Volume 3 No.6, June 2013 ISSN 2224-3577

International Journal of Science and Technology

http://www.ejournalofsciences.org

The Status of the Commercial Chambo (*Oreochromis* Species) fishery in Malaŵi: A Time Series Approach

Matthews Lazaro, Wilson Wesley Lazaro Jere
Aquaculture and Fisheries Science Department, Bunda College of Agriculture, P. O. Box 219, Lilongwe, Malaŵi.

ABSTRACT

Fisheries forecasting is a very important tool for fisheries managers and scientists to enable them to decide on sustainable management issues Time series models have been used to forecast catches in fisheries sectors in different countries but to the contrary, Malaŵi has lagged behind in using time series model in forecasting. In this paper time series models have been used to forecast commercial *Oreochromis* species locally known as chambo based on data on Lake Malawi fish catch during years from 1976 to 2010. The study considered Autoregressive (AR), Moving Average (MA) and Autoregressive Integrated Moving Average (ARIMA) processes to select the appropriate stochastic model for forecasting annual commercial chambo catch from Lake Malawi. Based on ARIMA (p, d, q) and its components ACF, PACF, Normalized BIC, Box-Ljung Q statistics and residuals estimated, ARIMA (1, 1, 0) was selected. Based on the selected model, it could be forecasted that the commercial chambo catch would increase to 854 tonnes in 2020 from 437 tonnes in 2010.

Keywords: *Forecasting, ARIMA, Management, Oreochromis species, Lake Malawi.*

I. INTRODUCTION

The tilapiine fish species of the genus *Oreochromis* Günther 1984 described from Lake Malaŵi are of great economic importance [1, 2]. The species which have males with tasselled genitalia belong to the subgenus *Nyasalapia* Thys van den Audenaerde 1968 while the rest belong to the subgenus *Oreochromis* [3]. Members of the subgenus *Nyasalapia* are collectively known as 'chambo'. Chambo tops every fish menu in Malaŵi [4, 5] and commands a higher market price than most other common species, which is said to encourage more entrants into the fishery [2]. After decades of research into members of the subgenus *Nyasalapia*, their taxonomy is still in limbo. This subgenus traditionally comprised the following species: *Oreochromis karongae* [6], *Oreochromis lidole* [6], *Oreochromis squamipinnis* and *Oreochromis saka* [7, 8]. Subsequent work has shown that *O. saka* was a junior synonym of *O. karongae* [9].

The chambo fishery supports both artisanal and commercial fisheries in Lake Malaŵi [10, 11]. Artisanal fisheries are open access, highly complex, scattered in all water bodies and mainly operate between 0-20 m in Lake Malaŵi while in other water bodies all depth ranges are covered. On the other hand commercial fisheries are mechanized, capital intensive and use mainly trawling and purse seining ('ring net') and are mostly confined to the southern part of Lake Malaŵi.

Over the past years, the total annual catch of the chambo had declined markedly from a record high of over 9400 metric tons in 1985 to a low level of about 1400 tons by 1999 [10]. To improve management and restoration of the chambo fishery, there is need for determining accurate forecasts of chambo catches. Since fisheries resources are renewable, good management issues should be taken into consideration to manage these fisheries resources. From the management point of view, fish forecasting is a very important tool for fisheries managers and scientists to enable them to decide on sustainable management issues [12]. Box-Jenkins models have been demonstrated to be appropriate for forecasting fisheries catches. They are specifically designed for estimating and testing models in the presence of autocorrelated errors [13]. Box-Jenkins models usually refer to autoregressive integrated moving average (ARIMA) models [14] and are attractive from the standpoint of data requirements in that only one series of data is required [15]. With ARIMA models, data are assumed to be the output of a stochastic process, generated by unknown causes, from which future values can be predicted as a linear combination of past observations and estimates of current and past random shocks to the system [16].

Time series models have been used to forecast catches in fisheries sectors in different countries. However fish catch forecasting in Malaŵi has been neglected [17]. This study therefore develops and tests ARIMA for forecasting annual commercial chambo catches from Lake Malaŵi.

II. LITERATURE REVIEW

In Malaŵi, few ways have been employed in assessing catch trend for a number of fish species in water bodies. The mostly commonly method in use is catch per unit effort (CPUE). However over the past years, the use of time series model have been recommended as the best in determining both trends and forecasts in other countries.

The time series analysis technique is used because it is easy to interpret and allows a detailed assessment of seasonal effects. It has ability to forecast production trends and seasonal fluctuations.

For example, time series models were used to forecast subsequent milk production on the dairy farm that had maize silage not been introduced, and these forecasts were compared with actual production after the introduction of maize silage. The models appeared to be accurate as there were almost no significant differences between actual and forecasted value for the entire forecasting period [18]. A nother related study was done by Iqbal *et al*. (2005) to forecast the production of wheat in Pakistan up to 2022 using last thirty years data of area and production on wheat for modeling purpose. The ARIMA model showed that production of wheat would be 29774.8 thousand tonnes in 2022. In Malaysia for example, Shitan .et. al (2004) carried out a research where the maximum likelihood and bootstrap method were used to forecast the total Malaysian marine fish production. Using time series models, demersal and pelagic marine fish production was forecasted and it was found that ARIMA models appeared to better than other method. The better performance of ARIMA could be due to a fact that ARIMA model bases on short memory process and hence pelagic fish are more subjective to variations caused by anthropogenic activities and other processes that may occur at the surface [12].

Furthermore, Pierce and Boyle (2003) used time series methodology to model and forecast the landings and catch per unit effort of marine fish and invertebrates species and suggested that ARIMA models are the most appropriate to forecast fishery landings in Hellenic marine waters, since systematic biological time series data sets from explanatory variables are lacking.

III. METHODOLOGY

The data used in this study are univariate time series of total commercial chambo fish catch from Malaŵi waters from 1976 to 2010 obtained from Fisheries Research Unit of the Malaŵi Fisheries Department. The unit of measurement refers to the weight of fish at the time of removal from water in metric tons. From this data set, the last 10 observations were kept for the purposes of checking the accuracy of post sample forecasting [12].

Time Series Modeling Procedure

Time series plots were made for each fishery to inspect the presence or absence of trend. If the trend was present, first order differencing was used to achieve stationarity; a requirement in stochastic modeling.

Differencing was achieved using the formula below:

$$\Delta x_t = x_t - x_{t-1}$$

Where Δx_t is differenced series, x_t and x_{t-1} are original series

In general, by correlating previous time lags, time series models were expressed as a combination of autoregressive (AR) and moving average (MA) parameters.

The AR process of order *p*, denoted as AR (*p*) is defined as follows:

$$X_t = \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \cdots\cdots\cdots + \emptyset_p X_{t-p} + w_t$$

Where $\emptyset_1, \emptyset_2,$ …….., $\emptyset_p$ are constants (parameters) and w_t is a random uncorrelated noise component (residuals).

The MA process of order *q*, denoted as MA (*q*) is defined as:

$$X_t = w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2} + \cdots\cdots\cdots + \theta_q w_{t-q}$$

Where θ_1,θ_2, …….., θ_q are constants (parameters) and w_t is a random uncorrelated noise component (residuals).
In cases where differencing occurred, the ARIMA model shown below was used

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2} + \cdots + \theta_q w_{t-q}$$

Where x_t is the original data series or differenced data at time t, w_t is the white noise at time t, ϕ are the AR parameters, p is the autoregressive order, θ are the MA parameters and q is the moving average order.

The generalized additive ARMA (*p, q*) model therefore allowed inclusion of both AR and MA terms. Estimating the number of AR and MA parameters to include in the time series model was accomplished through simultaneous inspection of the autocorrelation function (ACF) and the partial autocorrelation function (PACF). Fitting the model involved the Box-Jenkins three-step procedure as outlined by Chatfield (2004), which involves: identification, estimation and diagnosis. For a parsimonious representation of the time series model, as few terms as possible were used.
Where the process of differencing occurred, the basic ARMA model was extended to the auto-regressive integrated moving average (ARIMA) (p, d, q).

Trend Fitting

In this study the Box-Ljung Q statistics was used to transform the non-stationary data in to stationarity data and to check the adequacy for the residuals. For the purposes of evaluating the adequacy of AR, MA and ARIMA processes, various model fitting statistics like Root Mean Square Error(RMSE), Mean Absolute Percentage Error (MAPE), and Bayesian Information Criterion (BIC) were employed. Based on Normal BIC, the principle is that the lower the value the better the model.

Fit statistics such as MAPE, MAE and RMSE were calculated as shown below.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_i - \hat{Y}_i}{Y_i}\right|$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|Y_i - \hat{Y}_i\right|$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}$$

Where

Y_i and $\hat{Y}_i$ are actual observed and predicted values respectively while n is number of predicted values
Upon identification of optimum model, forecasts of the catches for ten years were made.

All analyses in this study were performed using International Business Management Statistical Package for Social Scientists software (IBM SPSS 20)

IV. RESULTS AND DISCUSSION

Identifications of models

In a move to test the stationarity of the commercial chambo catch data plot was made. ARIMA model was designed after assessing that transforming the variable under study was a stationary series. According to Sankar 2011 the stationary series is the set of values that vary over time around a constant mean and constant variance and the most common method used to check the stationarity is through explaining the data by graph and hence Figure 1 below.

Volume 3 No.6, June 2013 ISSN 2224-3577

International Journal of Science and Technology

http://www.ejournalofsciences.org

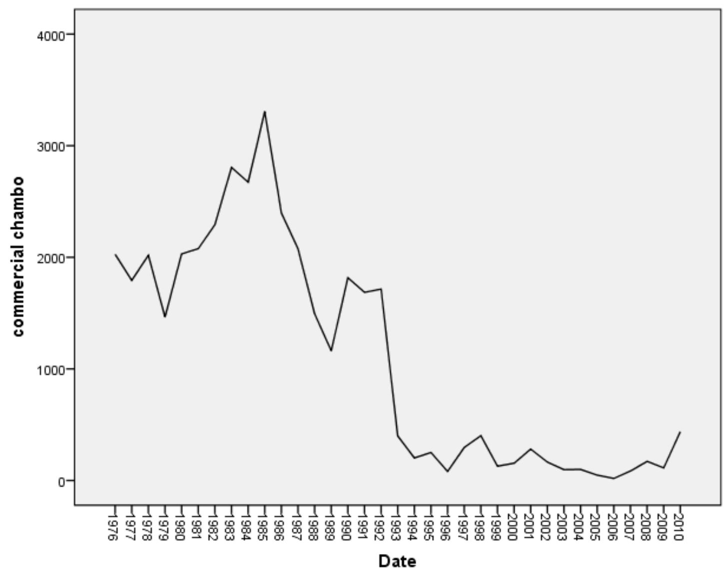

Figure 1: Catch of commercial chambo from Lake Malawi for period 1976 to 2010

From the graph in figure 1 above, it is apparent that the data is not stationary due to presence of trends in some sections throughout the 1976 to 2010 period. This observation in this study is in total agreement with what Zuur and Pierce 2004 observed that most fisheries data is non-stationary. The problem of non-stationary was overcome through first order differencing of the data and stationary test were made on newly constructed series of the data. Since the new series was stationary, then the next step was to determine the values of p and q in the ARIMA models by plotting autocorrelogram and partial autocorrelogram. The plot of autocorrelation function showed a seventh-order moving average model while a plot of partial autocorrelation function showed fourth-order autoregressive model as shown in Figure 2 below.

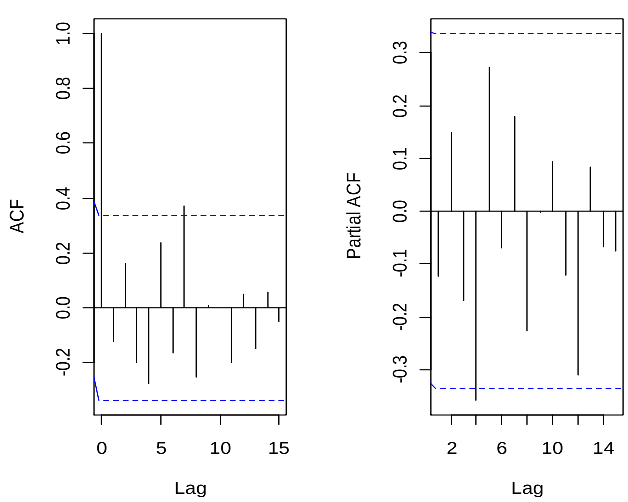

Figure 2: ACF and PACF of first order differenced data

For the purposes of clear presentation, the autocorrelation and partial autocorrelation coefficients (ACF and PACF) of various orders of differenced series of the data were computed and presented in Table 1.below.

Table 1: ACF and PACF for first order differenced data of commercial chambo

Lag	ACF	Std. Error	Box-Ljung Statistic Value	df	Sig.	PACF	Std Error
1	0.124	0.164	0.567	1	0.452	-0.124	0.171
2	0.163	0.162	1.578	2	0.454	0.150	0.171
3	-0.198	0.159	3.126	3	0.373	-0.169	0.171
4	-0.276	0.157	6.233	4	0.182	-0.357	0.171
5	0.239	0.154	8.640	5	0.124	0.274	0.171
6	-0.164	0.151	9.811	6	0.133	-0.069	0.171
7	0.373	0.149	16.121	7	0.024	0.178	0.171
8	-0.253	0.146	19.141	8	0.014	-0.226	0.171
9	0.008	0.143	19.144	9	0.024	-0.002	0.171
10	0.001	0.140	19.144	10	0.038	0.094	0.171
11	-0.198	0.137	21.233	11	0.031	-0.121	0.171
12	0.050	0.134	21.375	12	0.045	-0.309	0.171
13	-0.149	0.131	22.664	13	0.046	0.084	0.171
14	0.056	0.128	22.858	14	0.063	-0.067	0.171
15	-0.051	0.125	23.028	15	0.084	-0.076	0.171
16	0.025	0.121	23.070	16	0.112	-0.068	0.171

Using these coefficients, various competing hesitant model were identified and the models together with their corresponding fit statistics are given in Table 2. The ARIMA model with the lowest Normal BIC value was selected .The values of the NBIC of the selected ARIMA model was 12.302 as given in the table. Hence, the most suitable model for commercial chambo catch is ARIMA (1, 1, 0), as this model had the lowest normalized BIC value. According to Czerwinski et al 2007 the best model should have adequate accuracy measures (RMSE, MAE).for it to have accurate forecasts and thus why the model was chosen since it had meet the requirements.

Table 2: Fit statistics for various competing ARIMA models

ARIMA (p, d, q)	RMSE	MAPE	MaxPE	MAE	NBIC	MaxAE	NBIC
ARIMA (1, 1, 0)	423.875	86.325	686.520	292.793	**12.302**	1267.278	**12.302**
ARIMA (0 ,1, 1)	664.231	273.713	2584.411	548.099	13.200	1395.432	13.200
ARIMA (1, 1, 2)	416.164	106.312	1034.406	292.597	12.468	1163.564	12.468
ARIMA (0, 1, 2)	526.723	186.858	1446.067	379.738	12.838	1120.274	12.838
ARIMA(3, 1, 3)	400.653	77.163	578.836	274.932	12.697	894.404	12.697

Model Estimation

Having obtained suggested model the estimates for the parameters were found as given in table 3. The coefficients parameters of the model were found to be statistically significant; a requirement for forecasting models (27). Czerwinski *et al*. 2007 indicated that the Normalized BIC test reveals that the model with the least Normalized BIC is better in terms of forecasting performance than the one with a large Normalized BIC. Based on above findings, the most suitable model for commercial chambo was confirmed to be ARIMA (1, 1, 0). Model parameters of the selected model were estimated using IBM SPSS and the results of estimation are presented in Tables 3.

International Journal of Science and Technology

http://www.ejournalofsciences.org

Table 3: Commercial chambo catch Estimated ARIMA model

	Estimate	SE	T-value	P-value
constant	1144.219	633.972	1.8050	0.080
AR 1	0.91500	0.06600	13.951	0.000

Diagnostic Checks

Before forecasting, there is need to do diagnostic checks on the proposed best model. This involves checking the residuals of the model to see if they contained any systematic structure which still could be removed to improve the selected ARIMA.in this study diagnostic checking was achieved by examining the autocorrelations and partial autocorrelations of the residuals of various orders. To accomplish this, various autocorrelations up to 24 lags were computed and plotted as indicated in figure 3 below. As the plots of ACF and PACF residuals indicate, none of autocorrelations was significantly different from zero at any reasonable level. This proved that the selected ARIMA model was an appropriate model for forecasting fish catch from Lake Malawi. From the plots it is clear that the autocorrelation coefficients of the residuals are within 95% confidence interval. This proved that the selected model was optimum for commercial chambo catch forecasting and the fitted ARIMA was

$$x_t = 1144.219 - 0.915x_{t-1} + w_t$$

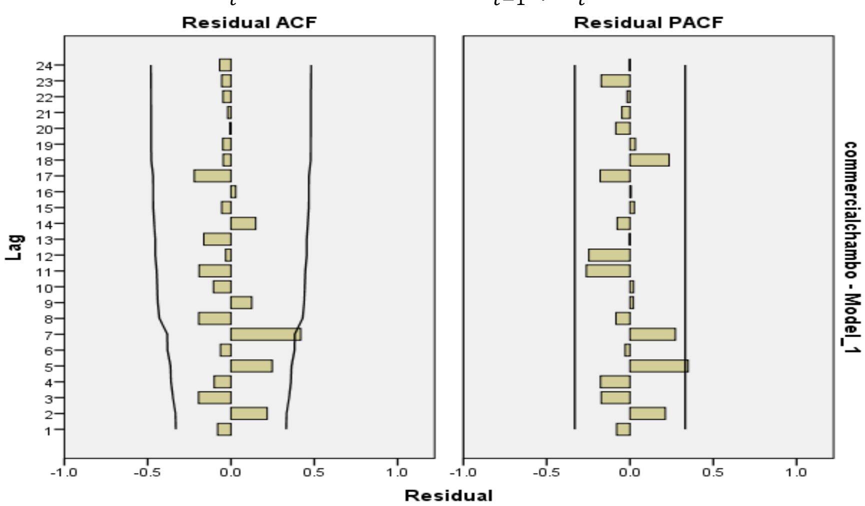

Figure 3: ACF and PACF residuals

Forecasting

Using the fitted model, forecasts for commercial chambo were made from the year 2011 to 2020. In order to assess the ability of the model in forecasting, actual catches were kept for the purposes of checking the accuracy of post sample forecasting. For the clear representation, only the last 10 observations have been compared with the forecasted values as shown in Table 4. However, all observations and forecasted values together with their 95 % confidence interval are given in Figure 4.This is in concordance with what Shitan et al, 2008 did in their study. In this study the forecasted and actual values were close meaning the forecasting error was low which is good for a model. It is argued that a good model has a low forecasting error, therefore when the distance between the forecasted and actual values are low then the model has a good forecasting power [23].

Table 4: Forecast of commercial chambo catch (in tonnes) together with 95 % Confidence Interval.

Year	Actual Catch	Predicted catch	95 % Confidence Interval
2001	281	239	(-574,1053)
2002	165	355	(-459,1168)
2003	97	249	(-565,1062)
2004	99	186	(-627,1000)
2005	48	188	(-625,1002)
2006	18	142	(-672, 955)
2007	85	114	(-699, 928)
2008	171	175	(-638,989)
2009	113	254	(-560,1068)
2010	437	201	(-613,1015)
2011	-	497	(-316,1311)
2012	-	553	(-550,1655)
2013	-	603	(-693,1899)
2014	-	649	(-788,2087)
2015	-	692	(-855,2238)
2016	-	730	(-901,2362)
2017	-	766	(-934,2465)
2018	-	798	(-957,2627)
2019	-	827	(-972,2627)
2020	-	854	(-981,2690)

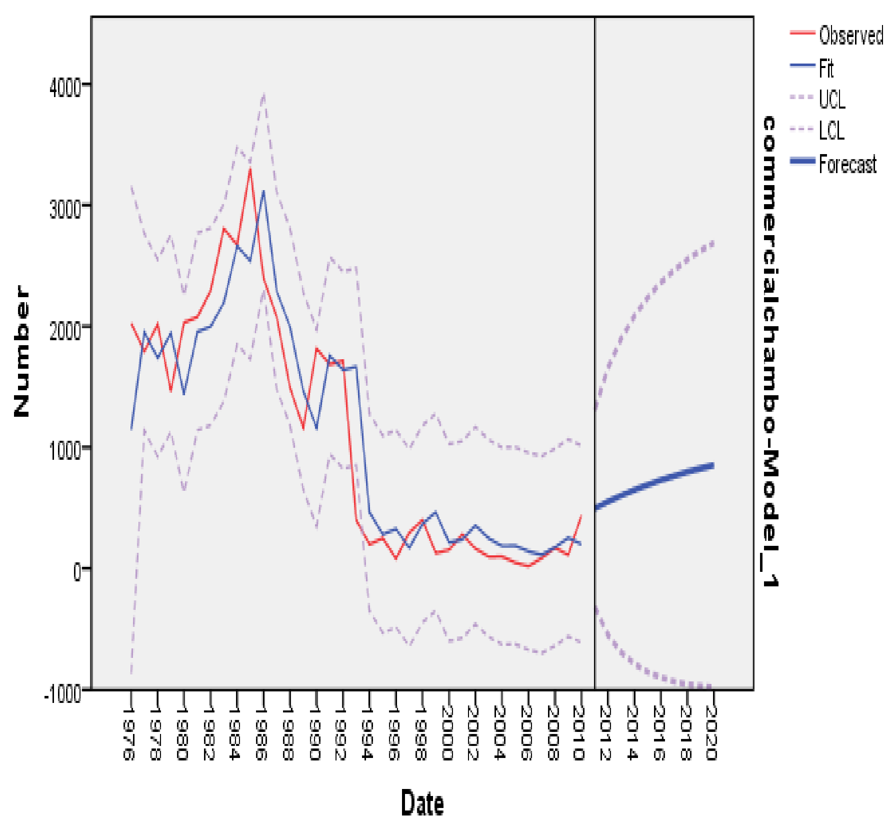

Figure 4: Actual and forecasted commercial chambo catch from Lake Malawi

International Journal of Science and Technology

http://www.ejournalofsciences.org

The mean commercial chambo forecasts in Lake Malaŵi show that the catches will continue to increase to approximately 800 metric tonnes by the year 2020. However, the 95 % confidence intervals for the commercial chambo fishery include a zero showing that we cannot reject the hypothesis that the commercial catches have reached their zero point. These results are a clear demonstration that chambo fishery has completely collapsed. The chambo have been known to be a declining component of the Lake Malaŵi fishery. The difficulties in identifying the species reliably, except as sexually mature adults, have been a major constraint in the rational management of the resource [8, 3]. The minimum size for accurate identification of the species of the genus *Nyasalapia* is 15 to 20 cm [9]. Catch statistics recorded by the Malaŵi Fisheries Department also fail to distinguish the species [8]. Furthermore, the biology of these species complicates the management of the chambo fishery. All the species are dimorphic maternal mouth brooding.

The decline in Chambo for the past years and the next ten years comes at a time when various strategies have been put in place to overcome this problem. As reported by Banda et al (2005) the Malaŵi Government's Department of Fisheries launched a National Save the Chambo Campaign which aimed at mobilizing all Malaŵians towards a program of sustainable exploitation of the Chambo and attracting more foreign and domestic funding to revive the Chambo fishery by 2015 by restoring depleted stocks to the pre-1990 level. By looking at Chambo catch forecasts, it is evident that this objective will not be achieved as such based on these findings policy makers should make necessary strategies in order to manager this fishery.

Acknowledgement

Authors would like to express gratitude to the Malaŵi Government's Department of Fisheries for providing the data, which has been used in this study.

REFERENCES

[1]. Mvula, P.M., Njaya F.J. and Mkoko,. B.J. 2003. Socio-economic factors influencing the decline of the chambo. In: Banda, M., Jamu, D., Njaya, F., Makuwira, M. and Maluwa, A. (Eds.).

[2]. Turner, G.F. 1996. Offshore cichlids of Lake Malaŵi. Cichlid Press.

[3]. Sodsuk, P.K., B.J. McAndrew and G.F. Turner. 1995. Evolutionary relationship of the Lake Malaŵi *Oreochromis* species: evidence from allozymes. Journal of Fish Biology 47: 321 – 333.

[4]. Kaunda, E., Maliro, D. Mphepo, M. Khaila, S. Kamanga, L. Mwangwela, A. and Chapotoka, O.2003. Chambo: Is it taste or something else? Preference and acceptability of roasted, fried and boiled *Oreochromis* spp (Chambo), *Ramphochromis* spp (Ncheni) and *Bagrus meriodionalis* (Kampango) from Lake Malaŵi. Aqua-Fish technical Report 2: 6 – 9.

[5]. Konings, A. (2007). Malaŵi cichlids in their natural habitat. Cichlid Press. Forth edition.

[6]. Trewavas, E. (1941). A revision of the cichlid fishes of the genus *Lethrinops*, Regan. Annals and Magazie of Natural History 10, (7), 133-152.

[7]. Lowe R.H (1952). Report on Tilapia and other Fish and Fisheries of Lake Nyasa 1945-47. Colonial Office Fishery Publication Vol. No. 2. Colonial Office. London

[8]. Turner, G.F. Pitcher, T.J and Grimm. A.S. 1989. Identification of the Lake Malaŵi *Oreochromis* (*Nyasalapia*) spp. using multivariate morphometric techniques. Journal of Fish Biology 35: 799 – 812.

[9]. Snoeks, J. and M. Hanssens. 2004. Identification guidelines to other non-mbuna. *In*: Snoeks, J. (ed.). The cichlid diversity of Lake Malaŵi/Nyasa/Niassa: identification, distribution and taxonomy. Cichlid Press. pp 266-310

[10]. Banda, M. C., Kanyerere, G.Z. and. Rusuwa, B.B 2003. The status of the chambo in Malaŵi: fisheries and biology. In: Banda, M., Jamu, D., Njaya, F., Makuwira, M. and Maluwa, A. (Eds.). The chambo restoration strategic plan: proceedings of the national workshop held on 13 – 16 May 2003 at Boadzulu Lakeshore Resort, Mangochi, Malaŵi, pp. 1 – 7.

[11]. Kafumbata, D.R. and. Ambali. A.J.D 2003. Genetic diversity of the *Oreochromis* (*Nyasalapia*) species in Lake Malaŵi. In: Banda, M., Jamu, D., Njaya, F., Makuwira, M. and Maluwa, A. (Eds.). The chambo restoration strategic plan: proceedings of the national workshop held on 13 – 16 May 2003 at Boadzulu Lakeshore Resort, Mangochi, Malaŵi, pp. 12 – 19.

[12]. Shitan, M. Wee, P.M.J. Chin, L.Y and Siew, L.Y. 2008. Arima and integrated Arfima Models for Forecasting Annual Demersal and Peragic marine fish production in Malaysia. Malaysian Journal of mathematical sciences 2 (2): 41-54 (2008)

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

[14]. Chatfield, C. 2004. The analysis of time series: an introduction. Chapman and Hall.

[15]. Myer, G. L. and. Yanagida, J. F. 1984. Combining annual econometrics forecasts with quarterly ARIMA forecasts: a heuristic approach. Western Journal of Agricultural Economics 9: 200 – 206.

[16]. Prista, N., Diawara, N. Costa, M. J. and Jones. C. 2011. Use of SARIMA models to assess data-poor fisheries: a case study with a sciaenid fishery off Portugal. Fishery Bulletin 109: 170 – 185.

[17]. SADC Secretariate (2009). Regional early warning system for food security; selected technical papers on methodology of food crop forecasting in SADC, Gaborone, Botswana

[18]. Mkhabela T.S (2004). Estimates of the increase in milk production due to the introduction of maize silage to a dairy farm in Kwazulu-Natal: A time series approach. Agrekon, 43 (4): 487 – 490.

[19]. Iqbal N, Maqbool A and Shoaib A (2005) Use of the ARIMA Model for Forecasting Wheat Area and Production in Pakistan, *Journal of Agriculture and Social Sciences*, Vol. 1(2) 120-122.

International Journal of Science and Technology

http://www.ejournalofsciences.org

[20]. Pierce, G.J., Boyle, P.R., 2003. Empirical modeling of interannual trends in abundance of squid (Loligo forbesi) in Scottish waters. Fisheries Research 59 (3), 305-326.

[21]. Sankar, J.2011. Stochastic Modeling Approach for forecasting fish product export in Tamilnadu. Journal of Recent Research in Science and Technology 3(7): 104-108, 2011.

[22]. Zuur, A.F and Pierce, G.J. 2004. Common trends in northeast Atlantic squid time series. Journal of Sea Research 52 (2004) 57– 72.

[23]. Czerwinski, I.A. Guti ´errez-Estrada, J.C and. Hernando-Casal, J.A 2007. Short-term forecasting of halibut CPUE: Linear and non-linear univariate approaches. Fisheries Research 86: 120–128, 2007.

[24]. Guti´errez-Estradade Pedro-Sanz, J.C. Opez-Luque, E. L and Pulido-Calvo, I. 2004. Comparison between traditional methods and artificial neural networks for ammonia concentration forecasting in an eel (*Anguilla Anguilla* L.) Intensive rearing system. Aquat. Eng. 31, 183–203.