J. mar. biol. Ass. India, 1995, **37** (1 & 2) : 171 - 178

TIME SERIES ANALYSIS OF MARINE FISH LANDINGS IN INDIA

T. V. SATHIANANDAN AND M. SRINATH
Central Marine Fisheries Research Institute, Kochi - 682 031

ABSTRACT

Catch data for the period 1950-91 on all India landings for three important marine fish species and total landings were analysed to develop suitable ARIMA models for time series forecast. An initial recurssive regression analysis showed that the models suitable are ARIMA (2, 2, 1) for total landings, ARIMA (1, 2, 1) for penaeid prawn landings and ARIMA (3, 2, 1) for both catfish and silverbellies landings. The parameters of these models were then estimated by a maximum likelihood estimation procedure following the algorithm of Godolphin (1984). A test by using Chi-square given by Lung and Box (1978) showed that the fitted models are adequate to explain the data. With the fitted models forecasts for the next year were computed along with their confidence limits.

INTRODUCTION

MARINE FISH landings in India has increased tremendously over the past four decades. From a mere six lakh tonnes in the early fifties it has increased to about twenty lakh tonnes during the nineties. This was possible through better capture techniques and increased effort. The marine fishery in India is a multispecies-multigear system and analysis of such a complex system is quite demanding. There had been attempts to analyse fish landings in India with a view to understand the status of the fish stocks and to propose suitable harvesting strategies (Alagaraja, 1984; Srinath and Datta, 1985). The Central Marine Fisheries Research Institute following a suitable sampling design, has been estimating marine fish landings in India to arrive at species wise estimates since 1950. The present study made use of this data is an attempt to forecast the marine fish landings of certain important groups through time series analysis. Reliable forecasts of catch are essential for fisheries management and the time series analysis is an economical method for forecasting catches. Classical methods are available in literature, which make use of fishery dependent and independent factors to explain and forecast the fishery. However, these involve lot of effort and it is time consuming. The time series analysis has been successfully applied to several fisheries (Jensen, 1976, 1985; Van winkle *et al.*, 1979; Saila *et al.*, 1980; Mendelssohn 1981. Stocker and Hilborn, 1981).

The authors thank Dr. K. Alagaraja, CMFRI for critically going through the manuscript and suggesting improvements.

DATABASE AND METHODS

The data for the study were obtained from reports of CMFRI on marine fish landings. In the present study, the total landings and the landings of catfish, silverbellies and penaeid prawns on all India basis were considered. Time series of catch data can be analysed by different methods (Box and Jenkins, 1976). Two widely applied methods are autocorrelation analysis and spectral analysis. Spectral analysis assumes that the underlying process can be described in terms of sine and cosine functions. Autocorrelation analysis makes few assumptions about the underlying process and seems more appropriate for ecological analysis (Moran, 1953). In the approach to time series analysis proposed by Box and Jenkins (1976), both moving average terms and autoregressive terms are tested to seek the best fitting model. The resulting model usually has no physical interpretation and provides little or no understanding, but it is a powerful tool for forecasting which can easily be updated. The order of moving average terms and autoregressive terms are usually determined by observing autocorrelations and partial autocorrelations. Methods of autocorrelation analysis are applicable only to stationary time series. In a stationary time series, the data fluctuate about some mean level and the mean, variance and autocovariance are not dependent on time. In practice, the identification of appropriate Autoregressive Integrated Moving Average (ARIMA) process is rather difficult. There are some objective methods of identifying the orders of ARIMA(p, d, q) process with the mathematical form as given in equation (1)

$$(1 - B)^d \, (y_t + \varphi_1 y_{t-1} + \cdots + \varphi_p y_{t-p}) = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} \quad \ldots \quad (1)$$

where p and q are the order of the autoregressive (AR) and moving average (MA) terms respectively, d is the number of differences to be applied on the basic series to make it stationary, B is the shift operator such that $B^k y_t = y_{t-k}$, φ_i (i = 1, .. P) and θ_j (j = 1, .., q) are the coefficients to be estimated, $\{y_t\}$ is the transformed series and ε_t are the error terms which are assumed to be distributed independently as normal with mean zero and variance σ^2. In this paper, the appropriate orders p and q in ARIMA (p, d, q) were determined by the method given by Hannan and Kavalieris (1984), which is briefly discussed below.

First an autoregression of suitable higher order is fitted to the data (stationary series obtained by applying differencing if necessary). The suitable higher order is determined by minimising a criterion namely AIC, given by equation (2)

$$\mathrm{AIC}(v) = \log(\sigma^2) + 2v/T \quad \ldots \quad (2)$$

where v is the order of autoregression fitted, T is the size of the series and $\sigma^2 = \sum_t \varepsilon_t^2/T$ is the estimated error variance for the fitted autoregression where the error term ε_t is given by the equation (3)

$$\varepsilon_t = \sum_j b_j y_{t-j} \text{ where } b_o = 1,\ y_t = o \text{ for } t \le o \quad \ldots (3)$$

where b_j's are the estimated regression coefficients. Now the initial estimates of the AR and MA coefficients are obtained by regressing y_t on y_{t-j} (j = 1, . . . , p) and $\in_{t-j}$ (j = 1, . . . , q). Then the orders of the AR and MA coefficients are obtained by minimising a criterion BIC given by equation (4)

$$\mathrm{BIC}(p, q) = \log(\sigma^2_{p,q}) + (p + q) \log(T)/T \quad .(4)$$

where $\sigma^2_{p,q}$ is the estimated error variance with orders (p, q). This has to be attempted for a sufficient order combinations (p, q) and choose the one which yields minimum BIC value. After estimating the orders of autoregression and moving average terms by the above method, the maximum likelihood estimates of the coefficients and the error variance were computed by following the procedure given by Godolphin (1977, 1978, 1984). The adequacy of the fitted model was tested by computing the quantity Q(r) in equation (5), which is a chi-square with (m-p-q) degrees of freedom (Lung and Box, 1978)

$$Q(r) = T\,(T + 2) \sum_{k-1}^{m} [r_k^2 / (T - k)] \quad \ldots (5)$$

$$r_k = \sum_{t=k+1}^{T} \varepsilon_t \, \varepsilon_{t-k} \Big/ \sum_{t=1}^{T} \varepsilon_t^2 \quad . . . (6)$$

where p and q are the orders of autoregressions and moving averages, m is the maximum lag of error autocorrelations r_k, given by equation (6), used for computing Q.

to a maximum of 2.23 million t in 1989. The average catch in 1950-60 period was 0.6566 million t with a CV of 18.8%, in 1961-70 period the average catch rose to 0.8331 million t with CV 15.6%, in 1971-80 period it again increased to 1.27 million t with CV 10.1% and in 1981-91 period the average catch reached

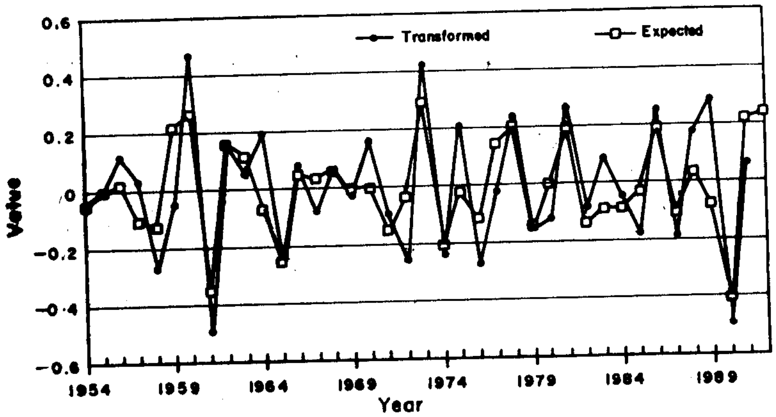

FIG. 1 a. Plot of transformed values, model : ARIMA (2, 2, 1) and their expectations of all India total marine fish landings.

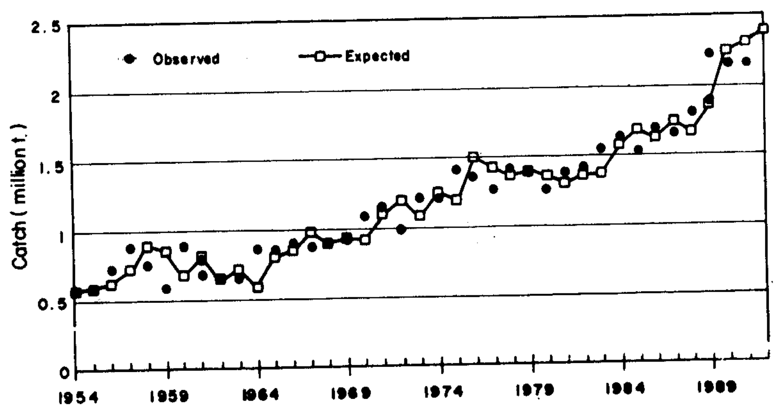

FIG. 1 b. Plot of all India total landings, model : ARIMA (2, 2, 1) and predicted landings.

RESULTS AND DISCUSSION

Total landings

There is an increasing trend in total landings from 1950 onwards. From a catch of mere 580,022 tonnes in 1950 it has increased

1.75 million t with CV 16.6%. This increase in catch is due to many factors like increased effort, improved technology, increased demand and so on.

Yearly landings data for the period 1950-1991 was used for the time series analysis.

The series was made stationary (by viewing the plots) by taking a simple second order difference and then centered to the average. By applying the method suggested by Hannan and Kavalieris (1984), the orders of AR and MA terms were estimated as 2 and 1 respectively. Hence, the model found suitable method of Godolphin (1984) are

$$\hat{\varphi}_1 = 0.614769 \quad \hat{\varphi}_2 = 0.430769 \quad \hat{\theta}_1 = -0.421130$$

$$\hat{m} = -0.00056023 \quad \hat{\sigma}^2 = 0.020208$$

Hence, the mathematical form of the model can be written as in equation (7) and in terms

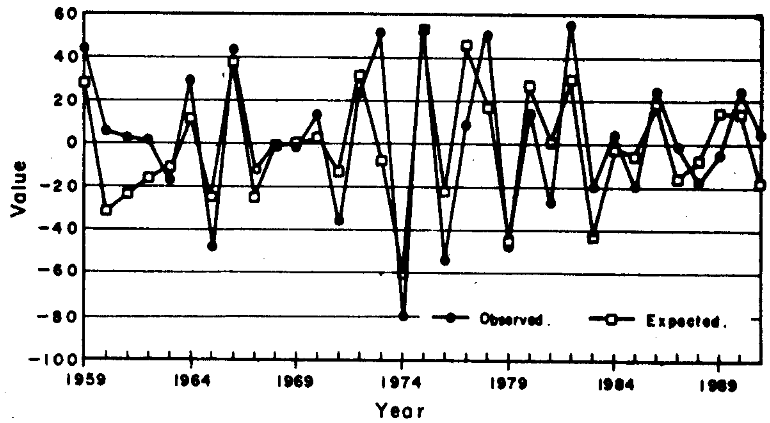

FIG. 2 a. Plot of transformed values, model : ARIMA (1, 2, 1) and their expectations of all India penaeid prawn landings.

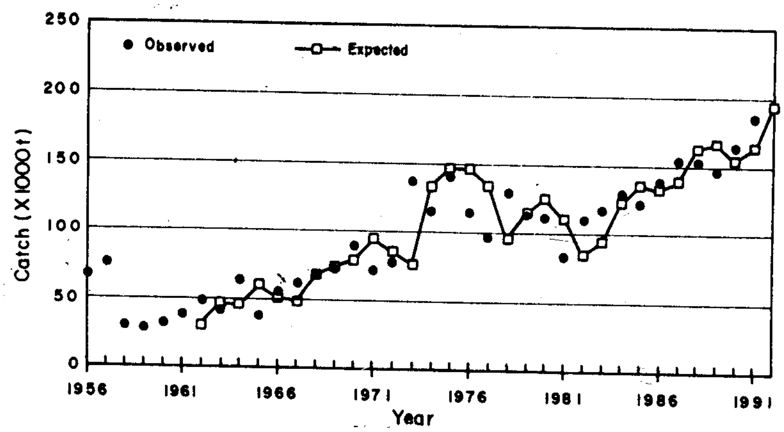

FIG. 2 b. Plot of all India penaeid prawn landings, model : ARIMA (1, 2, 1) and predicted landings.

is ARIMA (2, 2, 1) with $y_t = (1 - B)^2 z_t - m$, where Z_t is the catch in year t, as the input series, m is a constant to be estimated and ε_t is the error fluctuation in year t. The maximum likelihood estimates of the parameters of the above model estimated by following the of the catches (z_t) the estimated model takes the form as in equation (8).

$$y_t + 0.614769\ y_{t-1} + 0.430769\ y_{t-2} = \varepsilon_t - 0.421130\ \varepsilon_{t-1} \quad \ldots (7)$$

$$(1-B)^2 (z_t + 0.614769\ z_{t-1} + 0.430769\ z_{t-2}) = \varepsilon_t - 0.421130\ \varepsilon_{t-1} - 0.0011459 \quad . . . (8)$$

The differenced series and their predictions with the estimated model are shown in Fig. 1 a and the actual landings with the predicted catch are shown in Fig. 1 b. The forecast for 1992 is 2.39 million t with 95% confidence limits 2.39 ± 0.29 million t. By using the estimated error terms of the model, to test the adequacy of the model, the Lung-Box chi-square computed is 27.5 with 17 degrees of freedom which is not significant. Hence the model fitted is suitable for the data.

Penaeid prawn landings

In all India penaeid prawn landings also there is an increasing trend over years. From a catch of 66,910 t in 1956 it has increased to 186,330 t in 1991 which is about three times. The average landings in the period 1961-70 was 57,884 t with CV 27.4%, that in the period 1971-80 was 110,965 t with CV 19.6% and the average landings during 1981-91

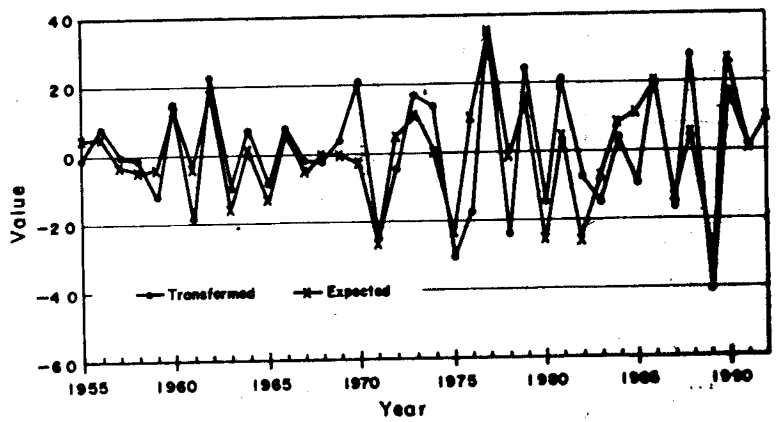

FIG. 3 a. Plot of transformed values, models : ARIMA (3, 2, 1) and their expectations of all India Catfish landings.

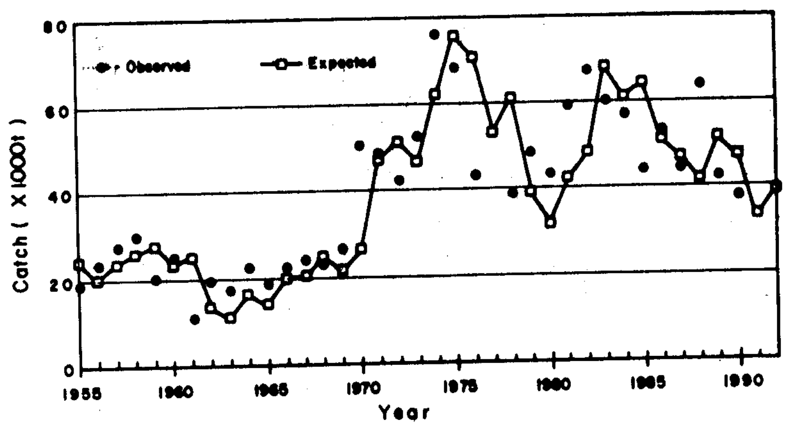

FIG. 3 b. Plot of all India Catfish landings, model : ARIMA (3, 2, 1) and predicted landings.

was 137,115 t with CV 19.7 %. The increase in landings may be due to increased effort targetted to this species which has very high export value.

Data for the period 1956-91 was used for the analysis. It was made stationary by taking second order difference and then centered to the mean. The orders of AR and MA coefficients estimated as 1 and 1 respectively resulting the model ARIMA (1, 2, 1) with $y_t = (1 - B)^2 z_t - m$ as the input series where z_t is the catch at time t and m is a constant. the maximum likelihood estimates of parameters of the model are

$$\hat{\varphi}_1 = 0.525380 \quad \hat{\theta}_1 = -0.560824$$
$$\hat{m} = 0.3097270 \quad \hat{\sigma}^2 = 471.47$$

With these estimates of parameters, the model in terms of the original catch series can be written as in equation (9).

$$(1-B)^2 (z_t + 0.525380\ z_{t-1}) =$$
$$\varepsilon_t - 0.560824\ \varepsilon_{t-1} - 0.472451 \quad . . . (9)$$

The transformed inputs for the analysis and their predictions using the estimated model are shown in Fig. 2 a. Actual catch and their predictions are shown in Fig. 2 b. Forecast made for 1992 with the above model is 1.94 lakh t with 95% confidence limits 1.94 ± 0.44 lakh t. The computed value of Lung-Box chi-square is 14.4, with 16 degrees of freedom, with is not significant showing the adequacy of the fitted model.

Catfish landings

There is an increasing trend in all India catfish landings from 1950 onwards with more fluctuations during 1961-'70 period. Average landings during 1950-'60 was 21,585 t with a CV of 22.2%, which rose to 23,762 t with CV 41.7% in 1961-'70 period, almost doubled then to an average of 51,767 t with CV 21.9% and remained steady in 1981-'91 period with an average of 51,526 t and CV 20.7%. From a catch of 11,779 t in 1950 it rose to the maximum of 76.196 t in 1974. Thereafter it was found to fluctuate over years. It reduced to 39,231 t in 1978, again reached 67,664 t in 1982 and 64,216 t in 1988. Thereafter it is found to decrease year after year and in 1991 the catch was 34,110 t.

A second difference of the catch sequences was found to be stationary. This differenced series was then centered to the mean. The orders of AR and MA terms were then estimated as 3 and 2 respectively, leading to the model ARIMA (3, 2, 1), with $y_t = (1-B)^2 z_t - m$, as the input series, where z_t is the catch in year t and m is a constant. The maximum likelihood estimates of the parameters of the above model are

$$\hat{\varphi}_1 = 0.772433 \quad \hat{\varphi}_2 = 0.590124$$
$$\hat{\varphi}_3 = 0.503512 \quad \hat{\theta}_1 = -0.383400$$
$$\hat{\sigma}^2 = 98.437 \quad \hat{m} = -0.25979$$

The estimated model in terms of the catches (z_t) is given in equation (10).

$$(1 - B)^2\ (z_t + 0.772433\ z_{t-1}$$
$$+ 0.590124\ z_{t-2} + 0.503512\ z_{t-3}) =$$
$$\varepsilon_t - 0.383400\ \varepsilon_{t-1} + 0.744576 \quad . . . (10)$$

Figure 3 a. Shows the plots of the transformed input series and its forecast using the estimated model. The actual landings and their forecasts are ploted in Fig. 3 b. The forecast made for 1992 by using the above model is 39.4 thousand tonnes with 95% confidence limits 39.4 ± 20.2 thousand tonnes.

The adequacy of the fitted model was tested by using the estimated error terms and the value of the chi-square computed is 16.7 with 16 degrees of freedom which is nonsignificant. This shows that the fitted model is adequate.

during 1950-60 period was 16,030 t with CV 30.0%, in 1961-70 period it was 31,850 t which CV 36.3% in 1971-80 period it rose to 43,106 t with CV 19.2% and in 1981-91 period the average catch reached 63,782 t with CV 18.6%.

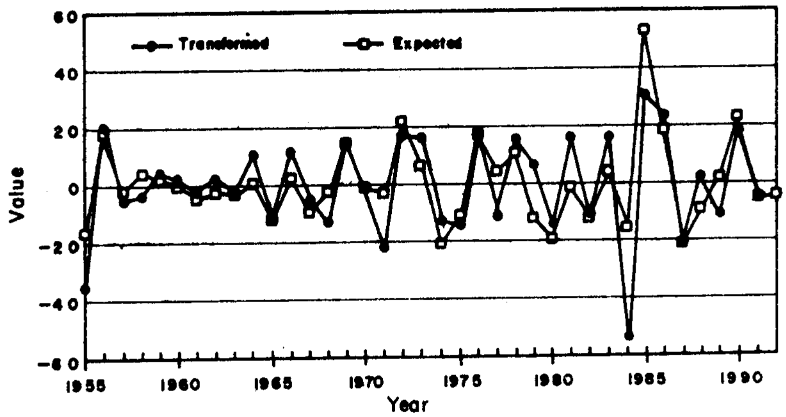

FIG. 4 a. Plot of transformed values, model : ARIMA (3, 2, 1) and their expectations of all India Silverbellies landings.

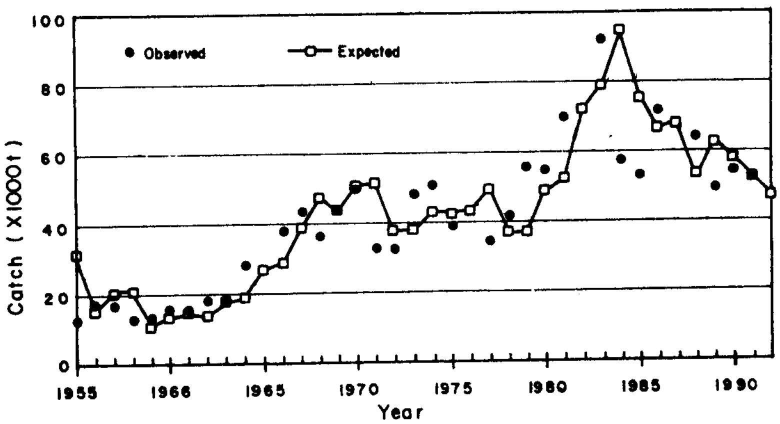

FIG. 4 b. Plot of all India Silverbellies landings, model : ARIMA (3, 2, 1) and predicted landings.

Silverbellies landings

Over years there is increase in all India silverbellies landings. From a catch of 15,274 t in 1950 increased to 91,733 t in 1983. Thereafter it showed a slight decrease and in 1991 the catch was 52,832 t. The average catch

From these figures it is can be seen that catch is more or less steady and maximum in the last decade.

Catch data for the period 1950-'91 was made stationary by applying second order difference and used for the analysis after centering it to the average. The orders of AR and MA coefficients were then estimated as 3 and 1 respectively. So the model found suitable is ARIMA (3, 2, 1) with $y_t = (1 - B)^2 z_t - m$ as the input series where z_t is the catch in year t and m is a constant. The maximum likelihood estimates of the parameters of the above model are

$$\hat{\varphi}_1 = 0.809937 \qquad \hat{\varphi}_2 = 0.726779$$
$$\hat{\varphi}_3 = 0.390921 \qquad \hat{\theta}_1 = -0.434474$$
$$\hat{\sigma}^2 = 102.06 \qquad \hat{m} = -0.00551$$

The model in terms of the actual catch (z_t) can then be written as in equation (11)

$$(1-B)^2 (z_t + 0.809937\ z_{t-1} + 0.726779\ z_{t-2} + 0.390921\ z_{t-3}) = \varepsilon_t - 0.434474\ \varepsilon_{t-1} + 0.060985 \qquad \ldots (11)$$

Figure 4 a. Shows the plot of the transformed input series and its predictions with the estimated model. The actual landings and their forecasts are ploted in Fig. 4 b. The estimated catches for 1992 is 46.8 thousand tonnes with 95% confidence limits 46.8 ± 20.6 thousand tonnes. The value of the chi-square computed by using the estimated error terms is 19.4 with 16 degrees of freedom which is not significant and shows the adequacy of the fitted model.

The above study shows that marine fish landings can be modelled by using ARIMA techniques, which in the absence of other auxiliary informations is a powerful tool for forecasting. The forecast can be improved by considering the marine landings of other species as a multiple time series which may reveal their inter influence if any. But their analysis is highly complex and costly.

References

Alagaraja, K. 1984. Simple method for estimation of parameters for assessing exploited fish stocks. *Indian J. Fish.*, **31** (2) : 177-208.

Box, G. E. P. and G. M. Jenkins 1976. *Time series analysis. Forecasting and control.* San Francisco; Holden-Day.

Godolphin, E. J. 1977. A direct representation for the maximum likelihood estimator of a Gaussian moving average process. *Biometrika*, **64** (2) : 375-384.

——— 1978. Modified maximum likelihood estimation of Gaussian moving average using a pseudoquadratic convergence criterion. *Ibid.*, **65** (1) : 203-206.

——— 1984. A direct representation for the large sample maximum likelihood estimation of a Gaussian autoregressive moving average process. *Ibid.*, **71** (2) : 281-289.

Hannan, E. J. and L. Kavalieris 1984. A method for autoregressive moving average estimation. *Ibid.*, **72** (2) : 273-280.

Jensen, A. L. 1976. Time series analysis and forecasting of Atlantic menhaden catch. *Chesapeake Sci.*, **17** : 305-307.

——— 1985. Time series analysis and the forecasting of Menhaden catch and CPUE. *North American J. Fish. Management*, **5** : 78-85.

Lung, G. M. and G. E. P. Box 1978. On a measure of lack of fit in time-series models. *Biometrika*, **65** (2) : 297-303.

Mendelssohn, R. 1980. Using Box-Jenkins models to forecast fishery dynamics : Identification estimation and checking. *Fish. Bull. U.S.*, **78** : 887-896.

Moran, P. A. P. 1953. The statistical analysis of the Canadian lynx cycles. I– structure and prediction. *Aust. J. Zool.*, **1** : 163-173.

Saila, S. B., M. Wighbout and R. J. Lermit 1980. Comparison of some time series models for the analysis of fisheries data. *J. Cons. Int. Explor. Mer.*, **39** : 44-52.

Srinath, M. and K. K. Datta 1985. Forecasting marine products exports – Time series analysis. *Indian J. Fish.*, **32** (2) : 264-267.

Stocker, M. and R. Hilborn 1981. Short-term forecasting in marine fish stocks. *Can. J. Fish. Aquat. Sci.*, **38** : 1247-1254.

Van Winkle, W., B. L. Kirk and B. W. Rust 1979. Periodicitis in Atlantic Coast striped bass (*Morone saxatilis*) commercial fisheries data. *J. Fish. Res. Bd. Canada*, **36** : 54-62.