Asian Fisheries Science **13**(2000): 75-85
Asian Fisheries Society, Manila, Philippines
https://doi.org/10.33997/j.afs.2000.13.1.009

Modelling and Forecasting the Catch of the Scads (*Decapterus macrosoma, Decapterus russellii*) in the Javanese Purse Seine Fishery Using ARIMA Time Series Models

M. POTIER and L. DRAPEAU

ORSTOM- IRD
911, Avenue Agropolis
BP 5045, 34 032 Montpellier cedex 1
France

Abstract

The standard of living of the purse seine fishers operating in the northern coast of the Java Island is closely related to scad resources. Scads landings of the semi-industrial purse-seine fishery were analyzed using Seasonal Auto Regressive Integrated Moving Average (SARIMA) techniques. Several models were found to be suitable for describing the temporal fishery pattern and for forecasting the landings two years ahead. Forecasts were matched to actual data.

It appears that such models can describe and forecast the dynamics of the scads fishery in the Java Sea. Until now, such results are difficult to predict owing to the strong influence of the year to year changes in the monsoon regime.

Forecasting is more reliable with neritic species such as *Decapterus russellii,* since their spatial distribution in the Java Sea is relatively independent of the monsoon variability. On the contrary, oceanic species such as *Decapterus macrosoma* are more difficult to forecast as their distribution is closely related to the environmental changes. In that case, transfer function models including environmental parameters would be more appropriate.

Introduction

In the Java Sea, Potier et Boely (1990) , and Potier (1998) have shown that the annual fluctuation of the catch is closely related to environmental factors. The intensity of the monsoon which is the main factor can be measured by rainfalls or wind speed. Unfortunately, in Indonesia, reliable long series of environmental data are scarce. Analytical and regression techniques or transfer function models which take into account such variables were hardly used. Modelling based on the time series method seems to be the most effective way to forecast the catch.

Introduced for economic studies (Chatfield, 1984), the univariate Auto Regressive Integrated Moving Average (ARIMA) models have been widely used in fisheries studies (Boudreault et al. 1977; Saila et al. 1979; Mendelssohn, 1981; Fogarty et al. 1986; Mendelssohn and Cury, 1987; Jeffries et al. 1989; Stergiou et al. 1997). 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. They apply to stationary time series. The time series (X_t) is second order stationary only if M_t=E(X_t) = m»t and its autocovariance function only depends on the distance between the two observations and not of the time (t) (Box and Jenkins, 1976).

Since the 19th century, pelagics have been the main resources of the Java Sea. Catches are landed at fishing places located along the north coast of the Java island. In the first half of the 20th century, the exploitation remained artisanal. In the seventies, drastic changes occured and the exploitation became a semi-industrial one. Today, the semi-industrial flottillas account for more than 60% of the landings in the Java Sea. Two species of scads, the Indian scad (*Decapterus russellii*) and the shortfin scad (*Decapterus macrosoma*) are the targets of javanese purse seiner fleets. They form the bulk of the catch and show a well marked seasonal pattern, a decreasing trend from 1984 to 1988 and an increasing one since 1988 (Fig. 1). Today, the scads landings reach 50,000 to 75,000 tons a year and they account for more than 50% of the pelagics caught by the fishery. They are of great importance to the economics of the northern coast of Java. The forecasts, one or two years in advance of the level of the catch maybe useful to management. SARIMA method which is a particular form of ARIMA method has been used to achieve this goal. In the SARIMA approach the variation in the time series X(t) is modelled by a combination of ARIMA with seasonal operators. The seasonal component is allowed to behave as an ARIMA process.

Materials and Methods

Materials

Landings of the semi-industrial fishery are available since 1976 when the national statistical plan was launched (Yamamoto, 1980). First, the data were recorded by commercial category and by month. The category « *layang* » gathered the two species of scads and it was impossible to know the percentage of each one in the landings. In 1984, a joint French-Indonesian research program, focused on the semi-industrial fishery was started. At that time, the purse seiners were targeting the whole fishing area. Nine fishing grounds were defined corresponding to the islands scattered in the Java Sea and the China Sea and which were serving as leading marks for the fishing vessels (Fig. 2). In 1991, a second research program funded by the European Union (PELFISH) followed. Since 1984, the catch is known through the vessels, fishing ground, and species. The data analyzed in this work are the monthly catches from January 1984 to December 1995 gathered from the PELFISH project. The data covering the 1984-1993 period were used for modelling and the 1994-1995 data for forecasting (Table 1).

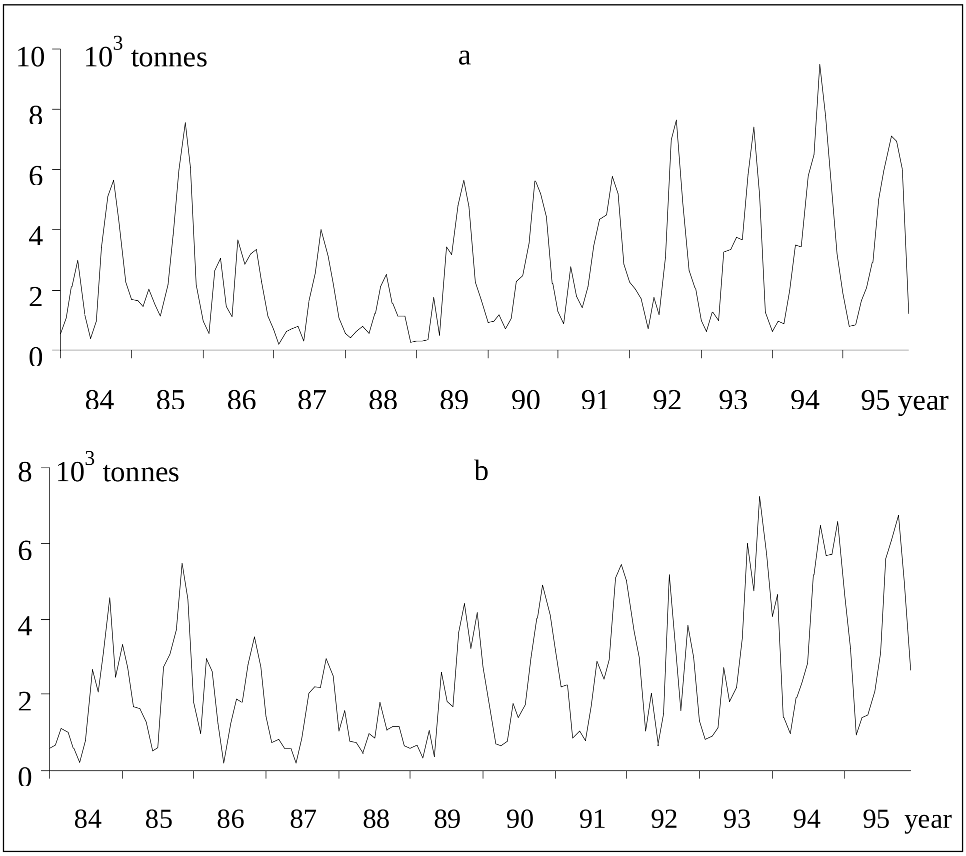

Fig. 1. Monthly catches of the Indian scad (*Decapterus russellii* - a) and the shortfin scad (*Decapterus macrosoma* - b) landed by the semi-industrial purse seiners from January 1984 to December 1995.

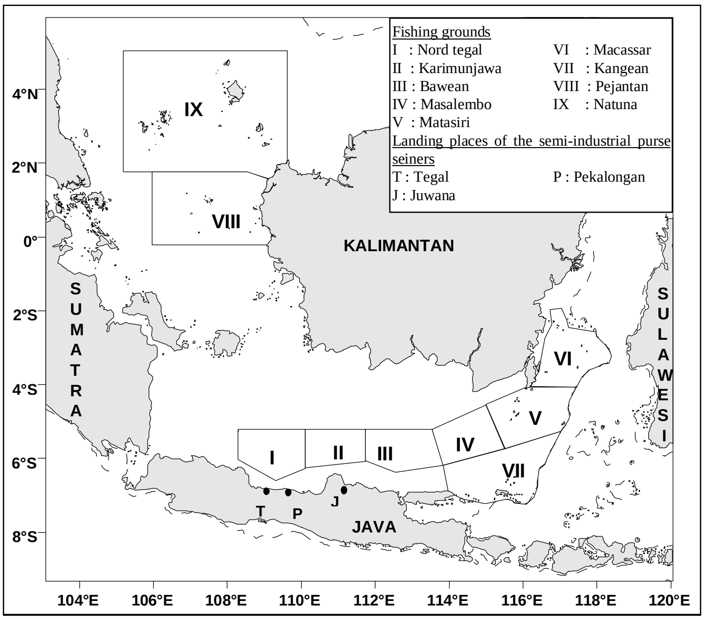

Fig. 2. Map showing the location of the fishing grounds of the semi-industrial Javanese purse seiners. The main landing places of the fleet along the north coast of Java are reported.

Table 1. Landings of *Decapterus macrosoma* and *Decapterus russellii* (in tonnes) by the semi-industrial Javanese purse seine fishery, 1984-1995 period.

	J	F	M	A	M	J	J	A	S	O	N	D
Decapterus macrosoma												
1984	564	666	1,100	994	585	222	784	2,663	2,062	3,113	4,562	2,474
1,985	3,307	2,720	1,665	1,632	1,278	517	611	2,748	3,066	3,712	5,471	4,535
1,986	1,800	62	2,959	2,623	1,245	197	1,235	1,899	1,788	2,812	3,536	2,731
1,987	1,433	740	824	564	563	191	848	2,044	2,233	2,181	2,943	2,485
1,988	1,029	1,575	764	719	466	976	863	1,805	1,066	1,166	1,158	652
1,989	576	684	346	1,056	376	2,573	1,813	1,682	3,653	4,412	3,214	4,181
1,990	2,736	1,921	686	639	772	1,765	1,392	1,745	2,974	4,002	4,900	4,107
1,991	3,186	2,207	2,264	845	1,049	792	1701	2,889	2,392	2,907	5081	5,443
1,992	5,019	3,685	2,988	1,027	2,045	637	1,502	5,158	3,191	1,578	3,824	2,996
1,993	1,311	832	905	1,133	2,700	1,829	2,188	3,496	5,991	4,739	7,238	5,713
1,994	4,091	439	1,395	987	1,915	2,321	2,840	5,185	6,483	5,690	5,714	6,565
1,995	4,631	3,229	930	1,401	1,466	2,109	3,108	5,592	6,092	6,749	4,999	2,658
Decapterus russellii												
1,984	547	1,082	2,082	2,995	1,194	388	972	3,390	5,120	5,638	4,315	2,257
1,985	1,685	1,635	1,469	2,009	1,499	1,160	2,166	3,829	6,000	7,542	6,058	2,180
1,986	952	577	2,618	3,040	1,448	1,097	3,676	2,881	3,193	3,375	2,297	1,134
1,987	693	207	594	740	805	291	1,615	2,565	3,996	3,142	2,218	1,079
1,988	578	437	606	802	571	1,233	2,090	2,532	1,555	1,151	1,141	254
1,989	303	300	360	1,739	488	3,452	3,157	4,822	5,646	4,767	2,257	1,691
1,990	921	965	1,193	736	1,030	2,280	2,468	3,548	5,629	5,181	4,415	2,214
1,991	1,296	883	2,766	1,800	1,422	2,112	3,464	4,354	4,487	5,770	5,201	2,849
1,992	2,261	2,022	1,708	743	1,766	1,196	3,106	6,985	7,651	4,868	2,622	2,060
1,993	973	593	1,245	982	3,239	3,357	3,731	3,651	5,859	7,418	5,161	1,268
1,994	610	958	866	1,986	3,495	3,426	5,794	6,474	9,516	7,803	5,509	3,195
1,995	1,890	806	822	1,635	2,069	2,956	5,004	5,954	7,105	6,961	5,987	1,209

Methods

The data are univariate time series. Various statistical modelling methods can be applied to forecast; Garch, STL, averaging, exponential smoothing and ARIMA. ARIMA models are developed from historical time series analysis, and based on well articulated statistical theory (Box and Jenkins, 1976). These models capture the historic autocorrelations of the data and extrapolate them into the future.

The general form of SARIMA models can be described by the following equation:

$$\Phi(B^s)\varphi(B)\nabla_s^D\nabla^d X_t = \Theta(B^s)\Theta(B)\varepsilon_t \qquad (1)$$

where :

X_t = value of variable at time t
$\Phi(B^S)$ = seasonal autoregressive coefficients
$\Theta(B^s)$ = seasonal moving average
∇_s^D = seasonal d-fold difference operator
$\varphi(B)$ = Nonseasonal component
$\Theta(B)$ = Nonseasonal moving average

∇^d = Nonseasonal d-fold difference operator

ε_t = white noise

The general form of SARIMA models is referred to as:

SARIMA (p,d,q)(P,D,Q)^S where the first term takes into account the nonseasonal and the second one the seasonal effects.
p = order of the autoregressive term (AR term)
d = degree of differencing involved to achieve stationarity (I term)
q = order of the moving average term (MA term)
S = seasonality (number of periods)
P,D,Q = order of terms corresponding to seasonality

The approach underlying the Box-Jenkins models (Box and Jenkins, 1976) is to empirically remove as much structure from the data as possible, with the ultimate goal of having the residuals as ‘ white noise’. Empirical representation of the response variable time series is desirable for forecasting.

Time series analysis is concerned with stationary series which means:

$\forall t \in N$ or Z $E(X_t^2) < \infty$; the process is a second order one
$\forall t$ $E(X_t)=m=$constant; the mean is independent of t
$\forall t$ $\forall n$ cov $(X_t,X_{t+h})=\gamma(h)$; the covariance is independent of t.

Fitting the model involves the Box Jenkins three step procedure of:

(1) identifying which terms are to remain in equation
(2) estimating the parameters
(3) checking the adequacy of the model by looking at the residuals.

Identification include examination of the Auto-Correlation Function (ACF) and Partial Auto-Correlation Function (PACF) of the transformed series. ACF gives information on the variability and temporal relations (AR terms). PACF measures the linear relation between X_t et X_{t-h} when we withdraw the linear relation brought by X_{t-1}, X_{t-p},......X_{t-h+1} (MA terms). The two functions give a clue for MA and AR type models.

To estimate the parameters and fit the best model, comparison of various models is performed. Significance of parameters is estimated using the Quenouille (1950) and Bartlett and Dianda (1950) statistics.

Checking is based on the examination of the ACF and PACF functions of the residuals. No significant lag must be present.

Comparison of forecasted values with actual monthly catches allows to arrive at a precise degree of confidence of the fitted model. The coefficient of determination (r2) is calculated as:

$$r^2 = 1 - \frac{\text{Var. (residuals)}}{\text{Var. (obs. values)}} \qquad (2)$$

In the study, SARIMA models were built using the approximate maximum likelihood algorithm of Mc Leod and Sales (1983) and performed using the Statistica software.

Results

The landings of the two species show a tendency for the variance to increase with the catch, so the need for transformation is real. The *Decapterus russellii* landings series has been log transformed, while better stationarity is reached for the *Decapterus macrosoma* series when power was transformed. For the two transformed series, examination of the ACF show that seasonality persisted (Fig. 3). Then, seasonality was removed by taking a differencing of order 12. Different SARIMA models were fitted to the original data. Results gathered are presented in Table 2. Parameters were estimated using backscating with length of 13.

For *Decapterus russellii* the model which presented the highest degree of confidence (r^2) was the SARIMA $(1,0,0)(1,1,0)^{12}$

$$(1-\varphi_1 B)(1-\Phi_1 B^{12})(1-B^{12})\ X_t = \varepsilon_t \qquad (3)$$

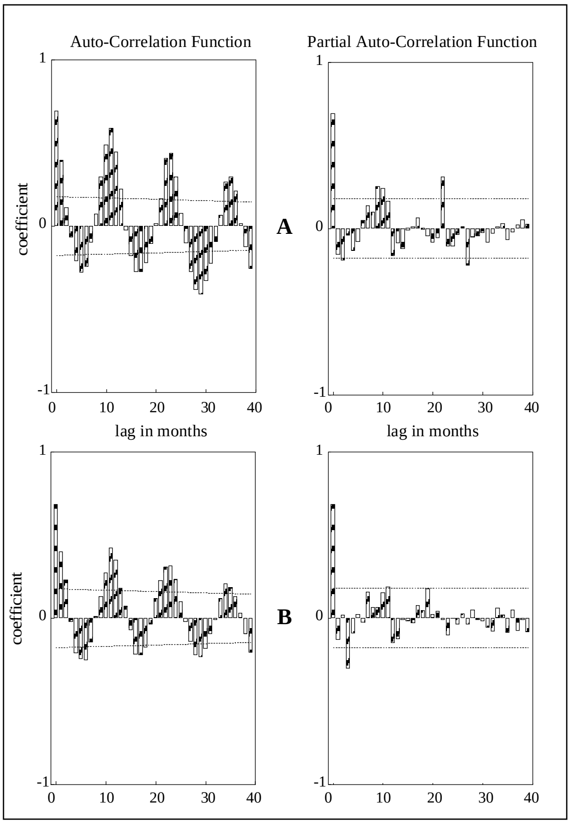

Fig. 3. Auto-correlation and partial auto-correlation functions of the *Decapterus russellii* (A) and *Decapterus macrosoma* (B) landings: tranformed series.

Table 2. Parameter estimates for different SARIMA models.

	$(1,0,1)(1,1,0)_{12}$		$(0,1,1)(0,1,1)_{12}$		$(1,0,0)(1,1,0)_{12}$	
Parameter	Estimate	SE	Estimate	SE	Estimate	SE
Decapterus russellii						
φ_1	0.810	0.096			0.548	0.082
Φ_1	-0.425	0.150			-0.370	0.097
φ_1^{12}	0.390	0.098	0.456	0.112		
Φ_1^{12}			0.517	0.090		
r2 (1994)	0.852		0.783		0.874	
r2 (1995)	0.894		0.839		0.897	
r2(24 month)	0.854		0.800		0.869	
Decapterus macrosoma						
φ_1	0.645	0.152			0.632	0.077
Φ_1	-0.428	0.108			-0.427	0.107
φ_1^{12}	0.021	0.214	0.396	0.102		
Φ_1^{12}			0.687	0.095		
r2 (1994)	0.704		0.680		0.708	
r2 (1995)	0.422		0.435		0.421	
r2(24 month)	0.562		0.469		0.563	

The autoregressive coefficients were estimated to be j_1 = 0.548 et F_1= -0.370. Hence the model becomes:

$$X_t = 0.548X_{t-1}+0.63X_{t-12}-0.345X_{t-13}+0.370X_{t-24}-0.203X_{t-25}+\varepsilon_t \quad (4)$$

The unit of time (t) is one month. The ACF and PACF of the residuals show a similar pattern (Fig. 4). They do not indicate any inadequacy in the model as the residuals are independent of each other. The normal probability plots of the residuals show that they are normally distributed (Fig. 4). The coefficient of determination was found to be r^2=0.869 during the 24 month forecasting. From year to year, the coefficient varies from 0.874 for 1994 to 0.897 for 1995 and shows a good stability.

For *Decapterus macrosoma* the model which fits best is the same as the one for *Decapterus russellii*. With coefficients equal to φ_1= 0.632 and Φ_1= -0.427 the model becomes.

$$X_t = 0.632\ X_{t-1}+0.572X_{t-12}-0.362X_{t-13}+0.427X_{t-24}-0.270X_{t-25}+\varepsilon_t \quad (5)$$

Figure 5 shows the ACF, the PACF and the normal probability plots of the residuals. The hypothesis of the validity of the model cannot be rejected. The coefficient of determination is equal to r^2= 0.569 for the two year forecasts. This coefficient varies from 0.708 in 1994 to 0.421 in 1995 and seems relatively unstable.

Forecasts and actual catches for 1994 and 1995, which were not used in the development of either models are plotted in Figure 6.

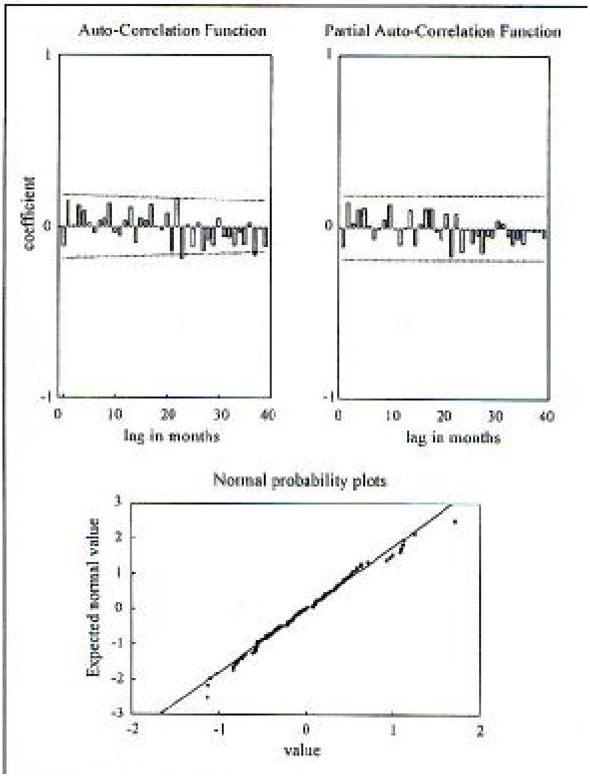

Fig. 4. Auto-correlation, partial auto-correlation functions and normal distribution of the residuals for the SARIMA (1,0,0)(1,1,0) model fitted on the *Decapterus russellii* landings series.

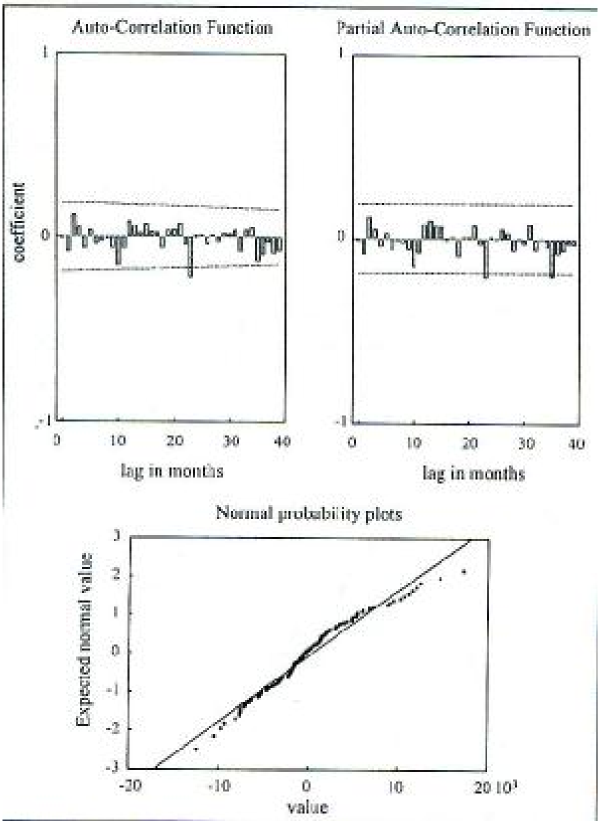

Fig. 5. Auto-correlation, partial auto-correlation functions and normal distribution of the residuals for the SARIMA (1,0,0)(1,1,0) model fitted on the *Decapterus macrosoma* landings series.

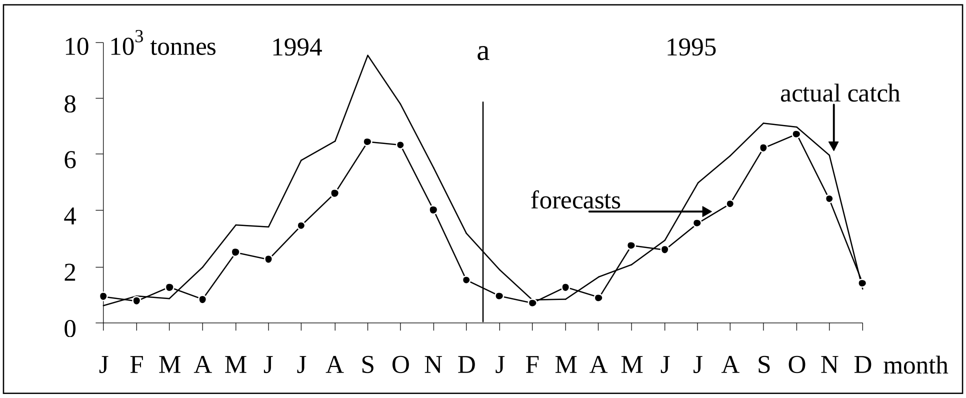

Fig. 6. Comparison during January 1994-December 1995 period between the actual monthly catches of *Decapterus russellii* (a) and *Decapterus macrosoma* (b) and the forecasts determined from the SARIMA (1,0,0)(1,1,0) model.

Discussion

The time series used in the study cover 10 years of landings which means 120 data points. The models suffer limitations as the minimal number of data needed to perform such a study must not be less than 100.

For the two species, there is no single model which fits the data well but a set of models, with similar coefficients of determination. The choice of the best model is made according to the Akaike criteria. The examination of the residuals of each one shows no inadequacy with the data. The seasonal difference term which appears in every model indicates that the fishing activity in the Java Sea is highly seasonal. The behaviour of the two species explain such results which have already been reported on species such as tuna in Hawaii (Mendelssohn, 1981) and anchovy in the Greek waters (Stergiou et al. 1997)

The hydrology of the Java Sea depends entirely of the monsoon regime. During the wet season (December-April) the Java Sea is occupied by a large amount of low salinity water. During the dry season (May-October), oceanic waters coming from the Flores Sea and Macassar Straight enter the region and the salinity increases (Wyrtki, 1597, 1962). These two contrasting periods determine the distribution of the fish populations over the whole area. The different fish species which inhabit the Java Sea are gathered in three populations (coastal, neritic and oceanic). Within the year, these populations exhibit different behaviour according to their ecological needs. According to the season, their distribution may overlap or be distinct.

Decapterus russellii belongs to the neritic population. This species stays in the Java Sea throughout the year and presents limited movements. Its interannual fluctuations are not much related to the environmental pressure. Every year, the degree of confidence of the models remains stable and forecasts are relatively good. Most of the deviation occurs from July to November in the second part of the south-east monsoon and during the inter-monsoon, the period that corresponds to the fishing season.

Decapterus macrosoma is an oceanic species and a seasonal inhabitant of the Java Sea. It enters the sea with the flow of oceanic waters and its migratory pattern is strongly dependent on the intensity of the moonson. In some years, the migration do not reach the Java Sea. These last years, the node of the exploitation shifted from the central to the eastern part of the Java Sea and the disponibility of that species increased a lot. The SARIMA model explains only 56 % of the variance and, from year to year, the degree of confidence varies a lot. The model seems not sufficient to forecast the catch of that species.

Conclusion

The fisheries resources of the Java Sea are heavily exploited (Widodo, 1988; Potier, 1998). SARIMA models allow short forecasts which are of great interest. However, the scads fishery seems difficult to forecast reasonably on a monthly basis because landings are greatly dependent on the year-to-year changes in oceanographic conditions particularly for the oceanic species.

For these last ones, transfer functions taking into account environmental parameters seem more appropriate. But such models are meaningful only when reliable forecasts of the independent variables are available.

Acknowledgment

The present study was supported by the Pelfish Project (Java Sea Pelagic Fishery Assessment project, ALA/INS/87/17) financed by the EU. We wish to thank our Indonesian colleagues who were part of the Pelfish Project; S.B. Atmaja, D. Nugroho, S. Nurhakim, B. Sadhotomo. We are also grateful to the two anonymous referees for the valuable comments on the earlier version of the manuscript.

References

Bartlett M.S. and P.H. Dianda. 1950. Extensions of Quenouille's test for autoregressive schemes. Journal Royal Statistics Society serie B 12:108-115.

Boudreault, F.R., J.N. Du Pont and C. Sylvain. 1977. Modèles linéaires de prédiction des débarquements de homard aux îles de la Madeleine (Golfe du Saint-Laurent). Journal Fisheries Research Board Canada 34:379-383.

Box, G.E.P. and G.M. Jenkins. 1976. Time series analysis, forecasting and control. Holden day, San Francisco. 575pp.

Chatfield, C. 1984. The analysis of time series : an introduction. Chapman and Hall, New-York. 286 pp.

Fogarty, M.J., J.S. Idoine, F.P. Almeida and M. Pennington. 1986. Modelling trends in abundance based on research vessel surveys. International Council Exploration Sea Statistics committee C.M. 1986/G:92: 13pp.

Jeffries, P., A. Keller and S. Hale. 1989. Predicting winter flounder (*Pseudopleuronectes americanus*) catches by time series analysis. Canadian Journal Fisheries Aquatic Science 46:650-659.

Mc Leod, A.I. and P.R.H. Sales. 1983. An algorithm for approximate likelihood calculation of ARMA and seasonal ARMA models. Algorithm AS 191. Applied Statistics 1983:211-223.

Mendelssohn, R. 1981. Using Box-Jenkins models to forecast fishery dynamics : Identification, estimation and checking. Fisheries Bulletin U.S. 78:887-896.

Mendelssohn, R. and P. Cury. 1987. Fluctuations of a fortnightly abundance index of the Ivoirian coastal pelagic species and associated environmental conditions. Canadian Journal Fisheries Aquatic Science 44:408-421.

Potier, M. and T. Boely. 1990. Influence de facteurs climatiques sur la pêcherie à la senne tournante et coulissante en mer de Java. Aquatic Living Resources 3:193-205.

Potier, M. 1998. Pêcherie de layang et senneurs semi-industriels javanais : Perspective historique et approche système. PhD dissertation, University of Languedoc, Montpellier, France. 280pp.

Quenouille, M.H. 1950. Introductory statistics. Butterworth Springer. 248pp.

Saila, S.B., M. Wigbout and R.J. Lermit. 1979. Comparison of some time series models for the analysis of fisheries data. Journal International Council Exploration Sea 39:44-52.

Stergiou, K.I., E.D. Christou and G. Petrakis. 1997. Modelling and forecasting monthly fisheries catches : comparison of regression, univariate and multivariate time series methods. Fisheries Research 29:55-95.

Widodo, J. 1988. Population dynamics and management of « ikan layang », scad mackerel, *Decapterus* spp. (Pisces : Carangidae) in the Java Sea. PhD dissertation, School of fisheries, University of Washington, Seattle, USA. 150pp.

Wyrtki, K. 1957. Die Zirkulation and der Oberfläche der südostasiatischen Gewässer. Deutsche Hydrogrographie Zeitschrift 10:1-18.

Wyrtki, K. 1962. Physical oceanography of the south-east asian waters. Scripps Institute Oceanography University California, Naga Report 2:195pp.

Yamamoto, T. 1980. A standard statistical system for current fishery statistics in Indonesia. A report preapred for the fisheries development and management project, Indonesia. FAO, Rome. FI:DP/INS/72/064, field document 7:79pp.

Manuscript received 30 April 1999; Accepted 05 January 2000.