.

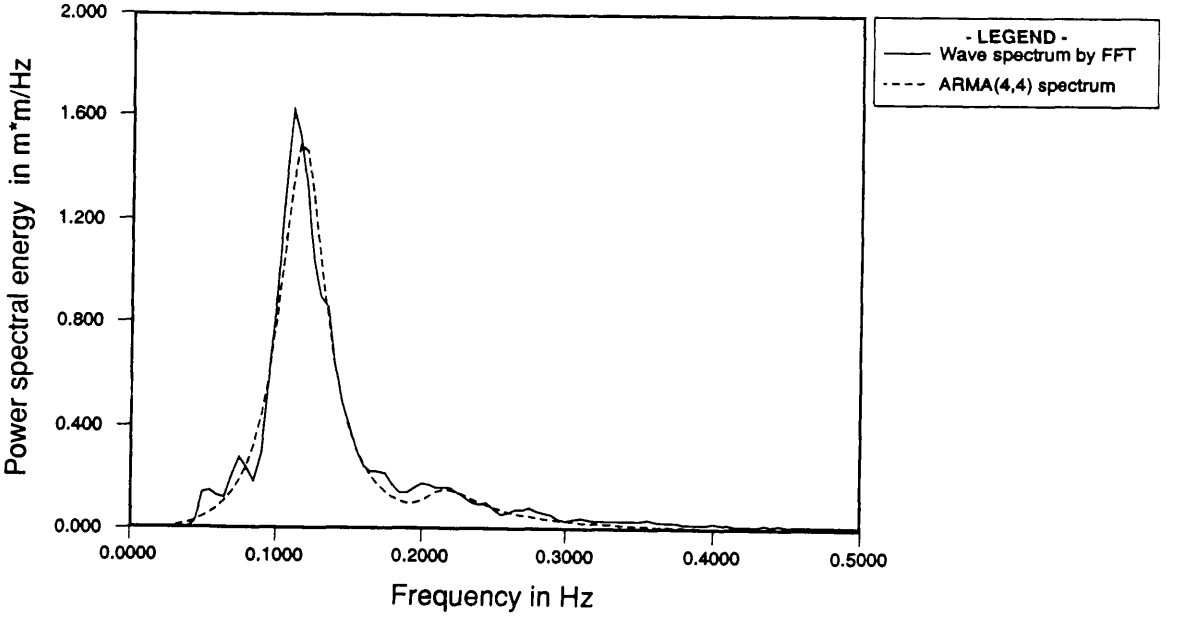

Figure 4.9b FFT wave (K15) spectrum and ARMA(4,4) spectrum estimated from AR(40) using MYW equations and reduction method.

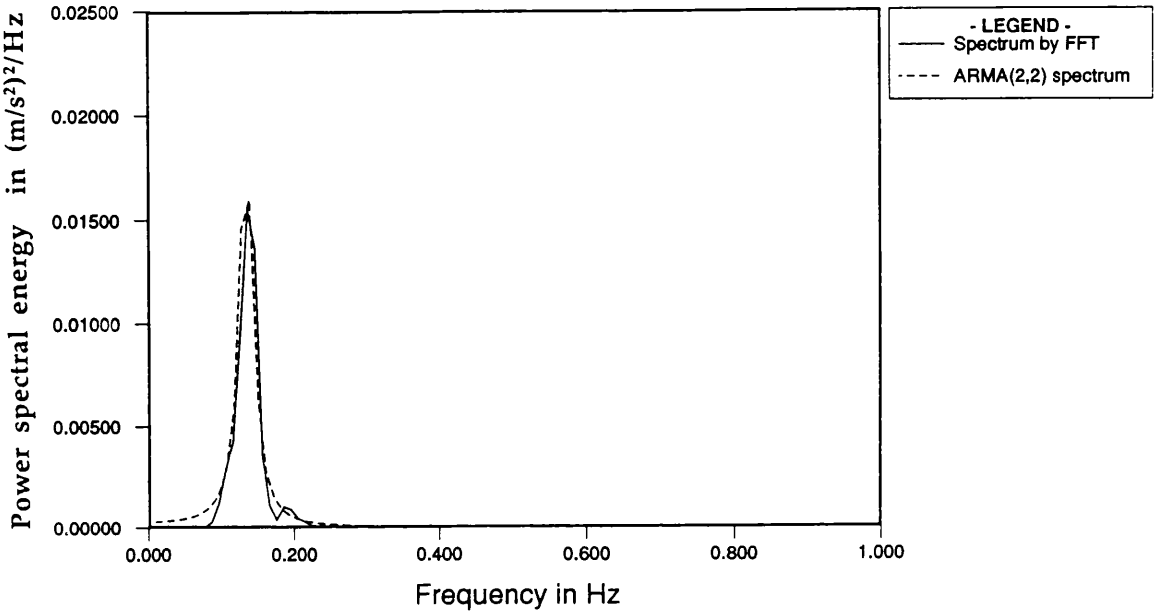

Figure 4.10 Comparison between FFT heave acceleration (SF28) spectrum of semisubmersible (Santa Fe Rig135) and ARMA(2,2) spectrum estimated from AR(98) using MYW equations and reduction method.

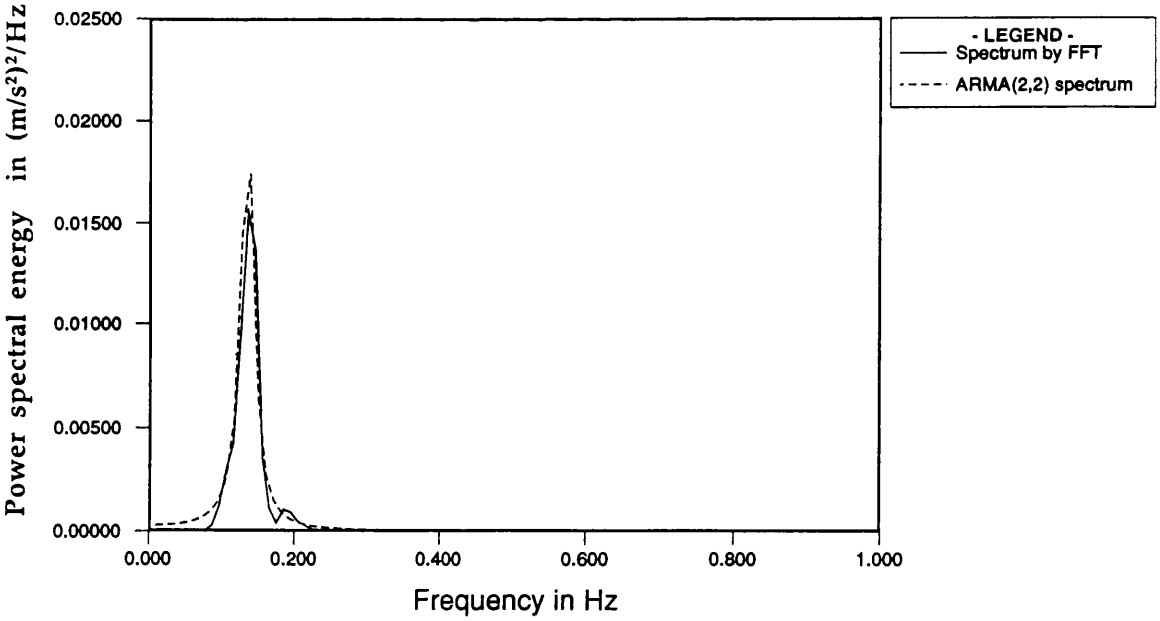

Figure 4.11 Comparison between FFT heave acceleration (SF28) spectrum of semisubmersible (Santa Fe Rig135) and ARMA(2,2) spectrum estimated from AR(100) using MYW equations and reduction method.

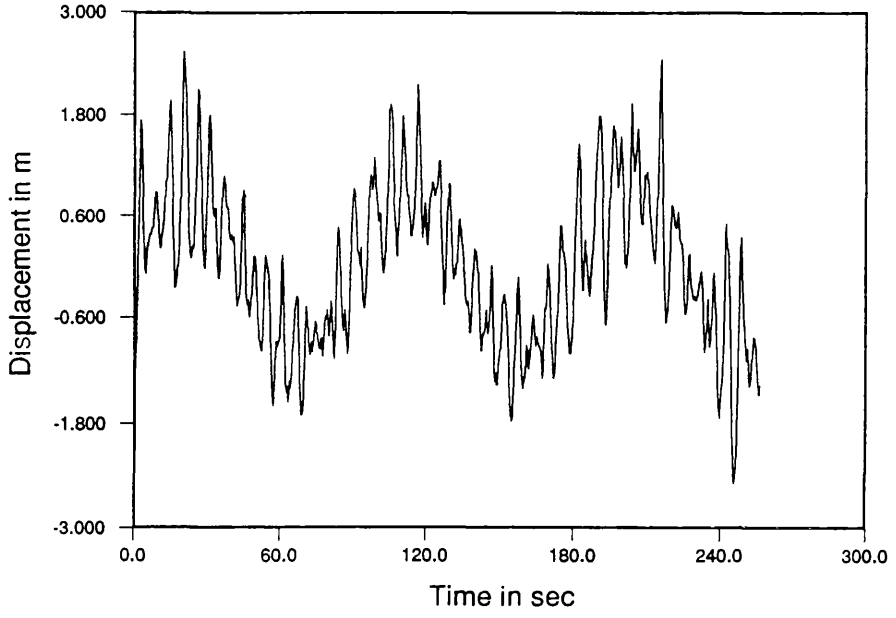

Figure 5.1 Generated nonstationary time series ocean waves.

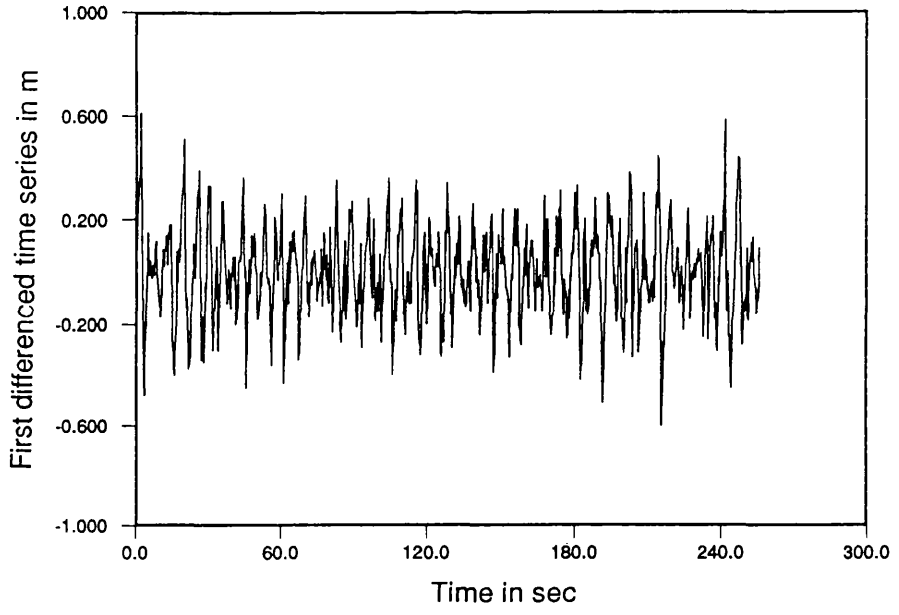

Figure 5.2 First differenced time series estimated from the non-stationary time series (Figure-5.1) ocean waves.

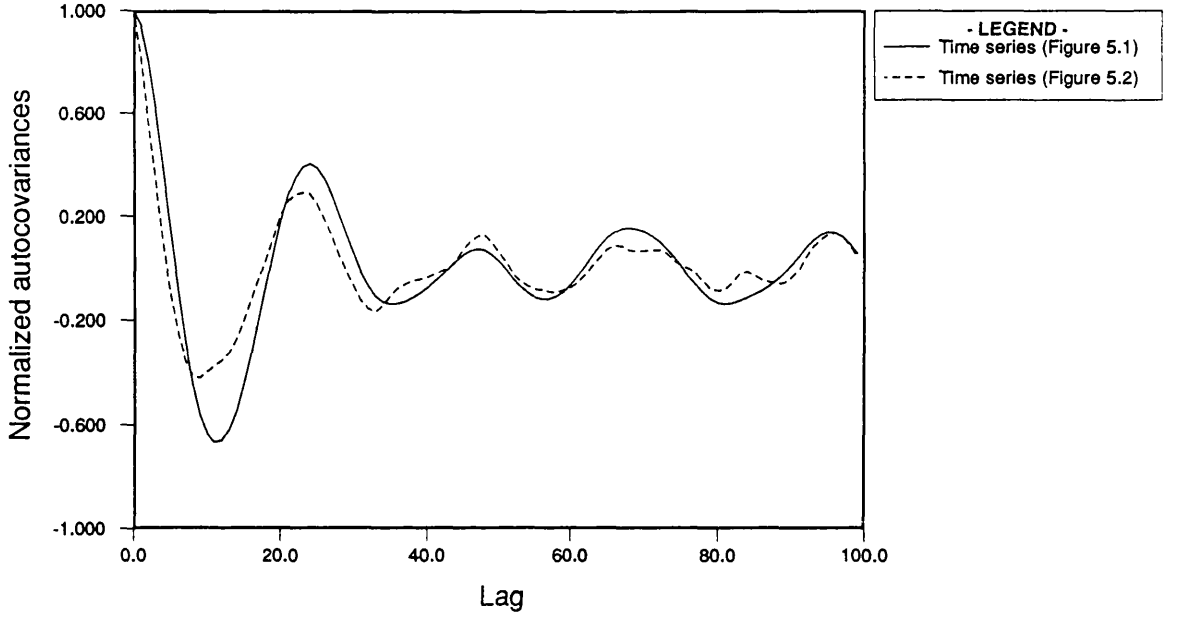

Figure 5.3 Normalized autocovariances of two time series (Figures 5.1 and 5.2).

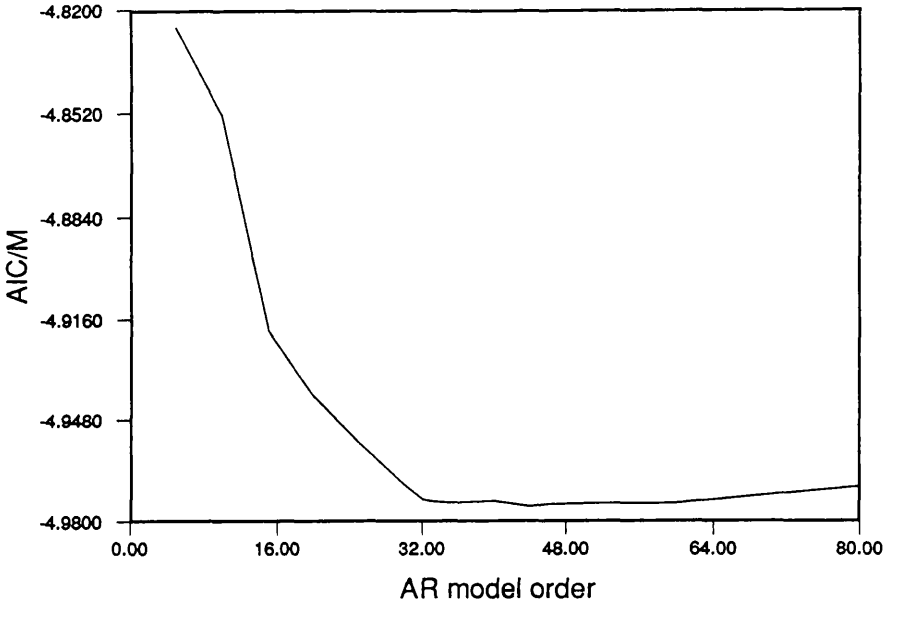

Figure 5.4 AIC variation for the AR process of the first differenced time series.

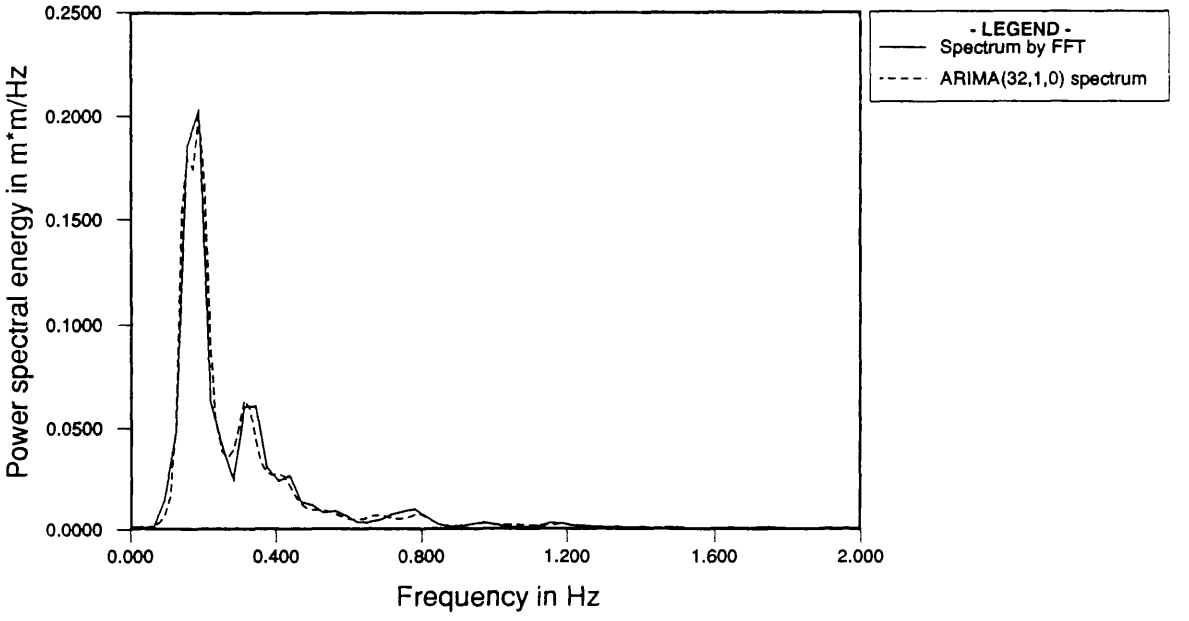

Figure 5.5 Power spectral energies of the time series (Figure 5.2) calculated by FFT and ARIMA(32,1,0) model.

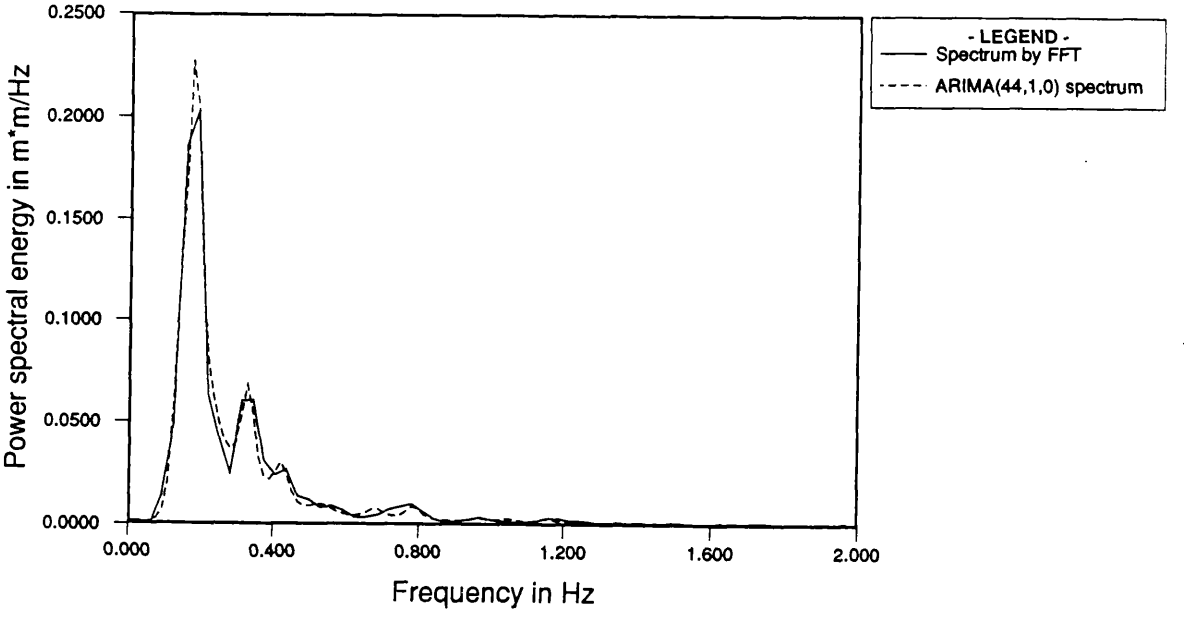

Figure 5.6 Power spectral energies of the time series (Figure 5.2) calculated by FFT and ARIMA(44,1,0) model.

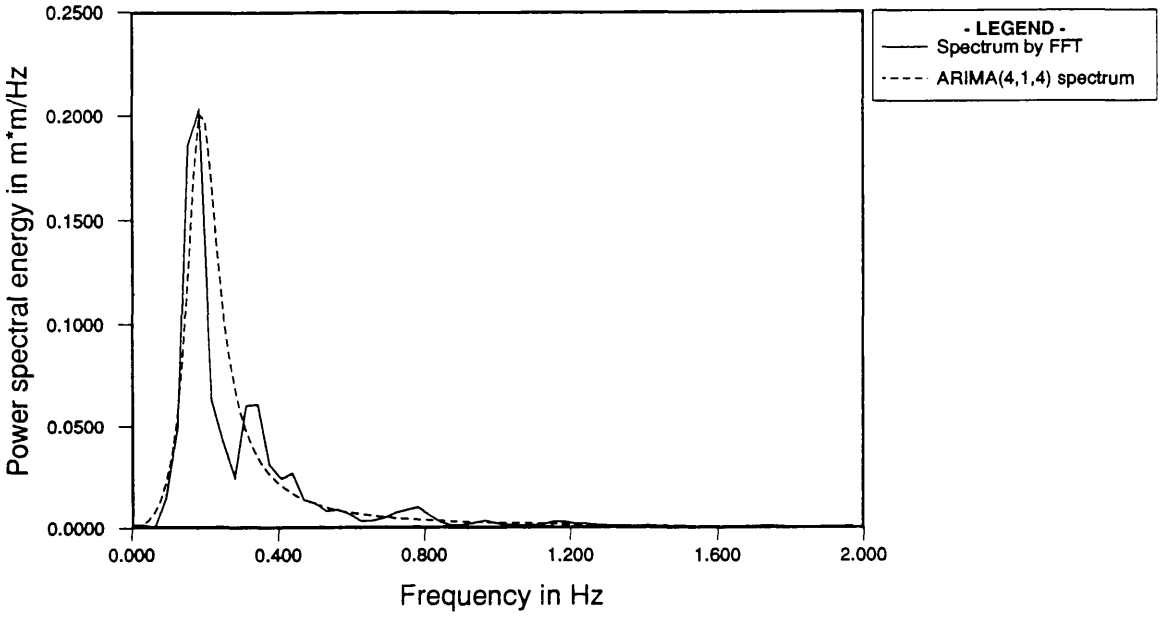

Figure 5.7 Comparison between first differenced time series (Figure 5.2) FFT spectrum and ARIMA(4,1,4) spectrum estimated from ARIMA(44,1,0) using MYW equations and reduction method.

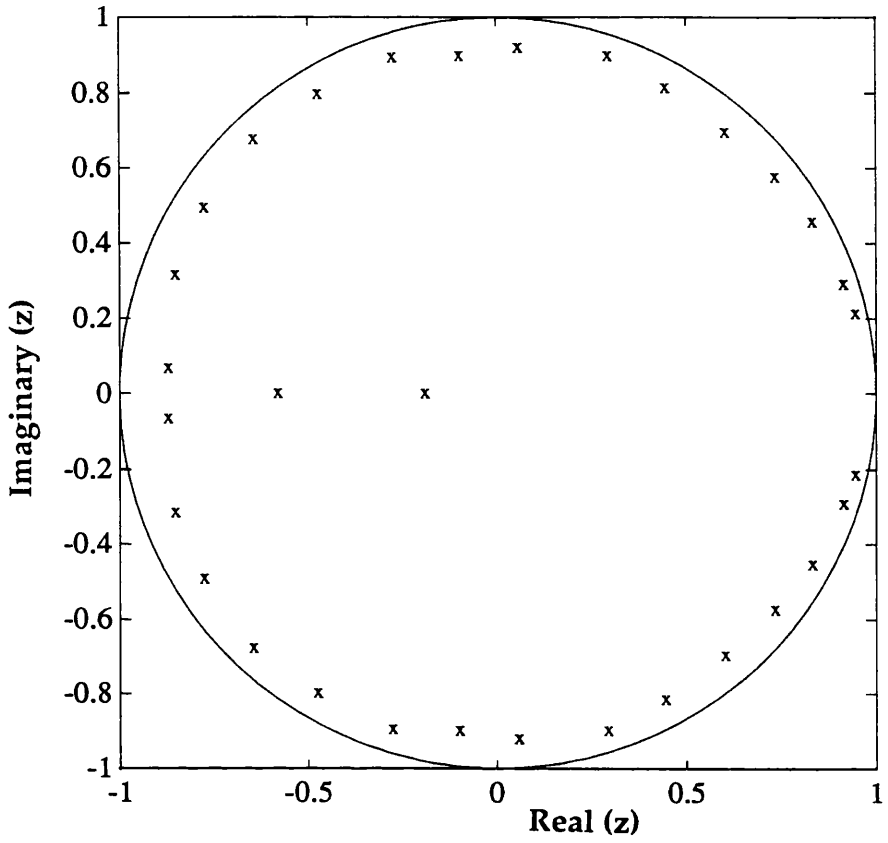

Figure 5.8a Location of poles of ARIMA(32,1,0) model

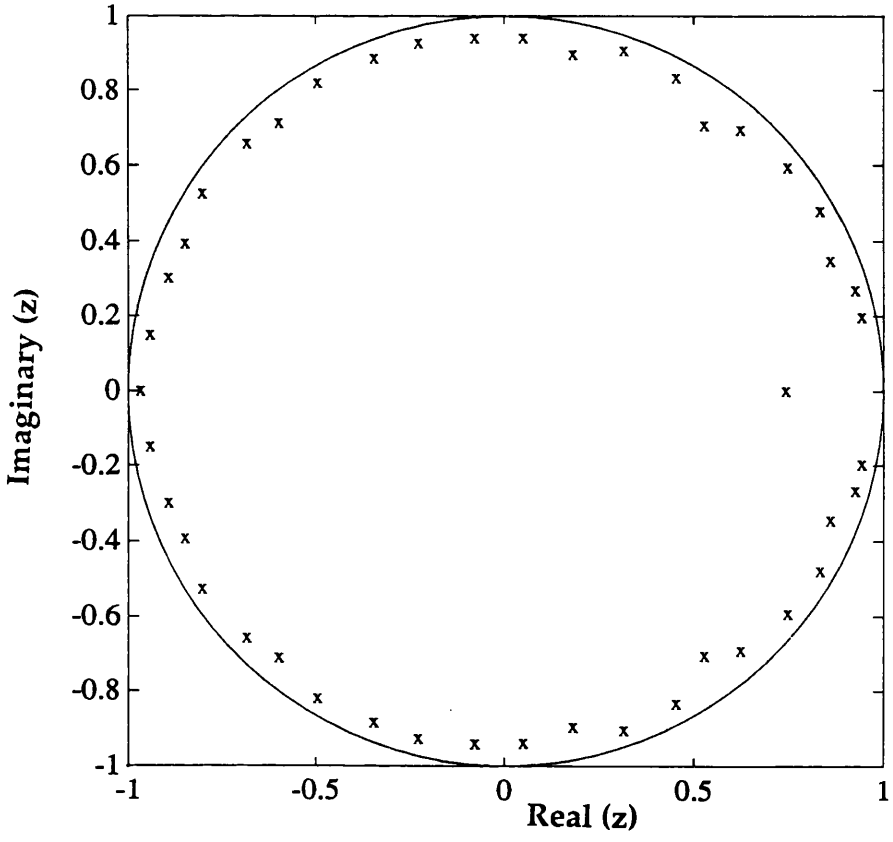

Figure 5.8b Location of poles of ARIMA(44,1,0) model

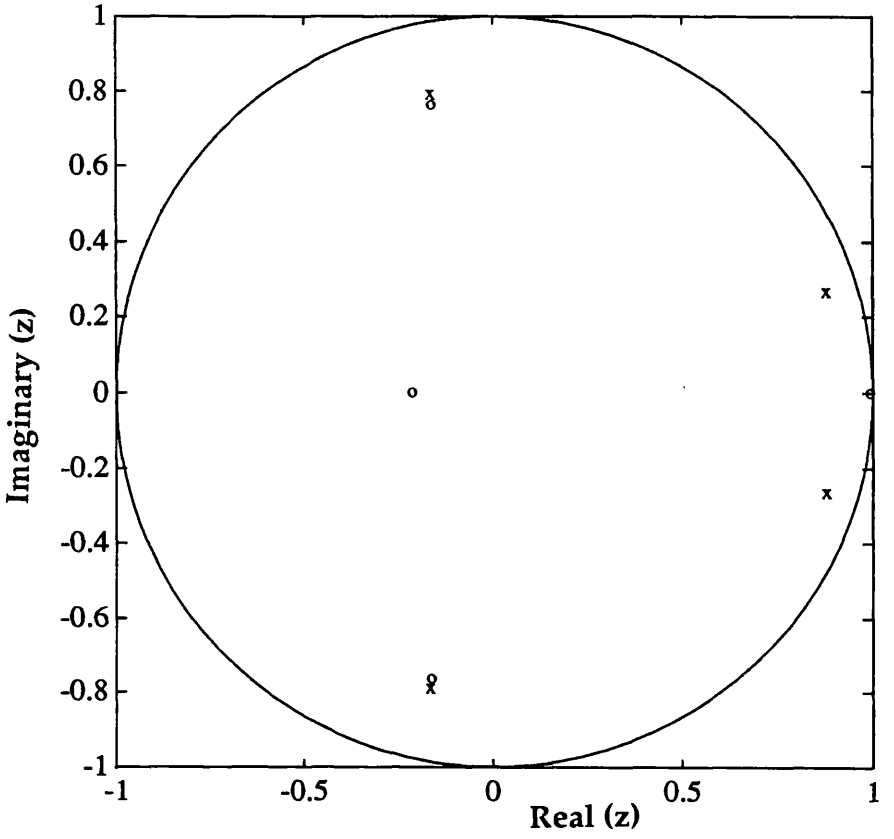

Figure 5.9a Location of poles and zeros of ARIMA(4,1,4) model

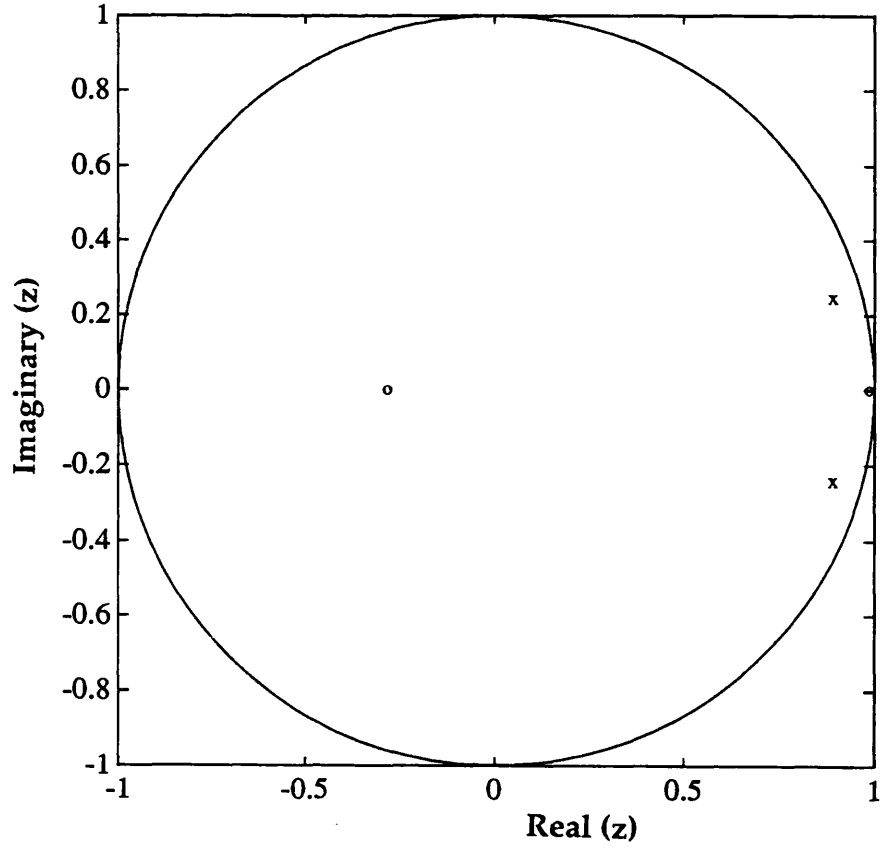

Figure 5.9b Location of poles and zeros of ARIMA(2,1,2) model

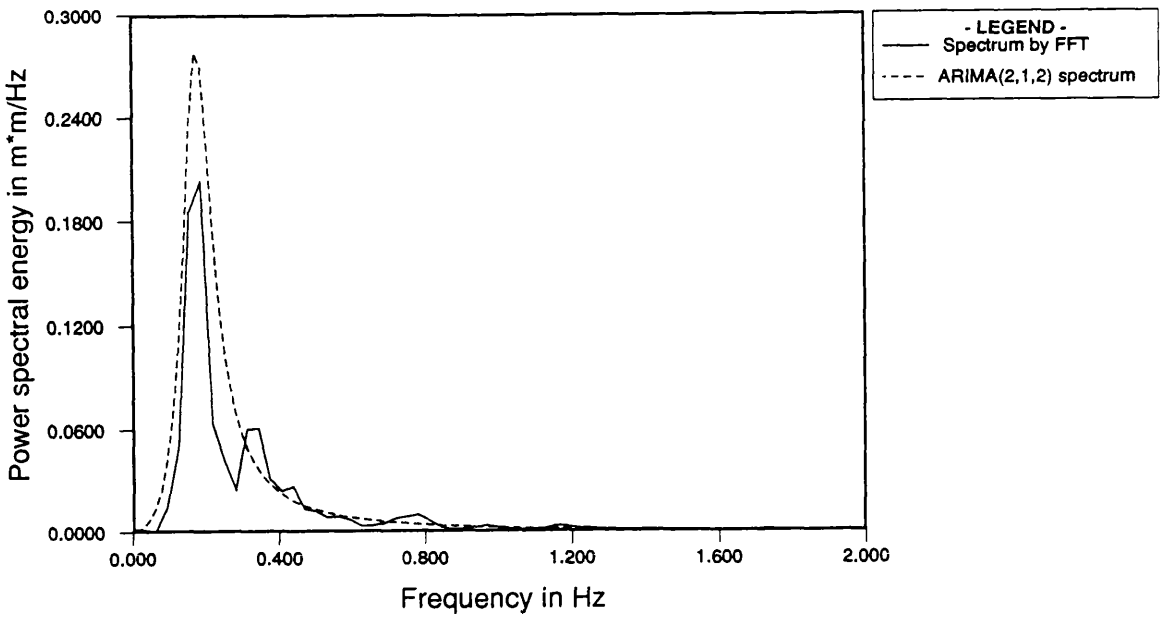

Figure 5.10 Comparison between first differenced time series (Figure 5.2) FFT spectrum and ARIMA(2,1,2) spectrum estimated from ARIMA(44,1,0) using MYW equations and reduction method.

Figure 5.11 Jacket platform (Magnus) in North Sea.

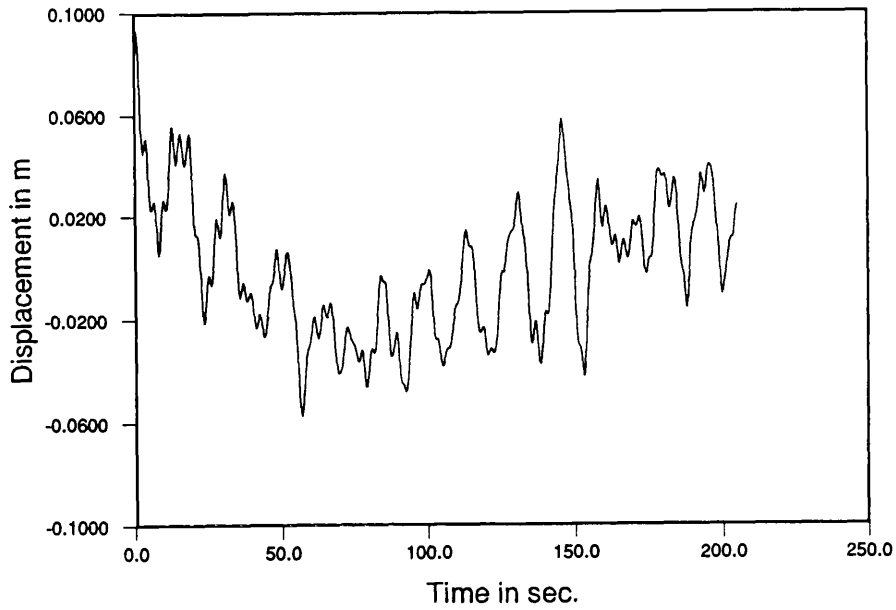

Figure 5.12 Time series offshore platform deck displacements.

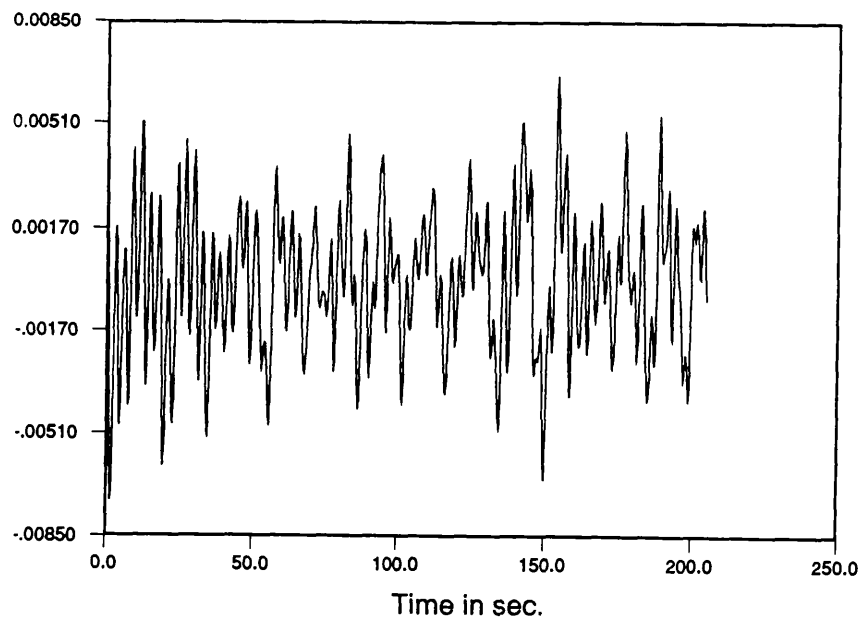

Figure 5.13 First differenced time series estimated from the nonstationary time series (Figure 5.12) offshore platform deck displacements.

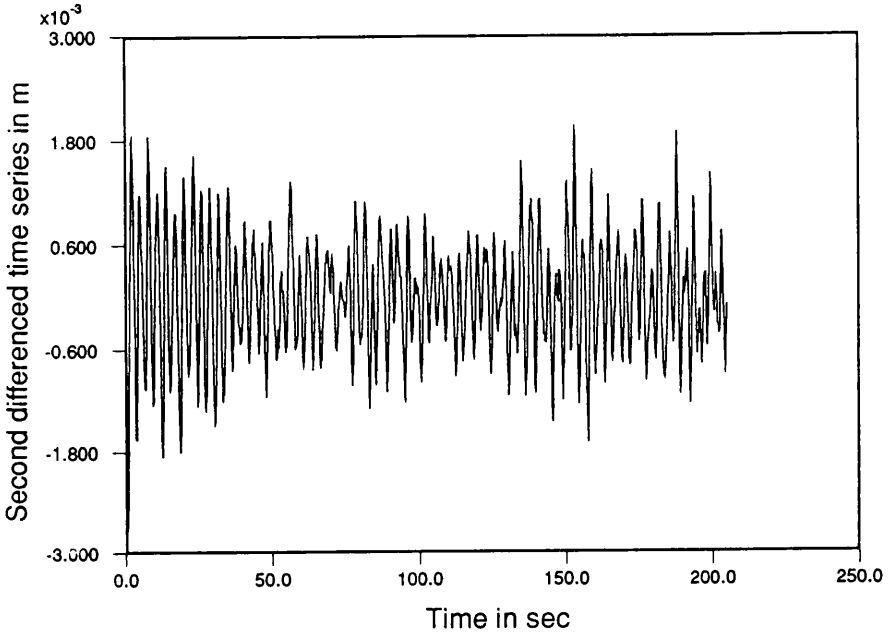

Figure 5.14 Second differenced time series estimated from the nonstationary time series (Figure 5.12) offshore platform deck displacements.

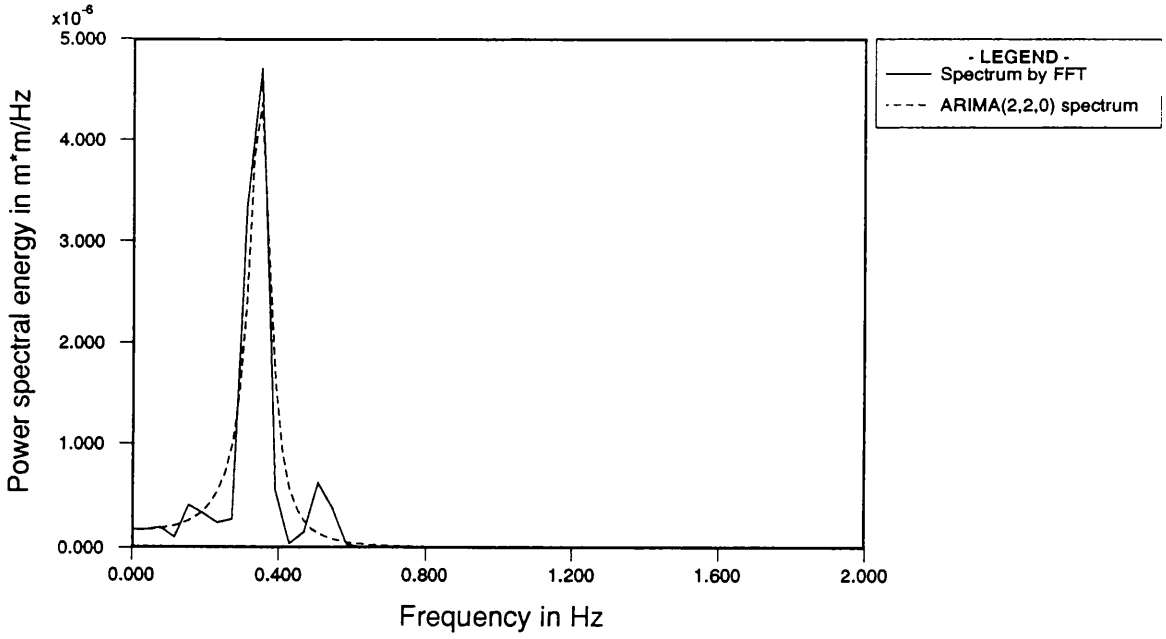

Figure 5.15 Comparison of second differenced time series (Figure 5.14) spectra calculated by FFT and ARIMA(2,2,0) model.

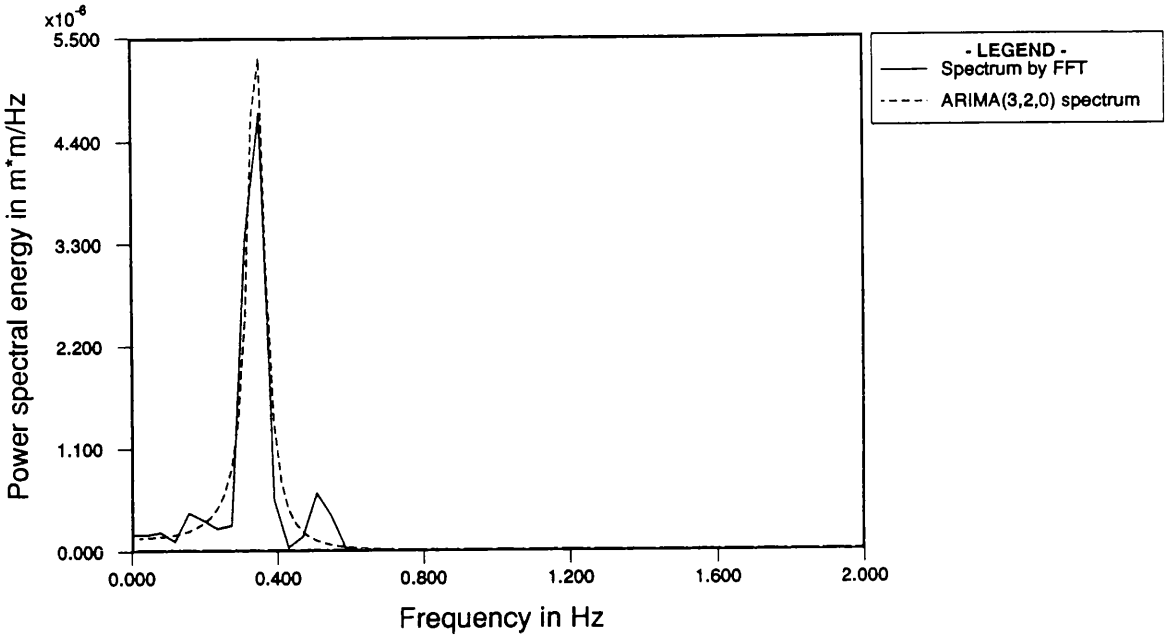

Figure 5.16 Comparison of second differenced time series (Figure 5.14) spectra calculated by FFT and ARIMA(3,2,0) model.

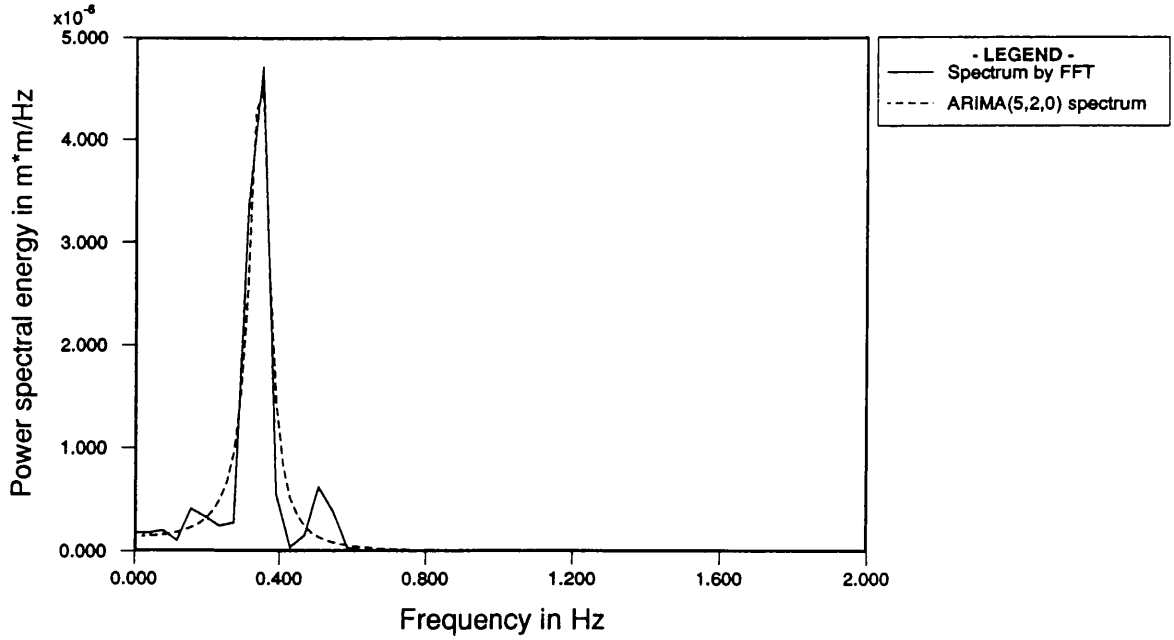

Figure 5.17 Comparison of second differenced time series (Figure 5.14) spectra calculated by FFT and ARIMA(5,2,0) model.

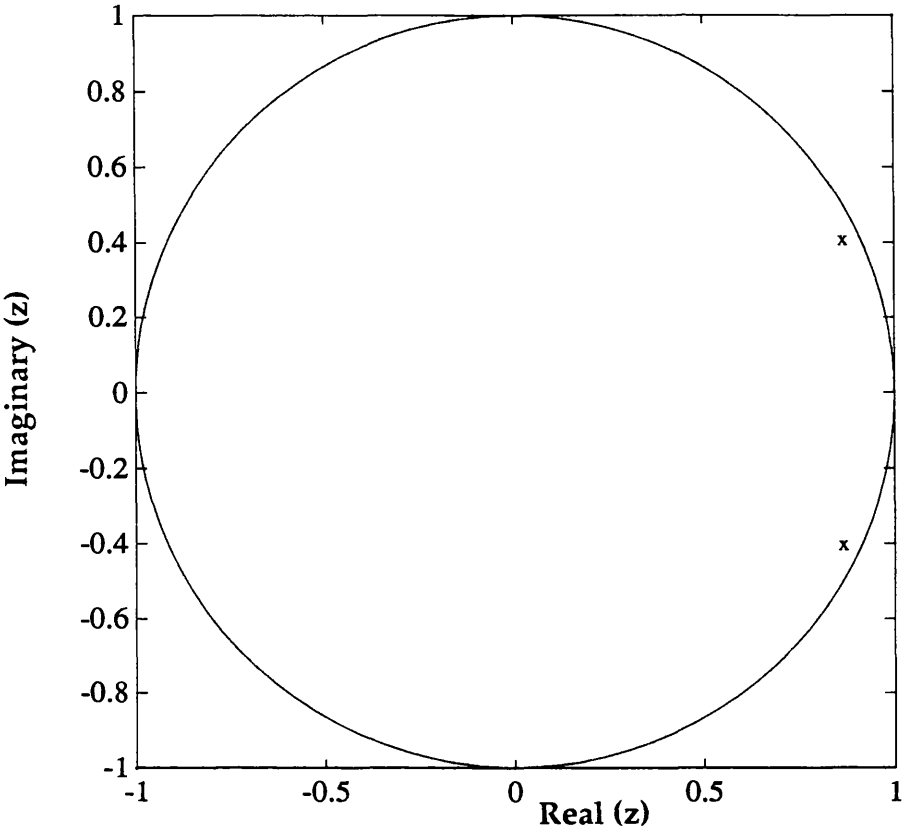

Figure 5.18a Location of poles of the ARIMA(2,2,0) model.

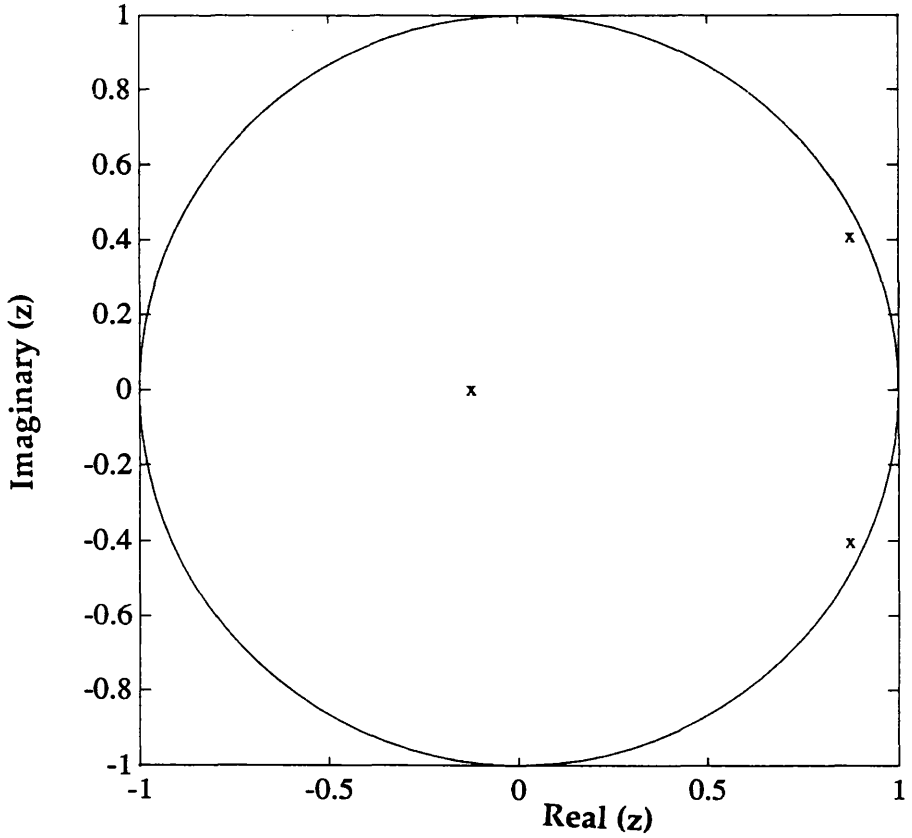

Figure 5.18b Location of poles of the ARIMA(3,2,0) model.

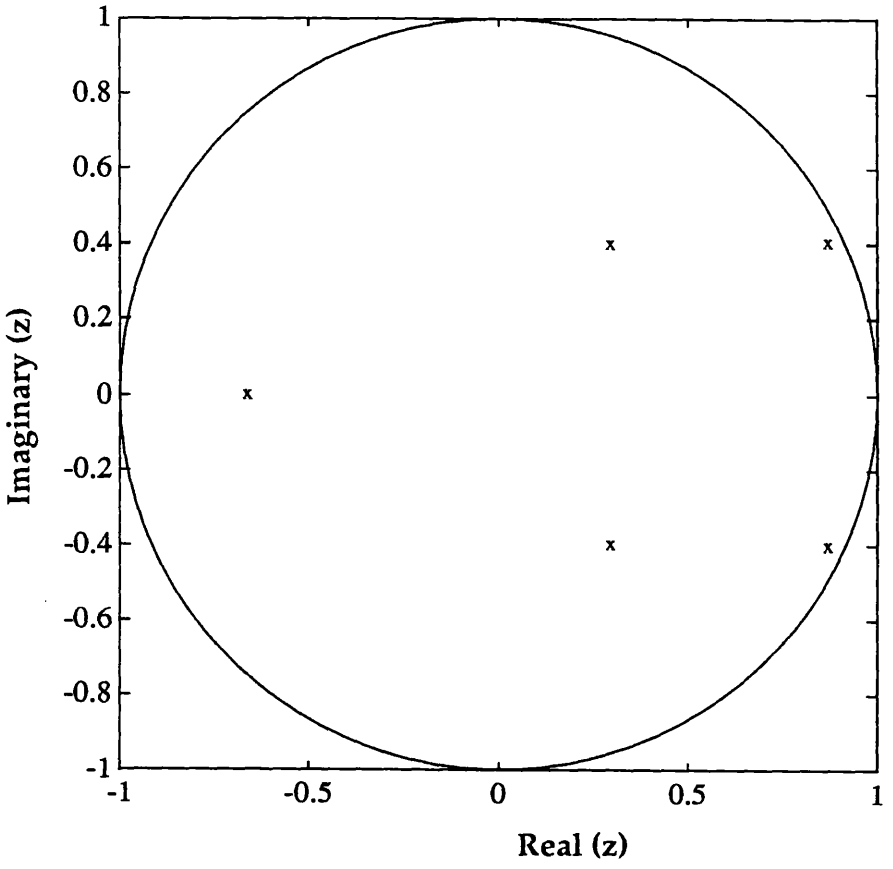

Figure 5.18c Location of poles of the ARIMA(5,2,0) model.

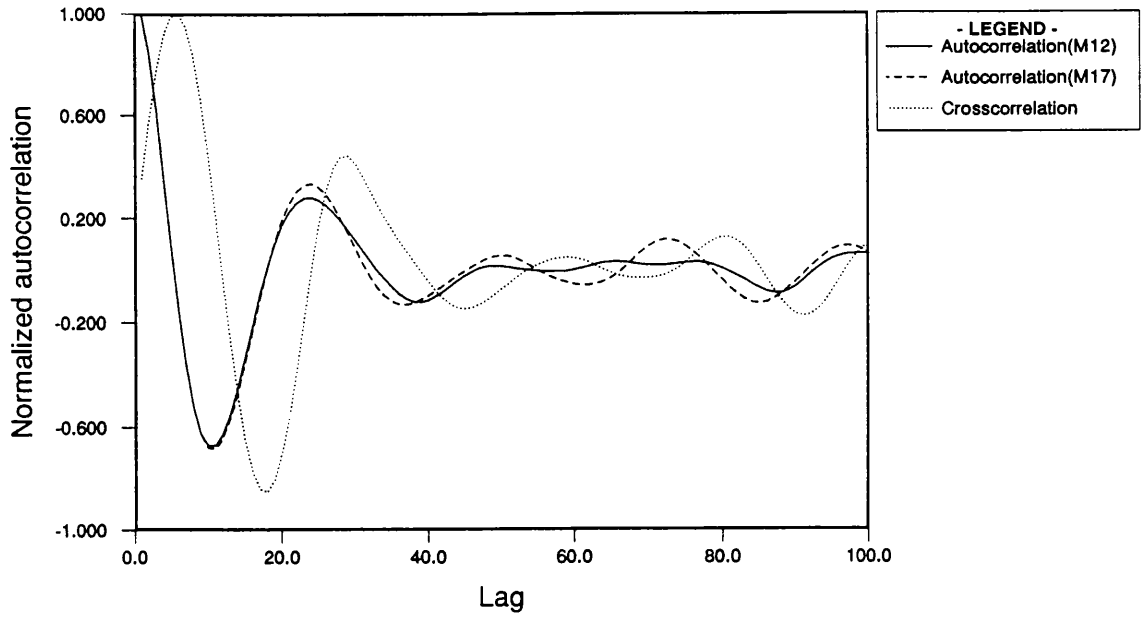

Figure 6.1a Normalized auto and cross correlations of two time series North Sea wave data sets: M12 and M17.

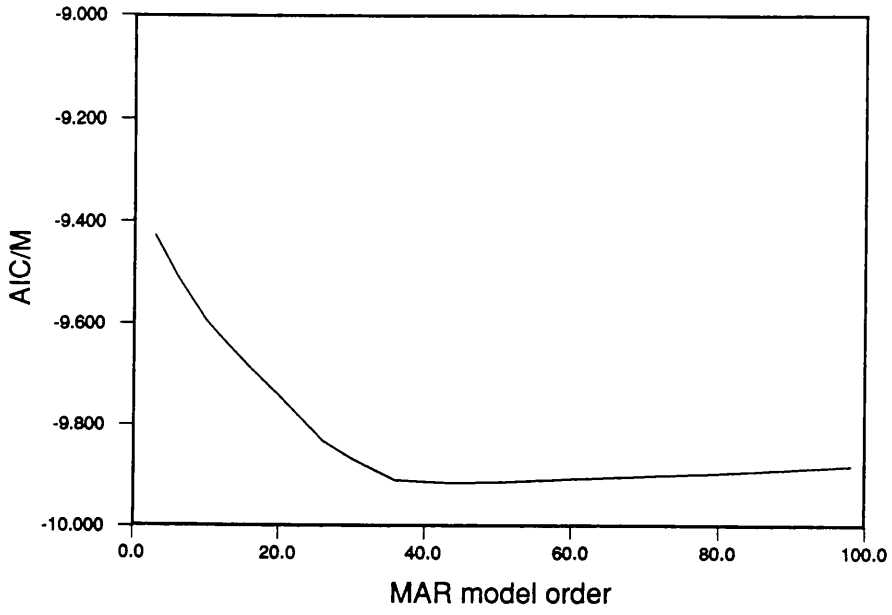

Figure 6.1b AIC variation for MAR processes of two time series data sets.

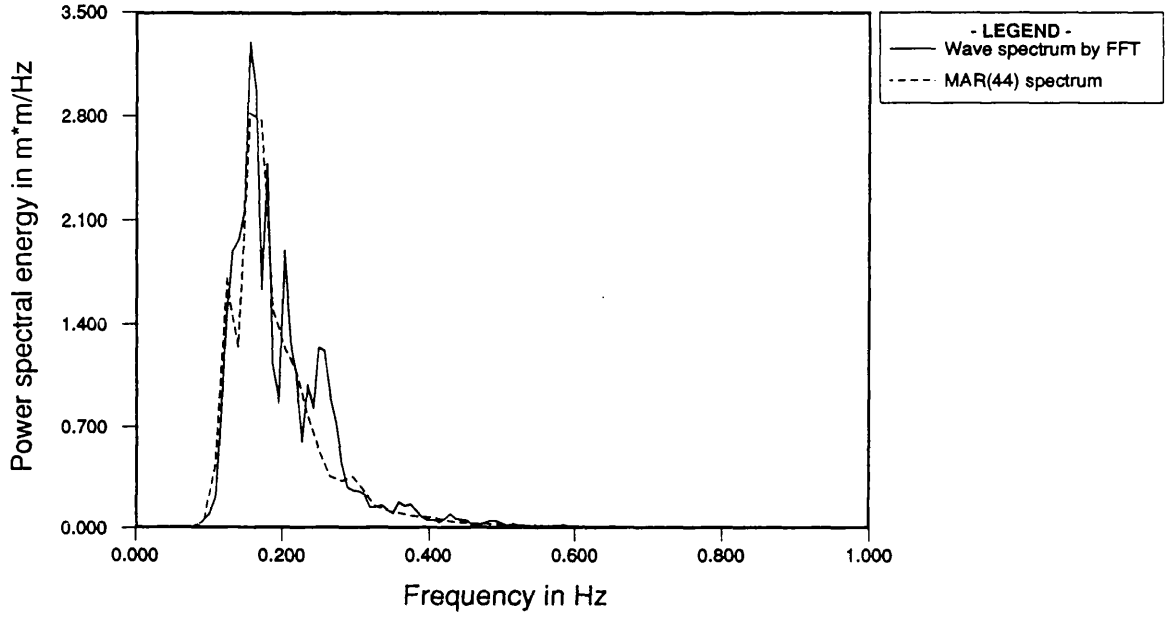

Figure 6.2 Comparison between FFT auto spectrum of M12 data and MAR(44) spectrum.

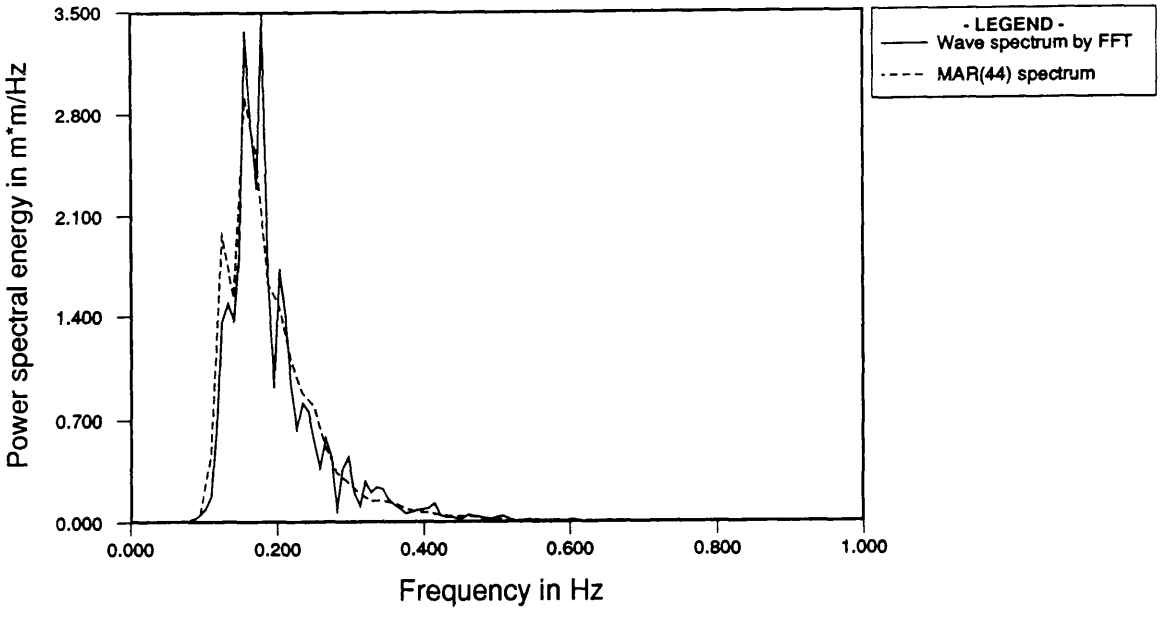

Figure 6.3 Comparison between FFT auto spectrum of M17 data and MAR(44) spectrum.

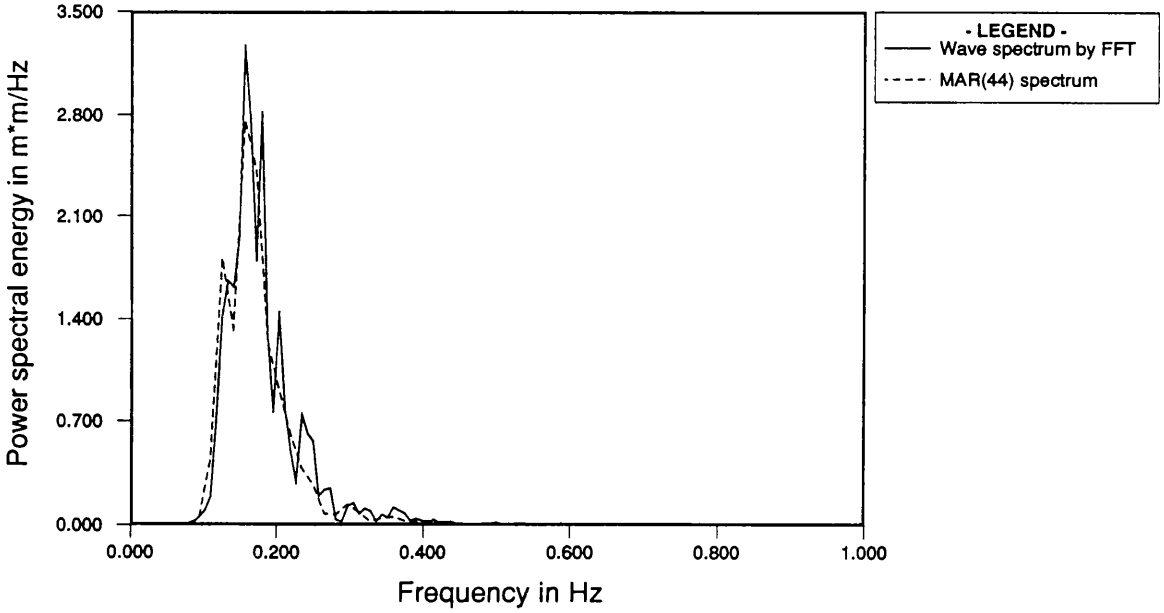

Figure 6.4 Comparison between FFT cross spectrum of M12 and M17 data sets and MAR(44) spectrum.

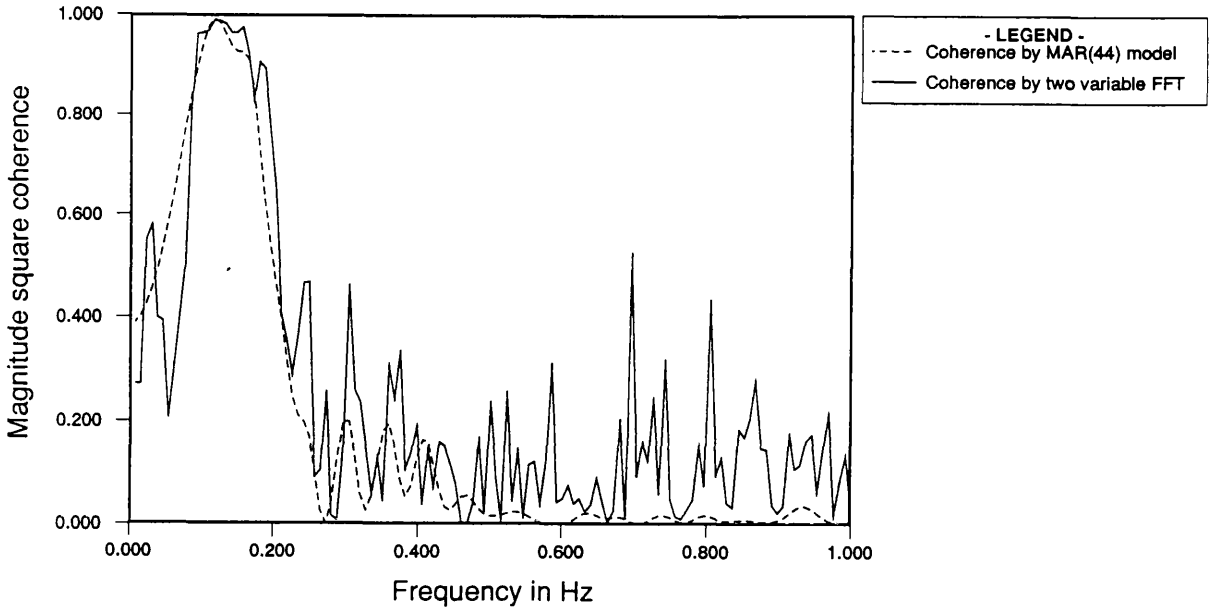

Figure 6.5 Coherence spectra of M12 and M17 data sets estimated by two variable FFT and MAR(44) model.

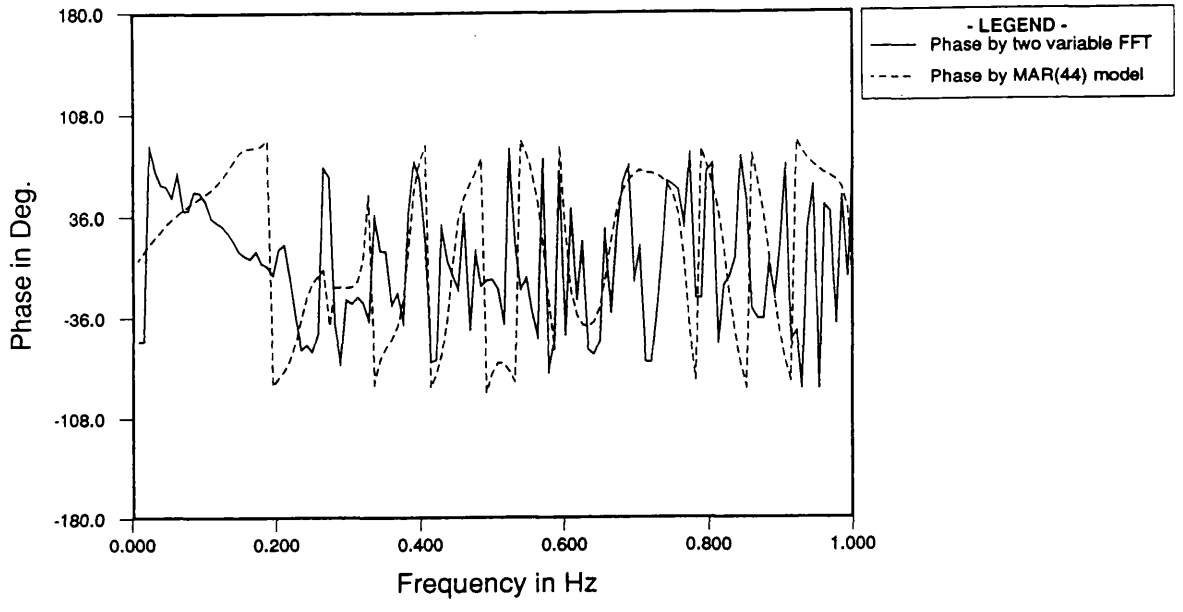

Figure 6.6 Comparison of phase of M12 and M17 data sets estimated by two variable FFT and MAR(44) model.