

## Modelling and forecasting fish catches : Comparison of regression, univariate and multivariate time series methods

R. VENUGOPALAN AND M. SRINATH

*Central Marine Fisheries Research Institute, Cochin - 682 014, India*

### ABSTRACT

Three different statistical modelling procedures, viz. deterministic regression modelling, univariate time series and multivariate time series modelling approaches were evaluated on the basis of their efficiency with a view to modelling and providing accurate operational forecasts of the quarterly commercial landings of seven species of marine fishes along with the total landings of Tamil Nadu. The forecasts were based on the database of 1975-'96. Sound statistical techniques were utilised to select an appropriate model. The selected models were then used to forecast species-wise landings for the next two more years.

### Introduction

Statistical modelling essentially consists of developing a model to adequately represent the salient features of the problem under study. Subsequently, it is used to forecast future values of the underlying phenomenon which may be for example, commercial landings of some important fish species. To model such a dynamic system, methods used in the literature (Stergiou *et al.*, 1997) are oriented towards the following.

a) Modelling on the basis of deterministic, regression techniques that explain changes in fishery variables (e.g. catch, catch per unit of fishing effort) in terms of changes in various biotic (e.g. spawning stock, predators, competitors) and/or abiotic variables (e.g. fishing effort, climate).

b) Modelling on the basis of univariate time series techniques that treat the system as a black box, viewed as an unknown generating process and forecasting is based on projecting past values of a variable and/or past errors into the future.

c) Models that synthesise the above mentioned two general approaches (multivariate time series).

In the present study under the first approach of regression modelling, polynomials of suitable orders and time varying seasonal regression models (TVS) are developed. The second part consists in developing Winters' (Winters, 1960) Exponential Smoothing model (WES) and Auto Regressive Integrated Moving Average (ARIMA) time series models. On the other hand, an

attempt has been made to fit harmonic multivariate regression (HREG) models, under the last approach. All these techniques are employed to seven different commercially important fish species. A suitable model in each of the above eight situations is selected not only on the basis of computing the frequently used measure of model accuracy, namely,  $R^2$ , the coefficient of determination, but also on testing the randomness or white noise assumption of the model generated residuals. Furthermore, an attempt has been made to examine the accuracy of the resultant model by comparing the last two observations with the values obtained by fitting the selected model in all the data sets after ignoring the last two data values. Finally forecast of fish landings is done utilising the chosen model for next two more years. The STATISTICA (Release 4.5) package available at CMFRI, Cochin was used for data analysis.

**Materials and methods**

The quarterly landings of seven commercially important fish groups, viz. elasmobranchs, other sardines, perches, silverbellies, croakers, carangids, and penaeid prawns along with total landings from 1975-'96 in Tamil Nadu obtained through a stratified multistage random sampling procedure developed by the Central Marine Fisheries Research Institute, Cochin form the database for this study. The above mentioned statistical modelling procedures, as discussed below, are employed to these data sets, with a view to selecting the most appropriate model for each of these eight data sets.

(i) *Regression modelling* : At the first instance, polynomial function of suitable degree is fitted. Denoting the

dependent variable of interest as,  $y_t$  (here, it refers to the fish species catches at time period, t) the polynomial function of  $k^{th}$  degree is given by :

$$Y_t = b_0 + b_1t + b_2t^2 + \dots + b_k t^k + e_t \dots \dots \dots (1)$$

where  $b_i$ 's are the unknown parameters to be estimated and  $e_t$  denotes the error term following some continuous statistical distribution. Eq.(1) of suitable degree will be selected by observing the decrease in the residual sum of squares at each stage. Specifically, eq.(1) will be chosen to represent the data set only when there is an insignificant decrease in the residual sum of squares due to fitting of  $(k+1)^{th}$  degree polynomial.

A special case of the above regression model is the time varying seasonal (TVS) regression model. For such models a modified version of eq. (1) can be used that copes with seasonal cycles. This is done by introducing  $(s-1)$  dummy variables (where  $s$  is the length of seasonality; in our case 4 quarters),  $D_1$  to  $D_3$ :  $D_1 = 1$ , if the quarter is first, and zero otherwise;  $D_2 = 1$ , if the quarter is second, and zero otherwise; and  $D_3 = 1$ , if the quarter is third, and zero otherwise.

Each of these three dummy variables is equivalent to a new regression :

$$y_t = b_0 + a_1 D_1 + a_2 D_2 + a_3 D_3 + b_j t^j + \dots + b_k t^k + e_t \dots \dots \dots (2)$$

and the set of three dummy variables identifies all four quarters (Makridakis *et al.*, 1983).

(ii) *Univariate time series models*: Under this approach two different categories of univariate time series models, namely, exponential smoothing and Auto Regressive Integrated Moving Average (ARIMA) models are fitted with a view

to describing the data effectively. Exponential smoothing models apply unequal exponentially decreasing weights for the averaging of past observations. In contrast, ARIMA models capture the historic autocorrelation of the data and extrapolate them into the future. They usually outperform the exponential smoothing models when the time series of data is long, not highly irregular and the autocorrelations are strong (Stergiou *et al.*, 1997). In the present study, Winters' exponential smoothing (WES) (Winters, 1960) models which can handle both trend and seasonality as well as randomness, are used. They are based on three smoothing equations such as trend, stationarity and seasonality (Stergiou *et al.*, 1997) In the multiplicative WES model it is assumed that each observation is the product of a deseasonalised value and a seasonal index:

$$S_t = \alpha(y_t / I_{t-L}) + (1 - \alpha)(S_{t-1} + b_t J)$$

$$b_t = \gamma(S_t - S_M) + (1 - \gamma) b_w$$

$$I_t = \beta(y_t / S_t) + (1 - \beta) I_t,$$

and the forecasts are computed based on:

$$F_{t+m} = (S_t + mb) I_t \dots \dots \dots 4$$

Here  $\alpha$ ,  $\beta$  and  $\gamma$  are the general smoothing, seasonal smoothing and trend smoothed coefficients respectively, taking values between the range 0 and 1;  $L$  is the length of seasonality;  $b_t$  is the trend component;  $I_t$  is the seasonal adjustment factor;  $S_t$  is the smoothed series (smoothed value at time  $t$ ) that does not include seasonality and  $F_{t+m}$  is the forecast  $m$  periods ahead. The computation of smoothing coefficients is based on the minimisation of mean squared error (MSE) and the approach to estimate these values is trial and error. In the present study, built-in-

options available in STATISTICA (Release 4.5) package is used to formulate these WES models.

ARIMA models formulated by Box - Jenkins (Box and Jenkins, 1976) assume that a time series is a linear combination of its own past values and current and past values of error terms. These models necessitate the stationarity assumption of the time series, a series for which mean and variance are constant over a period of time. It may be pointed out that differencing of the original series upto suitable order, makes the original series stationary. The general form of an ARIMA model can be written as :

$$U - \sum_{j=1}^p \alpha_j B^j - (1 - B)^d (1 - B^s)^s (1 - B^q)^q y_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) e_t \dots \dots \dots (5)$$

where  $B^p$  is the backward shift operator for which  $B^p y_t = Y_{t-p}$ ;  $\alpha_j$ ,  $\theta_j$  and  $\theta_j$  are the arithmetic coefficients; and  $e_t$  is the error term at time  $t$ . The general form of ARIMA model is referred to as:

$$ARIMA(p, d, q) (P, D, Q)^s \dots \dots \dots (6)$$

where  $p$ ,  $q$ , and  $d$  respectively denote the order of auto regression (AR), moving average (MA) and the degree of differencing needed to achieve stationarity. Moreover  $s$  is the seasonality (number of periods per season); and  $P$ ,  $D$ ,  $Q$  are seasonal terms corresponding to  $p$ ,  $d$ ,  $q$  respectively. In this approach at first instance, possible candidate models are selected to represent the data set by observing two plots, viz. auto-correlation function (ACF) and partial auto-correlation function (PACF) of the stationary series. Consequently, these chosen models are tested for improvement (by considering alternative models and/or overfitting) in terms of Bayesian Information Criterion (BIC),

MSE and  $R^2$  values, whereas the over-fitted term(s) had to have coefficients that are more than two standard errors. Maximum likelihood algorithm available in the STATISTICA package is utilised for building these ARIMA models.

*(Hi) Multivariate time series models:*

In this section the aim is to develop a multivariate time-series model, viz. Harmonic Regression (HREG) model, which synthesises the above discussed two general approaches. HREG models (Bulmer, 1974) incorporates sine and cosine terms to account for periodic variations existing in the time-series data. The general form of HREG model incorporating a set of sine and cosine waves, with known frequencies,  $c_i$  (for  $i = 1(1)k$ ), is :

$$y_t = \sum_{i=1}^k [a_{1i} \sin((a_i -) 2\pi c_i t) + a_{2i} \cos((a_i -) 2\pi c_i t)] + e_t \dots \dots \dots (7)$$

where  $a_{1i}$  and  $a_{2i}$  are the arithmetic coefficients and  $e_t$  is the error term at time  $t$ . The application of HREG requires that the frequencies  $c_i$  are known ahead of time. In this study the frequencies are estimated using Fast Fourier Transform (FFT), applied to the raw data. The five maximum periodogram values ( $od/n$ ) are observed. Thereafter, a multiple regression involving eq. (7) as a part of explanatory variables is fitted to obtain the parameter estimates.

*(iv) Measures of model adequacy :* To assess the goodness-of-fit of the models, the coefficient of determination ( $R^2$ ) statistic value is computed (Kvalseth, 1985):

$$R^2 = 1 - \frac{\sum_{j=1}^n (y_j - \bar{y})^2}{\sum_{j=1}^n (y_j - \bar{y})^2}$$

where  $n$  is the number of observations. The higher the value of  $R^2$ , better is the fitted model. In addition to this, before taking any final conclusion about the appropriateness of the fitted model, randomness assumption regarding the error term is tested. This is equivalent of testing the independency of residual autocorrelations of the fitted model. To test this assumption of residuals, the Box-Ljung statistic (Q) (Box and Jenkins, 1976) is utilised.

$$Q = \sum_{j=1}^{(n-d)} r_j^2(a)$$

where  $r_j^2(a)$  denotes the squared ACF corresponding to residuals of  $j$ th lag and  $n$  denotes the number of parameters estimated in the model. The hypothesis of adequacy of the model is not rejected at 5 % level, if the calculated Q value at various lags have the corresponding probability values (P) more than 0.05. Otherwise, the Q values indicate that the residual autocorrelations as a set are significantly different from zero at 5 % level.

**Results and discussion**

*i) Model fitting*

As discussed earlier, the three different statistical modelling procedures are applied to the quarterly data on seven different groups of commercially important fish landed along with the total landings observed in the state of Tamil Nadu during the period 1975-1996. Polynomial functions of suitable degree for describing these data sets are formulated by observing the reduction in the

TABLE 1. Results of 2nd degree polynomial fit

Species statistic	Elasmo- branches	Other sardines	Perches	Croakers	Carangids	Silver bellies	Penaeid prawns	Total
$b_0$	5001.874	8157.516	1887.626	3605.044	1690.363	6066.79	2609.94	56291.9
$b_1$	-64.441	-107.566	-18.938	-11.108	18.592	86.003	-6.640	-250.820
$b_2$	0.534	1.780	1.236	0.070	0.373	-1.576	0.674	9.950
$R^2$	0.131	0.165	0.755	0.009	0.468	0.164	0.642	0.682

residual sum of square at each model fit. It was observed that none of the models fit above 2nd degree yields effective reduction in residual sum of square. Accordingly, the results of 2nd degree polynomial fit to eight different data sets are presented in Table 1 along with the goodness-of-fit statistics. A cursory look at this Table reveals among other things that the  $R^2$  value of the model ranged from 0.009 for croakers catches, to 0.755 for perches. Furthermore, the Q statistic value at various lags for each

data fit, is computed and are presented in Table 2. A perusal of this Table indicates that except for perches, croakers and the total landings, the randomness (or independency) assumption of the error in other data sets is met.

In the next stage, time-varying seasonal (TVS) regression models are fitted to the data sets. The results obtained by this method are presented in Table 3, along with the goodness-of-fit statistic values. A perusal of this table indicates that the  $R^2$  values of the TVS models are

TABLE 2. Box-Ljung (Q) statistic along with the probability values

Model Species	Polynomial fit	TVS	ARIMA	WES	HREG
Elasmobranchs	19.660 (0.185)	24.600 (0.056)	11.0700 (0.689)	17.380 (0.299)	13.720 (0.5469)
Other sardines	17.680 (0.280)	10.440 (0.791)	08.900 (0.883)	13.350 (0.575)	10.730 (0.772)
Perches	66.41* (0.000)	22.450 (0.097)	60.59* (0.000)	34.05* (0.002)	15.330 (0.428)
Croakers	27.870* (0.022)	26.780* (0.031)	22.280 (0.101)	20.68* (0.017)	19.680 (0.185)
Carangids	9.940 (0.824)	11.110 (0.745)	18.100 (0.257)	20.470 (0.155)	9.610 (0.843)
Silverbellies	15.480 (0.418)	15.100 (0.444)	13.790 (0.542)	24.060 (0.068)	10.720 (0.772)
P. Prawns	14.700 (0.513)	11.080 (0.747)	31.80* (0.008)	14.190 (0.500)	15.660 (0.405)
Total	39.840* (0.0005)	19.210 (0.205)	7.480 (0.943)	15.53* (0.006)	11.130 (0.743)

+ Significant error autocorrelation at 5 % level.

TABLE 3. Results of TVS model fit

Species	Elasmo- branches	Other sardines	Perches	Croakers	Carangids	Silver bellies	Penaeid prawns	Total
Statistic								
$b_0$	4354.0	9429.0	519.0	3633.0	1502.0	4991.0	2342.0	50293.0
$b_1$	-64.295	-106.57	-18.449	-10.635	18.23	186.632	-6.938	-251.7
$K$	0.535	1.78	1.238	0.069	0.373	-1.575	0.675	9.96
$a_1$	505.334	834.35	1418.23	617.485	-76.716	1281.51	-226.78	5171.90
$a_2$	904.944	-2881.11	1545.872	-353.236	-209.292	1410.87	506.30	1798.84
$a_3$	1140.348	-3163.94	2402.26	-451.685	1095.432	1458.52	829.975	17067.97
$R^2$	0.219	0.25	0.83	0.11	0.539	0.192	0.697	0.785

quite high, as much as 0.830 in case of perches and as low as 0.110 in case of croakers. Comparing these  $R^2$  values with those corresponding values of 2nd degree polynomial fit, it may be noticed that eventhough there is a marginal increase in the latter case, they are lower  $R^2$  values, in case of elasmobranchs, other sardines, croakers and silver bellies. On the other hand, the  $R^2$  values of total landings and of perches are quite large, indicating the suitability of these fits in comparison with the polynomial models. But, before making any final conclusion, the randomness assumption of errors is tested in all the TVS models fit by computing the Q-statistic values at various lags and are presented in Table 2 along with their respective probability values. It may be noticed from this Table that, except for the croakers, the independency assumption of the residual autocorrelations is met in case of all the data sets.

In the next stage, Winters exponential smoothing (WES) models with multiplicative error term are constructed to describe the data sets. For all the WES models built in this study, the smoothing coefficients (not shown here) are computed based on the minimiza-

tion of MSE. The final values of the smoothing coefficients ( $\alpha, \beta, \gamma$ ) are presented in Table 4 along with the goodness-of-fit statistics. The  $R^2$  values due to this model fit, ranged from 0.810 to 0.954, showing that WES model, fits all the data sets quite well. Before making any final conclusion, the Q-statistic values are computed for different data sets and are presented in Table 2. A cursory look at these values indicates that the independency assumption of errors is met in all the data sets, except for perches, croakers and the total landings.

TABLE 4. Summary statistics of Winters exponential smoothing (WES) model fit

Data set	Alpha	Beta	Gamma	$R^2$
Elasmobranchs	0.37	0.40	0.10	0.860
Other sardines	0.60	0.50	0.10	0.810
Perches	0.50	0.40	0.10	0.954
Croakers	0.20	0.10	0.10	0.831
Carangids	0.45	0.20	0.10	0.863
Silverbellies	0.50	0.30	0.10	0.902
Penaeid prawns	0.50	0.10	0.05	0.911
Total	0.28	0.40	0.10	0.948

Univariate time-series modelling procedure is employed to select ARIMA models of suitable orders to all the data sets. Based on the ACF, the stationarity of the original series is checked. It may be noted here that in case of elasmobranchs, other sardines, croakers and silverbellies, the original series themselves are stationary. In the remaining data sets first difference of the original series is taken to achieve stationarity. Based on spikes at ACF and PACF obtained for the stationary data set, different candidate models are selected for further analysis. A final choice between different candidate models is made by comparing their respective Bayesian Information Criteria (BIC) and MSE (not reported here) values. Models with lesser BIC and MSE values are selected and the summary statistics for only these models are reported in Table 5. Perusal indicates that the R<sup>2</sup> values are quite high as much as 0.868 in case of total landings, to the lowest value of 0.407 in case of carangids data set. As a final test of adequacy of the

selected models, the Q-statistic values presented in Table 2 indicate that except for the perches and penaeid prawns data sets, the assumption of independency of errors is met in all the data sets.

Finally, Harmonic regression models (HREG) are formulated with a view to describing the data sets appropriately. Under this procedure the arithmetic coefficients (frequencies) of sine and cosine terms are initially identified using spectral analysis. The arithmetic coefficients of the HREG models fitted to the quarterly catches, using the estimated frequencies (as well as time  $t$  and  $t^2$ ) as independent variables are shown in Table 6, along with the R<sup>2</sup> values. The R<sup>2</sup> values of HREG are high ranging from 0.885 for perches to a lesser value of 0.448 for elasmobranchs. For all models, the Box-Ljung test indicates non-significant error autocorrelations (Table 2).

(ii) Comparative study

From the above model fitting exer-

TABLE 5. Summary statistics of ARIMA model fit

Data set	ARIMA model	Parameter estimates				Constant	
Elasmo-branches	(1,0,0)	0.249	-	-		3536.966	0.84
Other sardines	(1,0,0) (1,0,0) <sup>d</sup>	0.307	-	-	0.473	8444.5	0.84*
Perches	(2,1,1) (0,1,1) <sup>n</sup>	-0.871	-0.321	0.989		0.575	0.863
Croakers	(1,0,0)• (1,0,0) <sup>d</sup>	0.355	-	-	0.292	3280.4	0.891
Carangids	(0,1,1)	-	-	0.836			0.407
Silver bellies	(1,0,0)	0.394	-	-		10185.2	0.904
Penaeid prawns	(1,1,0)	0.693	-	-		3951.1	0.916
Total	(2,1,1) (1,1,0) <sup>n</sup>	-0.761	-0.275	0.948			0.868

TABLE 6. Summary statistics of HREG model fit

Species Statistic	Elasmo- branches	Other sardines	Perches	Croakers	Caran- gids	Silver- bellies	Penaeid prawns	Total
Constant	5180.660	7842.230	2343.340	206.834	1028.000	31.830	225000	5800000
<sup>a</sup> u	-0.118	0.347	-0.002	0.018	-0.020	0.324	-0.207	-0.065
a <sub>12</sub>	-0.196	-0.079	0.040	0.297	-0.210	-0.221	-2.210	-43.050
<sup>a</sup> ia	-0.228	0.142	-0.121	-0.088	-0.110	-0.239	-0.050	-0.204
<sup>a</sup> H	-0.121	-0.285	0.082	0.561	0.030	-0.195	0.130	0.048
<sup>a</sup> i5	-0.181	0.001	-0.115	-0.230	-0.070	0.093	0.138	-0.072
a <sub>21</sub>	0.278	0.251	-0.204	-0.140	-0.160	0.614	-0.105	-0.253
<sup>a</sup> 22	-0.205	0.272	-0.114	0.085	0.037	0.194	11.138	-42.560
<sup>a</sup> 23	0.080	0.199	-0.201	0.238	0.147	0.074	-0.146	-0.032
<sup>a</sup> 24	0.200	-0.211	-0.105	0.981	0.180	0.210	0.080	0.195
<sup>a</sup> 25	0.142	0.172	-0.043	-0.020	0.190	0.180	-0.062	0.903
t	-74.300	-1.121	-0.543	3.920	0.952	3.990	3.020	52.450
t <sup>2</sup>	0.660	0.990	1.469	-3.740	-0.340	-3.710	11.140	-51.800
R <sup>2</sup>	0.448	0.568	0.885	0.462	0.609	0.473	0.776	0.800

cise, the following conclusions may be drawn:

(a) The R<sup>2</sup> values for TVS models, in comparison with those of polynomial fit are invariably high in case of elasmobranchs, other sardines, croakers and silverbellies and in the remaining the corresponding increase is only marginal. However, these R<sup>2</sup> values which are still low, clearly indicating the inadequacy of these fits to the data sets. On the other hand, the R<sup>2</sup> values computed through the three time-series methods are quite satisfactory except in some odd situations.

(b) Irrespective of the satisfactory increase achieved in R<sup>2</sup> values by employing the time-series methodologies, a final decision on the model adequacy is done based on examining whether the model generated residuals are white

noise or not. From the Q-statistic values presented viz. Table 2, computed with the help of all the three statistical modelling procedures, one can notice that in case of elasmobranchs, other sardines, carangids and silverbellies, the above assumption is met. On the other hand, for perch catches, only TVS and HREG generated residuals are satisfying the randomness assumption. But the ARIMA and HREG generated residuals of croakers are independent. In case of penaeid prawns except the ARIMA, the other models generated white noise residuals. However, in case of total landings the assumption of independency of residual is satisfied in case of all but, polynomial and WES model fits.

(c) Combining these two measures of model adequacies it is suggested that, ARIMA model seems to fit four categories namely, total landings, other sar-



TABLE 7. *Forecasted landings (in '000 tonnes)*

Year		1996				1997+				1998+			
Quarter		I	II	III	IV	I	II	III	IV	I	II	III	IV
Elasmo- branchs (WES)	Obs.	4.439	3.167	4.313	3.805								
	Pred.	3.287	3.841	3.942	3.809	4.190	3.814	4.501	4.102	4.509	4.099	4.831	4.397
Other sardines (ARIMA)	Obs.	26.632	9.519	10.918	13.411								
	Pred.	17.902	11.526	8.448	15.421	23.291	16.696	19.660	22.825	23.211	21.509	21.492	22.990
Perches (HREG)	Obs.	10.885	7.636	13.845	5.869								
	Pred.	10.169	10.036	11.109	8.304	13.930	14.193	14.458	14.726	14.998	15.272	15.549	15.829
Croakers (ARIMA)	Obs.	4.319	2.546	3.298	3.244								
	Pred.	3.778	3.580	3.135	3.405	3.507	3.048	3.279	3.269	3.346	3.216	3.282	3.229
Carangids (WES)	Obs.	6.057	4.654	7.308	5.146								
	Pred.	5.323	5.721	7.450	5.715	5.152	4.960	7.159	5.463	5.215	5.021	7.247	5.530
Silverbellies (ARIMA)	Obs.	9.536	9.230	11.701	15.581								
	Pred.	10.837	9.940	11.819	10.798	12.337	11.051	10.541	10.339	10.258	10.227	10.214	10.209
P. prawn (WES)	Obs.	6.296	6.620	7.611	7.001								
	Pred.	6.225	8.064	7.795	6.329	6.175	7.873	8.422	6.972	6.490	8.208	8.776	7.262
Total (ARIMA)	Obs.	118.642	88.268	127.724	101.039								
	Pred.	117.132	106.397	123.804	109.081		94.774	121.505	106.537	112.652	94.635	124.038	104.496

+ = forecast periods, obs. = observed, pred. = predicted.

B  
a.

g

Ja-  
rs  
a

|"

dines, croakers and silverbellies effectively. While the WES model efficiently accounted for the elasmobranchs, carangids and penaeid prawn catches, the HREG models explain the perch landings in Tamil Nadu quite well as compared to the other statistical methods considered for data analysis.

#### (Hi) *Forecasting*

An attempt has been made to examine the accuracy of the selected model by comparing the last two data observations with the values obtained by fitting the selected model in all the data sets after ignoring the last two data values. These values along with the forecast of fish landing for next two years are presented in Table 7. It is observed that the predicted values, for all the data sets are quite near the actual values during 1996. This further strengthens the appropriateness of the different selected models.

#### (iv) *Biological explanations of models*

In the ARIMA models fitted to the other sardines, croakers and silverbellies and total landings data sets, the catches of a category at period  $t$  is partially predicted by the autoregressive terms at time periods  $(t-4)$  and  $(t-5)$ , and to a lesser extent at time-periods  $(t-1)$  and  $(t-6)$ . Hence, these models predicted persistence of catches. In other words, everything being equal, once catches are high they tend to remain high for 2-3 successive time-periods. Persistence may indicate that environmental conditions favouring the formation of good year classes affecting the fisheries of the species (or group of species) of concern tend to persist. In particular, silverbellies catches are purely autoregressive of

order one indicating apart from random fluctuations, the present period catch is predicted solely by the last period catch. A seasonal autoregressive model of order one with the above arguments, represents the other sardines and croakers catches. In the WES models selected to represent the elasmobranchs, carangids and prawns catches, the recent period catches of these fish species are given relatively more weights in forecasting than the older observations. As expected these models capture the highly irregular catches more effectively. HREG models developed for the perch catches account for the periodic variations present in the data sets.

#### (v) *Further scope*

The above study is only a beginning of modelling and forecasting fisheries catches in our country. It is a challenging task to construct statistical models, which not only handle the economic aspects but also take into account the highly fluctuating trends, which are highly in need to explain the present day fisheries. Studies incorporating the above aspects need to be taken on priority in order to assess our country's marine fishery resources effectively, thus helping in framing suitable management strategies for the future.

#### References

- Box, G. E. P. and G. M. Jenkins 1976. *Time Series Analysis, Forecasting and Control*. Holden-Day, San Francisco, CA.
- Bulmer, M. G. 1974. A statistical analysis of the 10-year cycle in Canada. *J. Anim. Ecol.* 43 : 701-718.
- Kvalseth, T. O. 1985. Cautionary note about  $R^2$ . *The American Statistician*, 39(4) : 279-285.

Makridakis, S., S. Wheelwright and V. McGee  
1983. *Forecasting : Methods and Applications*. John Wiley and Sons.

Stergiou, K. I., E. D. Christou and G. Petrakis  
1997. Modelling and forecasting monthly fisheries catches :

Comparison of regression, univariate and multivariate time series methods. *Fish. Res.*, 29 : 55-95.

Winters, P. R. 1960. Forecasting sales by exponentially weighted moving averages. *Manage. Sci.*, 6 : 324-342.