Predicting Monthly Catch for Some Western Australian Coastal Finfish Species with Seasonal ARIMA – GARCH Models

E. K. M. Lai[a], Y. W. Cheng[b] and M. McAleer[c]

[a]*WA Marine Research Laboratories, Department of Fisheries, Western Australia (elai@fish.wa.gov.au)*
[b]*WACEIO, Curtin University of Technology,* [c]*Department of Economics, University of Western Australia*

Abstract: In recent years, the allocation of fish resources between the commercial and recreational fishing sectors has become a high priority issue for fisheries management in Western Australia. Recreational anglers are concerned about the effect of commercial fishing activities on stocks of many key recreational finfish species of estuaries and the nearshore areas of WA. The lack of biological and economic information for some of these species makes them favourable targets for the use of time series modelling in fish stock assessment. Time series techniques were applied to monthly commercial catch data for four finfish species from 1976 to the present. The species were King George whiting (*Sillaginodes punctata*), red emperor (*Lutjanus sebae*), sea mullet (*Mugil cephalus*) and yellow eye mullet (*Aldrichetta forsteri*). Seasonal variations and trends in the catch for these species were observed. Seasonal autoregressive integrated moving average (ARIMA) models were identified by analysing the autocorrelation function (ACF) and partial autocorrelation function (PACF). Akaike information criterion (AIC) was used for model selection. After fitting the seasonal ARIMA models to the data, trends could be observed in the time series of the noise. The conditional variance for the time series of the noise was not always constant over time. A generalized autoregressive conditional heteroscedasticity (GARCH) model was then used to model the noise. The Ljung-Box and McLeod-Li tests were used to test the randomness of the noise. It was found that the GARCH effect exists in the catch data of most of these species. Based on fitting data from 1976 to 1998, predictions of monthly catch for 1999 and 2000 were compared with the actual figures. The results showed that ARIMA-GARCH models can describe the catch data and yield more accurate predictions than ARIMA models.

Keywords: Seasonal ARIMA; GARCH; ACF; AIC

1. INTRODUCTION

The principal objectives for fisheries management are to ensure the biological sustainability of fish stocks, establish a firm basis for a sustainable and profitable commercial fishing industry, and to allocate fish catch fairly among commercial fishing and recreational fishing and other sectors. Prediction of future commercial catch of fish in terms of whole weight is important for management decision making and for general public reference. For fisheries in Western Australia (WA), the traditional methods for prediction are based on biological and environmental factors such as spawning stock [Mendelssohn, 1988; Hall, 1997]. Unfortunately, the biological and environmental data are very expensive and difficult to collect. These methods are especially difficult to apply to smaller and less valuable finfish fisheries in which the stocks of the main species are fully exploited. Various studies indicate that time series modelling is appropriate for predicting catches for those fisheries where biological data are lacking [Mendelssohn, 1981; Stergiou and Christou, 1996]. In this paper, we study the application of the seasonal autoregressive integrated moving average (ARIMA) model with generalized autoregressive conditional heteroscedasticity (GARCH) errors for four finfish species with over twenty years of commercial catch data.

Autoregressive integrated moving average (ARIMA) models assume that a time series is a linear combination of its own past values and current and past values of a random error term [Box and Jenkins, 1976], and capture the historic autocorrelation of the data to extrapolate them into the future. In classical ARIMA time series models, the conditional variance is assumed to be a constant, which may not be a sensible assumption in practice. Among the models which take this into consideration, the generalized autoregressive conditional heteroscedasticity (GARCH) models are both popular and useful [Bollerslev, 1986]. Definitions of these models will be given in the next section.

Commercial fishermen are required by the Department of Fisheries, Government of Western Australia, to report their monthly catch under the Fish Resources Management Act (1994) regulations. The data are entered into the Catch and Effort Statistics (CAES) System administered by the Research Division of the Department (WAMRL). Monthly catch data for four finfish species collected since 1976 were obtained from the CAES system for this study. They were King George whiting (*Sillaginodes punctata*), red emperor (*Lutjanus sebae*), sea mullet (*Mugil cephalus*) and yellow-eye mullet (*Aldrichetta forsteri*).

King George whiting are popular recreationally-sought fish, as well as a targeted species for some small commercial fisheries located around the Perth Metropolitan region, Bunbury and Albany (Figure 1). In 1998, the total commercial catch in WA was the highest for the last 25 years, most probably as a result of high juvenile recruitment into estuarine and coastal nursery habitat approximately 2-3 years earlier [Penn, 2000:pp.87-89]. Since then, the commercial catch has gradually decreased.

Red emperor are demersal fish found in waters from Shark Bay to the WA/Northern Territory border (Figure 1), and are one of the major species taken by the commercial trap, line and trawl fisheries in this area [Penn 2000:pp.60-67]. The catch of red emperor increased steadily from 1992 to 1996, followed by a gradual decline. This decrease was partly due to the implementation of management controls and considerable latent effort available in the fisheries was not utilized. The size of the recreation catch for this species is not known, but it is likely that the recreational effort increases each year.

Sea mullet are coastal fish found in coastal bays and estuaries, from Port Hedland to Esperance in WA (Figure 1). They are principally commercial species and are taken throughout the year in estuaries. There was a trend of slow reduction in the commercial catch for recent years, which may relate to the decrease in commercial fishing effort in West Coast estuaries.

Yellow-eye mullet are schooling fish inhabiting bays, estuaries and open coastlines, from Shark Bay to the southern coast in WA (Figure 1). They have been sold traditionally as rock lobster bait. As there are now other bait sources being used, the demand for them has gradually decreased, leading to a decline in commercial catch [Lenanton et al., 1984]. On the other hand, the species has become more popular for recreational fishing.

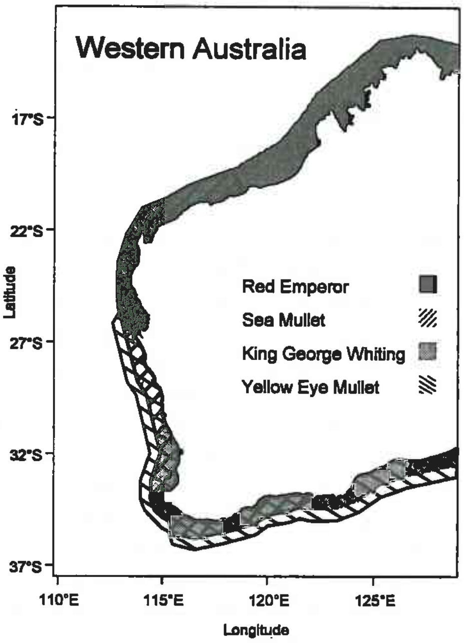

Figure 1. Geographic distribution of the four species in Western Australia.

2. METHODS

A hierarchical approach is used to fit a seasonal ARIMA model with GARCH errors to the time series of the mentioned finfish species. The catch data from 1976 to 1998 are fitted with a seasonal ARIMA model. Resulting residuals are modelled by a GARCH model if necessary. Predictions of monthly catch for 1999 and 2000 are calculated and compared with the actual values.

The appropriate model is identified by examining the autocorrelation (ACF) and partial autocorrelation (PACF) functions of the time series. Model selection can also be based on the minimization of the Akaike information criterion (AIC) [Akaike, 1974]. If the resulting residuals are found to be volatile over time, which is often called the "GARCH effect", then a GARCH model can be applied to smooth the conditional variance and provide better predictions.

Definition 1. Seasonal Autoregressive Integrated Moving Average (ARIMA) model

Let $\{X_t\}$ be a set of observations X_t, each one being related to a specific time t. Then a time series $\{X_t\}$ is a seasonal ARIMA $(p,d,q)\times(P,D,Q)_s$ process with period s if it satisfies a difference equation of the form

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D X_t = \theta(B)\Theta(B^s)Z_t \ , \quad \{Z_t\} \sim \mathrm{N}(0,\sigma^2)$$

where p, d, q, P, D and Q are nonnegative integers; $\phi(z)=1-\left(\sum_{i=1}^{p}\phi_i z^i\right)$, $\Phi(z)=1-\left(\sum_{i=1}^{P}\Phi_i z^i\right)$, $\theta(z)=1+\left(\sum_{j=1}^{q}\theta_j z^j\right)$ and $\Theta(z)=1+\left(\sum_{j=1}^{Q}\Theta_j z^j\right)$; B is the backward shift operator $(B^j X_t = X_{t-j}, B^j Z_t = Z_{t-j}, j=0,1,\cdots)$ and Z_t is the error term. The parameters $\phi_1,\ldots,\phi_p$ are the autoregressive coefficients, $\Phi_1,\ldots,\Phi_P$ are the seasonal autoregressive coefficients, $\theta_1,\ldots,\theta_q$ are the moving-average coefficients and $\Theta_1,\ldots,\Theta_Q$ are the seasonal moving average coefficients. d is the degree of differencing required to achieve stationarity.

Definition 2. Modelling volatility with GARCH

The generalized autoregressive conditional heteroscedasticity GARCH (m,n) process $\{X_t\}$ is a solution of the equations

$$X_t = \sigma_t Z_t, \qquad \{Z_t\} \sim \mathrm{NID}(0,1)$$

where σ_t is the function of $\{X_s, s<t\}$, defined by

$$\sigma^2{}_t = \alpha_0 + \sum_{i=1}^{m}\alpha_i X^2{}_{t-i} + \sum_{j=1}^{n}\beta_j \sigma^2{}_{t-j},$$

with $\alpha_0 > 0$ and $\alpha_i, \beta_j \geq 0$.

Checking model accuracy

There are several methods to validate an ARIMA model, such as examining the autocorrelation function of the estimated residuals and calculating the Ljung-Box *portmanteau* statistic Q [Ljung and Box, 1978] for the estimated residuals. If the correct ARIMA model is fitted and the estimated residuals are Gaussian, then Q is approximately distributed as a chi-squared χ^2 random variable with K degrees of freedom, where K is the number of lags.

The existence of a "GARCH effect" can be checked with the McLeod-Li test statistic $\tilde{Q}$ for the squared estimated residuals [McLeod and Li, 1983]. If the data are identically and independently normally distributed (NID), then $\tilde{Q}$ is approximately a chi-squared distribution with K degrees of freedom.

3. RESULTS

King George whiting

There was a seasonal pattern in the catch time series, with the peak in April each year. The data were fitted with a seasonal ARIMA $(1,1,1)\times(1,1,1)_{12}$ model identified by AIC. The estimated parameter values were $\phi_1 = 5.05\times10^{-1}$ $(p=0.00)$, $\theta_1 = 8.95\times10^{-1}$ $(p=0.00)$, $\Phi_1 = 6.70\times10^{-2}$ $(p=0.00)$ and $\Theta_1 = 8.94\times10^{-1}$ $(p=0.00)$. The Ljung-Box portmanteau statistic $Q=19.04$ $(p=0.52)$ indicated that the selected model was appropriate for the data. The McLeod-Li statistic $\tilde{Q}=19.65$ $(p=0.47)$ indicated that the "GARCH effect" did not exist. The plot of the estimated residuals showed a random pattern. Hence, the conditional variance could be assumed to be a constant, so that further modelling of the noise was not necessary. The fitted and predicted catch values are shown in Figure 2. About 96% of the real monthly catch for 1999 and 2000 was within the 95% confidence interval (Figure 3).

Red emperor

There was a seasonal pattern in the catch time series, with the peak around August each year. The commercial catch increased rapidly from an average of 10 tonnes in 1988 to 50 tonnes in 1996, but gradually decreased after 1996. A seasonal ARIMA $(2,1,1)\times(2,1,1)_{12}$ model was fitted to the data. The estimated parameter values were $\phi_1 = 4.02\times10^{-1}$ $(p=0.00)$, $\phi_2 = -7.10\times10^{-2}$ $(p=0.00)$, $\theta_1 = 8.27\times10^{-1}$ $(p=0.00)$, $\Phi_1 = -1.94\times10^{-1}$ $(p=0.00)$, $\Phi_2 = -2.30\times10^{-1}$ $(p=0.00)$ and $\Theta_1 = 5.20\times10^{-1}$ $(p=0.00)$. The Q statistic of 15.75 $(p=0.73)$ indicated that the selected model was appropriate. The fitted and predicted catch values are shown in Figure 2. The $\tilde{Q}$ statistic of 1.08×10^2 $(p=0.00)$ indicated that a "GARCH effect" existed. Estimated residuals

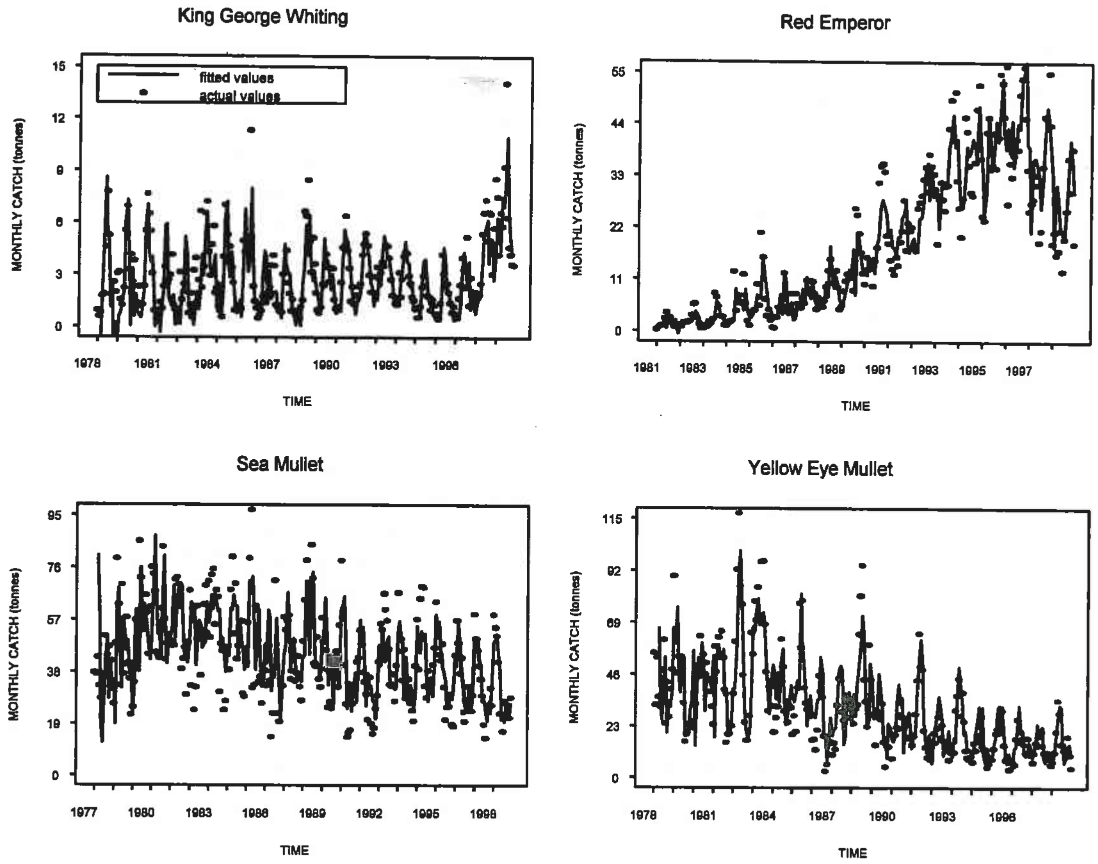

Figure 2. Graphical presentation of the fitted and actual monthly catch from a seasonal ARIMA model for the four finfish species time series over the last 20 or more years.

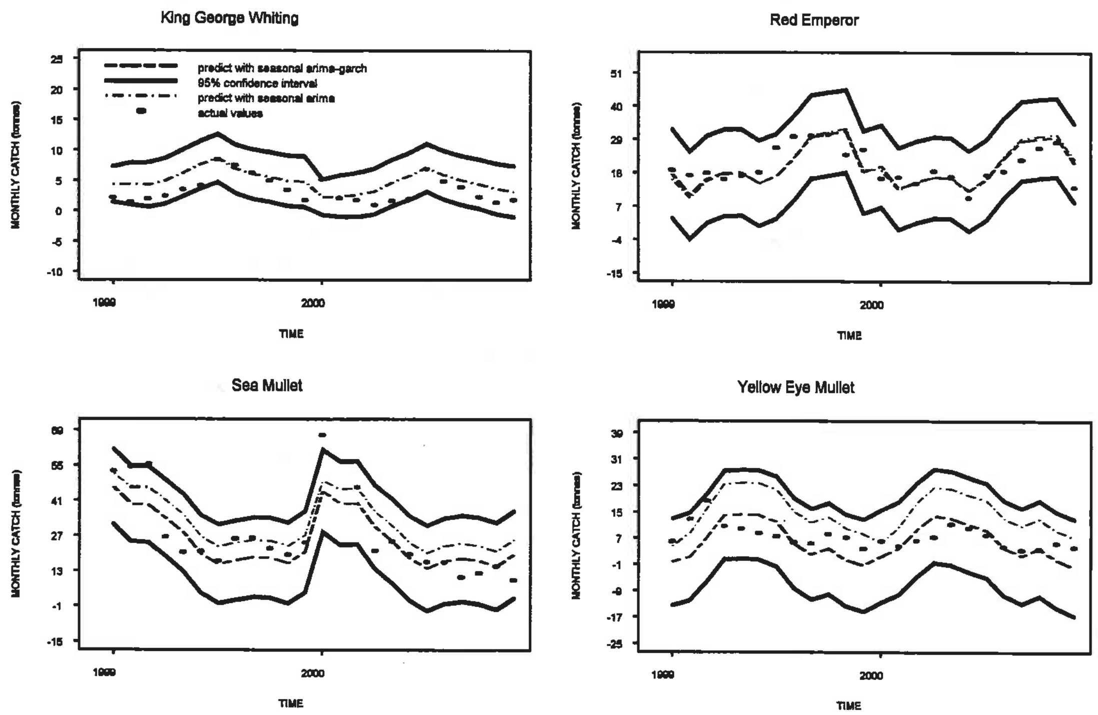

Figure 3. Graphical presentation of the predicted monthly catch for 1999 and 2000 from a seasonal ARIMA model with and/or without GARCH errors for the four finfish species time series.

were found to be gradually increasing over time, so a GARCH $(1,1)$ model was selected by AIC to fit the estimated residuals. The resulting parameter values were $\alpha_0 = 1.85 \times 10^6$ $(p = 0.004)$, $\alpha_1 = 1.20 \times 10^{-1}$ $(p = 0.001)$ and $\beta_1 = 8.32 \times 10^{-1}$ $(p = 0.00)$. Predictions of monthly catch for 1999 and 2000 are shown in Figure 3. These predictions from using the GARCH model were similar to those using the seasonal ARIMA model. The Q statistic for the estimated standard residuals and the $\tilde{Q}$ statistic for the squared standardized residuals were 5.98 $(p = 0.92)$ and 8.31 $(p = 0.76)$, respectively. These results showed that the selected GARCH model was appropriate for the data. About 98% of the real monthly catch for 1999 and 2000 was within the 95% confidence interval (Figure 3).

Sea mullet

There was a seasonal pattern in the catch time series, where most of the catch was taken in the winter months (June-August). In addition, there was a trend of slow reduction which might be related to the decrease in commercial fishing effort. The data were fitted with a seasonal ARIMA $(1,1,1) \times (0,1,1)_{12}$ model and the estimated parameter values were $\phi_1 = 5.01 \times 10^{-1}$ $(p = 0.00)$, $\theta_1 = 9.85 \times 10^{-1}$ $(p = 0.00)$ and $\Theta_1 = 8.14 \times 10^{-1}$ $(p = 0.00)$. The fitted and predicted catch values are shown in Figure 2. The Q statistic of 24.14 $(p = 0.06)$ indicated that the selected model was appropriate, and the $\tilde{Q}$ statistic of 43.92 $(p = 0.00)$ indicated that there existed a "GARCH effect". Estimated residuals were found to be gradually decreasing over time, and a GARCH $(1,1)$ model was selected by AIC to fit the estimated residuals. Resulting parameter values were $\alpha_0 = 2.30 \times 10^6$ $(p = 0.22)$, $\alpha_1 = 7.01 \times 10^{-2}$ $(p = 0.03)$ and $\beta_1 = 9.10 \times 10^{-1}$ $(p = 0.00)$. Predictions of monthly catch for 1999 and 2000 are shown in Figure 3. Better predictions for months with low catch in a year were obtained by applying the GARCH model. The Q statistic for the estimated standard residuals and the $\tilde{Q}$ statistic for the squared standardized residuals were 9.07 $(p = 0.69)$ and 16.81 $(p = 0.16)$ respectively, which showed the selected GARCH model was appropriate. About 88% of the actual monthly catch for 1999 and 2000 was within the 95% confidence interval (Figure 3).

Yellow-eye mullet

In general, there was a seasonal pattern in the catch time series. Most of the catch was taken in the winter months. The catch has followed a decreasing trend since 1976, along with the number of fishers in commercial fishing. The data were fitted with a seasonal ARIMA $(3,0,0) \times (1,1,1)_{12}$ model, and the estimated parameter values were $\phi_1 = 6.77 \times 10^{-1}$ $(p = 0.00)$, $\phi_2 = 1.81 \times 10^{-1}$ $(p = 0.00)$, $\phi_3 = -1.68 \times 10^{-1}$ $(p = 0.00)$, $\Phi_1 = -9.70 \times 10^{-2}$ $(p = 0.00)$ and $\Theta_1 = 7.10 \times 10^{-1}$ $(p = 0.00)$. Fitted and predicted catch values are shown in Figure 2. The Q statistic of 10.95 $(p = 0.95)$ indicated that the selected model was appropriate, and the $\tilde{Q}$ statistic of 6.66×10^1 $(p = 0.00)$ indicated that there existed a "GARCH effect". Estimated residuals were found to be gradually decreasing over time, so that a GARCH $(1,1)$ model was fitted to the estimated residuals, with resulting parameter values being $\alpha_0 = 1.27 \times 10^6$ $(p = 0.19)$, $\alpha_1 = 5.03 \times 10^{-2}$ $(p = 0.06)$ and $\beta_1 = 9.34 \times 10^{-1}$ $(p = 0.00)$. Predictions of monthly catch for 1999 and 2000 are shown in Figure 3. The GARCH model had improved the predictions for most of the months with low catch, with the Q statistic for the estimated standard residuals and the $\tilde{Q}$ statistic for the squared standardized residuals as 8.88 $(p = 0.71)$ and 9.57 $(p = 0.65)$, respectively. The selected GARCH model was appropriate. About 92% of the actual monthly catch for 1999 and 2000 was within the 95% confidence interval (Figure 3).

4. DISCUSSION

In this paper, only the commercial catch history of four species over twenty years was used to predict the catch of two consecutive years. However, many fishers and biologists believe that factors such as fishing effort or fishing power also have a strong impact on the catch. Unfortunately, these relevant biological and environmental data are very difficult and expensive to record properly and precisely, in particular, for multi-species commercial fisheries. Hence, we have chosen only to use monthly catch data.

Modelling time series of catch data in a hierarchical approach was found to be most satisfactory for sea mullet and yellow-eye mullet. The conditional variance of the model was found to be volatile over time for these two species, but the GARCH model has addressed this phenomenon well.

In recent years, the stocks of red emperor were found to be fully exploited in the managed fishing grounds. Catch increased steadily between 1992 and 1996, then decreased from 1997 in response to the implementation of management controls and to latent effort resulting from fishers choosing not to fish in some fisheries. The models have addressed this change and give accurate predictions, but in this case the GARCH model had no discernible impact on the predictions.

Prior to 1997, the catch of King George whiting was steadily decreasing before a sudden jump occurred in 1998, after which the catch fell again. The very high 1998 catch resulted from high juvenile recruitment into Wilson Inlet several years earlier. The seasonal ARIMA model has addressed this change and gives accurate predictions. The volatility of the time series was found not to be significant.

Some problems were encountered when calculating the 95% confidence intervals. The lower bound for the estimated monthly catch was found to have some negative values, which might be unrealistic. This happened primarily because the noise was not normally distributed. Nonparametric bootstrapping can be used to overcome this problem as it is distribution free, but it involves intensive computation and is time consuming. Further investigation would be worth.

If biological and environmental data are available, they can be incorporated into the ARIMA model as a transfer function. A further investigation using different types of GARCH models would be useful. Multivariate time series modelling could be used to study the relationship among different species in the same region.

5. ACKNOWLEDGEMENTS

This study is supported by the Fisheries Research and Development Corporation (project number 99/155). The authors wish to thank the staff from WAMRL, who reviewed this manuscript. The third author wishes to acknowledge the financial support of the Australian Research Council.

6. REFERENCES

Akaike, H., A new look at the statistical model identification. *IEEE Transactions on Automatic Control,* 19, 716-723, 1974.

Bollerslev, T., Generalized autoregressive conditional heteroscedasticity. *Journal of Econometries*, 31, 307-327, 1986.

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

Hall, N.G., Delay-difference model to estimate the catch of different categories of the western rock lobster (*Panulirus cygnus*) for the two stages of the annual fishing season. *Marine Freshwater Research*, 48, 949-958, 1997.

Lenanton, R.C.J., I.C. Potter, N.R. Loneragan and P. J. Chrystal, Age structure and changes in abundance of three important species of teleost in a eutrophic estuary (Pisces: Teleostei). *Journal of Zoology*, 203, 311-327, 1984.

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

McLeod, A.I. and W.K. Li, Diagnostic checking ARMA time series models using squared-residual autocorrelations. *Journal of Time Series Analysis*, 4, 269-273, 1983.

Mendelssohn, R., Using Box-Jenkins models to forecast fishery dynamics: identification, estimation and checking. *Fishery Bulletin*, 78, 887-896, 1981.

Mendelssohn, R., Some problems in estimating population sizes from catch-at-age data. *Fishery Bulletin*, 86, 617-630, 1988.

Penn, J., *State of Fisheries Report 1999/2000.* Western Australian Marine Research Laboratories, 87-89, 2000.

Stergiou, K.I. and E.D. Christou, Modelling and forecasting annual fisheries catches: comparison of regression, univariate and multivariate time series methods. *Fisheries Research*, 25, 105-138, 1996.