

Time series analysis and forecasting techniques applied on loliginid and ommastrephid landings in Greek waters

S. Georgakarakos^{a,*}, D. Koutsoubas^{b,1}, V. Valavanis^{c,2}

^a Fisheries and Sonar Laboratory, Department of Marine Sciences, University of the Aegean, University Hill, 81100 Mytilini, Lesvos Island, Greece

^b Department of Marine Sciences, University of the Aegean, University Hill, 81100 Mytilini, Lesvos Island, Greece

^c Institute of Marine Biological Resources, Hellenic Centre of Marine Research, P.O. Box 2214, 71003 Heraklion, Crete, Greece

Abstract

Time series analysis techniques (ARIMA models), artificial neural networks (ANNs) and Bayesian dynamic models were used to forecast annual loliginid and ommastrephid landings recorded from the most important fishing ports in the Northern Aegean Sea (1984–1999). The techniques were evaluated based on their efficiency to forecast and their ability to utilise auxiliary environmental information. Applying a “stepwise modelling” technique, namely by adding stepwise predictors and comparing the quality of fit, certain inferences concerning the importance of the predictors were made.

The ARIMA models predicted the test data very precisely (high R^2), especially if the target time series contained a strong autoregressive character, after they were first differenced to obtain stationarity ($R^2 > 0.96$). The disadvantage of the ARIMA, as with most statistical models, is their assumption that the relationships and system parameters remain the same across the observation and forecasting periods.

The influence of temperature on catches was mainly investigated by applying neural models, which predicted the monthly landings with high precision ($R^2 = 0.89$), even when incorporating in the model exclusively monthly SST descriptors. Similarly, ANN models of annual landings containing monthly mean temperatures provided high precision ($R^2 = 0.87$) and valuable inference concerning the possible effect of the SST in certain months.

Bayesian dynamic models also provided a high precision ($R^2 = 0.96$). They combined the information of both environmental and landing time series, namely the monthly mean temperatures and the monthly seasonality of the landings. The impact factors estimated from the model have the form of time series representing the temperature effect.

The results reveal that both the monthly and the annual landings can be predicted and that the Bayesian model is the best performer overall, characterised by a higher number of stable forecasts, and forecasts with higher precision and accuracy, than the other methods. It is evident, from application of the “stepwise modelling” technique, that the incorporation of temperature descriptors can significantly improve the model performance.

© 2006 Elsevier B.V. All rights reserved.

Keywords: ARIMA models; Artificial neural networks; Bayesian models; Cephalopods; Forecasting

Abbreviations: ACF, sample autocorrelation function; AIC, information criterion; ANN, artificial neural networks; ARIMA, autoregressive integrated moving average; AVHRR, advanced very high resolution radiometer; BDM, Bayesian dynamic models; SAF, seasonal adjustment factor, which contains the seasonality of the time series; SAS, seasonal adjusted series (the original series minus the seasonality); SBC, Schwartz’s Bayesian criterion; STC, de-seasoned trend and cycle; SST, sea surface temperature; PACF, sample partial autocorrelation function; RCF, relative contribution factor

* Corresponding author. Tel.: +30 22510 36822.

E-mail addresses: stratisg@aegean.gr (S. Georgakarakos), drosos@aegean.gr (D. Koutsoubas), vasilis@imbc.gr (V. Valavanis).

¹ Tel.: +30 22510 36814.

² Tel.: +30 2810 337817.

1. Introduction

In fisheries management, as well as in other cases, forecasting strategies are based on the development of either descriptive or explanatory models. The first category, usually a synonym of “black box” models (Haddon, 2001), is a necessary approach, whenever important information for the construction of explanatory models fails. In fisheries, explanatory models often obey certain assumptions (Fox, 1970; Schaefer, 1984) or are susceptible to the completeness and the quality of the data (Hilborn and Walters, 1992). Therefore, and due to their parsimony,

monious character, descriptive models gained in popularity at least for forecasting purposes (Cohen and Stone, 1987). In addition to the forecasting character, the multivariate descriptive models have the advantage that by “stepwise modelling”—namely by adding stepwise predictors and comparing the quality of fit, certain inferences concerning the importance of the predictors can be made.

Descriptive models used to predict and analyse time series data attempt to decompose up the dependent variable into four main components. Simple time trends, periodic fluctuations, predictors’ effect and the error component. A common realisation of this approach is the development of the multivariate ARIMA models (Box and Jenkins, 1976).

In the present work, three modelling and forecasting techniques were evaluated on the basis of their efficiency to forecast and their ability to utilise auxiliary environmental information: ARIMA models, artificial neural networks (ANNs) and dynamic models (DMs).

ARIMA and ANNs have been fitted to the data using Frequentist techniques (e.g. maximum likelihood technique) while the dynamic models have been fitted using the Bayesian estimation technique.

In particular, neural networks are very sophisticated non-linear techniques and have been chosen from other non-linear tools, because of their capability to model extremely complex functions.

The models have been applied to the same data set, namely the landings of long-finned squids (*Loliginidae*) and short-finned squids (*Ommastrephidae*), recorded in the northern part of the Aegean Sea during the 1984–1999 time period. Previous attempts to model cephalopod landings in Greece (Stergiou, 1987, 1989; Stergiou and Christou, 1996; Georgakarakos et al., 2002) showed that sea temperature could be an important environmental predictor. The aim of the present work is to investigate further the environmental impact on such biological resources, like the cephalopods in Greece, where fishing effort and biological data are lacking or limited, rendering the development of stochastic, non-linear and dynamic models more attractive.

Readers interested in the biological and geographic characteristics of the target species may refer to previous contributions; about a general aspect of the invertebrates (Caddy, 1989), their distribution in the Mediterranean and eastern Atlantic (Roper et al., 1984) or the Aegean Sea (D’onghia et al., 1996; Koutsoubas et al., 1999; Arvanitidis et al., 2002; Moreno et al., 2002). A comparison of cephalopod landings in the Northeast Atlantic and Mediterranean, over the period 1980–1992, supported the hypothesis of a large-scale climatic influence on the cephalopod abundance (Pierce et al., 1995). Additionally references for the development of fisheries forecasting models are included in Caddy (1989).

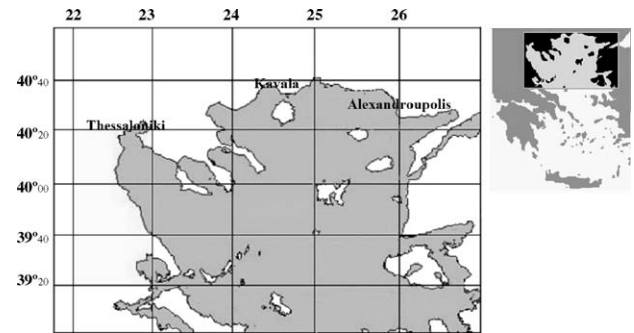


Fig. 1. Northern Aegean Sea. The map shows the three most important fishing ports in the area. The geographic rectangles were used for the analysis of the SST data.

2. Materials and methods

2.1. Study area, source of data and pre-processing

More than 50% of the cephalopod production in Greek Seas is concentrated in the North Aegean Sea (Koutsoubas et al., 1999). The present statistical analysis is based on the monthly landings from the three most important fishing ports of this area, namely from Thessaloniki, Kavala and Alexandroupolis (Fig. 1). The landing data have been provided by the Fishing Development Corporation in Greece (ETANAL S.A).

Daily satellite images, provided from the National Oceanic and Atmospheric Administration (NOAA), were transformed into Sea Surface Temperature (SST) in Celsius by using GIS tools. The area under study was divided in statistical rectangles ($1^{\circ}\text{N} \times 1/3^{\circ}\text{E}$) and in each rectangle the spatially averaged temperature means, minima, maxima, and the spatio-temporal standard deviation in each month were estimated. Finally, the same descriptors were estimated for each landing port, weighted according to the number of valid pixels per image (Valavanis et al., 2002).

All data sets, landings and SST descriptors cover the period between the years 1984 and 1999. The time series were tested for stationarity, evaluating arithmetic mean, variance, and autocorrelation through time. Time series used in the ARIMA models were first differenced until they were stationary in the mean, and log transformed to stabilise the variance. For the development of the other models, the landings were previously log transformed.

The statistical procedures mentioned in the present study were implemented, if not otherwise mentioned, using the SPSS for Windows software, SPSS Trends routines (SPSS Inc., 2002). Sample autocorrelations (ACF) served also as a measure of the periodicity in the time series and were plotted systematic for both landings and environmental descriptors. Seasonal dependency was proved by comparing the results of the Partial Autocorrelation Function (PACF), which considerably reduces the dependence on the intermediate elements, within the lag, and the results of the ordinary ACF (Box and Jenkins, 1976; see also McDowall et al., 1980).

2.2. Seasonal decomposition procedure

The goal of the seasonal decomposition procedure was to extract a seasonal component and a combined trend—cycle component from the observed signal, assuming the residual as the “error” component. In the present study, the seasonality of the time series was represented by the 12 months periodicity. The following components or combinations of them were estimated by the SPSS Trend routines:

- SAF: seasonal adjustment factor, which contains the seasonality of the time series,
- SAS: seasonal adjusted series (the original series minus the seasonality),
- STC: de-seasoned trend and cycle (the trend component plus the cycle) and
- ERR: the residuals.

The residuals were tested to ensure pure randomness (white noise properties) and that they were not correlated to other potential predictors. The other generated components were further analysed, checking their distributions and their auto- and cross-correlograms. The characteristics of the STCs, which include the trend and the non-seasonal periodicity, were also investigated for autocorrelation, trends, and possible correlation with other parameters. Finally, all these parameters were crosschecked as potential ARIMA predictors.

2.3. Development of models

2.3.1. ARIMA models

ARIMA models were developed following the standard three-step procedure described in Box and Jenkins (1976), namely model identification, parameter estimation and finally the diagnosis of the simulation and its verification (see also, Makridakis, 1990).

The general form of the model can be summarised by the commonly used three types of parameters: the autoregressive parameters (p), differencing passes (d), and moving average parameters (q). In the standard notation “ p – d – q ”, a model described as (0, 1, 2) means that it contains 0 (zero) autoregressive (p) parameters and 2 moving average (q) parameters which were computed for the series after it was differenced once ($d=1$). For the seasonal ARIMA process the required parameters sp , sd and sq are also determined interpreting the results of the corresponding ACF and PACF. The identification of the tentative model starts with the estimation of the seasonal model and afterwards with the analysis of its residuals. This can provide a clearer view of the futures of the non-seasonal model. In a well-identified seasonal model, its residuals show the non-seasonal portion of it.

The parameters of the tentative model were estimated by applying maximum-likelihood techniques. The derived residuals should be without any pattern. Therefore, the residuals of all presented models were tested to be normal distributed and non-autocorrelated (non-significant Box–Ljung Q statistic).

Finally, the following results were summarised: the parameter estimates, standard errors, estimate of residual variance, standard error of the estimate, log likelihood, Akaike’s Information Criterion (AIC), and Schwartz’s Bayesian Criterion (SBC). The minimising of SBC and AIC were used, taking into account both how well the model fitted the observed series, and the number of parameters used in the fit (SPSS manual, Trends, Release 6.0).

These procedures were applied separately to the landings and the SST time series, in a univariate analytical approach. After the estimation of their parameters, new multivariate ARIMA models were developed, taking as covariates the SST parameters (mean, minimum and maximum values) as well as STC and the residuals of the seasonal decomposition.

Finally, the multivariate model was compared to the univariate one in order to test the significance of the covariate contribution.

A more detailed description of the development procedure can be found in Georgakarakos et al. (2002) included. In all cases, if not mentioned otherwise, the data from 1984 to 1998 were used for evaluating (fitting) the model, while the data of the last 12 months (1998–1999) were used for testing the forecasting performance of the established models.

2.3.2. Neural network development

Artificial neural networks (ANNs) comprise algorithms for mapping the input vector (predictors) to an output vector (responses), through processing elements called ‘neurons’, via a training process. The most common type of ANNs is a three-layer backpropagation network, which consists of three basic groups of neurons:

- The input neurons collect the external information and send it to the hidden neurons in the middle layer.
- The hidden neurons accumulate the previous signals, adjusting each input by certain weights and applying certain threshold functions.
- The output neurons accumulate in a second step the input from the hidden neurons, applying again a set of weight parameters and threshold functions.

A three-layer backpropagation network, which includes i input, j hidden and k output neurons, can be represented in mathematics by the following equation (Ripley, 1996):

$$y_k = f_k \left(a_k + \sum_{j \rightarrow k} w_{jk} f_j \left(a_j + \sum_{i \rightarrow j} w_{ij} x_i \right) \right) \quad (1)$$

where y_k is the neural output of the k neuron, f_k and f_j are activation or threshold functions.

The intercepts a_k and a_j are known as *biases* and w_{jk} or w_{ij} are *weights* which multiply the signals travelling from j to k or from i to j correspondingly.

The activation functions are used to introduce non-linearity at the hidden layer, and are often taken to be

$$\text{sigmoid } \sigma(z) = \frac{1}{1 + e^{-z}} \quad (2a)$$

or

$$\text{logistic } f_h(x) = \frac{e^x}{1 + e^x} \quad (2b)$$

Standard three-layer backpropagation networks were chosen for modelling the monthly landings, due to their good generalisation efficiency. All ANNs were constructed and supervised following the same procedures and initial settings, in order to have comparable results by their application. The networks were developed in the NeuroShell 2 environment, provided by the Ward Systems Group.

The monthly landings of the last year have been selected for building the test set, leaving the remainder as the training set. After each complete learning pass through the training data, the software automatically predicts equivalent values to the test data and estimates the error factor of the prediction. This comparison controls the training process and avoids network overtraining. Once, the computed error factor ceases to improve, the constructed network can be applied to the total data or only to the test data set.

The performance of the ANNs was measured by estimating the R^2 -adjusted coefficient of determination, an indicator usually applied to multiple regression analysis (Zar, 1999). It compares the accuracy of the model to the accuracy of a trivial benchmark model wherein the prediction is just the mean of all of the samples. It is adjusted for the different degrees of freedom or amounts of replication in the data set. Note that it is different from the squared multiple correlation coefficient. A detailed description of the method is contained in Twomey and Smith (1996).

2.3.3. Bayesian dynamic system analysis

Forecasting in a Bayesian approach is the quantification of the probability distribution, which describes the nature of the uncertainty, taking into account existing knowledge. The dynamic models (DMs) were based on sequential Bayesian learning and forecasting procedures across the whole time space, according to the well-known relation between prior and posterior information:

$$\text{posterior} \sim \text{prior} \times \text{likelihood}$$

Using the standard notation of a linear DM (West and Harrison, 1989) the vector of response variables Y is related to the regressor matrix X as:

$$Y = X\theta + v \quad (3a)$$

where θ is the vector of the unknown parameters, which should be estimated and v is the vector of the uncorrelated stochastic errors.

In accordance, for a single observation the above equation is:

$$Y_t = x_t'\theta_t + v_t \quad (3b)$$

where x is a column vector and x' is its transpose.

Consequently, the univariate linear DM that describes the monthly landings can be written:

$$\text{landings}_t = F_t'\theta_t + v_t, \quad v_t \sim N[0, V_t] \quad (4)$$

$$\theta_t = G_t\theta_t + w_t, \quad w_t \sim N[0, W_t] \quad (5)$$

where landings_t are the observed landings value at time t , F_t a vector of known constants (the regression vector), θ_t denotes the vector of model state parameters, v_t a stochastic error term having a normal distribution with zero mean and variance V_t , G_t a matrix of known coefficients that defines the systematic evolution of the state vector across time and w_t is a stochastic error term having a normal distribution with zero mean and covariance matrix W_t .

The prior information on the state vector for time $t+1$ is represented by a normal distribution with mean $= a_{t+1}$ and covariance R_{t+1} ,

$$\theta_{t+1}|D_t \sim N[a_{t+1}, R_{t+1}] \quad (6)$$

where D_t denotes the state of knowledge at time t .

In order to forecast landings in the next time period (landings_{t+1}) a linear combination of the normally distributed variables $\theta_{t+1}|D_t$ and the v_{t+1} is used. The forecast mean will be equal to:

$$\begin{aligned} E[Y_{t+1}|D_t] &= E\{F_{t+1}'\theta_{t+1} + v_{t+1}|D_t\} = E\{F_{t+1}'\theta_{t+1}|D_t\} \\ &\quad + E\{v_{t+1}|D_t\} = F_{t+1}'E\{\theta_{t+1}|D_t\} + E\{v_{t+1}\} \\ &= F_{t+1}'a_{t+1} \end{aligned} \quad (7)$$

while the forecast variance will be equal to:

$$\begin{aligned} V[Y_{t+1}|D_t] &= V\{F_{t+1}'\theta_{t+1} + v_{t+1}|D_t\} = V\{F_{t+1}'\theta_{t+1}|D_t\} \\ &\quad + V\{v_{t+1}|D_t\} = F_{t+1}'V\{\theta_{t+1}|D_t\}F_{t+1} \\ &\quad + V\{v_{t+1}\} = F_{t+1}'R_{t+1}F_{t+1} + V_{t+1} \end{aligned} \quad (8)$$

In the previously described equations, it is assumed that the stochastic time series are not autocorrelated or cross-correlated. Data sets providing models with correlated stochastic terms can be transformed in order to remove the correlation (Box and Jenkins, 1976). Furthermore, we assume that the system under study is invariant and therefore G_t , the matrix of known coefficients that defines the systematic evolution of the state vector across time is constant ($G_t \equiv G$).

In addition, both the observation variance and the system covariances are unknown and in the Bayesian approach should be specified as probability distributions, which will be updated with new observations according to the Bayes Theorem. The same is valid for any other prior, as for instance the priors describing the seasonal components. Priors either

are specified by the BATS user through intervention, as a pure subjective hypothesis, or are the result from the previous period’s posterior. This procedure, which assigns some starting values to the model parameters, is called reference analysis. It calculates one-step-ahead point forecasts, based on the assumed parameter values deemed most appropriate for one-step forecasts.

Finally, the resulted residuals should be tested and especially forecast residuals are tested on-line during monitoring to ensure that they are not autocorrelated.

A detailed description of the method is contained in Box and Tiao (1973) and West and Harrison (1989). Reference to the applied software for the Bayesian algorithm in DMs can be found in Pole and West (1990).

2.3.4. Step by step DM development

For the time series of landings, we can split up the response variable into three components: an underlying level independent from the regressor ($level_t$), a second part determined by the regressor and a third component, that represents the stochastic error. The Observation Equation for the model that describes the landings is written:

$$landings_t = level_t + \beta_t x_t + v_t, \quad v_t \sim N[0, V] \quad (9)$$

where x_t is the regressor value at time t , v_t is a stochastic error at time t having a normal distribution with zero mean and variance V .

The two uncertain parameters $level_t$ and β_t may be expressed by the two system equations, which have the form of a simple random walk, as:

$$level_t = level_{t-1} + \Delta level_t, \quad \Delta level_t \sim N[0, W_1] \quad (10)$$

$$\beta_t = \beta_{t-1} + \Delta \beta_t, \quad \Delta \beta_t \sim N[0, W_2] \quad (11)$$

In standard notation, the above observation and system equations should be written:

$$Y_t = F_t' \theta_t + v_t, \quad v_t \sim N[0, V] \quad (12)$$

$$\theta_t = G_t \theta_{t-1} + w_t, \quad w_t \sim N[0, W] \quad (13)$$

where Y_t is the time series of the observed monthly landings, F_t a vector of known constants (the regression vector), θ_t denotes the vector of model state parameters ($level, \beta$), v_t the observation stochastic error normal distributed with zero mean and variance V , G_t a matrix of known coefficients that defines the systematic evolution of the state vector across time and w_t is the system stochastic error normal distributed with zero mean and covariance matrix W .

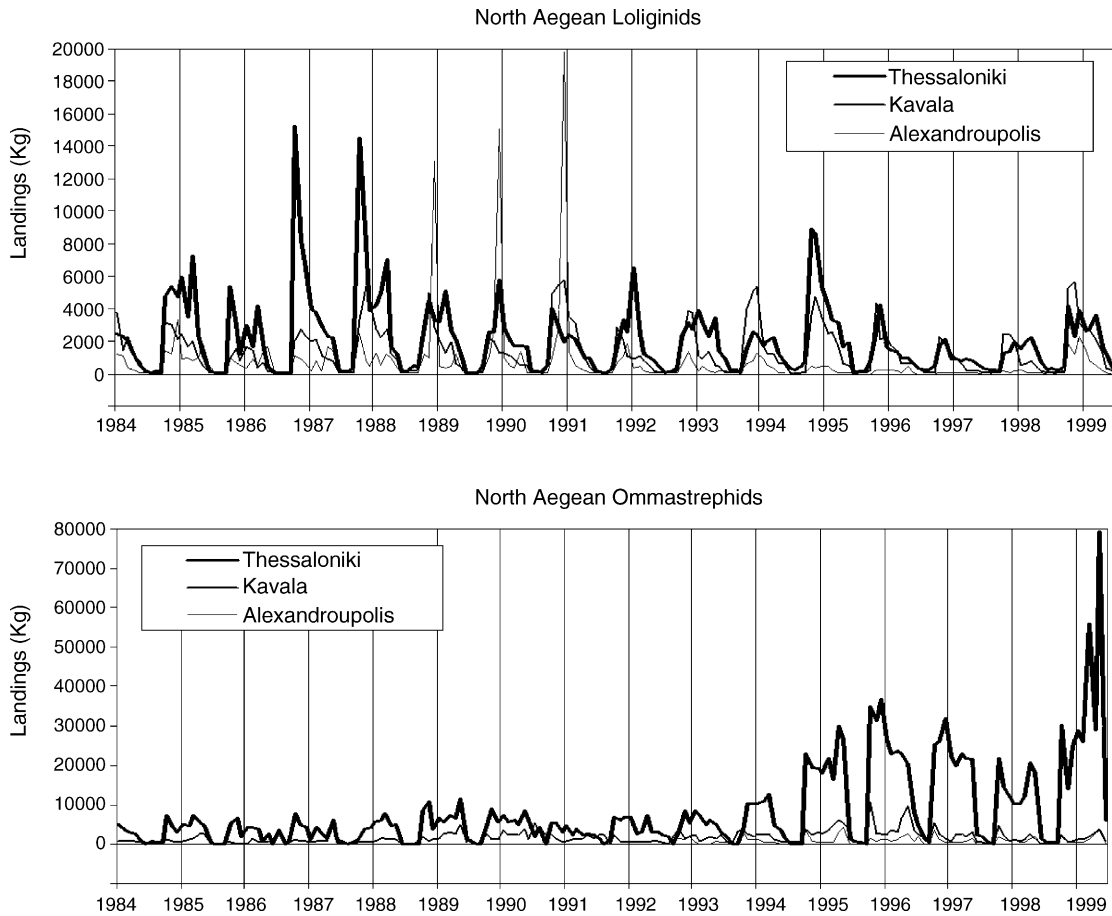


Fig. 2. Time series of loliginid (above) and ommastrephid (below) monthly landings (Kg) for the major three fishing ports of North Aegean Sea (source: ETANAL S.A).

All models were started from an initial steady state model, which was developed according to the forms:

$$\text{landings} = \text{level}_t + v_t, \quad v_t \sim N(0, V_t) \quad (14)$$

$$\text{level}_t = \text{level}_{t-1} + w_t, \quad w_t \sim N(0, W_t) \quad (15)$$

This initial model assumes that the mean remains roughly constant over short time periods, shifting by increments obeying given distributions, with the observation variance (V_t) and the system variance (W_t). The starting conditions for the analysis are based on the default value provided by the reference analysis.

In a next step, a series of models was developed by adding certain explanatory environmental variables (SST descriptors) as regressors and testing the performance against the steady state models. The new models include as dynamic parameter certain SST descriptors according to the form:

$$\text{landings} = \text{level}_t + b_t \text{SST}_t + v_t \quad (16)$$

where

$$b_t = b_{t-1} + w'_t, \quad w'_t \sim N(0, W'_t) \quad (17)$$

Note that by adding regression components, the level is effectively playing the role of a dynamic regression intercept, while the b_t is a varying slope of the regression (Eq. (7)). Bearing this interpretation in mind there is a strong argument for the distribution of the level to be quite close to the regression distribution.

At the end of the development phase all models were expanded by incorporating the monthly seasonality, namely by adding 12 seasonal factors, 1 for each month.

The performance of all models was tested by three tools, by estimating the coefficient of determination R^2 between the predicted versus the observed values, by analysing the statistics of their residuals and finally by cross-validating their predictions with unused, during the modelling procedure, test data.

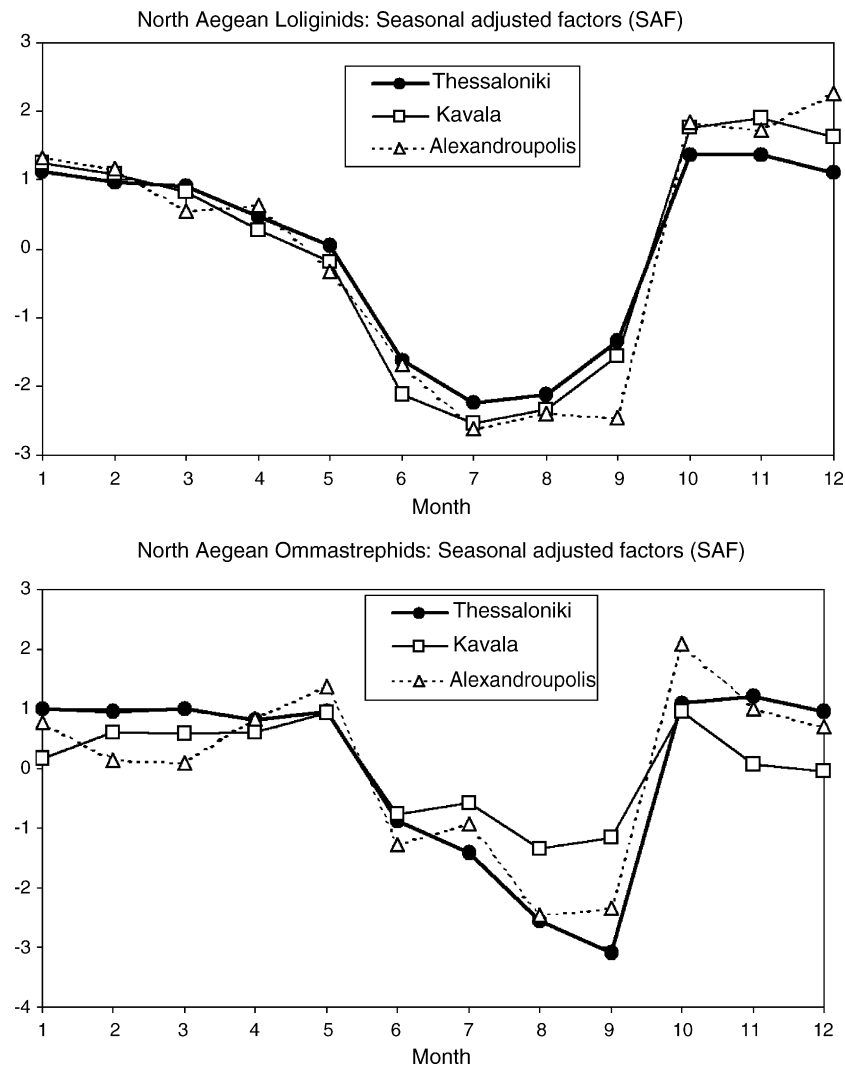


Fig. 3. Seasonal adjusted factors (SAFs) after the seasonal decomposition of Loliginids (above) and Ommastrephids (below) series, for the three major fishing ports of the northern Aegean Sea.

3. Results

3.1. Analysis based on classical time series models

Monthly loliginid and ommastrephid landings in all fishing ports in Greece oscillate with an annual periodicity, generally showing their minima during the summer time (Fig. 2). Several temperature spatio-temporal estimates (mean, minimum and maximum of SST) show a strong seasonal periodicity, in step with the landings, which may complicate any approach to deconvolute interactions between the environmental and fishery data. As it was demonstrated in a previous investigation, seasonality did not increase with the level of the landings and additive models can be used successfully for standard Seasonal Decomposition analysis (Georgakarakos et al., 2002).

Time series of landings by species group and fishing ports were pre-processed calculating the first-order differenced seasonal transformed series, in order to fulfil the stationarity requirements.

3.1.1. Seasonal decomposition

The time series of loliginid and ommastrephid monthly landings of the target areas were decomposed into a seasonal factor, a cycle component, a combined trend and the remaining “error”. The estimated adjustment factor (SAF) in Fig. 3 portrays the well-known pattern of the rapidly decreasing landings during the summer time. Loliginid landings showed less variability among the different areas compared to the ommastrephid data. The STC component, which contains the trend component plus the cycle, provided indications for trends longer than 12 months, as it is demonstrated, for example, in the ACF plots of the loliginid landings in Thessaloniki (Fig. 4).

3.1.2. ARIMA models

The ARIMA models cover the period 1984–1999, while the last 12 months (July 1998 to June 1999) were kept out of fitting and used only for the cross-validation of the forecasting. The model development proceeded by the following steps:

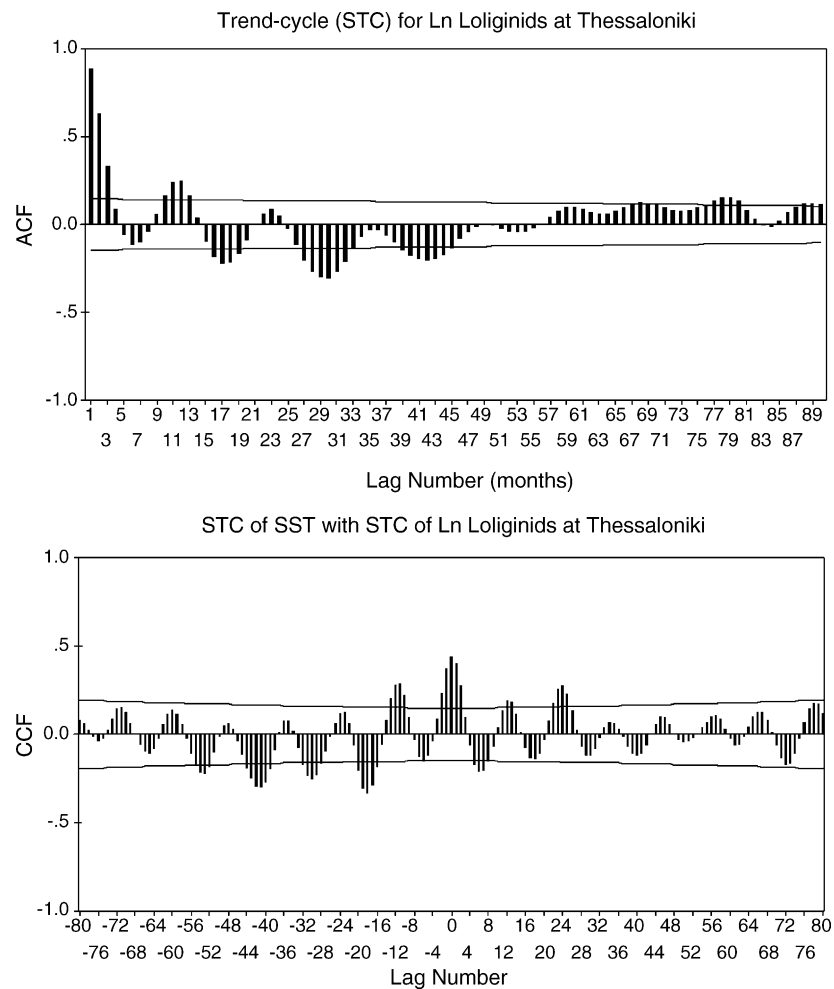


Fig. 4. Autocorrelation function (ACF) of the trend-cycle component (STC) after seasonal decomposition of the Thessaloniki landings of Loliginids (above). Cross-correlation function (CCF) between SST (STC) and landings (STC) (below). In addition to the correlation coefficients, the ACF and CCF plots show the 95% confidence limits.

The ACF and PACF of the natural log transformed series showed large values at lags 12, 24 and 36. The slowness with which values at those seasonal lags declined confirmed that seasonal differencing was required to achieve a stationary mean.

The seasonal ACFs and PACFs smoothed out the rapid seasonal fluctuations, while the ACFs still showed significant correlation, with a single seasonal spike emerging at lag 12

for loliginids and ommastrephids. The PACFs showed a large spike at lag 12 and a smaller one at lag 24.

The pattern 'one spike in ACF, rapidly declining PACF' was interpreted as a MA(1) (Moving Average of order 1) process, here a seasonal MA(1) process, since the pattern appeared at the seasonal lags. The tentative seasonal model was (0, 1, 1) since the data were from seasonal differenced series.

Table 1

Coefficients* and summary statistics of ARIMA modelling of the loliginid and ommastrephid landings (natural log transformed) at the three ports of Northern Aegean Sea

Species	Loliginids			Ommastrephids		
	Thessaloniki	Kavala	Alex/polis	Thessaloniki	Kavala	Alex/polis
A. Univariate ARIMA						
ARIMA model:	(1–0–1) (0–1–1) ¹²	(1–0–1) (0–1–1) ¹²	(1–0–1) (0–1–1) ¹²	(1–1–1) (0–1–1) ¹²	(1–0–1) (0–1–1) ¹²	(1–0–1) (0–1–1) ¹²
Number of residuals	162	162	162	161	162	57
Standard error	0.646	0.656	0.985	1.009	1.086	1.023
<i>B</i> (coefficient)						
AR1	0.6502	0.6044	0.5915	0.3161	0.75105	0.25717
MA1	0.2626	0.1737	0.1146	0.9749	0.36091	–0.58512
SMA1	0.5978	0.7938	0.4506	0.9044	0.77247	0.26203
Standard error of <i>B</i>						
AR1	0.1280	0.1290	0.1241	0.0753	0.09637	0.18566
MA1	0.1620	0.1582	0.1562	0.0420	0.13499	0.15156
SMA1	0.0703	0.0696	0.0771	0.0997	0.06978	0.15633
Approximated probability						
AR1	0.0000	0.0000	0.0000	0.0000	0.00000	0.01717
MA1	0.0107	0.0274	0.0464	0.0000	0.00829	0.00030
SMA1	0.0000	0.0000	0.0000	0.0000	0.00000	0.09949
Adjusted <i>R</i> ²						
Fit	0.9912	0.9890	0.9636	0.9838	0.9739	0.9721
Forecast	0.9985	0.9961	0.9798	0.9934	0.9902	0.9684
Covariate	STC of SST-max	STC of SST-min	STC of SST-max	SAS of SST-min	STC of SST-max	SST-max
B. Multivariate ARIMA						
Number of residuals	162	162	162	161	162	57
Standard error	0.631	0.652	0.970	0.939	1.075	0.996
<i>B</i> (coefficient)						
AR1	0.6132	0.5328	0.4953	0.2970	0.7825	0.3358
MA1	0.2687	0.1017	0.0331	0.9738	0.4035	–0.5791
SMA1	0.6313	0.8199	0.4940	0.8892	0.7800	0.1242
Covariate	0.3260	0.1073	0.4929	–0.3034	–0.3621	–0.2398
Standard error of <i>B</i>						
AR1	0.1498	0.1440	0.1467	0.0767	0.0886	0.1770
MA1	0.1831	0.1678	0.1730	0.0486	0.1284	0.1525
SMA1	0.0685	0.0698	0.0765	0.0908	0.0696	0.1635
Covariate	0.1128	0.0881	0.1963	0.0896	0.1815	0.1007
Approximated probability						
AR1	0.0001	0.0003	0.0009	0.0002	0.0000	0.0633
MA1	0.0444	0.0545	0.0486	0.0000	0.0020	0.0004
SMA1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0510
Covariate	0.0044	0.1103	0.0130	0.0009	0.0477	0.0209
Significance level of covariate	0.9956	0.8897	0.9870	0.9991	0.9523	0.9791
Adjusted <i>R</i> ²						
Fit	0.9916	0.9890	0.9649	0.9843	0.9745	0.9745
Forecast	0.9978	0.9952	0.9833	0.9910	0.9897	0.9634

*AR1: autoregressive coefficient of order 1, MA1: moving average coefficient of order 1, SMA1: seasonal moving average coefficient of order 1.

After the coefficients of the seasonal model had been estimated, the ACF and PACF of their residuals were examined. Their ACFs started large and then died out, while their PACFs also died out somewhat more quickly. The non-seasonal model could therefore be ARIMA (1, 0, 1) in all cases except for ommastrephids of Thessaloniki where the non-seasonal model of ARIMA (1, 1, 1) performed best.

The combined tentative models incorporating both non-seasonal and seasonal parameters in the ARIMA($p-d-q$) format were ARIMA(1, 0, 1)(0, 1, 1)₁₂ or ARIMA(1, 1, 1) (0, 1, 1)₁₂. The coefficients and the summary statistics of these univariate ARIMA models are given in Table 1.

In order to test whether the SST descriptors could improve the model performance multivariate ARIMA models were developed, incorporating certain SST parameters (mean, minimum and maximum values as well as STC and the residuals of their seasonal decomposition). The covariate that best improved each of the initial (univariate) ARIMA model in terms of reduction of the standard error and the residual variance is presented in Table 1. The STC and the SAS of the maximum value of SST gave the best results for loliginids in all fishing ports. For ommastrephids, different parameters of SST (min, max and mean) for each port model gave the best results.

The observed values (transformed with natural logarithm) along with their model fit and confidence intervals were plotted (as an example, the models of the Thessaloniki fishing port are shown in Fig. 5). The ACF and PACF of the model residuals (not shown) did not provide any significant pattern; i.e. the residuals had a stochastic character.

The observed values of the forecasting period 1998–1999 fall within the confidence intervals of the ARIMA models.

The predicted values, for both the fit and forecast periods, were compared to the observed ones and the estimated adjusted R^2 varied between 0.96 and 0.99 (Table 1).

3.2. Time series analysis via artificial neural networks

3.2.1. Stepwise modelling the monthly landings by applying environmental predictors

Standard backpropagation ANNs were applied for modelling the monthly landings of the target species using different combinations of regressors. Firstly, simple ANNs incorporating as inputs only the year and the month of the observed landings were developed, in order to absorb any trend and seasonality of the time series, not included in the “environmental variables”. The incorporation of a “dummy” variable (e.g. year) was used as a tool to absorb long-term variability.

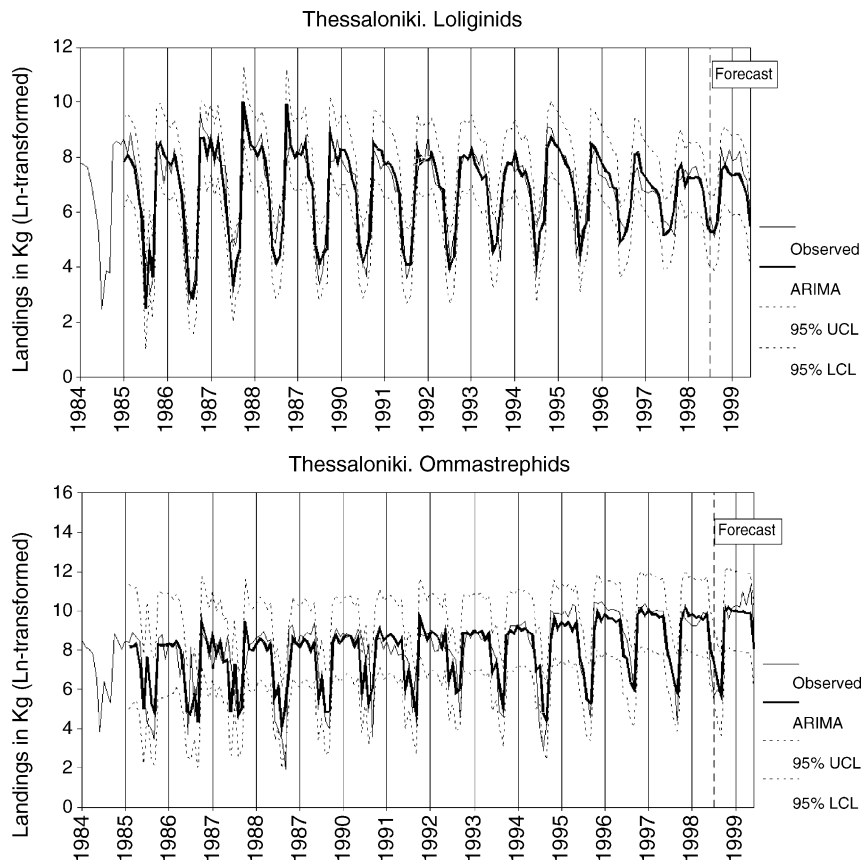


Fig. 5. ARIMA fitted models with 95% upper (UCL) and lower (LCL) confidence levels. Monthly landings, natural logarithmic transformed, from the Thessaloniki area (above: Loliginids, ARIMA(1–0–1)(0–1–1)₁₂ and below: Ommastrephids, ARIMA(1–1–1)(0–1–1)₁₂). The dotted vertical line indicates the start of the forecasting period (July 1998).

In the next step the ANN models were extended by adding different combinations of the spatio-temporal SST descriptors, such as the average, the minimum, maximum and the spatio-temporal deviation of SST acquired from the same geographic area as the landings data.

All ANNs were based on the same architecture and were trained under the same conditions. The goal of the training was not to achieve the best fit of the trained data set, but to quantify the contribution of the regressors to the prediction performance of the neural model. The performance was measured by comparing the neural output to the observed values and estimating the adjusted coefficient of determination (R^2). The estimated coefficients, for the loliginid and ommastrephid species in the main ports of the Northern Aegean Sea are summarised in Table 2. The R^2 estimates based only on the year and month variables, varied from 0.24 to 0.67 and were used as the reference (R_0^2) for testing if the incorporation of the SST descriptors could increase the model predictability. Each adjusted coefficients R_j^2 estimated by a new combination of added environmental descriptors were compared to the reference R_0^2 , where the H_0 hypothesis was that both coefficients were equal ($R_j = R_0$). The H_0 was rejected by different levels of significance, namely by $*p < 0.05$, $**p < 0.01$ or $***p < 0.001$.

The same stepwise modelling procedure was repeated in order to compare the effect of the SST descriptors to those of the previous 12 months. The environmental descriptors were used without (Lag = 0) or with a backward time shift (Lag = 12). In the most analysed cases the addition of the concurrent or shifted SST regressors increased significantly

the R^2 values and therefore the H_0 -hypothesis of the R^2 equality was rejected. The ommastrephid landings of Kavala, however, did not provide significant improvements of the coefficient of determination R^2 . Also shifting backwards the SST data by 12 months did not significantly improve the model. The only case with a significant improvement in R^2 after backward shifting was the $T_{\min}-T_{\max}$ model of the loliginid landings in Kavala ($p = 0.011$).

3.2.2. The monthly pattern for different SST conditions

The seasonal pattern of the landings, identified by simply plotting the data or by analysing the Seasonal Adjusted factors of the ARIMA models (Fig. 3), seems to vary by changing the SST descriptors. Moreover, the ANN models have predicted similar patterns and incorporating certain SST regressors were able to demonstrate the effect of temperature on the monthly patterns. In general all models predicted fewer landings by an increase of the minimum observed SST. It is noticeable, that loliginid landings (Fig. 6a and c) behaved differently from the ommastrephid landings (Fig. 6b and d). The former showed a higher sensitivity to the environmental change during the first months of the year (Thessaloniki) or later during the middle months (Kavala). On the contrary, ommastrephid landings showed much less sensitivity to the SST variations, while in the case of the Kavala data set the landings were relatively stable for the different months.

The interpretation of the landing changes in relation to the observed monthly SST, in Fig. 6, can be facilitated by referring to the monthly minima and maxima of SST summarised in Table 3.

Table 2
Prediction of monthly landings using different regressors

Time shift in SST data	Species			
	Loliginids		Ommastrephids	
	Lag = 0	Lag = 12	Lag = 0	Lag = 12
Thessaloniki				
Year, month	0.38	0.59	0.28	0.32
Year, month, $T_{\min}-T_{\text{diff}}$	0.83***	0.78***	0.73***	0.73***
Year, month, $T_{\min}-T_{\text{max}}$	0.69***	0.74*	0.75***	0.75***
Year, month, mean T	0.73***	0.73*	0.76***	0.78***
Kavala				
Year, month	0.67	0.58	0.24	0.26
Year, month, $T_{\min}-T_{\text{diff}}$	0.85***	0.85***	0.31	0.21
Year, month, $T_{\min}-T_{\text{max}}$	0.74	0.84***	0.28	0.38
Year, month, mean T	0.89***	0.83***	0.31	0.28
Alexandroupolis				
Year, month	0.53	0.64		
Year, month, $T_{\min}-T_{\text{diff}}$	0.76***	0.76*		
Year, month, $T_{\min}-T_{\text{max}}$	0.81***	0.77*		
Year, month, mean T	0.76***	0.78**		
	$N_{(\text{lag}=0)} = 188$	$N_{(\text{lag}=12)} = 175$	$N_{(\text{lag}=0)} = 188$	$N_{(\text{lag}=12)} = 175$

The models are based on trained ANNs. Adjusted coefficients of determination (R^2) were used as indicators for the “Goodness of fit” of the predicted vs. observed values. The environmental descriptors used in the models were from the same month (Lag = 0) or from 1 year before (Lag = 12).

* $p < 0.05$.

** $p < 0.01$.

*** $p < 0.001$.

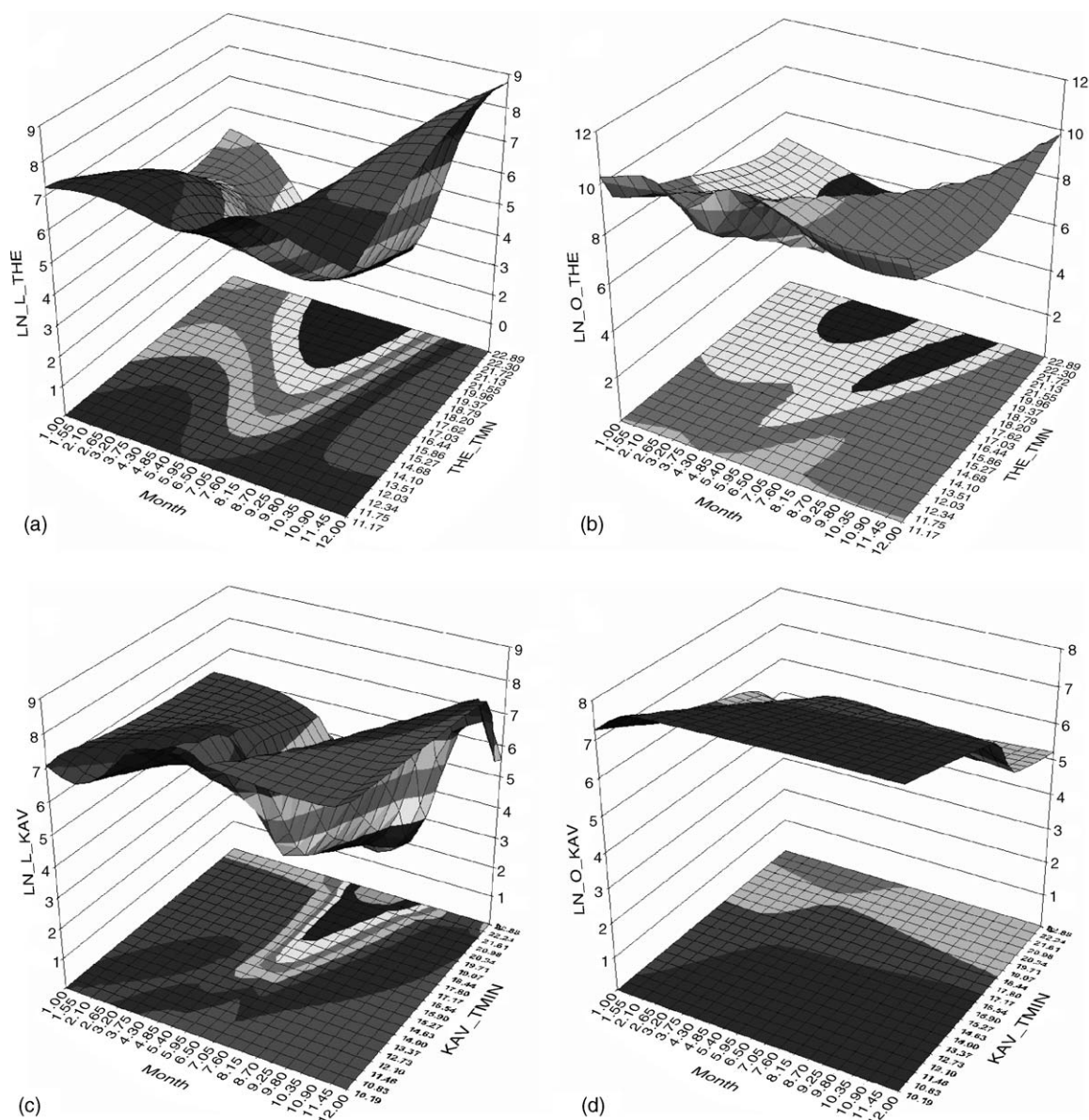


Fig. 6. Pattern of monthly landings by different T_{min} (minimum observed SST). In general, the increase of the T_{min} (y-axis) coincides with lower landings (z-axis). Significant differences were observed among the patterns of the different areas and species (see details in text).

Table 3
Monthly minimum and maximum SST observed in the three main fishing ports

	1	2	3	4	5	6	7	8	9	10	11	12
Thessaloniki												
Minimum	12.83	12.36	11.80	12.35	15.55	18.82	20.70	21.41	20.40	18.05	14.56	13.32
Maximum	15.95	15.42	15.11	16.00	17.50	21.22	22.83	23.73	22.75	21.16	18.91	17.09
Kavala												
Minimum	12.62	12.02	11.11	11.98	15.53	18.86	20.68	21.41	19.12	17.26	14.52	13.33
Maximum	16.18	15.43	15.17	16.02	17.51	20.84	22.81	23.20	22.44	21.26	19.12	17.16
Alexandroupolis												
Minimum	13.19	12.39	11.80	12.18	15.21	18.90	20.61	21.33	18.17	17.01	14.96	13.68
Maximum	16.27	15.44	15.19	16.03	17.52	20.67	22.81	23.22	22.44	21.32	19.21	17.22

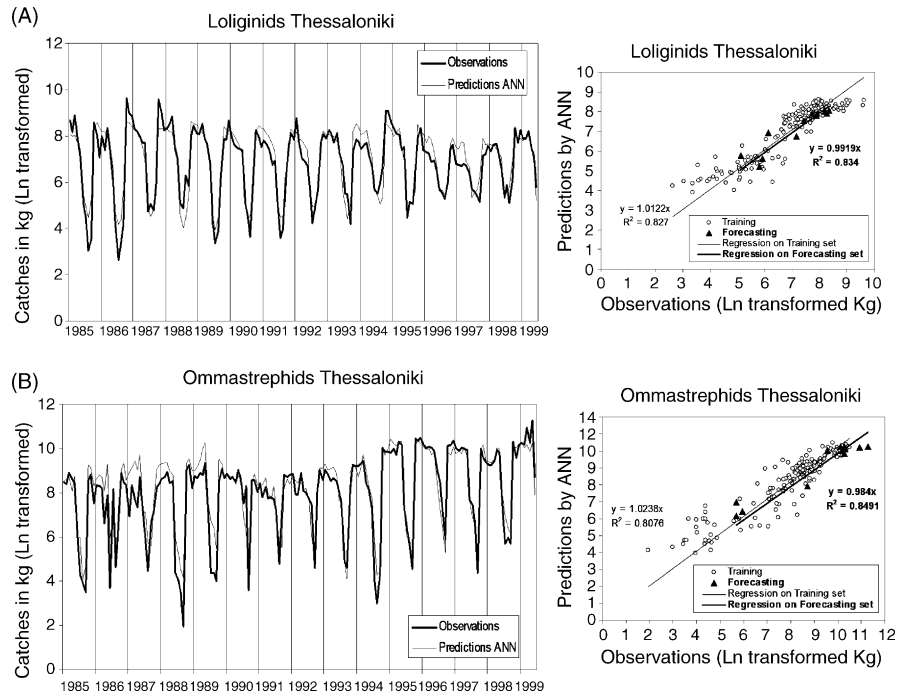


Fig. 7. Models based on trained Backpropagation ANNs. Observed vs. predicted monthly landings in Thessaloniki. Estimated R^2 for forecasting cases (July 1998–June 1999). Above: Loliginids ($R^2 = 0.834$). Below: Ommastrephids ($R^2 = 0.849$).

3.2.3. Prediction of the monthly landings

Based on the experience of the previous results new back-propagation ANNs were developed in order to predict the landings utilising as regressors the monthly SST values of the previous 12 months. The modelled and predicted landings of the loliginid and ommastrephid species are displayed in Fig. 7. Summary results of regression statistics between the predicted and the observed landings for the two species are contained in Table 4. The forecasting performance of the time series of both species was similar ($R^2 = 0.84$), not however the Relative Contribution Factors, namely the contribution to the model of each monthly mean temperature from the pre-

Table 4
Prediction of monthly landings using as regressors the SSTs of the previous 12 months

Summary results	Loliginids	Ommastrephids
Correlation coefficient r	0.9226	0.9287
R squared:	0.8443	0.8432
r squared	0.8511	0.8625
Mean squared error	0.338	0.567
Mean absolute error	0.457	0.535
Min. absolute error	0	0.002
Max. absolute error	1.84	2.353

Summary results of the final ANNs taking into account the total data set i.e. both training and testing (forecasting) cases.

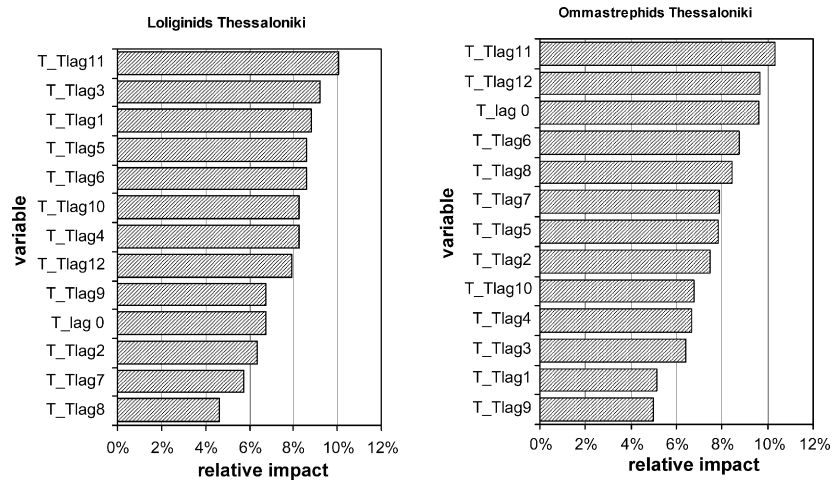


Fig. 8. Relative impact factors of the backward time-shifts in the ANN models. Monthly landings in Thessaloniki. Left: loliginids, Right: ommastrephids. Each bar represents the relative importance of the monthly SST value lag \times months before ($0 \leq x < 12$).

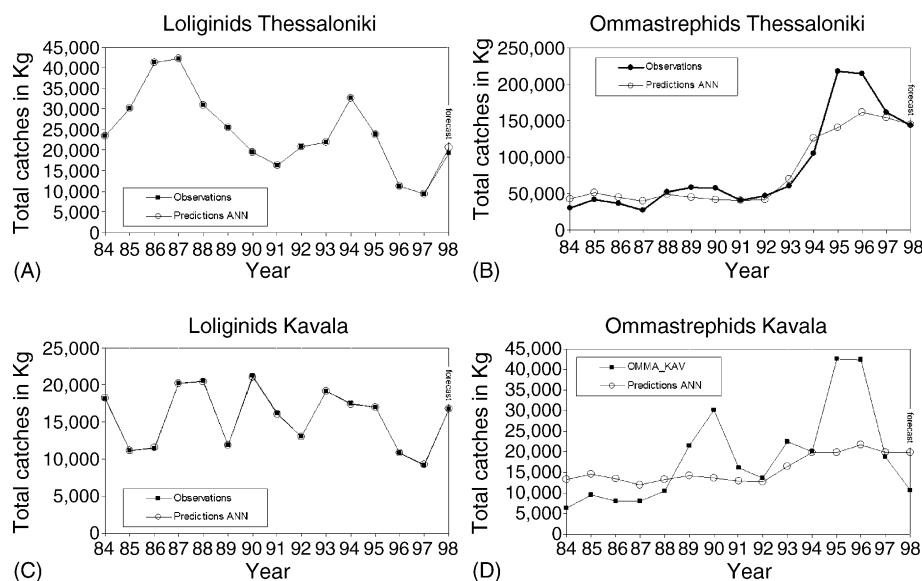


Fig. 9. Results based on backpropagation ANNs. The curves display observed vs. predicted yearly landings (estimated R^2 for forecasting cases). (A) Loliginids in Thessaloniki ($R^2 = 0.99$), (B) Ommastrephids in Thessaloniki ($R^2 = 0.87$), (C) Loliginids in Kavala ($R^2 = 0.99$) and (D) Ommastrephids in Kavala ($R^2 = 0.29$).

vious 12 months. It is remarkable that 1 or 3 months before the (mean) SST was relatively important for the loliginid but not for the ommastrephid landings (Fig. 8).

3.2.4. Prediction of the annual total landings based on monthly SST descriptors

The annual total landings of the target species were modelled by developing new backpropagation ANNs and utilising twelve regressors, namely the monthly SST means for each given year (Fig. 9). Although loliginid models performed very well, providing high R^2 estimates between predicted and observed values (Table 5), ommastrephid models were less successful in the Thessaloniki area ($R^2 = 0.87$) and did not work in the remaining geographic areas.

The importance of each month for the model prediction varied between species and areas. The Relative Contribution Factors (RCFs) showed in general higher standard errors in the ommastrephid data, confusing their importance. On the contrary, loliginid data showed clearly higher precision and clear differences among the RCFs. For instance, the relative importance of the months January, June–July and

September–October was always above the average 8% limit for the loliginids of Thessaloniki.

3.3. Bayesian approach in forecasting

Loliginid and ommastrephid monthly landings of Thessaloniki and Kavala were used for developing Bayesian models incorporating among other components monthly SST descriptors as environmental factors.

A steady state model was developed, based on a simple dynamic level according to the equations described in methodology (forms 5 and 6). The starting conditions of the models were based on the default values estimated by the reference analysis. In all models, the incorporation of temperature increased significantly their predictability.

The stepwise development is demonstrated, using as example the case of the loliginid landings of Thessaloniki, as follows. The observed versus the predicted values from the steady state model provided, as expected, low coefficient of determination (R^2). By adding in the steady state model an explanatory environmental variable (mean SST) the pre-

Table 5
Prediction of yearly landings using as regressors the SSTs of each month per year

Summary results	Thessaloniki		Kavala		Alexandroupolis
	Loliginids	Ommastrephids	Loliginids	Ommastrephids	Loliginids
Correlation coefficient r	0.9990	0.9415	0.9999	0.5848	0.9612
R squared	0.9979	0.8738	0.9997	0.2923	0.9131
r squared	0.9981	0.8865	0.9998	0.3420	0.924
Mean squared error	0	0.058	0	696272	0.103
Mean absolute error	0.005	0.203	0.003	786	0.086
Min. absolute error	0	0.009	0	928	0
Max. absolute error	0.072	0.434	0.01	18393	1.244

Summary results of the final ANNs taking into account the total data set i.e. both training and testing (forecasting) cases.

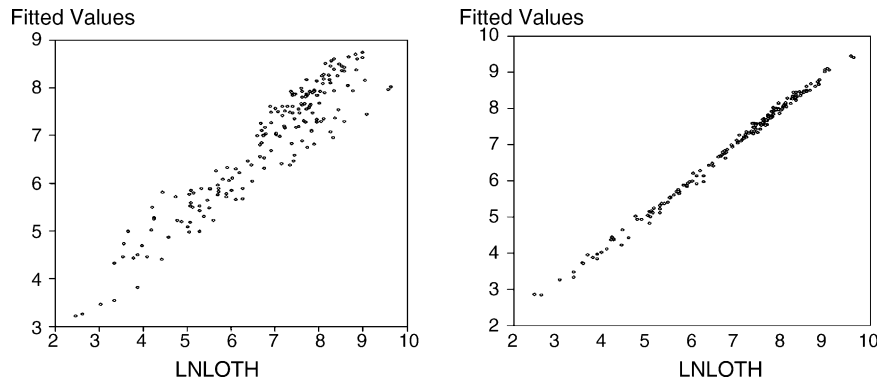


Fig. 10. Predicted vs. observed monthly loliginid landings in Thessaloniki based on Bayesian Models. Left: univariate model, the average SST of the area under study has been used as Regressor. Right: multivariate model, both average SST of the area under study and monthly seasonality have been used as Regressors.

dictability increased achieving a coefficient of determination of $R^2 = 0.889$ (Fig. 10). Further analysis of the residuals, estimated applying the above dynamic regression model, still showed the existence of seasonal patterns. The residual seasonality (which was not included in the SST covariate) was absorbed after adding 12 seasonal factors, which had a total sum of 1 over a full year. The “full” model achieved a higher R^2 (0.996, Fig. 10). The residuals were randomly distributed about the zero mean.

The final model was cross-validated with forecasting test data, which had not been used in the analysis. Forecasted values from our example of the loliginid landings in Thessaloniki is showed in Fig. 11, together with their 90% margin limits. The estimated R^2 between forecasted and observed monthly data was equal to 0.96.

The estimated coefficients of determination of the step-wise dynamic models are summarised in Table 6. Not all steady state models provided predictions with R^2 more than 0.22. The incorporation of the temperature regressor increased rapidly their performance to R^2 varying from 0.80 to 0.89. The final models, which also incorporated the seasonal periodicity, provided the highest R^2 (0.99). The results of the cross validation procedure showed very good performance for the loliginid model ($R^2 = 0.96$) and less precise predictions for the others ($R^2 = 0.82–0.83$). All cases were inside the 90% margin limits.

Table 6
Prediction of monthly landings using Bayesian models developed in three steps

		Thessaloniki loliginids	Thessaloniki ommastrephids	Kavala loliginids
1	Simple model	0.22	0.19	0.22
2	Adding mean SST	0.89	0.80	0.84
3	Adding mean SST and seasonality	0.99	0.99	0.99
4	Cross-validation data	0.96	0.83	0.82

Summary results of the estimated coefficients of determination R^2 between predicted and observed values, in each trial.

The incorporation of all four SST regressors improved the R^2 of the output of all final models, without changing significantly their margin limits.

4. Discussion

4.1. The physical environmental

The northern part of the Aegean Sea represents the most productive area for cephalopods in Greece and for

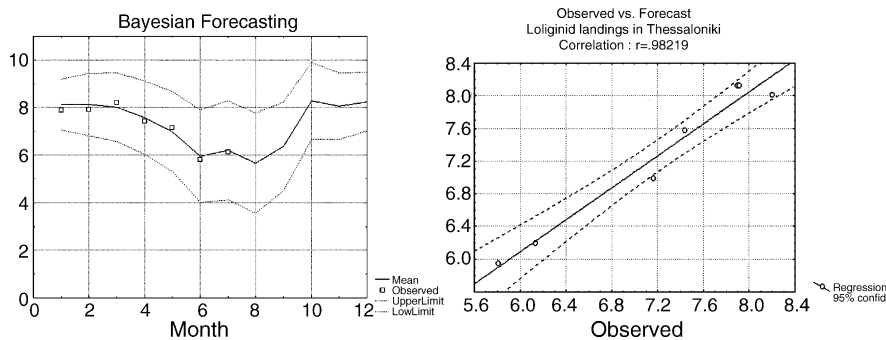


Fig. 11. Left: cross-validation of the Bayesian model based on data not used during the analysis. Right: Forecast vs. observed cross-validation data. Estimated coefficient of determination $R^2 = 0.96$ (Dotted lines represent the 95% confidence intervals).

this reason has been the target of the present investigation. Obviously modelling the fishery landings using environmental parameters is an attractive approach towards a better management of the biological resources. Environmental predictors have been repeatedly used to uncover the role of the environment in the distribution and abundance of several cephalopod species (Cushing, 1981; Stergiou, 1987, 1989). Pierce et al. (1995) comparing fishery landings of the Northeast Atlantic and Mediterranean loliginid landings, over the period 1980–1992, suggested the possibility of a large-scale climatic influence on the cephalopod abundance.

The marine environment in the studied area is affected by two water influxes, a significant water flow from a large system of rivers (Balopoulos, 1982) and a cold, less saline, surface water originating from the Black Sea (Georgopoulos, 1984). The inflow is noticeable by simple scrutinising the NOAA-SST or SeaWiFS satellite images. During the first months of the year, the first water influx dominates especially in the Thermaikos Gulf in Thessaloniki, while the second influx affects mainly the areas near Kavala and Alexandroupolis. The extended use of several SST parameters in the present analysis relies on the hypothesis that SST can provide adequate information concerning the water masses and their environmental status. Since, the target area is mainly enriched in nutrients by the Black Sea surface inflow, a 5–25 m thin surface layer coming out through the Straits of Dardanelles, the satellite information has been considered as adequate.

4.2. The modelling approach

The hypothesis that sea temperature may play an important role in the evolution of the landings in the target area and that SST parameters are adequate to describe the environmental impact on this mechanism was tested in the present work by applying stepwise model development.

From the results demonstrated mainly in Tables 2 and 6 it is evident that the incorporation of temperature descriptors can significantly improve the model performance. The estimated very low p -levels ($p < 0.001$) support the validity of our hypothesis. The results in Table 2 gives the impression that certain descriptors are more effective than others; however such differences are not statistically significant and furthermore their order varied among the different species and places. Replacing the SST descriptors with the shifted ones did not significantly improve the fits, with the exception of the loliginid landings in Kavala, which achieved a significant ($p = 0.011$) improvement in R^2 after the backward time-shift. It should be recognised however, that due to the strong autocorrelation of the SST data it may be hard to distinguish differences between the raw and the 12-month shifted temperature descriptors.

The same problem was investigated applying ANN models utilising as inputs “scrolled” the 12 previous months in each case. This structure tests not for the importance of certain months but for the effect of the temperature 1, 2, 3, etc. months before a given month. It is the equivalent to a

sample cross-correlogram implemented by non-linear functions. The performance of the new models was comparable to those developed in the previous steps (Tables 2 and 3). As is suggested from Fig. 8, the estimated relative impact factor of the backward time-shifted mean SST varied significantly. Apart from the 11 and 12 months before, which may be introduced due to the strong autocorrelation of the data, other components (3, 1 and 5 months before) showed higher relative impact factors. The cross-correlation-like graph may be compared to that in Fig. 4 (below), which also suggested correlations in periods shorter than 12 months. These results may be difficult to be interpreted biologically, but they can improve the predictability of models.

The existence of a strong seasonal pattern in the monthly landings is easily seen in the original data (Fig. 2) or the seasonal adjusted factors (SAFs) displayed in Fig. 3. Further analysis of the seasonal patterns under various SST conditions confirmed the initial hypothesis that the SST can explain a lot of the landing seasonality, in the same way that SST explained the variability in the time series of the landings. The model outputs portrayed in Fig. 6 suggest that the monthly landings are not equally sensitive to the SST changes all year around. In certain months, a minor SST change coincides with a significant gradient in the model output. Fig. 6 may be better interpreted taking into account the data summarised in Table 3. For instance, the mean SST observed during June and July in Thessaloniki varied between 18 and 23 °C, a range, which in plot of Fig. 6A, can significantly affect the amount of the loliginid landings. Similar “hot conditions” during September and October (mean SST varied between 20 and 23 °C), can explain order of magnitudes more than four in logarithmic scale. The sensitivity shape portrayed in the 3D-graphs varies among the different species and areas. The graph representing the ommastrephid landings in Kavala may be interpreted with caution due to the very low R^2 provided from this model (0.38). In Kavala the loliginid landings showed their sensitive period during July and August, where the expected mean SST varied between 20 and 24 °C.

The modelling of the annual landings is practically more attractive and seems to be easily interpretable. However, it is obvious that the prediction of annual landings is only a cumulative monthly prediction and in any case should be the final approach. The results of the ANNs developed for investigating this approach are displayed in Fig. 9 and are summarised in Table 5. The estimated R^2 of the models concerning the loliginid landings varied between 0.91 and 0.99, and was significantly better than the corresponding estimate of ommastrephid, which achieved $R^2 > 0.87$ in Thessaloniki and failed in Kavala ($R^2 < 0.30$).

The impact of the monthly SST on the annual prediction varied between species and areas. For instance, the relative importance of the months January, June–July and September–October is always above the average 8% limit for the loliginids of Thessaloniki.

The importance of the observed SST in certain months may be compared to the seasonal patterns portrayed in Fig. 6.

Despite the fact that they represent different models, namely annual and monthly landings, they showed similarities. The importance of the SST for the loliginid landings during June–July, September–October and July–August in Thessaloniki and Kavala, respectively seems to be in accordance with the modelled sensitivity proposed in the graphs of Fig. 6. Such results may suggest further research on certain time periods of the life history of the target species and elucidate the relations among the biological and environmental parameters.

4.3. Comparison of the final models

It is well known that most statistical models are restricted on a static character. They assume that the quantified relationships and system parameters remain the same across the observation and forecasting time period. Therefore, the ability to apply and test more flexible and sophisticated algorithms in order to achieve better forecasting performance seems to be very promising.

ARIMA models show a relative good performance in case of time series with a strong autoregressive character. However, their generalisation capability, namely the ability to predict well if a strong deviation from the previous pattern occurs, is limited. The univariate ARIMA models provided a relatively high R^2 , which could not be improved significantly by adding any other regressor in the multivariate version. This limitation probably leads to the inability of the ARIMA to show any environmental impact on landings. As a prediction tool, it provides good results, if the time series is stationary with a strong autoregressive character.

The neural models, applied on both monthly and annual time series, investigated the importance of the SST descriptors for the prediction approach as well as for uncovering the sensitivity of the monthly patterns. The estimated R^2 of all monthly ANN models were lower compared to the ARIMA. However, it should be emphasised that the ANNs included exclusively SST variables and no direct seasonal patterns. It is remarkable, that the SST descriptors alone can predict the landings at such high levels of R^2 (0.89). The prediction of the annual landings was possible only by applying neural models, other methods being limited by the limited data set. Again, the annual models did not contain seasonal patterns and the prediction was based exclusively on the temperatures encountered in each month. The high scores encountered in the prediction allow the application of this method for practical purposes.

Species differences were better judged in closed areas like the Thermaikos Gulf near the fishing port of Thessaloniki. The strange behaviour of the ommastrephid models in Kavala could imply that SST could not express adequately the environmental conditions in this area. This should be attributed to the fact that ommastrephids live in deeper pelagic zones than the loliginids (Roper et al., 1984).

The last category of models was implemented according to the Bayesian approach. The observed monthly mean and the

monthly seasonality were the main model components. In the Bayesian models, the parameters, which would be assumed constant in the classical frequentist approach, are allowed to vary slowly over time if there is evidence in the data for such changes. The estimated coefficients of determination were the highest of all other models (0.99), even by utilising only the mean SST descriptor.

The models provided also high performance on the cross-validation test. The advantage of the Bayesian approach over the structural time series approach is that the applied methods are exact even in the non-linear case and can provide detailed impact factors of each variable in a given moment. The Bayesian model may provide an appropriate tool, combined with the study of the life history of the two target species, for the investigation of possible relationships among the estimated fluctuations of the monthly components and the concurrent environmental changes in the target areas.

Acknowledgements

The authors wish to thank the Greek General Secretary of Research and Technology (GSRT) for the financial support of the work, under the IMAS project: Development of an Integrated Management System to support the sustainability of Greek fisheries resources (Contract No: Natural Environment-25); the Commission of the European Communities for the financial support of the work, under the research projects: CEPHVAR, FAIR-PL-1520/1999–2001 and CEPH-STOCK, QoL-2001/2002–2004; and the German Aerospace Agency for the distribution of AVHRR data through the freely available GISIS (<http://isis.dlr.de/>).

References

- Arvanitidis, C., Koutsoubas, D., Robin, J.P., Pereira, J., Moreno, A., Da Cunha, M.M., Valavanis, V., Eleftheriou, A., 2002. A comparison of the Fishery Biology of three *Illex coindetii* VÉRANY, 1839 (Cephalopoda: Ommastrephidae) Populations from the European Atlantic and Mediterranean Waters. *Bull. Mar. Sci.* 71, 129–146.
- Balopoulos, E.T., 1982. Circulation and Mixing in the Water Masses of the N.W. Aegean Sea. Ph.D. Thesis, University of Wales, UK.
- Box, G.E.P., Jenkins, G.M., 1976. *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco.
- Box, G.E.P., Tiao, G.C., 1973. *Bayesian Inference in Statistical Analysis*. Addison-Wesley, Massachusetts.
- Caddy, J.F., 1989. *Marine Invertebrate Fisheries: Their Assessment and Management*. John Wiley and Sons, New York.
- Cohen, Y., Stone, J.N., 1987. Multivariate time series analysis of the Canadian fisheries system in Lake Superior. *Can. J. Fish. Aquat. Sci.* 44 (Suppl. 2), 171–181.
- Cushing, D.H., 1981. Temporal variability in production systems. In: Longhurst, A.R. (Ed.), *Analysis of Marine Ecosystems*. Academic Press, New York, 741 pp.
- D'onghia, G., Matarese, A., Tursi, A., Maiorano, P., 1996. Cephalopods collected by bottom trawling in the North Aegean Sea (Eastern Mediterranean). *Oebalia* 22, 33–46.
- Fox, W.W., 1970. An exponential surplus-yield model for optimising exploited fish populations. *Trans. Am. Fish. Soc.* 99, 80–88.

- Georgakarakos, S., Haralabous, J., Valavanis, V., Arvanitidis, C., Koutsoubas, D., 2002. Prediction of fishery exploitation stocks of loliginid and ommastrephid squids in Greek waters (Eastern Mediterranean) using uni- and multivariate time series analysis techniques. *Bull. Mar. Sci.* 71 (1), 269–288.
- Georgopoulos, D., 1984. Sea surface temperature distribution in the Aegean Sea using IR Satellite Imagery. In: Proceedings of the First Panhellenic Symposium in the Oceanography and Fisheries, Athens, pp. 105–112 (in Greek).
- Haddon, M., 2001. *Modelling and Quantitative Methods in Fisheries*. Chapman and Hall/CRC.
- Hilborn, R., Walters, C.J., 1992. *Quantitative Fisheries Stock Assessment*. Chapman and Hall, New York, London.
- Koutsoubas, D., Arvanitidis, C., Valavanis, V.D., Georgakarakos, S., Kapantagakis, A., Magoulas, A., Kotoulas, G., 1999. Cephalopod Resources in the Eastern Mediterranean with particular emphasis in Greek Seas: Present and Future Perspectives. ICES CM 1999/G:4.
- Makridakis, S.G., 1990. *Forecasting Planning and Strategy for the 21st Century*. Free Press, London.
- McDowall, D., McCleary, R., Meidinger, E.E., Hay, R.A., 1980. *Interrupted Time Series Analysis*. Sage Publications, Beverly Hills, CA.
- Moreno, A., Pereira, J., Arvanitidis, C., Robin, J.P., Koutsoubas, D., Pelales-Raya, C., Da Cunha, M.M., Balguerias, E., Denis, V., 2002. Biological Variation of *Loligo vulgaris* (Cephalopoda: Loliginidae) in the Eastern Atlantic and Mediterranean. *Bull. Mar. Sci.* 71, 515–534.
- Pierce, G.J., Collins, M.A., Cunha, M.M., Guerra, A., Piatkowski, U., Porteiro, F., Robin, J.P., 1995. Correlation analysis of interannual variation in Cephalopod landings from European waters. ICES CM 1995/K:22.
- Pole, A., West, M., 1990. Efficient Bayesian Learning in non-linear dynamic models. *J. Forecast.* 9, 119–136.
- Ripley, B.D., 1996. *Pattern Recognition and Neural Networks*. Cambridge University Press.
- Roper, C.F.E., Sweeney, M.J., Nauen, C.E., 1984. *FAO Species Catalogue. Cephalopods of the World. An annotated and illustrated catalogue of species of interests to fisheries*. FAO Fish. Synop. 3 (125).
- Schaefer, M.B., 1984. Some aspects of the dynamics of populations important to the management of commercial marine fisheries. *Bull. Inter-Am. Trop. Tuna Comm.* 1, 27–56.
- SPSS Inc., 2002. *SPSS Base 11.5. User's guide*, 613 pp.
- Stergiou, K.I., 1987. Cephalopod abundance in Greek waters in relation to environmental fluctuations. *Biol. Gallo-Hell.* 13, 25–34.
- Stergiou, K.I., 1989. Assessment of the state and management of the cephalopod trawl fisheries in Greek waters. *Toxicol. Environ. Chem.* 20/21, 233–239.
- Stergiou, K.I., Christou, E.D., 1996. Modelling and forecasting annual fisheries catches: comparison of regression, univariate and multivariate time series methods. *Fish. Res.* 25, 105–138.
- Twomey, J.M., Smith, A.E., 1996. Artificial neural networks for civil engineers. In: Kartam, N., Flood, I., Garrett, J. (Eds.), *Fundamentals and Applications*. ASCE Press.
- Valavanis, V.D., Georgakarakos, S., Koutsoubas, D., Arvanitidis, C., Haralabous, J., 2002. Development of a Marine Information System for Cephalopod Fisheries in the Greek Seas (Eastern Mediterranean). *Bull. Mar. Sci.* 71 (2), 867–882.
- West, M., Harrison, P.J., 1989. *Bayesian Forecasting and Dynamic Models*. Springer-Verlag, New York.
- Zar, J.H., 1999. *Biostatistical Analysis*. Prentice Hall International, Inc., New Jersey, USA.