ELSEVIER

Fisheries Research 25 (1996) 105-138

FISHERIES RESEARCH

Modelling and forecasting annual fisheries catches: comparison of regression, univariate and multivariate time series methods

K.I. Stergiou*, E.D. Christou

National Centre for Marine Research, Agios Kosmas, Helliniko, 16604, Athens, Greece

Accepted 8 February 1995

Abstract

In the present work, eight forecasting techniques are evaluated on the basis of their efficiency to model and provide accurate operational forecasts of the annual commercial landings of 16 species or groups of species in the Hellenic (Greek) marine waters. The development of operational forecasts was based on the following four general categories of forecasting techniques: (a) deterministic simple or multiple regression models incorporating different exogenous variables (time-varying regression, TV; multiple regression models incorporating time, fishing effort, wholesale value of catch and climatic variables, MREG); (b) univariate time series models (Brown's one parameter exponential smoothing, BES; Holt's two parameter exponential smoothing, HES; and AutoRegressive Integrated Moving Average (ARIMA)); (c) multivariate time series techniques (harmonic regression, HREG; dynamic regression, DREG; and vector autoregressions, VAR); and (d) the 'biological' exponential surplus-yield model, FOX. Fits (for 1964–1987) and forecasts (for 1988–1989) obtained by the different models were compared with each other and with those of a naive method (NM) and an empirical one (i.e. combination of forecasts, EMP) using 32 different measures of accuracy.

The results revealed that HREG and MREG models outperformed, in terms of fitting accuracy, the remaining eight models (NM, TV, BES, HES, FOX, ARIMA, VAR and EMP). They were both characterised by: (a) higher accuracy in terms of all, or most, standard and relative statistical measures; (b) unbiased fits; (c) much better performance than NM; (d) transformed errors which were essentially white noise. In addition, HREG and MREG models: (e) explained from 79% to 97% of the variability of 13 transformed annual catches and from 31% to 61% for the remaining ones; (f) produced fits with MAPE values ranging from 3.4% to 21.2%; (g) in all, or most cases, predicted the between year variations during the fitting period, 1964–1987. In terms of forecasting performance, however, not a single best approach was found for the 16 annual catches. In general, BES and, to a lesser extent, HES, NM (which actually is an ARIMA (0,1,0)), EMP and HREG models were among the best performers more often, produced the worst forecasts more rarely and were generally characterised by the higher number of stable forecasts and of forecasts with MAPE <20% and <10%.

* Corresponding author. Tel. (301)941 8075; Fax. (301) 983 3095.

SSDI 0165-7836(95)00389-4

TV was also efficient for some annual series. Conversely, the poorest performers (FOX, MREG and ARIMA) rarely did better than average. The biological FOX models produced the least accurate and biased fits, bad forecasts ($>34.1\%$) in two out of four cases, and were characterised by transformed errors that were significantly ($P<0.05$) autocorrelated. Some of the empirical models also had interesting explanations. Hence, the univariate ARIMA and multivariate VAR time series models predicted persistence of catches. The multivariate VAR and HREG time series models also predicted cycles in the variability of the catches with periods of 2–3 years. Moreover, FOX and MREG models indicated that fishing effort, wholesale value of catch and climate may, in a synergetic fashion, affect long-term trends and short-term variation in the catches of, at least, some species (or groups of species). Finally, MREG and VAR models predicted that variability and replacement of anchovy by sardine catches are not due to chance, and wind activity over the northern Aegean Sea may act as a forcing function.

Keywords: Forecasting; Annual catches; Multiple regression; Time-varying regression; Surplus-yield; Exponential smoothing; Dynamic regression; Harmonic regression; Vector autoregression

1. Introduction

Forecasting is an interesting subject, partly because it is difficult to do it accurately, inasmuch as forecasters are confronted by all sorts of uncertainty (Hilborn, 1987; Getz et al., 1987), but mainly because it plays a central role in management: it precedes planning which, in turn, precedes decision making (Makridakis et al., 1983). Policy makers establish goals and objectives, seek to forecast uncontrollable events, then select appropriate actions that, hopefully, will result in the realisation of the goals and objectives.

Apart from methods based on biological principles (e.g. Fox, 1970; Shepherd, 1984; Pope and Shepherd, 1985; Borges, 1990), a variety of statistical techniques have also been used/adapted to fisheries forecasting. These methods are oriented towards (Stergiou, 1991a): (a) modelling on the basis of deterministic, regression techniques that explain changes in fishery variables (e.g. catch, recruitment) in terms of changes in various biotic (e.g. spawning stock) 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; and (c) models that synthesise the above mentioned two general approaches (multivariate time series). Various studies indicate that although time series models do not have built-in stock structure, they should not be dismissed (e.g. Stocker and Noakes, 1988; Noakes et al., 1990; Stergiou, 1991b). This is especially true of cases for which time series of biological data (e.g. catch-at-length/age) on various species are lacking, as is the case in Hellenic and eastern Mediterranean waters, in general, a fact rendering the application of such 'biological' forecasting models impossible (with the exception of surplus-yield models).

In the present study, eight forecasting techniques are evaluated on the basis of their efficiency to provide accurate fits and operational forecasts (sensu Bocharov (1989): 1 or 2 years ahead) of the annual commercial landings of 16 species or groups of species in the Hellenic marine waters. Modelling the catches of a single species is important because it may cast light on the factors affecting its fishery dynamics and may provide forecasts of its future catch level. In addition, modelling the combined catch of a group of species is also of primary importance because: (a) it reflects the situation from the fishers' and fishing industry's viewpoints (Mendelssohn and Cury, 1987); and (b) the combined catch derived from a marine region possibly reflects the carrying capacity of the region.

The development of operational forecasts was based on the following four general categories of forecasting techniques: (a) deterministic simple or multiple regression models incorporating different exogenous variables (time, fishing effort, wholesale value of catch and climatic variables); (b) univariate time series models (Brown's (1963) one parameter exponential smoothing; Holt's (1957) two parameter exponential smoothing; and Auto-Regressive Integrated Moving Average (ARIMA)); (c) multivariate time series techniques (harmonic regression, dynamic regression and vector autoregressions); and (d) the 'biological' exponential surplus-yield model (Fox, 1970). Fits (for 1964–1987) and forecasts (for 1988–1989) obtained by the different forecasting techniques were compared with each other and with those of a naive and an empirical one (i.e. combination of forecasts) using 32 different measures of accuracy. Overall, 104 annual models were built (without taking into account the naive and empirical ones). Such a large scale comparison is conducted for the first time in fisheries forecasting. Analogous comparisons have been undertaken in other fields (e.g. hydrology: three techniques applied to 30 monthly riverflow Series; Noakes et al., 1985) with the most famous and large one being the M-Competition (Goodrich, 1989) in which 1001 yearly, monthly and quarterly time series of microeconomic, macroeconomic, industrial and demographic data were collected and analysed using 20 different forecasting methods.

2. Materials and methods

2.1. Sources of data

Fisheries statistics for the Hellenic waters have been recorded since January 1964 by the National Statistical Service of Hellas (NSSH Bulletins, 1965–1992). For a better evaluation of the available data, the waters fished by the Hellenic vessels have been divided into 18 statistical fishing subareas (Fig. 1). Fishing subareas 1 and 2 (not shown in Fig. 1) refer to the Atlantic Ocean and the northern coast of Africa, respectively. Catch data are collected directly from a sample of fishing vessels that are surveyed by local customs authorities (stratified random sampling). For each vessel surveyed, a statistical questionnaire is completed showing the quantities of each major fish species (or group of species) caught during the previous month (or that the vessel did not work during that period).

In general, the Hellenic fishing fleet includes: (a) fishing vessels operating in distant waters (Atlantic Ocean and northern African coast, and thus of no concern to the present study); (b) trawlers operating in Hellenic open-sea waters; (c) purse seiners operating in Hellenic open-sea and coastal waters; (d) beach seiners operating along the Hellenic coasts; and (e) 'other coastal boats' (including small ring netters, drifters, liners, etc.) operating along the Hellenic coasts. Since 1969 the catches of the smaller inshore ring netters, drifters and liners (i.e. boats with engine horsepower of less than 20 HP) have not been recorded by the local customs authorities.

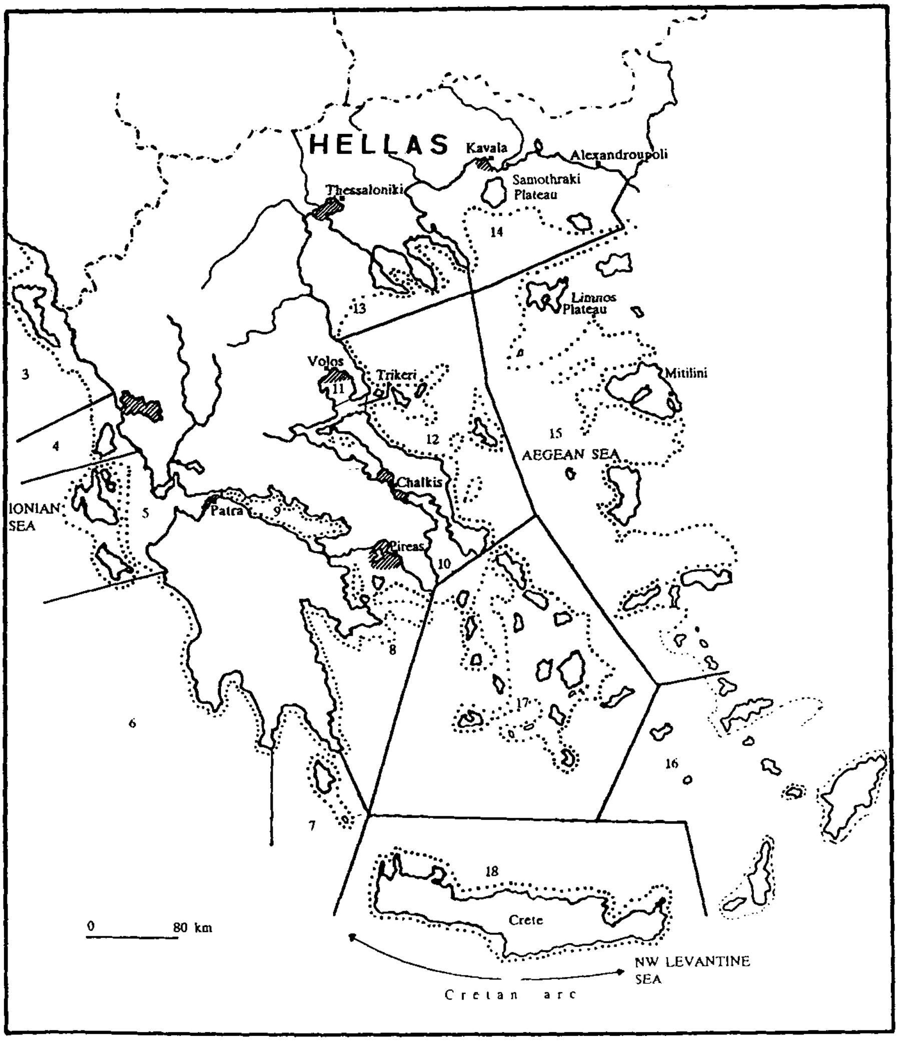

Fig. 1. Map showing the 16 Hellenic fishing subareas. The dotted line represents the 200 m isobath whereas hatched areas show areas where anthropogenic eutrophication is locally important.

In the present study an attempt is made to model and forecast the annual commercial catch weights of 16 species (or groups of species) in the Hellenic marine waters using historical fisheries time series for the years 1964–1989. The following time series were considered:

(1) total (i.e. all fishing subareas combined) annual commercial catches of trawlers, beach seiners, purse seiners and 'other coastal boats';

(2) total annual commercial catches of anchovy (*Engraulis encrasicolus*), sardine (*Sardina pilchardus*), bogue (*Boops boops*), red pandora (*Pagellus erythrinus*), gadiformes (hake *Merluccius merluccius* and blue whiting *Micromesistius poutassou*), *Trachurus* spp., *Scomber* spp., *Mullus* spp. and *Spicara* spp.; and

(3) total annual commercial catches of fishes, cephalopods and crustaceans.

Anchovy, sardine, bogue, *Scomber* spp., *Trachurus* spp., hake, blue whiting, red pandora, *Mullus* spp. and *Spicara* spp. are the most important fish species in the Hellenic marine waters either in terms of abundance (their combined annual catch represents more than 70% of the total Hellenic marine catch: Stergiou and Petrakis, 1993; Stergiou and Pollard, 1994) or in terms of commercial value.

The independent variables used for the development of the models are listed in Table 1. Monthly climatic data for the Hellenic marine waters are available since the late 1940s in

Table 1
Time series of independent variables used for the development of multivariate models. Hellenic waters, 1964–1989

Time series	Symbol	Source
Fishing effort		
Number of fishers (total)	FI	NSSH (1968–1992)
Trawlers	FIT	
Purse seiners	FIP	
Beach seiners	FIB	
Other coastal boats	FIC	
Demersal*	FID	
Number of boats (total)	BO	
Trawlers	BOT	
Purse seiners	BOP	
Beach seiners	BOB	
Other coastal boats	BOC	
Demersal*	BOD	
Engine horse power (total)	HP	
Trawlers	HPT	
Purse seiners	HPP	
Beach seiners	HPB	
Other coastal boats	HPC	
Demersal*	HPD	
Boat tonnage (total)	TO	
Trawlers	TOT	
Purse seiners	TOP	
Economic variables		
Value of catch		
Trawlers	VAT	
Purse seiners	VAP	
Beach seiners	VAB	
Other coastal boats	VAC	
Climatic variables		
Sea surface temperature	SST	COADS database
Air temperature	AIRT	
North–south wind	NSW	
Wind speed cubed	WISC	
Sea-level pressure	SLP	

*Sum of corresponding values of trawlers, beach seiners and 'other coastal boats'.

the COADS database, an extensive meteorological database of individual meteorological ship observations, currently re-processed within the framework of the CEOS project (Bakun et al., 1993). In the present study, we used the climatic time series from three 1° squares of the North Aegean Sea (38–39°N and 24–25°E; 38–39°N and 25–26°E; and 39–40°N and 25–26°E) because: (a) the North Aegean Sea represents the most important Hellenic fishing and spawning grounds for both demersal and pelagic fishes (Stergiou and Pollard, 1994; Stergiou et al., 1994); (b) the mean monthly number of observations in the selected squares was the highest (25, 16 and 18, respectively) whereas the number of missing values was the lowest (only three to four missing values per square) during the 1964–1989 period. The monthly sea surface temperature (SST) records in the three 1° squares were highly significant ($P<0.001$) correlated with each other (for all three combinations: $r>0.97$, $n=312$, $P<0.001$). The same was also true of air temperature (AIRT) ($r>0.95$, $n=312$, $P<0.001$), north–south wind component (NSW) ($r>0.62$, $n=312$, $P<0.001$), wind speed cubed (WISC) ($r>0.52$, $n=312$, $P<0.001$) and sea-level pressure (SLP) ($r>0.81$, $n=312$, $P<0.001$). Hence, for each climatic variable, the monthly records were averaged over the three 1° squares, weighted by the number of observations. Annual values were then computed from monthly averages.

2.2. *Forecasting models*

2.2.1. *Deterministic multivariate models*

The general form of a linear multiple regression (MREG) model in the time domain is:

$$X_t = \mathrm{a} + \mathrm{b}_1 Y_{1t} + \mathrm{b}_2 Y_{2t} + \ldots + {}_k Y_{kt} + e_t \quad \text{for } t=1 \text{ to } N$$

where X_t is the value of the dependent variable of interest at time t, Y_1 to Y_k are k independent variables and e_t is an error term, at time t, which is assumed to be sampled independently from a normal distribution. MREG models were developed for all annual catches using a variety of independent variables (Table 1). The variables that entered into the final models were selected through stepwise variable selection. All MREG models were developed using STATGRAPHICS/PLUS for DOS (STCS, 1993) with default criteria. MREG models in which the forecasted variable, X_t, is a function of one or more independent variables at the same time point, Y_t, are often of no real-time forecasting power when the independent variable(s) is not controllable (e.g. climate, value) as opposed to controllable ones (e.g. time, fishing effort). This is the result of the considerable time lag required for the exchange of statistics between the appropriate authorities. In contrast, all univariate time series models are of real-time forecasting power. Hence, in the present study MREG models were developed using lagged (by one year) independent variables.

A special case of regression models is the time-varying (TV) regression and the 'biological' exponential surplus-yield (FOX) model (Fox, 1970). In TV models the dependent variable of interest, X_t, is regressed against time t:

$$X_t = a + a_1 t + a_2 t^2 + \ldots + a_k t^k + e_t$$

where $t=1$ to N. The definition for $t=1$ for the first observed value is arbitrary. When $k=1$ the above equation refers to a linear trend (i.e. $X_t = a + bt + e_t$), when $k=2$ to a quadratic trend (i.e. $X_t = a + bt + ct^2 + e_t$) and so on.

In FOX model the catch/effort U, fishing effort F, and maximum sustained yield MSY, are related as follows (Fox, 1970):

$$U = U_\infty e^{-bF}$$

which after taking the logs is transformed to the linear form:

$$\ln U = \ln U_\infty - bF$$

and

$$Y = U_\infty F e^{-bF}$$

Although surplus-yield models have been widely employed for managing various species or groups of species in different marine regions (e.g. Pauly, 1989; Sparre et al., 1989), their application to multi-species fisheries, such as the Hellenic fisheries (Stergiou and Petrakis, 1993), is not without theoretical problems and practical pitfalls (Sparre et al., 1989). However, despite their shortcomings, surplus-yield models may be cautiously used when biological data on populations is not available (Pitcher and Hart, 1982; Sparre et al., 1989), as is the case in Hellenic waters. For multi-species fisheries, surplus-yield models can be applied to the total catch of all species combined (Sparre et al., 1989). FOX models were fitted to the annual catch weights and fishing effort per component fishery (trawl, purse seine, beach seine and 'other coastal boats').

2.2.2. Univariate time series models

The use of a 'naive' model (NM) helps us to decide whether or not the improvement achieved from going from a simple model to a sophisticated model is worth the time and cost involved (Makridakis et al., 1983). The model used in the present study uses as a forecast at time $t+1$ (F_{t+1}) the catch one time period ago (X_t). This method is actually an ARIMA (0,1,0) model (see below).

Two different exponential smoothing models were used in the present study: Brown's (1963) one parameter linear or quadratic exponential smoothing (BES) and Holt's (1957) double exponential smoothing (HES). BES and HES have never been applied before to fisheries forecasting. The linear BES model allows exponential smoothing with one general smoothing constant and can cope with trend. The equations used for the implementation of this technique are:

$$S'_t = X_t + (1-\alpha) S'_{t-1}$$

$$S''_t = \alpha S'_t + (1-\alpha) S''_{t-1}$$

$$a_t = S'_t + (S'_t - S'''_t) = 2S'_t - S''_t$$

$$b_t = (\alpha/(1-\alpha))(S'_t - S''_t)$$

and

$$F_{t+m} = a_t + mb_t$$

where α is the general smoothing coefficient with values ranging between 0 and 1; b_t is the estimated trend component at time t; S'_t is the single exponential smoothed series (smoothed value at time t); S''_t is the double exponential smoothed series; and F_{t+m} is the forecast m periods ahead. The linear BES model is generally used when the time series in concern exhibits a linear trend. The quadratic BES model incorporates an additional level of smoothing (triple smoothing), $S'''_t = \alpha S''_t + (1-\alpha) S'''_{t-1}$, and the equations for a_t, b_t and F_{t+m} are more complicated (see Brown, 1963). The quadratic BES model is generally used when the time series in concern exhibits a trend which can be characterised as a second order polynomial.

HES model is similar to the linear BES model with the exception that it uses two smoothing constants and smoothes the trend separately. The equations used for the implementation of HES are:

$$S_t = \alpha X_t + (1-\alpha)(S_{t-1} + b_{t-1})$$

$$b_t = \gamma(S_t - S_{t-1}) + (1-\gamma) b_{t-1}$$

$$F_{t+m} = S_t + m b_t$$

where α and γ are the general smoothing and trend smoothing coefficients, respectively, with values ranging between 0 and 1; b_t is the estimated trend component at time t; S_t is the smoothed series (smoothed value at time t); and F_{t+m} is the forecast m periods ahead.

For all exponential smoothing models built in the present study the computation of the smoothing coefficients was based on the minimisation of the Bayesian Information Criterion, mean squared error and mean absolute percentage error (described in next section), in that order of importance, and the approach to estimate these values was trial and error. BES models were developed using STATGRAPHICS/PLUS for DOS (STCS, 1993) whereas HES models using FORECAST/PRO for WINDOWS (Stellwagen and Goodrich, 1993).

ARIMA models assume that a time series is a linear combination of its own past values and current and past values of an error term (Box and Jenkins, 1976). ARIMA models capture the historic autocorrelations of the data and extrapolate them into the future. They apply to stationary time series (i.e. time series with no systematic change in mean and variance; Box and Jenkins, 1976; Makridakis et al., 1983). First or second order differencing handles problems of non-stationary mean, and logarithmic (or power) transformation of the raw data handles non-stationary variance. The general form of a non-seasonal ARIMA model can be described by the following equation:

$$(1-\phi_1 B^p)(1-B^d) X_t = (1-\theta_1 B^q) e_t$$

where X_t is the value of the variable of interest at time t; B^p is the backward shift operator for which $B^p X_t = X_{t-p}$; ϕ_1 and θ_1 are the arithmetic coefficients; and e_t is the error term at time t. The general form of the ARIMA models is referred to as ARIMA (p,d,q), where p is the order of the autoregressive term (AR term); d is the degree of differencing involved to achieve stationarity (I term); and q is the order of the moving average term (MA term). Identification of the appropriate model used (i.e. how many terms to be included in the model) was based on the examination of the autocorrelation (ACF) and partial autocorrelation (PACF) functions (not shown here) of the differenced, log-transformed time series (Box and Jenkins, 1976; Makridakis et al., 1983). Consequently, initial models were tested for improvement (by considering alternative models and/or overfitting) in terms of BIC, MSE, MAPE and r^2 values, in that order of importance, whereas the overfitted term(s) had to have coefficients that were > 2SE. All ARIMA models were built using the approximate maximum likelihood algorithm of McLeod and Sales (1983).

2.2.3. Multivariate time series models

The ordinary MREG models developed were subsequently examined for inclusion of Cochrane–Orcutt autoregressive error terms (Cochrane and Orcutt, 1949), autoregressive terms of the dependent variable and lagged terms of the independent variable(s). The general form of a dynamic regression (DREG) model including κ autoregressive terms of the dependent variable of interest, X_t, μ lags of an independent variable, Y_t, and λ Cochrane–Orcutt autoregressive error terms, e_t, is:

$$X_t = c + a_t X_{t-1} + \ldots + a_\kappa X_{t-\kappa} + b_0 Y_t + b_1 Y_{t-1} + \ldots + b_\mu Y_{t-\mu} + \theta_0 e_t + \theta_1 e_{t-1} + \ldots + \theta_\lambda e_{t-\lambda}$$

For all MREG models a Lagrange multiplier test (Engle, 1984) was performed for the first three lags of the Cochrane–Orcutt errors, dependent and independent variable(s). New terms were included in the models when the Lagrange multiplier test was significant at the level 0.01 (Stellwagen and Goodrich, 1993). All DREG models were developed using FORECAST/PRO for WINDOWS (Stellwagen and Goodrich, 1993).

Harmonic regression (HREG) models incorporate sine and cosine terms to account for periodic variations. This technique was applied for the first time to fisheries by Bulmer (1974). The general form of an HREG model incorporating a set of sine and cosine waves, with known frequencies, f_i (for $i = 1$ to k), is:

$$X_t = \sum_{i=1}^{k} [b_{1i}\sin[(f_i t/n)2\pi] + b_{2i}\cos[(f_i t/n)2\pi]] + e_t$$

where X_t is the value of the variable of interest at time t, b_{1i} and b_{2i} are the arithmetic coefficients, estimated using ordinary least squares regression techniques, and e_t is the error at time t. The application of HREG requires that the frequencies f_i are known ahead of time. In the present study the frequencies f_i were estimated using Fast Fourier Transform (FFT), applied to the log-transformed and detrended annual catches. The frequencies that entered into the final HREG models were selected through stepwise variable selection. FFT and HREG models were developed using STATGRAPHICS/PLUS for DOS (STCS, 1993).

Vector autoregression (VAR) models allow interdependence among a set of variables (called internal variables). Each internal variable is regressed against its own value in each of the n preceding periods, against the values in each of the n preceding periods of all other variables included in the model, and against a constant term and external variable(s) (Schlegel, 1985). The equations are estimated individually in order to compute the arithmetic coefficients and the constant. Consequently, the reduced form of the system is calculated and forecasts are produced. Although VAR models were applied to fisheries forecasting for the first time by Stergiou (1991a), they are of increasing interest to economic forecasters (Litterman, 1979; Todd, 1984; Schlegel, 1985). The general form of a VAR model including two internal variables, X and Y, n lags, and an external variable Z, is as follows (Schlegel, 1985):

$$X_t = c_1 + a_{11}X_{t-1} + a_{21}X_{t-2} + \dots$$
$$+ a_{n1}X_{t-n} + b_{11}Y_{t-1} + b_{21}Y_{t-2} + \dots + b_{n1}Y_{t-n} + d_1Z_t + e_1$$
$$Y_t = c_2 + a_{12}X_{t-1} + a_{22}X_{t-2} + \dots$$
$$+ a_{n2}X_{t-n} + b_{12}Y_{t-1} + b_{22}Y_{t-2} + \dots + b_{n2}Y_{t-n} + d_2Z_t + e_2$$

where X_{t-n} is the value of the variable of interest n periods before time t; c is a constant; a, b and d are the arithmetic coefficients of the model; and e_1 and e_2 are the error terms. The arithmetic parameters of the equations are estimated using ordinary least squares regression.

In the present study, VAR models were used to model and forecast the catches of the following two complexes: sardine/anchovy and trawl/purse seine. The above mentioned complexes were selected because: (a) VAR models take into account interactions between variables; (b) the replacement of sardine by anchovy has been observed throughout the Mediterranean Sea (Stergiou, 1992) and elsewhere (e.g. Daan, 1980); and (c) the replacement of pelagic fisheries by demersal fisheries has been observed in various regions of the world ocean (e.g. Caddy and Sharp, 1986). The tested VAR models included two internal variables (the above mentioned complexes) and involved lag lengths ranging from 1 to 3 lags as well as various external variables. VAR models were developed using TSP for DOS (Hall et al., 1990).

2.2.4. Empirical (EMP) models

An empirical approach to forecasting is by combining the forecasts from the different models used. In this case the forecast at time $t+1$, F_{t+1}, can be the unweighted mean of the F_{t+1} of all and/or of the best fitting/forecasting models. The success achieved by combining forecasts produced by different methods has been documented for economic (e.g. Armstrong and Lusk, 1983) as well as fisheries time series (Noakes et al., 1990).

2.3. Measures of accuracy

There are many measures of forecasting accuracy that one may use to compare different models (Makridakis et al., 1983). The following general categories of statistical measures were used in the present study: (a) standard statistical measures (mean and median error, ME and MDE, respectively; mean and median absolute error, MAE and MDAE, respectively; mean and median squared error, MSE and MDSE, respectively; root MSE, RMSE; minimum and maximum values of errors, absolute errors and squared errors; standard deviation of errors, absolute errors and squared errors, SDE, SDAE and SDSE, respectively); (b) relative statistical measures (mean and median percentage error, MPE and MDPE, respectively; mean and median absolute percentage error, MAPE and MDAPE, respectively; minimum and maximum values of percentage errors and absolute percentage errors; standard deviations of percentage errors and absolute percentage errors, SDPE and SDAPE, respectively; number of data points predicted with absolute percentage error (APE) smaller than an arbitrary percentage, here 10% and 20%); and (c) other statistical measures (such as Theil's (1966) U-statistic, U, and bias component, B; McLaughlin's (1975) batting average, MBA; and Ljung and Box's (1978) Q-statistic).

U is defined as follows:

$$U=\sqrt{\frac{\sum_{t=1}^{n-1}\left(\frac{e_{t+1}}{X_t}\right)^2}{\sum_{t=1}^{n-1}\left(\frac{X_{t+1}-X_t}{X_t}\right)^2}}$$

where e_{t+1} is the error at time $t+1$. U is a compromise between standard and relative measures (Makridakis et al., 1983). It provides a measure of comparison of a given method with a naive method that uses as a forecast at time $t+1(=F_{t+1})$ the catch at time t $(=X_t)$ (which is also used in the present study: NM). U attains values >0; $U>1$ indicates poor forecasting efficiency since the naive method produces better forecasts; $U<1$ indicates good forecasting efficiency (the smaller the value the better the method). MBA is estimated from U, MBA $=100(4-U)$, and ranges between 200 and 400 with a value of 300 having a similar interpretation to $U=1$.

B is defined as follows:

$$B=\frac{\left(\frac{\sum_{t=1}^{n}(F_t)}{n}-\frac{\sum_{t=1}^{n}(X_t)}{n}\right)^2}{\text{MSE}}$$

and is a measure of over- or underestimation of actual values. Low values of B indicate small bias.

The Ljung–Box Q is used to indicate whether the errors, e_t, are autocorrelated or not. It is computed as follows:

$$Q=t(t+2)\sum_{i=1}^{L}\frac{(r_i^2)}{(t-i)}$$

where t is the number of data points, r_i is the ith autocorrelation coefficient and L is the number of autocorrelation coefficients. Q is a sum of squared autocorrelations and, hence, attains a value of zero only when all autocorrelations are zero. It is weighted in order to obtain a statistic which is approximately x^2 with $L-n$ degrees of freedom, where n is the number of variables fitted to the model.

Apart from the above mentioned measures, the Bayesian Information Criterion, BIC, was also used to compare the different versions of HES, BES and ARIMA models fitted to the annual catches. BIC provides a combined measure of goodness-of-fit (as measured by MSE) and model parsimony (i.e. number of variables). The following equation was used for its computation (Stellwagen and Goodrich, 1993):

$$\text{BIC}=(\text{MSE})t^{n/2t}$$

where t is the number of data points and n is the number of variables fitted to the model. This version of BIC is scaled the same as the standard forecast error and can very loosely be interpreted as an estimate of out-of-sample error (Stellwagen and Goodrich, 1993). BIC is meaningful only for comparing different versions from the same family of models and for the same time series and the model with the minimum BIC value is selected as the most appropriate.

2.4. Model fitting and forecasting

All models were fitted to the 1964–1987 period and the different measures of accuracy were computed and compared for the untransformed time series. NM, HES and BES models

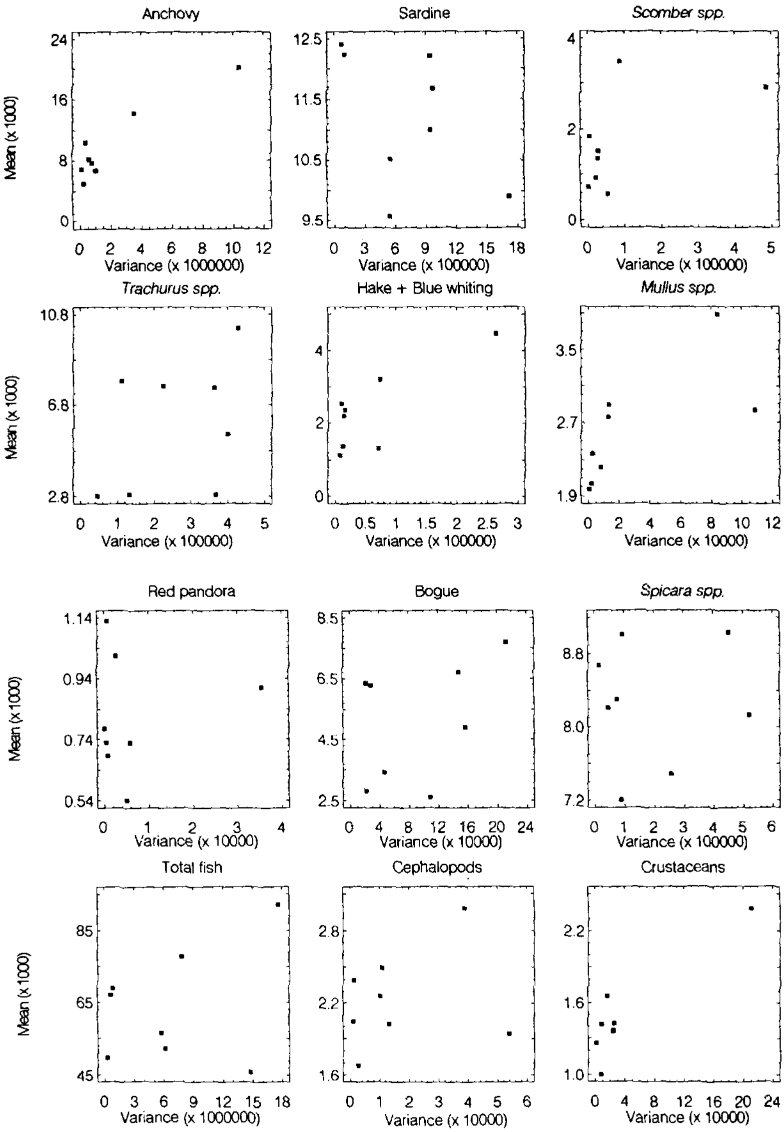

Fig. 2. Variance–mean plots for the 16 annual time series, Hellenic waters, 1964–1989.

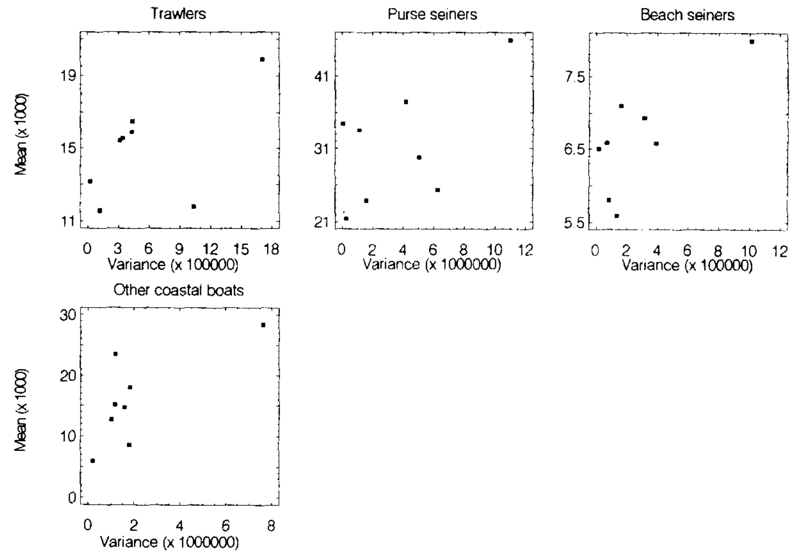

Fig. 2 (continued).

were fitted to the original catches. All remaining models were fitted to the log-transformed catches. Logarithmic transformation was justified by examining the mean-variance plots of the annual catches, obtained by dividing the series into periods of three years, calculating the variance and mean of each period and plotting the variance vs. the mean.

Since good fitting does not necessarily imply good forecasting, the ability of all models to produce forecasts was tested by using the arithmetic values of the parameters of the estimated models to develop one-step-ahead annual forecasts for the years 1988 and 1989. Forecasts were then compared with the actual catches in 1988 and 1989 using the APE and MAPE values as well as an empirical measure named 'stable forecasts' which was defined as follows: (higher APE) / (lower APE) <4 and individual APE values for 1988 and 1989 both ≤ 20.

The fitting (all methods) $\times$ (all measures) matrices were subjected to principal component analysis (PCA; based on correlation matrix). PCA identified three groups of methods of similar fitting performance. Each of the three groups was assigned an integer rank from 1 to 3 to indicate fitting from worst to best. Consequently, all models included in the groups were assigned the rank of the group and rank averages ($\times 100$) were then computed for each model. PCA was conducted using the PRIMER algorithms of Plymouth Marine Laboratory (Clarke and Warwick, 1989).

3. Results

The variance–mean plots of the 16 annual time series (Fig. 2) suggest that the variances of the catches generally increase with the level of the catch. Hence, all models, with the exception of NM, BES and HES, were fitted to the log-transformed catches.

Table 2
Surplus-yield (FOX) models between the log-transformed annual cath per unit of fishing effort and fishing effort, for the four component fisheries, Hellenic waters, 1964–1987 ($n=24$). The values of the r^2 and Ljung–Box test, LB(11), are also shown; significant ($P<0.05$) LB(11) values are marked with an asterisk

Fishery	FOX model	r^2	LB(11)
Trawl	Ln(C/HP) = 5.72 − 0.00000649(HP)	0.89	19.8*
Purse seine	Ln(C/HP) = 6.93 − 0.00000985(HP)	0.84	19.7*
Beach seine	Ln(C/HP/BO) = 13.14 − 0.02468(HP/BO)	0.90	37.5*
Other coastal boats	Ln(C/HP/BO) = 14.35 − 0.02845(HP/BO)	0.80	75.8*

FOX models between the annual log-transformed catch per unit of fishing effort and fishing effort for the 1964–1987 period (Table 2) were built only for those catch-effort combinations characterised by the highest r^2 values during the 1964–1989 period (Stergiou et al., 1994). The r^2 values of the FOX models were >0.80. For all models, the Ljung–Box test indicated significant ($P<0.05$) error autocorrelation (Table 2).

For all MREG models, the independent variables initially considered for inclusion into the models were: (a) the five climatic variables; and (b) fishing effort, defined according to the percentage of species' catch caught by the four component fisheries (Stregiou and Pollard, 1994). In addition, for the four component fisheries, the wholesale value of the catch was also considered for inclusion whereas for the time series participating in the two complexes, considered for the development of VAR models, the catch of the other member of the complex was also considered for inclusion. The MREG models fitted to *Mullus* spp. and *Spicara* spp. catches included only one independent variable, time (i.e. TV model), and, hence, are not shown here (Table 3). The r^2 values of the annual MREG models ranged from 0.31, for sardine catches, to 0.95, for bogue catches. The number of the independent variables that entered into the models ranged from one, for sardine and red pandora models, to four, for anchovy, bogue and trawl models (Table 3). The Ljung–Box test indicated significant ($P<0.05$) error autocorrelation for the models fitted to trawl and *Trachurus* spp. catches whereas for the 12 remaining models it did not indicate error autocorrelation ($P>0.1$; Table 3). The MREG models fitted to sardine and purse seine catches did not include environmental variables (Table 3). In addition, the MREG models fitted to gadiform, red pandora, cephalopod, beach seine and 'other coastal boats' catches did not include fishing effort variables (Table 3). From the five environmental variables tested for entry into the models, SST entered into seven MREG models, SLP into five, AIRT and NSW into two whereas WISC did not enter into any MREG model (Table 3). In addition, engine HP, number of boats and number of fishers each entered into three MREG models (Table 3).

MREG models (Table 3) were tested for significant improvement by including Cochrane–Orcutt autoregressive error terms, lagged dependent variables as well as lagged independent variables. All Lagrange multiplier tests performed indicated that the Cochrane–Orcutt autoregressive error terms at lags 1 to 3, the first two lags of the dependent variables, as well as the first two lags of the independent variables were not statistically ($P>0.05$) significant.

The r^2 values of the TV models ranged from 0.40, for *Spicara* spp. catches, to 0.93, for purse seine catches (Table 4). The Ljung–Box test did not indicate significant ($P>0.05$) error autocorrelation for four models (*Spicara* spp.; total fish; purse seine; and red pandora)

Table 3

Regression (MREG) models between the log-transformed annual catches of the 16 species or groups of species, Ln(C), and various independent variables, Hellenic waters, 1964–1987. All independent variables refer to the 1964–1986 period whereas all dependent variables refer to the 1965–1987 period (i.e. lagged by one year). For all regressions $n=23$. t is time ($t=1$ to 23). The r^2, BIC and Ljung–Box test (LB) values are also shown. For LB, numbers in parentheses show the degrees of freedom and significant ($P<0.05$) values are marked with an asterisk

Species	MREG model	r^2	BIC	LB
Anchovy	Ln(C) = −4.13 − 0.54Ln(*Trachurus*) + 4.17Ln(SST) + 0.57Ln(FIP) + 0.11t	0.93	0.14	8.8(8)
Sardine	Ln(C) = 13.61 − 0.54Ln(FIP)	0.31	0.11	15.6(11)
Scomber spp.	Ln(C) = −18.44 + 0.52Ln(SST) + 4.02Ln(BOP)	0.80	0.30	12.2(10)
Trachurus spp.	Ln(C) = −739.74 + 0.55Ln(sardine) + 106.12Ln(SLP) + 0.78Ln(HPP)	0.94	0.15	17.9(9)*
Gadiformes	Ln(C) = −765.37 + 111.56Ln(SLP) + 0.06t	0.94	0.12	12.8(10)
Red pandora	Ln(C) = −21.94 + 9.84Ln(SST)	0.46	0.19	5.1(11)
Bogue	Ln(C) = −745.42 + 5.53Ln(SST) + 106.88Ln(SLP) − 0.31Ln(BO) + 0.04t	0.95	0.11	12.0(8)
Total fish	Ln(C) = −399.00 + 3.42Ln(SST) + 57.11Ln(SLP) + 0.38Ln(HP)	0.93	0.08	10.2(9)
Cephalopods	Ln(C) = −7.98 + 3.98Ln(SST) − 1.38Ln(AIRT) + 0.73Ln(total fish)	0.86	0.08	11.6(9)
Crustaceans	Ln(C) = 5.46 − 0.21Ln(NW) + 0.19Ln(BO) + 0.03t	0.79	0.14	12.6(9)
Trawl	Ln(C) = 18.56 − 1.19Ln(AIRT) + 0.56Ln(HPT) + 0.62Ln(TOT) − 2.96Ln(BOT)	0.91	0.07	19.2(8)*
Purse seine	Ln(C) = 10.96 + 0.16Ln(VAP) − 0.34Ln(FIP)	0.89	0.09	9.9(10)
Beach seine	Ln(C) = −597.22 + 0.14Ln(NW) + 87.52Ln(SLP)	0.38	0.12	11.9(10)
Other coastal boats	Ln(C) = −27.27 + 8.28Ln(SST) + 0.27Ln(VAC) + 1.03Ln(seine)	0.82	0.26	8.8(9)

Table 4

Time-varying (TV) regression models between the log-transformed annual catches of the 16 species or groups of species, Ln(C), and time t ($t=1$ to 24), Hellenic waters, 1964–1987. The r^2 and Ljung–Box test, LB, values are also shown. For LB, numbers in parentheses show the degrees of freedom and significant ($P<0.05$) values are marked with an asterisk

Species	TV model	r^2	LB
Anchovy	$\mathrm{Ln(C)} = 8.383 + 0.057t$	0.82	26.8(12)*
Scomber spp.	$\mathrm{Ln(C)} = 6.380 + 0.071t$	0.72	26.2(12)
Trachurus spp.	$\mathrm{Ln(C)} = 7.710 + 0.068t$	0.84	53.4(12)*
Gadiformes	$\mathrm{Ln(C)} = 6.858 + 0.063t$	0.91	29.6(12)*
Bogue	$\mathrm{Ln(C)} = 7.805 + 0.052t$	0.81	33.6(12)*
Spicara spp.	$\mathrm{Ln(C)} = 9.121 - 0.008t$	0.40	19.3(12)
Total fish	$\mathrm{Ln(C)} = 10.655 + 0.031t$	0.86	17.7(12)
Cephalopods	$\mathrm{Ln(C)} = 7.447 + 0.020t$	0.60	31.7(12)*
Crustaceans	$\mathrm{Ln(C)} = 6.913 + 0.028t$	0.60	23.6(12)*
Trawlers	$\mathrm{Ln(C)} = 9.306 + 0.023t$	0.84	29.9(12)*
Purse seine	$\mathrm{Ln(C)} = 9.902 + 0.034t$	0.93	11.8(12)
Other coastal boats	$\mathrm{Ln(C)} = 8.979 + 0.047t$	0.45	46.4(12)*
Mullus spp.	$\mathrm{Ln(C)} = 8.136 - 0.089t + 0.004t^2$	0.75	20.9(11)*
Red pandora	$\mathrm{Ln(C)} = 7.194 - 0.082t + 0.003t^2$	0.46	18.6(11)*
Beach seine	$\mathrm{Ln(C)} = 8.907 - 0.096t + 0.011t^2 - 0.0003t^3$	0.51	21.3(10)*
Sardine	$\mathrm{Ln(C)} = 9.672 - 0.244t + 0.037t^2 - 0.002t^3 + 0.00003t^4$	0.56	32.3(9)*

Table 5

Arithmetic values of the coefficients of the Brown (α) and Holt (α,γ) smoothing models (BES and HES, respectively) fitted to the annual catches of the 16 species or groups of species, together with the type of BES model (L, linear; Q, quadratic), and the corresponding r^2 and BIC values, Hellenic waters, 1964–1987. For all models, the Ljung–Box(12) values were <18.3, $P>0.1$

Species	BES model				HES model			
	Type	α	r^2	BIC	α	γ	r^2	BIC
Anchovy	L	0.325	0.87	2405	0.212	0.998	0.87	1987
Sardine	L	0.380	0.31	1357	0.579	0.0008	0.07	1450
Scomber spp.	L	0.480	0.73	505	1	0.025	0.77	530
Trachurus spp.	L	0.492	0.88	915	0.997	0.028	0.90	953
Gadiformes	L	0.410	0.87	414	0.789	0.065	0.88	412
Mullus spp.	Q	0.276	0.69	363	1	0.001	0.68	388
Red pandora	Q	0.174	0.19	164	0.766	0.0005	0.33	175
Bogue	L	0.614	0.86	661	1	0.035	0.89	674
Spicara spp.	L	0.154	0.11	734	0.097	0.102	0.26	732
Total fish	L	0.329	0.89	5565	0.66	0.157	0.89	5514
Cephalopods	L	0.307	0.73	215	0.807	0.034	0.74	227
Crustaceans	L	0.420	0.68	257	1	0.041	0.69	267
Trawl	L	0.314	0.76	1276	0.787	0.035	0.80	1326
Purse seine	L	0.243	0.87	3385	0.146	0.233	0.88	3068
Beach seine	L	0.512	0.45	734	1	0.001	0.36	785
Other coastal boats	L	0.507	0.83	3172	0.976	0.032	0.82	3311

Table 6

Frequencies identified through FFT analysis applied to the log-transformed and detrended annual catches of the 16 species or groups of species, Hellenic waters, 1964–1987

Species	Frequencies (years)											
	24	12	8	6	4.8	4	3.4	3	2.7	2.4	2.2	2
Anchovy	+				+			+			+	
Sardine	+		+			+		+			+	
Scomber spp.	+		+		+							+
Trachurus spp.		+				+						
Gadiformes		+			+						+	
Mullus spp.			+				+				+	
Red pandora				+			+				+	
Bogue		+			+							
Spicara spp.	+			+			+					
Total fish		+				+					+	
Cephalopods	+			+			+					
Crustaceans	+				+				+			
Trawl		+			+					+		
Purse seine		+				+		+		+		
Beach seine	+			+			+					
Other coastal boats	+				+		+					

Table 7

Arithmetic coefficients of the independent variables of the harmonic multiple regression (HREG) models between the log-transformed annual catches of the 16 groups of species and the various frequencies identified in their variability, using FFT analysis, as well as time t ($t = 1$ to 24), Hellenic waters, 1964–1987. c is constant. The r^2 and Ljung–Box (LB) values are also shown. For LB, numbers in parentheses show the degrees of freedom and significant ($P < 0.05$) LB values are marked with an asterisk

Species		Arithmetic coefficients of HREG models									
		$x = 2\pi(0.04)t$		$x = 2\pi(0.08)t$		$x = 2\pi(0.12)t$		$x = 2\pi(0.17)t$		$x = 2\pi(0.21)t$	
		$\sin(x)$	$\cos(x)$	$\sin(x)$	$\cos(x)$	$\sin(x)$	$\cos(x)$	$\sin(x)$	$\cos(x)$	$\sin(x)$	$\cos(x)$
Anchovy	8.21	0.20	0.13								
Sardine	9.32	−0.09									
Scomber spp.	6.50		0.21			−0.25	−0.12				0.16
Trachurus spp.	7.75			0.11	0.17						
Gadiformes	6.86				0.13					0.08	
Mullus spp.	8.17					−0.08					
Red pandora	7.22							−0.11			
Bogue	7.72			0.17							
Spicara spp.	9.11							0.06			
Total fish	10.62			0.06							
Cephalopods	7.46		0.10								
Crustaceans	6.70	0.23	0.06							0.07	0.09
Trawl	9.31				0.07						
Purse seine	9.90				0.04						
Beach seine	8.79	−0.13							−0.06		
Other coastal boats	9.26	−0.23	0.39								

	Arithmetic coefficients of HREG models							
	$\cos[2\pi(0.25)t]$	$\sin[2\pi(0.33)t]$	$\sin[2\pi(0.46)t]$	t	t^2	t^3	r^2	LB
Anchovy		0.09	−0.07	0.07			0.93	17.0(7)*
Sardine							0.25	12.2(11)
Scomber spp.				0.06			0.90	18.3(7)*
Trachurus spp.	−0.06				0.009	−0.0003	0.97	26.2(7)*
Gadiformes				0.06			0.96	12.2(9)
Mullus spp.				−0.09	0.004		0.81	16.6(9)*
Red pandora				−0.09	0.003		0.55	18.8(9)*
Bogue				0.06			0.87	13.5(10)
Spicara spp.				−0.01			0.54	16.3(10)
Total fish				0.03			0.88	12.6(10)
Cephalopods				0.02			0.74	39.1(10)*
Crustaceans				0.05			0.89	14.8(7)*
Trawl				0.02			0.91	7.1(10)
Purse seine	−0.03	0.04		0.03			0.95	21.2(7)*
Beach seine							0.61	21.8(10)*
Other coastal boats				0.02			0.74	21.2(9)*

Table 8
Arithmetic coefficients of the autoregressive integrated moving average (ARIMA) models fitted to the log-transformed annual catches of the 16 species or groups of species, Hellenic waters, 1964–1987. n is number of data points fitted. The r^2, BIC and Ljung–Box, LB(12), values are also shown; significant ($P<0.05$) LB(12) values are marked with an asterisk

Species	ARIMA model	Arithmetic coefficients			n	r^2	BIC	LB(12)
		$\theta 1$	$\theta 2$	c				
Anchovy	(0,1,1)	0.563		0.064	23	0.83	1775	5.4
Sardine	(0,1,0)				23	0.00	1396	22.2
Scomber spp.	(0,1,0)				23	0.71	469	11.1
Trachurus spp.	(0,1,0)				23	0.91	859	7.8
Gadiformes	(0,1,0)				23	0.87	362	12.7
Mullus spp.	(0,1,0)				23	0.71	313	4.4
Red pandora	(0,1,0)			0.047	23	0.33	150	6.5
Bogue	(0,1,1)	−0.588			23	0.89	680	7.7
Spicara spp.	(0,1,2)	0.303	0.697	−0.010	23	0.47	581	10.0
Total fish	(0,1,0)				23	0.85	5754	10.5
Cephalopods	(0,1,0)				23	0.74	204	7.5
Crustaceans	(0,1,0)				23	0.65	222	18.8
Trawl	(0,1,0)				23	0.82	1111	6.1
Purse seine	(0,1,2)	0.515	0.405	0.035	23	0.90	2543	11.2
Beach seine	(0,1,0)				23	0.45	650	8.7
Other coastal boats	(0,1,0)				23	0.70	3974	3.6

whereas it indicated significant ($P<0.05$) error autocorrelation for the remaining 12 TV models (Table 4).

For the development of BES and HES models various smoothing coefficient values (for BES models) or combinations of smoothing coefficient values (for HES models) were tested and the final values of the smoothing coefficients (those leading to fits for 1964–1987 characterised by the smallest BIC, MSE and MAPE values, in that order of importance) are shown in Table 5. The quadratic BES model was used for *Mullus* spp. and red pandora catches whereas the linear BES model was used for the remaining 14 annual time series. With the exception of the BES and HES models fitted to *Spicara* spp., red pandora, sardine and beach seine catches, which had very low r^2 values (from 0.07 to 0.45), all remaining BES and HES models had high r^2 values, ranging from 0.68 to 0.90 (Table 5). For all BES and HES models the Ljung–Box test did not indicate significant ($P>0.01$) error autocorrelation (Table 5).

Overall, 12 frequencies were identified in the variability of the 16 catches (24, 12, 7, 8, 6, 4.8, 4, 3.4, 3, 2.7, 2.4, 2.2 and 2 years: Table 6). All catches displayed cycles with frequencies between 4 and 6 years, with the exception of *Mullus* spp. catches, and between 2 and 4 years, with the exception of *Trachurus* spp. and bogue catches (Table 6). In addition, *Trachurus* spp., gadiform, bogue, total fish, trawl and purse seine catches all displayed a 12-year cycle whereas sardine, *Scomber* spp. and *Mullus* spp. an 8-year cycle (Table 6). The arithmetic coefficients of the HREG models fitted to the 16 catches, using the estimated frequencies (as well as time t, t^2 and/or t^3), as independent variables, and stepwise variable selection, are shown in Table 7. With the exception of the HREG models fitted to sardine, red pandora, *Spicara* spp. and beach seine catches, which had low r^2 values, from 0.25 to 0.61, the remaining HREG models had r^2 values between 0.74, for cephalopods, and 0.97, for *Trachurus* spp. (Table 7). The Ljung–Box test indicated significant ($P<0.05$) error autocorrelation for nine HREG models (Table 7).

The final ARIMA models fitted to the catches of the 16 species or groups of species (Table 8) are characterised by the smallest BIC, MSE and MAPE values, the highest r^2 values and have arithmetic coefficients that are significantly ($P<0.05$) different from 0. Three general ARIMA models were fitted to the 16 annual catches (Table 8). Hence, the annual anchovy and bogue catches are described by the ARIMA (0,1,1) model, the *Spicara* spp. and purse seine catches by the ARIMA (0,1,2) model whereas all remaining catches by the ARIMA (0,1,0) model. The latter is a special case of the first order autoregressive model, ARIMA (1,0,0) ($X_t = c + aX_{t-1} + e_t$), in which the regression parameter $a=1$. This model implies that the catch at year t equals that at year $t-1$ (in other words is similar to NM). With the exception of the sardine ARIMA (0,1,0) model for which the r^2 value was 0, the r^2 values of the remaining models ranged from 0.33, for red pandora catches, to 0.91, for *Trachurus* spp. catches. The Ljung–Box test did not indicate significant ($P>0.05$) error autocorrelation (Table 8).

The two VAR models fitted did not include environmental or fishing effort variables, had high r^2 values, ranging from 0.67, for sardine catches, to 0.86 for trawl catches and the Ljung–Box test did not indicate significant ($P>0.05$) error autocorrelation (Table 9). It is worthy to point out that the r^2 value of the sardine model (Table 9) is higher than that of the remaining models fitted to sardine catches (see Tables 3–5, 7 and 8).

The initialising period ranged from zero years, for TV and HREG models, to three years, for the quadratic BES models of *Mullus* spp. and red pandora. Hence, the accuracy measures estimated for the fits of the *Mullus* spp. and red pandora models refer to the 1967–1987 period whereas those for the remaining 14 time series to the 1966–1987 period. In general, all models produced very good fits for all annual catches and the lower MAPE values achieved by any method ranged between 3.8%, for trawl catches, to 16%, for the 'other coastal boats' catches. (The different accuracy measures for the untransformed fits for the 16 annual catches obtained by the different models are not shown here.)

For all time series, there was an important improvement in fitting accuracy as indicated by the increase in MAPE values because of the use of different methods (anchovy: from 7.0%, for HREG, to 17.4%, for NM; sardine: from 5.0%, for VAR, to 32.5%, for TV; *Scomber* spp.: from 12.9%, for HREG, to 28.5%, for TV; *Trachurus* spp.: from 7.9%, for HREG, to 20.3%, for EMP; hake: from 7.2%, for HREG, to 20.9%, for EMP; *Mullus* spp.: from 6.5%, for HREG, to 30.2%, for EMP; red pandora: from 11.6%, for MREG, to 18.7%, for EMP; bogue: from 6.6%, for MREG, to 21.1%, for EMP; *Spicara* spp.: from 5.0%, for HREG, to 29.3%, for EMP; total fish: from 4.8%, for MREG, to 20.5%, for EMP; cephalopods: from 4.4%, for MREG, to 20.0%, for EMP; crustaceans: from 6.2%, for HREG, to 21.0%, for EMP; trawl: from 3.8%, for MREG, to 6.5%, for BES; purse seine: from 4.1%, for HREG, to 9.4%, for FOX; beach seine: from 6.1%, for HREG, to 12.8%, for FOX; and 'other coastal boats': from 16.0%, for MREG, to 38.1%, for FOX).

HREG produced the most accurate fits in terms of MAPE for nine annual series, MREG for six series and VAR for one series. In addition, EMP produced the worst fits for nine annual series, FOX for three series, TV for two series, and NM and BES for one series each.

Table 9
Vector autoregression (VAR) models between the log-transformed annual catches of (a) anchovy and sardine and (b) trawl and purse seine, Hellenic waters, 1964–1987. n is number of data points fitted. r^2 and Ljung–Box, LB(12), values are also shown; significant ($P<0.05$) LB(12) values are marked with an asterisk

VAR model	n	r^2	LB(12)
Anchovy/sardine			
$\mathrm{Ln(anchovy)} = 7.209 + 0.590\mathrm{Ln(anchovy)}_{t-2} - 0.664\mathrm{Ln(sardine)}_{t-2} - 0.0012\mathrm{Ln}(\textit{Trachurus}) + 0.387\mathrm{Ln}(\textit{Scomber})$	22	0.84	13.9
$\mathrm{Ln(sardine)} = 9.264 - 0.163\mathrm{Ln(anchovy)}_{t-2} - 0.033\mathrm{Ln(sardine)}_{t-2} + 0.378\mathrm{Ln}(\textit{Trachurus}) - 0.196\mathrm{Ln}(\textit{Scomber})$	22	0.67	19.5
Trawl/purse seine			
$\mathrm{Ln(trawl)} = 1.187 + 0.483\mathrm{Ln(trawl)}_{t-1} + 0.368\mathrm{Ln(purse\ seine)}_{t-1}$	23	0.86	11.0
$\mathrm{Ln(purse\ seine)} = -0.740 + 0.489\mathrm{Ln(trawl)}_{t-1} + 0.621\mathrm{Ln(purse\ seine)}_{t-1}$	23	0.85	17.9

Overall, HREG and MREG models were both characterised by higher accuracy (better performance with respect to all or most, measures with values generally tied together), unbiased fits (i.e. very low B values, usually zero) and transformed errors that were essentially white noise (Tables 3 and 7). In addition, HREG and MREG models outperformed NM (in terms of U and MBA values) for all species (or groups of species) with the exceptions of: (a) cephalopods, for which the HREG model had $U=0.99$ and MBA = 301; (b) beach seine, for which the MREG model had $U=1$ and MBA = 300; and (c) 'other coastal boats', for which the HREG model had $U=0.97$ and MBA = 303.

All ARIMA (0,1,0) models had, as expected, U and MBA values that were 1 and 300, respectively. The ARIMA models fitted to the catches of anchovy, bogue and purse seine all performed better than NM (U and MBA values were 0.80 and 320, 0.81 and 319, and 0.68 and 332, respectively) whereas that fitted to *Spicara* spp. catches performed as the NM model (U and MBA values 1 and 300, respectively).

The above mentioned observations are clearly depicted in the results of PCA performed on the 16 (methods) × (measures) fitting matrices (Fig. 3). PCA identified three general groups of methods, characterised by increasing fitting accuracy (general improvement in all or most measures; Fig. 3). For all PCAs, the percentage of variance explained by the first two axes was higher than 72.5%.

The average ranks (×100) per method are shown in Table 10. HREG and MREG models were more superior than the remaining ones. In addition, ARIMA, those differing from the (0,1,0) model, and VAR models performed better than NM (Table 10). HREG, MREG, ARIMA and VAR models had average ranks (×100) that were higher than the total average rank (×100). HES models had an average rank (×100) marginally higher than that of NM whereas the remaining models (TV, BES, EMP and FOX) had average ranks (×100) that were lower than that of NM (Table 10).

The forecasts developed for the years 1988–1989 are shown in Table 11 and the corresponding individual APE and MAPE values in Tables 12 and 13, respectively. In general, APE values are higher for 1989 than 1988 for most method/series combinations, a fact indicating that APE increases with the length of forecasting horizon.

For 1988, NM produced forecasts with the best APE for three annual series (sardine, bogue and *Trachurus* spp.), NM and EMP for one series (*Scomber* spp.), TV for three

Table 10
Average rank (× 100) of fitting performance of the models fitted to the annual catches of the 16 species or groups of species, Hellenic waters, 1964–1987

Model										
NM	TV	BES	HES	HREG	ARIMA	MREG	FOX	VAR	EMP	Average
206	194	200	213	288	250	286	100	225	163	217

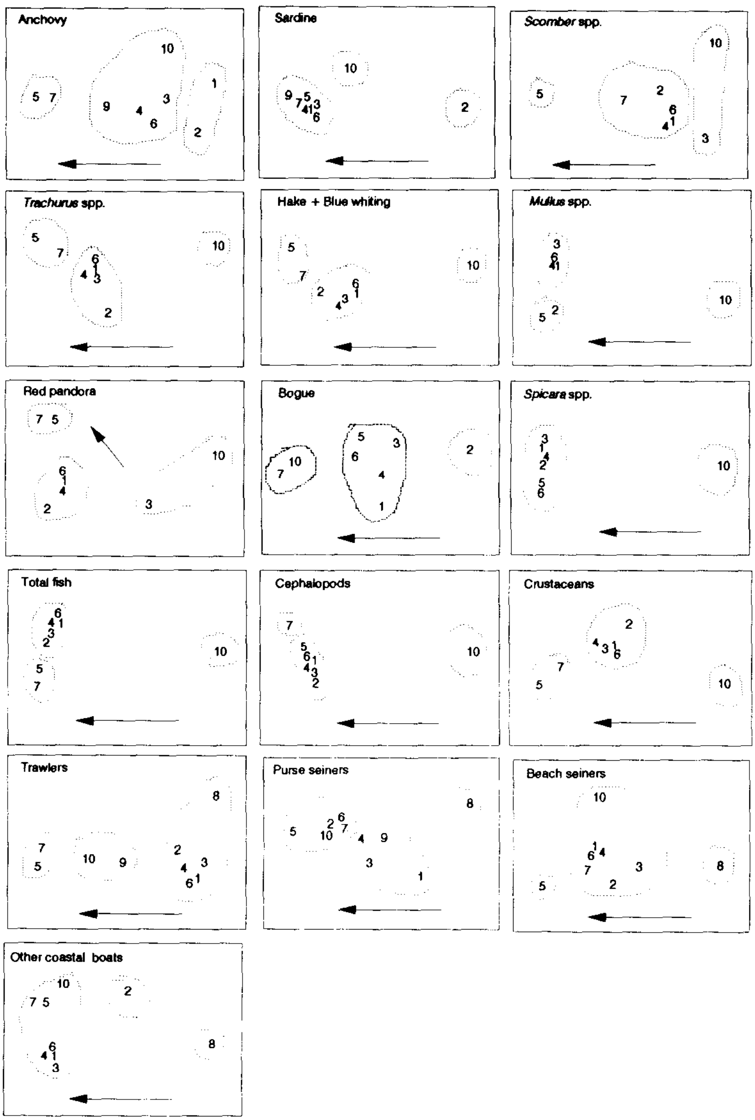

Fig. 3. Results of principal component analysis performed on (measures) × (methods) fitting matrices (fitting period: 1964–1987) for the 16 annual time series. 1, NM; 2, TV; 3, BES; 4, HES; 5, HREG; 6, ARIMA; 7, MREG; 8, FOX; 9, VAR and 10, EMP. Arrows indicate increasing fitting performance.

Table 11
Actual annual catches and forecasts estimated per model for the 16 species or groups of species, Hellenic waters, 1988–1989

Species	Year	Model										
		Actual	NM	TV	BES	HES	HREG	ARIMA*	MREG	FOX	VAR	EMP
Anchovy	1988	19182	24737	18179	24012	24089	26370	22807	24697		22575	22961
	1989	14791	19182	19245	25749	26780	28542	24318	31385		32766	24751
Sardine	1988	9831	9685	1294	9622	10107	10793		10369		9176	8951
	1989	10239	9831	990	9308	10104	10558		10744		7885	8816
Scomber spp.	1988	3723	3835	3481	3978	3928	2985		3129			3612
	1989	5498	3723	3737	4215	4023	2622		1772			3678
Trachurus spp.	1988	7488	9253	12210	10021	9446	11335		9501			9813
	1989	6935	7488	13069	10228	9635	10822		7754			9398
Gadiformes	1988	4814	4674	4596	5133	4903	5573		4650			4877
	1989	5710	4814	4895	5477	5081	5449		4175			5034
Mullus spp.	1988	4567	3633	4496	4293	3636	4223					4069
	1989	5149	4567	5044	4601	3638	4625					4465
Red pandora	1988	906	782	1118	877	849	754		719			848
	1989	1162	906	1200	910	848	791		815			927
Bogue	1988	7748	8259	9000	8704	8449	10741	8921	7933			8858
	1989	9489	7748	9481	9191	8639	12121	9348	6948			9068
Spicara spp.	1988	7511	7382	7488	7218	7499	7875	7627				7514
	1989	9105	7511	7428	7129	7436	7817	7241				7667
Total fish	1988	93238	97418	92042	99278	99629	97140		89154			95665
	1989	101012	93238	94940	103000	103091	102788		94151			98705
Cephalopods	1988	3753	3049	2827	3149	3107	3028		3090			3131
	1989	3783	3753	2884	3249	3148	3050		3006			3240
Crustaceans	1988	3304	2890	2024	2982	2966	3143		2221			2802
	1989	2905	3304	2082	3241	3042	3025		2284			2846
Trawl	1988	22461	19102	19555	20521	19878	20797		23028	17249	20593	20229
	1989	19350	22461	20010	21090	20158	20751		34968	17276	21129	21629
Purse seine	1988	41473	50588	46723	48431	45211	50343	49656	46520	38172	49154	46627
	1989	42662	41473	48339	50029	46582	49332	49258	47459	38094	50089	46332
Beach seine	1988	3710	5774	5970	5652	5772	6133		6958	5235		5664
	1989	7192	3710	5293	5310	5770	6297		6355	5246		5661
Other coastal boats	1988	33698	31343	25694	33186	32023	25779		24738	17587		28377
	1989	43345	33698	26930	35310	32746	23878		18710	17587		29283

*All remaining ARIMA models had forecasts same as those of NM.

series (anchovy, *Mullus* spp. and total fish), BES for three series (red pandora, cephalopods and 'other coastal boats'), FOX for two, out of four, series (purse seine and beach seine), EMP for two series (gadiformes and *Spicara* spp.), HREG for one series (crustaceans) and MREG for one series (trawl) (Table 12). In addition, HREG produced forecasts with the worst APE for five annual series, TV for four series, NM, HES and FOX for two series each, and MREG for one series (Table 12). For 1989, NM produced forecasts with the best APE for four annual series (anchovy, *Trachurus* spp., cephalopods and purse seine), TV for four series (*Mullus* spp., red pandora, bogue and trawl), BES for three series (*Scomber* spp., gadiformes and 'other coastal boats'), HREG for two series (total fish and *Spicara* spp.), MREG for one series (beach seine), HES for one series (sardine) and EMP for one

Table 12
APE of annual forecasts per model for the 16 species or groups of species, Hellenic waters, 1988–1989

Species	Year	Model									
		NM	TV	BES	HES	HREG	ARIMA*	MREG	FOX	VAR	EMP
Anchovy	1988	29.0	5.2	25.2	25.6	37.5	18.9	28.8		17.7	19.7
	1989	29.7	30.1	74.1	81.1	93.0	64.4	112.2		121.5	67.3
Sardine	1988	1.5	86.8	2.1	2.8	9.8		5.5		6.7	9.0
	1989	4.0	90.3	9.1	1.3	3.1		4.9		23.3	13.9
Scomber spp.	1988	3.0	6.5	6.8	5.5	19.8		16.0			3.0
	1989	32.3	32.0	23.3	26.8	52.3		67.8			33.1
Trachurus spp.	1988	23.6	63.1	33.8	26.2	51.4		26.9			31.1
	1989	8.0	88.4	47.5	38.9	56.0		11.8			35.5
Gadiformes	1988	2.9	4.5	6.6	1.9	15.8		3.4			1.3
	1989	15.7	14.3	4.1	11.0	4.6		26.9			11.8
Mullus spp.	1988	20.4	1.5	6.0	20.4	7.5					10.9
	1989	11.3	2.0	10.6	29.3	10.2					13.3
Red pandora	1988	13.7	23.3	3.3	6.4	16.8		20.7			6.4
	1989	22.0	3.3	21.7	27.0	31.9		29.8			20.2
Bogue	1988	6.6	16.2	12.3	9.0	38.6	12.0	15.1			14.3
	1989	18.3	0.1	3.1	9.0	27.7	8.6	1.5			4.4
Spicara spp.	1988	1.7	0.3	3.9	0.2	4.9	1.5				0.0
	1989	17.5	18.4	21.7	18.3	14.2	20.5				15.8
Total fish	1988	4.5	1.3	6.5	6.9	4.2		4.4			2.6
	1989	7.7	6.0	2.0	2.1	1.8		6.8			2.3
Cephalopods	1988	18.8	24.7	16.1	17.2	19.3		17.7			16.6
	1989	0.8	23.8	14.1	16.8	19.4		20.6			14.4
Crustaceans	1988	12.5	38.7	9.7	10.2	4.9		32.8			15.2
	1989	13.7	28.3	11.6	4.7	4.2		21.4			2.0
Trawl	1988	15.0	12.9	8.6	11.5	7.4		2.5	23.2	8.3	9.9
	1989	16.1	3.4	9.0	4.2	7.2		80.7	10.7	9.2	11.8
Purse seine	1988	22.0	12.7	16.8	9.0	21.4	19.7	12.2	8.0	18.5	12.4
	1989	2.8	13.3	17.3	9.2	15.6	15.5	11.2	10.7	17.4	8.6
Beach seine	1988	55.6	60.9	52.3	55.6	65.3		87.5	41.1		52.7
	1989	48.4	26.4	26.2	19.8	12.4		11.6	27.1		21.3
Other coastal boats	1988	7.0	23.8	1.5	5.0	23.5		26.6	47.8		15.8
	1989	22.3	37.9	18.5	24.5	44.9		56.8	59.4		32.4

*All remaining ARIMA models had APE values same as those of NM.

series (crustaceans) (Table 12). In addition, TV produced forecasts with the worst APE for four series, MREG for three series, NM, BES and HREG for two series each, and HES, FOX and VAR for one series each (Table 12). Finally, the number of stable forecasts with individual APE values for both years $\leq 20\%$ was nine for BES, eight for HREG, seven for HES and EMP, five for NM and TV, three for MREG, two, out of four, for ARIMA and VAR, and one, out of four, for FOX (Table 12).

With respect to both years combined, BES produced forecasts with the lower MAPE for five annual series (*Scomber* spp., gadiformes, red pandora, bogue and 'other coastal boats'), TV for two series (anchovy and *Mullus* spp.), NM for two series (*Trachurus* spp. and cephalopods), HES for two series (sardine and purse seine), HREG for two series (crus-

Table 13
MAPE of annual forecasts per model for the 16 species or groups of species, Hellenic waters, 1988–1989

Species	Model									
	NM	TV	BES	H-ES	HREG	ARIMA*	MREG	FOX	VAR	EMP
Anchovy	29.3	17.8	50.3	53.4	65.3	41.7	70.5		69.6	43.5
Sardine	2.8	88.6	5.6	2.1	6.5		5.2		14.9	11.5
Scomber spp.	17.8	19.3	15.1	16.2	36.1		41.9			18.1
Trachurus spp.	15.8	77.8	40.7	32.6	53.7		19.3			33.3
Gadiformes	9.3	9.4	5.4	6.5	10.2		15.2			6.6
Mullus spp.	15.9	1.8	8.3	24.9	8.9					12.1
Red pandora	17.9	13.3	12.5	16.7	24.4		25.3			13.3
Bogue	12.5	8.2	7.7	9.0	33.2	10.3	8.3			9.4
Spicara spp.	9.6	9.4	12.8	9.8	9.6	11.0				7.9
Total fish	6.1	3.7	4.3	4.5	3.0		5.6			2.5
Cephalopods	9.8	24.3	15.1	17.0	19.4		19.2			15.5
Crustaceans	13.1	33.5	10.7	7.5	4.6		27.1		8.8	8.6
Trawl	15.6	8.1	8.8	7.9	7.3		41.6	17.0	17.9	10.9
Purse seine	12.4	13.0	17.1	9.1	18.5	17.6	11.7	9.3		10.5
Beach seine	52.0	43.7	39.3	37.7	38.9		49.6	34.1		37.0
Other coastal boats	14.7	30.9	10.0	14.8	34.2		41.7	53.6		24.1

*All remaining ARIMA models had MAPE values same as those of NM.

Table 14
Various indices of forecasting performance of the different models fitted to the annual catches of the 16 species or groups of species, Hellenic waters, 1988–1989

Index	NM	TV	BES	HES	HREG	ARIMA	MREG	FOX	VAR	EMP
Best APE 1988	4	3	3	0	1	0	1	8	0	3
Best APE 1989	4	4	3	1	2	0	1	0	0	1
Worst APE 1988	2	4	0	2	5	0	1	8	0	0
Worst APE 1989	2	4	2	1	2	0	3	4	4	0
APE 1988 <10%	7	6	10	9	6	4	4	4	8	7
APE 1989 <10%	5	5	5	6	5	4	3	0	4	4
Best MAPE	2	2	5	2	2	0	0	4	0	2
Worst MAPE	2	4	1	1	2	0	5	4	0	0
Stable forecasts	5	5	9	7	8	8	3	4	8	7
MAPE 1988–1989 <20%	14	10	13	12	9	12	7	8	12	12
MAPE 1988–1989 <10%	5	6	6	8	6	0	3	4	4	5

taceans and trawl), EMP for two series (total fish and *Spicara* spp.) and FOX for one series ('other coastal boats') (Table 13). In addition, MREG models produced forecasts with the highest MAPE values for five annual series, TV for four series, NM and HREG for two series each, and BES, HES and FOX for one series each (Table 13). Finally, NM produced 14 forecasts with MAPE <20%, BES 13, EMP and HES 12, TV ten, HREG nine, MREG

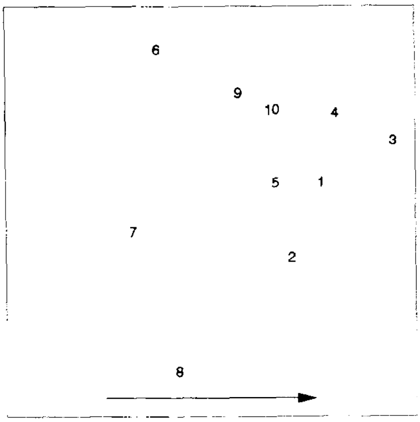

Fig. 4. Results of principal component analysis performed on the (index) × (method) forecasting matrix (Table 14) for the 16 annual times series. 1, NM; 2, TV; 3, BES; 4, HES; 5, REG; 6, ARIMA; 7, MREG; 8, FOX; 9, VAR and 10, EMP. Arrow indicates increasing forecasting performance.

seven, ARIMA and VAR three, out of four, and FOX two, out of four (Table 13).

The above mentioned facts are summarised in Table 14. The matrix shown in Table 14 was subjected to PCA (Fig. 4). The results indicate that BES and, to a lesser extent, HES and NM models, were the best overall performers, MREG, ARIMA and FOX the worst ones whereas the remaining models were of intermediate performance (Fig. 4). Yet, only BES and HES models seem to perform better than NM (Fig. 4).

The annual catches of the 16 species (or groups of species) for 1964–1989 and fits and forecasts for those years, produced by the best fitting and best forecasting model (other than NM or EMP) per species, are shown in Fig. 5. Overall, the amplitude and the duration of the between-year fluctuations are adequately described and forecasted by the models with few exceptions. Hence, all, or most, methods produced bad forecasts for beach seine (MAPE per model ranged between 37% and 52%) and, to a lesser extent, for anchovy (MAPE per model ranged between 17.8% and 70.5%) and *Trachurus* spp. catches (MAPE per model ranged between 14.4% and 77.8%) (Table 13). This is reasonable inasmuch as beach seine catches declined sharply from 5744 t in 1987, to 3710 t in 1988, and then increased to 7192 t in 1989 (Fig. 5). In addition, anchovy catches declined sharply from 24 737 t in 1987 to 19 182 t in 1988 and 14 791 t in 1989 (about 40% overall decline; Fig. 5) The same was also true of *Trachurus* spp. catches (10 696 t in 1986 to 6935 t in 1989; about 35% overall decline).

4. Discussion

In the present study, the ability of eight models to produce accurate forecasts for 16 annual fisheries time series was compared with each other and with those of a naive method and an empirical one. The fitting and forecasting performances of the different models was evaluated using 32 measures of accuracy. Overall 104 annual models were built (without taking into account the NM and EMP ones), belonging to four different families of forecasting techniques: deterministic regression, univariate and multivariate time series and 'biological' models. All models were built with equal care.

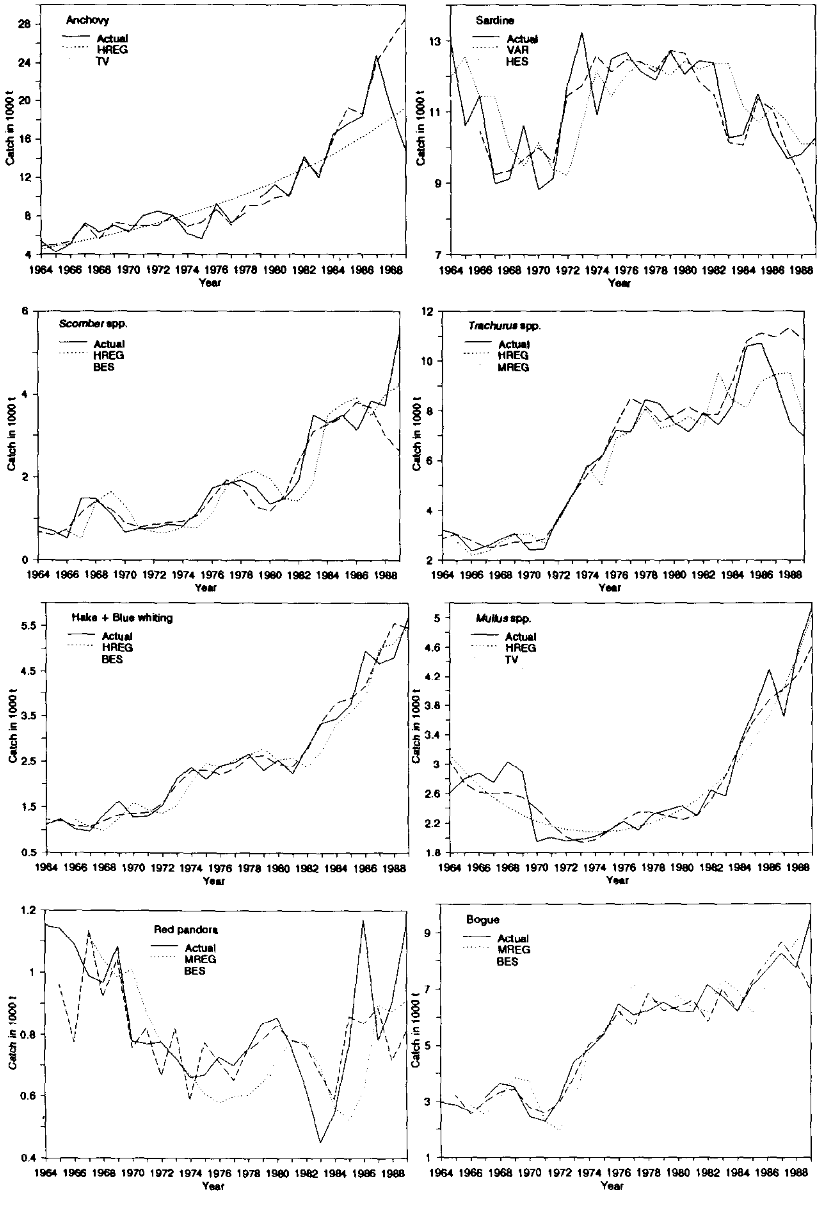

Fig. 5. Actual annual catches of the 16 species or groups of species during 1964–1989, fits (for 1964–1987) produced by the best fitting model per species and forecasts (for 1988–1989) produced by the best forecasting model per species (other than NM or EMP), Hellenic waters.

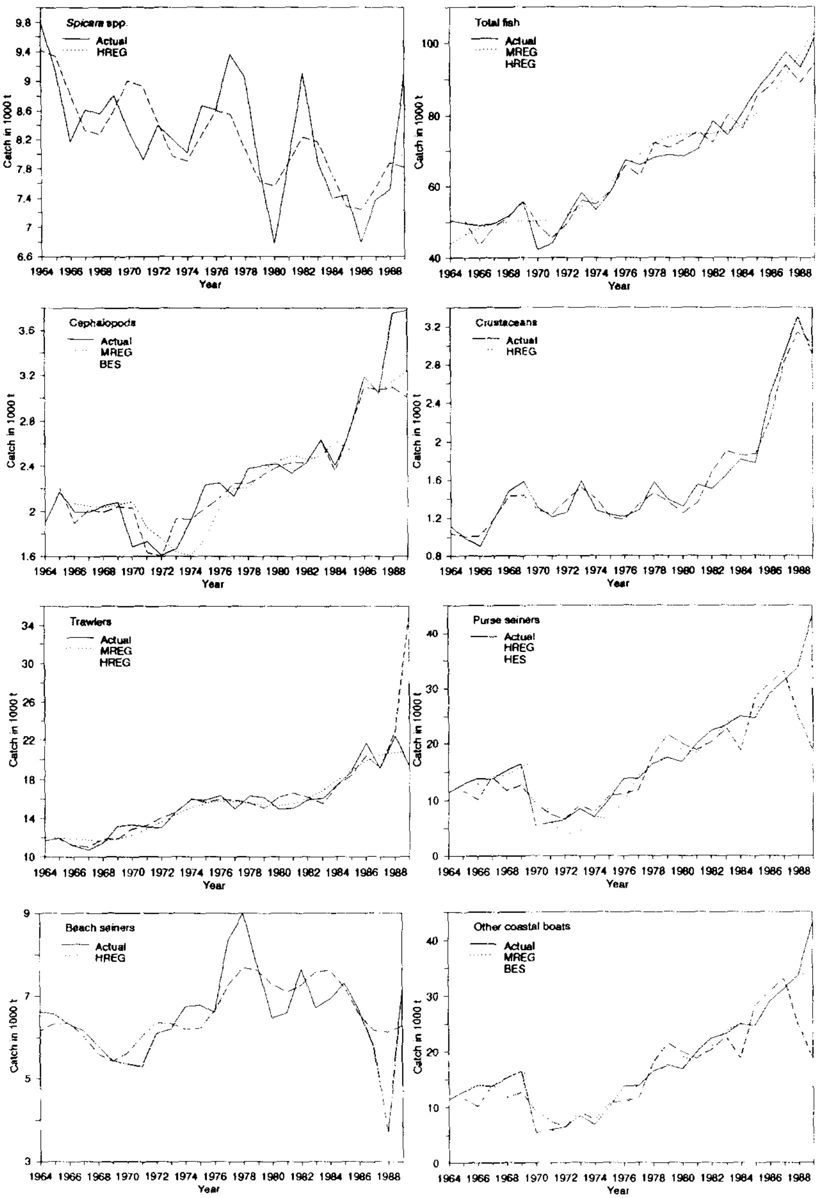

Fig. 5 (continued).

The results revealed that HREG and MREG models outperformed, in terms of fitting accuracy, the remaining eight models (NM, TV, BES, HES, FOX, ARIMA, VAR and EMP). They were both characterised by: (a) higher accuracy in terms of all, or most, standard and relative statistical measures; (b) unbiased fits; (c) much better performance than NM; and (d) transformed errors which were essentially white noise. In addition, HREG and MREG models: (e) explained from 79% to 97% of the variability of the transformed annual catches (Tables 3 and 7), with the exception of those fitted to sardine, red pandora and beach seine catches that explained from 31% to 61% of the variance; (f) produced fits with MAPE values ranging from 3.4% to 21.2%; and (g) in all, or most cases, predicted the between year variations during the fitting period 1964–1987.

In terms of forecasting performance, however, not a single best approach was found. In general BES and, to a lesser extent, HES, NM (which actually is an ARIMA (0,1,0)), EMP and HREG models were among the best performers (in terms of MAPE) more often, produced the worst forecasts more rare and generally were characterised by the higher number of stable forecasts and of forecasts with MAPE $<20\%$ and $<10\%$ (Table 14, Fig. 4). TV was also efficient for some annual catch series (e.g. anchovy, *Mullus* spp.; Tables 12 to 14). Conversely, the poorest performers (FOX, MREG and ARIMA) rarely did better than average. BES, HES, TV and NM are particularly adapted to handle trends and short and highly variable time series whereas HREG is adapted to handle cycles as well. Indeed, all 16 annual catches exhibited strong trends and cycles (Tables 4, 6 and 7).

ARIMA models were not found among the best performers because the yearly data were not characterised by strong autocorrelations at lags >1 year (ACFs not shown here) and, hence, most of the models fitted (12 out of 16 models) were of the ARIMA (0,1,0) type (Table 8), which, as mentioned above, is equivalent to the NM model. Hence, for annual time series of moderate length not displaying strong autocorrelations ARIMA models are probably outperformed by exponential smoothing models, a fact also stressed for economic time series (e.g. Goodrich, 1989; Stellwagen and Goodrich, 1993). In contrast, seasonal ARIMA models have been successful in modelling and forecasting the fisheries of a wide variety of species differing in their biology and behaviour (lobsters: Boudreault et al., 1977, Saila et al., 1979; tuna: Mendelssohn, 1981; sardine: Stergiou, 1989, Moura and Alfonso dos Santos, 1989; anchovy: Stergiou, 1990a) or groups of species (combined purse seine catch: Moura and Alfonso dos Santos, 1989; *Mullus* spp.: Stergiou, 1990b).

The biological FOX models produced biased and the least accurate fits (Table 2), bad forecasts ($>34.1\%$) in two out of four cases (Tables 12 and 13), and were characterised by transformed errors that were significantly ($P<0.05$) autocorrelated (Table 2). The above mentioned facts probably suggest that surplus-yield models should not be used for fisheries forecasting. Stocker and Noakes (1988) and Noakes et al. (1990) also found that time series techniques performed better than Ricker-type biological models for forecasting recruitment in Pacific herring and sockeye-salmon stocks, respectively. In contrast, EMP may incorporate the particular strengths of different models and, hence, lead to more accurate forecasts. The potential use of EMP for short-term forecasting has been stressed for fisheries (Noakes et al., 1990) and other time series (e.g. Armstrong and Lusk, 1983; McLeod et al., 1987).

It is worthy to point out that although the VAR models performed better than NM in terms of fitting performance (Table 10, Fig. 3), their forecasting performance was very poor (Fig. 4). Nevertheless, given their good fitting performance and the fact that VAR produced comparatively accurate forecasts (MAPE between 8.8% and 17.9%) for three (sardine, trawl and purse seine catches) out of four annual series, their application to fisheries modelling and forecasting should not be dismissed, especially for fisheries variables exhibiting strong and justified interactions (e.g. anchovy and sardine complex, predator–prey relationships).

In addition, MREG models were not among the best performers in terms of forecasting accuracy (Tables 12 to 14, Fig. 4), despite their very good overall fitting performance (Table 10, Fig. 3). The opposite was true of the annual BES and HES models (i.e. comparatively poor fitting and good forecasting performances; Table 10 and Figs. 3 and 4). In general the model that explains the historical data best does not necessarily produce the most accurate forecasts for many reasons: (a) the future may not be described by the same probability model as the past; (b) the model may involve too many independent variables (i.e. overfitting) which account for noise or other features in the data that are not likely to extend into the future; and (c) the errors involved in overfitted models may be damaging to forecasting accuracy (Goodrich, 1989). In addition, it must be pointed out that both model fitting and forecasting may fail after a certain period of time, even if the parameters of the model are readjusted each year, and different variables (external and/or lagged) may have to be taken into account from time to time (e.g. Dement'Eva, 1987).

It is clear that all models developed here suffer from certain limitations. Firstly, catches and fishing effort data refer to aggregate nationwide data whereas climatic variables to certain specific areas of the North Aegean Sea (which account, however, for the major part of the total Hellenic catch). Secondly, fishing effort may differ from fishing subarea to fishing subarea. Thirdly, for the development of the multivariate models we used the annual mean values of the five climatic variables and assumed that the relationship between climatic variables and catches is linear. It is known, however, that many biological processes are affected by environmental parameters that probably operate on either shorter or longer time scales. In addition, non-linear relationships allow for an 'optimal response' of a biological variable to an environmental variable (e.g. Mendelssohn and Cury, 1987, 1989; Roy et al., 1992).

Although many practitioners would argue that forecasting models must not always be 'biologically' meaningful, an empirically developed forecasting model is more appealing if it has a sound biological explanation. The ARIMA (and NM), VAR, MREG, FOX and HREG models presented here have interesting explanations (persistence, periodicity and trends of catches) that are discussed below.

In VAR models (Table 9), the response variables were the catches of the individual components at time t and the predictors were, apart from the external variables and the catches of the remaining components of the complex, the autoregressive terms: (a) at year $t-2$ for the annual anchovy and sardine catches; and (b) at year $t-1$ for the annual trawl and purse seine catches. In ARIMA (and NM) models (Table 8), the catch of a species (or groups of species) at year t was partially predicted by the autoregressive terms at year $t-1$. Hence, the above mentioned models predicted persistence of catches. In other words, all else being equal once catches are high they tend to remain high for one to two successive years. Persistence may indicate that environmental conditions favouring the formation of good year classes (and/or large schools) and/or other factors (e.g. microeconomics) affecting the fisheries of the species (or groups of species) in concern tend to persist.

The autoregressive terms at year $t-2$ in VAR models (Table 9) may also indicate a 2–3 year periodicity in the catches. This is consistent with the frequencies identified in the variability of the 16 annual time series (Tables 6 and 7) and the cycles with frequencies

< 8 years could be generally considered with some confidence since they are less than one-third of the length of the time series.

Cycles of 2–4, 4–6 and 8–12 years have also been identified in the air temperature in Athens and in different biotic (zooplankton, phytoplankton, fish egg and larval abundance, fish catches) and abiotic variables (air temperature and pressure, sea temperature and salinity) in different areas of the Mediterranean, Black and Azov Seas (Stergiou, 1992). Similar cycles have also been identified in the physical environment and marine populations in other areas of the world (e.g. Kort, 1970; Shuntov et al., 1981; Vasil'kov et al., 1981; Mysak, 1986; Colebrook and Taylor, 1984). The 11-year cycle in various climatic (e.g. Vibe, 1967; Shuntov et al., 1981; Vasil'kov el al., 1981) and marine biological variables (e.g. Vibe, 1967; Zupanovic, 1968; Regner and Ciacic, 1974; Grainger, 1979; Love and Westphal, 1981; Vasil'kov et al., 1981) has been generally related to the 11-year cycle in sunspot number. Shorter periodicities have been related to other short-term ocean–atmosphere interactions (e.g. surface heat-exchange phenomena: Zupanovic, 1968; Colebrook and Taylor, 1984; advection: Kort, 1970; Mysak, 1986).

The annual catches of the 16 species (or groups of species) displayed long-term trends and, as mentioned above, year-to-year variability. Theoretically, fishing effort, climate and microeconomics may in a synergetic fashion mediate such patterns. Long-term eutrophication of the Black Sea waters and/or locally of Hellenic waters may also partially account for long-term catch trends (Stergiou and Georgopoulos, 1993) but our data cannot justify for such an effect. The role of the remaining factors is discussed below.

There is a lack of long-term studies on the ecology and population dynamics of the studied species in the Hellenic waters. Hence, the factors responsible for their catch trends and variability cannot be conclusively determined. Changes in catches do not necessarily reflect changes in stock strength, especially so for schooling fishes such as the pelagic and semi-pelagic ones (e.g. anchovy, sardine, *Trachurus* spp., *Scomber* spp., bogue). However, annual fishing effort of the four component fisheries has increased considerably between 1964 and 1989 and the Hellenic fleet has been strongly modernised (Stergiou et al., 1994). The FOX models developed in the present study (Table 2) did indicate that fishing effort is strongly related to the long-term catch trends of the four component fisheries. In addition, MREG models suggested that this may also be true of various species or groups of species (Table 3). The catches of pelagic fishes in Hellenic waters are also affected by microeconomics because of their low commercial price and demand both of which result in a high discard rate. Wholesale catch values were only included in the models fitted to the total catches per component fishery. The MREG models pointed to probable catch-value relationships for purse seine and 'other coastal boats' fisheries, for which the catch at year t was found to be positively related to the wholesale value at year $t-1$ (Table 3). Although it is uncertain how catch values can mediate catch changes and, hence, these relationships must be justified from further thorough bioeconomic studies, the positive relationship seems reasonable and implies that high values achieved in the year before represent a motive for fishers to increase next year's catch.

MREG models also pointed to a probable catch–climate relationship for 14 out of 16 species or groups of species (Table 3). From the 14 catch–climate relationships identified through empirical modelling only those referring to anchovy and sardine have a biological explanation which is consistent with previous hypotheses and observations (see below).

For the remaining species (or groups of species) it is uncertain how these climatic variables mediate catch variations and, hence, these relationships must be also justified from further field, or other, studies.

The annual VAR model for the anchovy/sardine complex pointed to a negative relationship (negative coefficient) between sardine and anchovy catches at lag 2 years (Table 9). The above mentioned negative relationships imply a replacement of Hellenic sardine by anchovy catches and vice versa, a fact suggested by various authors for the Mediterranean Sea (Stergiou, 1992) and other areas of the world (e.g. Daan, 1980). Replacement of sardine by anchovy cannot be attributed to competition for food at the larval/juvenile stages since spawning and larval emergence periods for these species do not overlap in time (Stergiou, 1992). The same is also true of the California sardine and anchovy (MacCall, 1983). The fact that the anchovy and sardine catches were correlated with NSW and/or SLP over the Northern Aegean Sea (Table 3) suggests that wind activity may mediate such a replacement either by (a) involving changes in recruitment rates, and/or by (b) involving changes in the relative availability of anchovy and sardine to purse seiners, and/or in the relative fishing effort expended on each species, rather than to changes in the abundance of anchovy and sardine themselves. A detailed account of mechanisms through which NSW can mediate such changes has been given in Stergiou (1992).

Some general conclusions can be drawn from the discussion and facts presented so far. Hence, the fact that MREG and HREG models produced the best fits whereas BES and, to a lesser extent, HES, NM, EMP and HREG models produced the best forecasts, probably indicates that for annual fisheries time series not displaying strong autocorrelations MREG models may be better suited for capturing the longer term trends and the univariate time series models BES, HES and NM for capturing short-term variations, whereas HREG models are probably best for capturing both types of variation. In general, multivariate models should be ideal for use in the following two cases. (a) When independent explanatory variables have been shown, through experimental and/or long-term field studies, to significantly affect the fishery variable concerned (e.g. recruitment, catch, catch per unit effort, spawning biomass). When this is not true, as is the case of the Hellenic and many other Mediterranean fisheries, it is quite likely that spurious correlations may arise, especially so when the number of independent variables is high. We believe that when a sound biological hypothesis supporting the inclusion of one or more explanatory variables into the models is lacking, more credit should be given to variables entering into different empirical models in a consistent manner, i.e. appearance of a variable for a particular species in different multivariate models with a coefficient having a consistently negative or positive value (as was the case with sardine and anchovy in the present study). (b) Should real-time forecasting be the object, multivariate models are meaningful only when reliable forecasts of the independent variables are available or their values are provided to forecasters in real-time (e.g. real-time provision of climatic data through satellites). In cases that (a) and (b) are not satisfied, the univariate BES and HES and multivariate time series HREG models, along with EMP, are probably the most appropriate models for annual fisheries time series.

Acknowledgements

The senior author (K.I.S.) wishes to express his gratitude to the staff of the Pacific Fisheries Environmental Group (PFEG), Monterey Bay, CA, USA for their hospitality which rendered his stay at PFEG pleasant and productive. The senior author would like to thank, in particular, C. Roy for the extraction of the Hellenic climatic COADS data set and for discussions on the limitations of the data set, and P. Cury for fruitful discussions on time series analysis. This study was partially financed by the EU (Contract Number MED92/019). It does not necessarily reflect the views of the Commission, and in no way anticipates the Commission's future policy in this area.

References

Armstrong, J.S. and Lusk, E.J., 1983. Commentary on Makridakis time series competition (M-competition). J. Forecasting, 2: 259–311.

Bakun, A., Christensen, V., Curtis, C., Cury, P., Durand, M.H., Husby, D., Mendelssohn, R., Mendo, J., Parrish, R., Pauly, D. and Roy, C., 1993. The Climate Eastern Ocean System Project. Naga (ICLARM), 15(4): 26–30.

Bocharov, L.N., 1989. Problems of fishery forecasts in Far Eastern Seas from the point of view of systems analysis. Soviet J. Mar. Biol., 14: 1–8.

Borges, M.F., 1990. Multiplicative catch-at-age analysis of scad (*Trachurus trachurus*) from western Iberian waters. Fish. Res., 9: 333–353.

Boudreault, F.R., Dupont, N. and Sylvain, J.N., 1977. Modeles lineaires de prediction des debarquement de homard aux Iles-de-la-Madeleine (Golfe du Saint-Laurent). J. Fish. Res. Board Can., 34: 379–383.

Box, G.E.P. and Jenkins, G.M., 1976. Time Series Analysis, Forecasting and Control. Holden-Day, San Francisco, 575 pp.

Brown, R.G., 1963. Smoothing, Forecasting and Prediction. Prentice Hall, Englewood Cliffs, NJ, USA.

Bulmer, M.G., 1974. A statistical analysis of the ten-year cycle in Canada. J. Anim. Ecol., 43: 701–718.

Caddy, J.F. and Sharp, G.D., 1986. An ecological framework for marine fishery investigations. FAO Fish. Techn. Pap., (283), 152 pp.

Clarke, K.R. and Warwick, R.M., 1989. Lecture notes prepared for the training workshop on the statistical treatment and interpretation of marine community data. September 1989, Part II, FAO/IOC/UNEP, Athens, Hellas.

Cochrane, D. and Orcutt, G.H., 1949. Application of least squares regression to relationships containing autocorrelated error terms. J. Am. Stat. Assoc., 44: 32–61.

Colebrook, J.M. and Taylor, A.H., 1984. Significant time scales of long-term variability in the plankton and the environment. Rapp. P.-v. Reun. Cons. Int. Explor. Mer, 183: 20–26.

Daan, N., 1980. A review of replacement of depicted stocks by other species and the mechanisms underlying such replacement. Rapp. P.-v. Reun. Cons. Perm. Int. Explor. Mer, 177: 405–421.

Dement'Eva, T.F., 1987. A method for correlation of environmental factors and year-class strength of fishes. J. Ictchyol., 27: 55–59.

Engle, R.F., 1984. Likelihood ratio and Lagrange multiplier tests in econometrics. In: Z. Griliches and M.D. Intriligator (Editors), Handbook of Econometrics Vol. II. North-Holland, Amsterdam.

Fox, W.W., 1970. An exponential surplus-yield model for optimizing exploited fish populations. Trans. Am. Fish. Soc., 99: 80–88.

Getz, W.M., Francis, R.C. and Swartzman, G.L., 1987. On managing variable marine fisheries. Can. J. Fish. Aquat. Sci., 44: 1370–1375.

Goodrich, R.L., 1989. Applied Statistical Forecasting. Business Forecast Systems Inc., Belmont, MA, 254 pp.

Grainger, R.J.R., 1979. Herring abundance off the west of Ireland in relation to oceanographic variation. J. Cons. Int. Explor. Mer, 38: 180–188.

Hall, R.E., Johnston, J. and Lilien, D.M., 1990. Micro TSP user's manual for time series analysis, regression and forecasting. QMS Quantitative Micro Software Irvine, Ca.

Hilborn, R., 1987. Living with uncertainty in resource management. N. Am. J. Fish. Manage., 7: 1–5.

Holt, C.C., 1957. Forecasting seasonal trends by exponentially weighted moving averages. Office of Naval Research, Carnegie-Mellon University, Research Memorandum No. 52.

Kort, V.G., 1970. Large-scale interaction between the ocean and the atmosphere using the North Pacific as an example. Oceanology, 10: 171–183.

Litterman, R.B., 1979. Techniques of forecasting using vector autoregressions. Working paper 115, Federal Reserve Bank of Minneapolis, MI.

Ljung, G.M. and Box, G.E.P., 1978. On a measure of lack of fit in time series models. Biometrika, 67: 297–303.

Love, M.S. and Westphal, W.V., 1981. A correlation between annual catches of dungeness crab, *Cancer magister*, along the west coast of North America and mean annual sunspot number. Fish. Bull. US, 79: 794–796.

MacCall, A.D., 1983. Variability of pelagic fish stocks off California. FAO Fish. Rep., 291: 101–112.

Makridakis, S., Wheelwright, S. and McGee, V., 1983. Forecasting: Methods and Applications. Wiley, New York, 926 pp.

McLaughlin, R.L., 1975. The real record of the economic forecasters. Bus. Econ., 10: 28–36.

McLeod, A.I. and Sales, P.R.H., 1983. An algorithm for approximate likelihood calculation of ARMA and seasonal ARMA models. Algorithm AS 191. Appl. Stat., 1983: 211–223.

McLeod, A.I., Noakes, D.J., Hipel, K.W. and Thompstone, R.M., 1987. Combining hydrologic forecasts. J. Wat. Res. Plan. Manage., 113: 29–41.

Mendelssohn, R., 1981. Using Box–Jenkins models to forecast fishery dynamics: identification, estimation and checking. Fish. Bull. US, 78: 887–896.

Mendelssohn, R. and Cury, P., 1987. Fluctuations of a fortnightly abundance index of the Ivorian coastal pelagic species and associated environmental conditions. Can. J. Fish. Aquat. Sci., 44: 408–421.

Mendelssohn, R. and Cury, P., 1989. Temporal and spatial dynamics of a coastal pelagic species, *Sardinella maderensis*, off the Ivory Coast. J. Fish. Res. Board Can., 46: 1686–1687.

Moura, O. and Alfonso dos Santos, G., 1989. Purse seine fishery: a stochastic approach. ICES Symposium on Multispecies Models Relevant to Management of Living Resources. Netherlands Congress Centre, Hague, 2–4 October 1989, 30 pp.

Mysak, L.A., 1986. El Nino, interannual variability and fisheries in the Northwest Pacific Ocean. Can. J. Fish. Aquat. Sci., 43: 464–497.

Noakes, D.J., McLeod, A.I. and Hipel, K.W., 1985. Forecasting monthly riverflow time series. Int. J. Forecast., 1: 179–190.

Noakes, D.J., Welch, D.W., Henderson, M. and Mansfield, E., 1990. A comparison of preseason forecasting methods for returns of two British Columbia sockeye salmon stocks. N. Am. J. Fish. Manage., 10: 46–57.

NSSH (National Statistical Service of Hellas), 1965–1992. Statistical Year-Book. 26 issues (for the years 1964–1989). NSSH, Athens, Hellas.

Pauly, D., 1989. Biology and management of tropical marine fisheries. Res. Manage. Optim., 6: 253–271.

Pitcher, J.T. and Hart, P.J.B., 1982. Fisheries Ecology. Croom Helm, London, 414 pp.

Pope, J.G. and Shepherd, J.G., 1985. A comparison of the performance of various methods for tuning VPAs using effort data. J. Cons. Int. Explor. Mer, 42: 129–151.

Regner, S. and Gacic, J., 1974. The fluctuation of sardine catch along the eastern Adriatic coast and solar activity. Acta Adriat., 15: 1–15.

Roy, C., Cury, P. and Kifani, S., 1992. Pelagic fish recruitment success and reproductive strategy in upwelling areas: environmental compromises. S. Afr. J. Mar. Sci., 12: 135–146.

Saila, S.B, Wigbout, J. and Lermit, R.J., 1979. Comparison of some time series models for the analysis of fisheries data. J. Cons. Perm. Int. Explor. Mer, 39: 44–52.

Schlegel, G., 1985. Vector autoregressive forecasts of recession and recovery: is less more? Econ. Rev., IIQ/1985: 2–12.

Shepherd, J.G., 1984. Status quo catch estimation and its use in fisheries management. ICES CM, 1984/G:5: 14 pp.

Shuntov, B.P., Shurunov, N.A. and Vasil'kov, V.P., 1981. Periodicity in the reproduction of fish from the New Zealand plateau based on data from commercial statistics. J. Ichthyol., 22: 32–43.

Sparre, P., Ursin, E. and Venema, S.C., 1989. Introduction to tropical fish stock assessment. Part 1 – Manual. FAO Fish. Tech. Pap., 301(1), 337 pp.

STCS, 1993. Statgraphics plus reference manual. Manugistics, MD.

Stellwagen, E.A. and Goodrich, R.L., 1993. FORECAST/PRO for WINDOWS. BFS Business Forecast Systems, Belmont, MA.

Stergiou, K.I., 1989. Modelling and forecasting the fishery of pilchard, *Sardina pilchardus*, in Hellenic waters using ARIMA time series models. J. Cons. Int Explor. Mer, 46: 16–23.

Stergiou, K.I., 1990a. An autoregressive model of the anchovy *Engraulis encrasicolus* fishery in the eastern Mediterranean. Fish. Bull. US, 88: 411–414.

Stergiou, K.I., 1990b. Prediction of the Mullidae fishery in the eastern Mediterranean 24 months in advance. Fish. Res., 9: 67–74.

Stergiou, K.I., 1991a. Describing and forecasting the sardine–anchovy complex in the eastern Mediterranean using vector autoregressions. Fish. Res, 11: 127–141.

Stergiou, K.I., 1991b. Short-term fisheries forecasting: comparison of smoothing, ARIMA and regression techniques. J. Appl. Ichthyol, 7: 193–204.

Stergiou, K.I., 1992. Variability of the monthly catches of anchovy, *Engraulis encrasicolus* in the Aegean Sea. Fish. Bull. US, 90: 211–215.

Stergiou, K.I. and Georgopoulos, D., 1993. The distribution of phytoplankton pigments and the fishery of anchovy (*Engraulis encrasicolus*) in the Hellenic Seas. Naga (ICLARM) 16(2): 34–37.

Stergiou, K.I. and Petrakis, G., 1993. Description, assessment of the state and management of the demersal and inshore fisheries resources in the Hellenic Seas. Fresenius Environ. Bull., 2: 312–319.

Stergiou, K.I. and Pollard, D., 1994. A spatial analysis of the commercial fisheries catches from the Hellenic Aegean Sea. Fish. Res., 20: 109–135.

Stergiou, K.I., Christou, E. and Petrakis, G., 1994. Operational forecasting of Hellenic fisheries catches. Final Report (Contract No Med92/019, EU), Stergiou & Co., Athens, June 1994, 1, 184 pp.

Stocker, M. and Noakes, D.J., 1988. Evaluating forecasting procedures for predicting Pacific herring (*Clupea harengus pallasi*) recruitment in British Columbia. Can. J. Fish. Aquat. Sci., 45: 928–935.

Theil, H., 1966. Applied economic forecasting. North-Holland, Amsterdam.

Todd, R.M., 1984. Improving economic forecasting with Bayesian Vector Autoregressions. Quarterly Review, Federal Reserve Bank of Minneapolis, MI, Fall 1984, pp. 18–29.

Vasil'kov, V.P., Chupysheva, N.G. and Kolesova, N.G., 1981. The possibility of long-term forecasting from solar activity cycles of catches of the Saffron cod, *Eleginus gracilis*, in the sea of Japan. J. Ichthyol., 22: 26–33.

Vibe, C., 1967. Arctic animals in relation to climatic fluctuations. Meddr. Gronland, 170(5): 1–227.

Zupanovic, S., 1968. Causes of fluctuations in sardine catches along the eastern coast of the Adriatic Sea. Anali Jandranskog Instituta, 4: 401–489.