A time series analysis of Prochilodus nigricans landings caught by small-scale fisheries in the lower stretch of the Amazon River

Santana, IF. and Freitas, CEC.*

Departamento de Ciências Pesqueiras, Faculdade de Ciências Agrárias, Universidade Federal do Amazonas – UFAM, Av. Gen. Rodrigo Otávio, 3000, CEP 69077-000, Manaus, AM, Brazil
*e-mail: cefreitas@ufam.edu.br

Received October 7, 2011 – Accepted March 8, 2012 – Distributed February 28, 2013
(With 4 figures)

Abstract
We developed a time series analysis using data on curimatã (Prochilodus nigricans), which landed in Santarém, a small city located on the right banks of the Amazon River. A 10-year record of monthly average catches per day of P. nigricans was analyzed using forecasting procedures in the open-source software GRETL 1.7.8. We established two models from the identifications made with the correlograms of hyperparametrization and seasonal differences. The autoregressive terms of the model reach three years, indicating that individuals of the species are being caught around the age of three. This may indicate that the curimatãs in the landings at Santarém from 1992 to 2002 were more than two years old, potentially a sign of a lack of fishing pressure on the lower age groups.

Keywords: fish landings, curimatã, small-scale fishery, Rio Amazonas.

1. Introduction

The genus Prochilodus, in the family Prochilodontidae, has thirteen species that are widely distributed throughout South America, absent only in the coastal rivers of Ecuador (Castro and Vari, 2003). The curimatã (P. nigricans) is one of the most commercially important species for Amazon fisheries (Batista and Petrere Junior, 2007). In Manaus, Saint-Paul and Bayley (1979), it is estimated that the landings of curimatã accounted for 11% and 19% of the total fish catch in 1976 and 1977, respectively. The prevalence of curimatã in the landings remains high; between 1994 and 2003, curimatã was responsible for 18 to 30% of the total catch landed in Manaus (Batista and Petrere Junior, 2003). The landings of this species are also important in Santarém, located in the Lower Amazon River (Ruffino and Isaac, 1994); in Porto Velho, along the Madeira River (Boishio, 1992); and in small cities in the state of Amazonas (Batista and Petrere Junior, 2007). In the Central Amazon, curimatã are mainly captured using purse seine and gillnets (Petrere Junior, 1978; Batista and Petrere Junior, 2007).

Ruffino (2006) observed that the average fork length of curimatã remained constant in the landings caught in 2001 and 2002 and was approximately 34.0 ± 3.0 cm and 33.8 ± 2.9 cm, in the respective years. According to Ruffino (2006), as compared to curimatãs from 2001 and 2002 there was a slight decrease in the average length of landed curimatãs in 2003. However, this decrease in length was not indicative of an increase in overfishing. Because of its commercial importance, the curimatã is one of the primary species monitored by the fishery statistical systems that are deployed in the Amazon basin.

Despite the commercial importance of curimatã, only Freitas et al. (2007) have evaluated the state of stocks of curimatã in Tefé, Manaus, and Santarém based on
landing data, using surplus production models. They did not observe an occurrence of overfishing. However, the fit of the models used was poor, most likely as a result of the small extent of the data series. Moreover, the multi-gear, multi-specific nature of these fisheries poses challenges in identifying the actual fishing effort invested in the capture of a species, leading to the use of general effort units.

The use of surplus production models is based on a priori assumptions about the nature of population growth, the relationship between the fishing effort and population size, and the form of the relationship between yields and fishing effort (Saila et al., 1980). It is only possible to establish a maximum equilibrium yield and optimum applied fishing effort when adequate catch and fishing effort data are available. Except for some large catfish, such as *Brachyplatystoma vailantii* (Barthem and Petrere Junior, 1995), these types of data are absent for Amazonian fisheries because of the complexity generated by using several types of fishing gear to capture a range of species. Additionally, the data series are short and incomplete. In response to the lack of adequate data series of catch and fishing effort, this article developed a prediction model for fisheries that catch *Prochilodus nigricans*, using time series models based only on catch data, in an attempt to identify temporal patterns in the fish catches, which landed in Santarém.

2. Material and Methods

2.1. Data collection

The Amazon Institute of Environmental Resource Management collected daily data on the fish landings from 1992 to 2002. Data on the amount of fish landed, separated by species, were collected by interviewing the fishers immediately below the landings at two sites in the city of Santarém. This city is the main port of landing of freshwater fish in the Lower Amazon, and the fisheries operate in both the Amazon River and in a large floodplain area at the western edge of the river (Figure 1). A few boats from the Amazon fishing fleet landed fish when the price of fish was higher in Santarém than in Manaus or Parintins. The daily data on the amount of curimatã, which landed were calculated by month to obtain monthly landings for each year.

2.2. Data analysis

Data were log-transformed and the assumption of normality was tested using a Lilliefors test (software GRETL 1.7.8). The three stages of temporal interpolation (identification, estimation, and verification) were also performed using the open software GRETL 1.7.8. The models were built by forming combinations of parameters that appeared to match the correlograms; parameters that appeared to match the correlograms; parameters.

![Figure 1](image_url). Study area showing Santarém city, the rivers Amazonas and Tapajós, and the large floodplain areas placed at the left margin of the Rio Amazonas.
that were not found in the correlograms were included in order to generate hyperparametrized models. Then, the models were tested with and without regular and seasonal differences, until the best-fit model was found. The correlograms of the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) were built with 40 lags and 24 lags, respectively, to allow for better interpretation of the charts and observation of seasonal components.

After identifying and estimating the model, we assessed whether it was appropriate to describe the behavior of the data. If the model was not appropriate, the cycle was repeated, returning to the identification phase. Evaluating the model order using Akaike Information Criteria (AIC), Schwarz’s Bayesian Information Criteria (SBC) and Hannan-Quinn Information Criteria (HQC) carried out this step. The models were also evaluated by residual analysis using four tests: the function of residual autocorrelation; the Ljung-Box Q-statistics test; analysis of normality of residuals using the Jarque-Bera test; and residual variance.

The model that showed a normal residual distribution, nonsignificant autocorrelation of residuals, lowest value in the evaluation of the model order, and lowest residual variance was identified as the best model for making predictions.

Predicted values were compared with data on curimatã landings published for 2003 and 2004 by Ruffino (2006) and Thomé-Souza (2007), respectively. Initially, we graphically compared the predicted values with the observed, processed, and unprocessed data. We then used metric tools (mean square error prediction, MQE; root mean square error, RMSE; mean absolute error, MAE; absolute percentage error, APE; coefficient U-Theil; mean absolute percentage error, MAPE; and median absolute percentage error, MdAPE) to determine the errors between the observed and the predicted values in each of the comparisons (Souza and Camargo, 2004). These seven metric tools were used in evaluating the performance of the prediction models with horizons of 12 and 24 months.

3. Results

The data set achieved the best results in tests of normality after transformation with a logarithm and a seasonal difference (Lilliefors test = 0.06, p = 0.46, n = 132). In addition to making the series stationary (Table 1), the seasonal difference transformation removed the seasonality presented in the ACF correlogram of the log-transformed series, making the correlogram of the series suitable for interpretation.

After the difference of the transformed series was determined, the correlogram of ACF showed a correlation in the first lags, with a slow decrease to gap 6. This result is indicative of the absence of moving average terms at this moment. However, PACF fell sharply after the first lag, indicating terms of the moving average AR (1). ACF and PACF also showed high correlations in the seasonal lag 12, which indicates the presence of a moving average and seasonal autoregressive (Figure 2).

We define two models from the identification made in the correlograms of hyperparametrization and seasonal difference. Lower values of AIC, SBS, and HQC were observed for model (1,0,0) (1,1,1), followed by model (1,0,0) (2,1,0). However, model (1,0,0) (1,1,1) showed no significant terms in its composition (Figure 3). The residuals of all models were normally distributed.

The models are described by the following equations 1 and 2:

Model (1,0,0) (1,1,1):

$$Z_t = (1 - \Theta B^{12}) \varepsilon_t$$

$$Z_t = (1 - \Theta B^{12}) \varepsilon_t$$

Model (1,0,0) (2,1,0):

$$Z_t = (1 - \Theta B^{12} - \Phi B^{24}) (1 - B^{12}) \varepsilon_t$$

None of the models had significant residual correlations. Model (1,0,0) (1,1,1) was accepted due to its lower AIC value, even though one of its terms is not significant. On the other hand, model (1,0,0) (2,1,0) was included in the analysis because it has significant terms and a low AIC value.

The models showed a reasonable fit to the data (Figure 4), with the occurrence of three peaks for 2003 and 2004. The two models overestimated the fish landings compared to the observed values. The actual values for 2003 and 2004 are lower than the predicted values for the two models. This difference is also indicated by high values of MAPE and MdAPE. Although model (1,0,0) (1,1,1) had the lowest values of AIC, SBC, and HQC, this model also presented...
the highest value in all of the validation tests using the tools APE, MAPE, MdAPE, MQE, RSME, MAE, and U-Theil (Table 2). As a result, model (1,0,0) (2,1,0) was considered the model that best fits the data series.

4. Discussion

The city of Santarém is 22,887.08 km² and has a population of 262,538 inhabitants, approximately 8,000 of whom are fishermen (Almeida et al., 2003). The production of fish landed annually oscillates between 3,000 and 5,000 tons, mainly as a result of fisheries in the region of Lago Grande de Monte Alegre, which is one of the most productive regions in the Amazon (Almeida et al., 2003). Santarém is the third largest port of landing in the Brazilian Amazon, second only to Manaus and Belém (Ruffino, 2006), and has a fleet of approximately 150 boats (Almeida et al., 2003). Although approximately 100 different species of fish are sold here, only 10 species account for more than 80% of the sales (Ruffino, 2006; Thomé-Souza, 2007). Landings of curimatã (Prochilodus nigricans) ranged between 100 and 800 tons/year from 1992 and 2002, making it one of the primary species landed.

Despite the considerable empirical knowledge of fishermen, a high level of stochasticity characterizes data on catches of artisanal fisheries. Stergiou et al. (1997) used ARIMA models to analyze catches of anchovies and Scomber spp, with MAPEs of 26.5% and 33.3%, for the respective species. According to Stergiou (1989), estimates of MAPE approximately 14% are considered acceptable, taking into account the stochasticity of fisheries abundance data resulting from biological and environmental factors.

Figure 2. Residual autocorrelations and arithmetic coefficients of the autoregressive integrated moving average (ARIMA) models fitted to the monthly log-transformed data of landings of curimatã (Prochilodus nigricans): \( \phi \) = autoregressive parameter; \( \theta \) = moving average parameter; \( \sigma^2 \) = Chi-square; AIC = Akaike Information Criterion; SBC = Schwarz-Bayesian Information Criterion; HQC = Hannan-Quinn Information Criterion; \( \delta^2 \) = residual variance. Significant parameter values (\( P < 0.05 \)) are marked with an asterisk and A = model (1,0,0)(1,1,0), B = model (1,0,0)(1,1,1) and C = model (1,0,0)(2,1,0).
A time series analysis of Prochilodus nigricans landings

It is possible that the large differences between predicted and observed values (Figure 4) are a consequence of the characteristics of small-scale fisheries.

The small differences between predicted and observed values for Prochilodus nigricans landings in 2003 contributed to the low values of MAPE and MDAPE for this year, indicating that we obtained better predictions for 2003 than for 2004.

An increase in landings was observed in the month of June for the years 1993, 1996, 1999, and 2001, coinciding with periods of high water. In 1993, Mota and Ruffino (1997) argued that the increase in landings was due to the landings of curimatã from four boats originating from Manaus. This process may have also occurred in other years because the boats that usually land in Manaus can capture fish in other cities where the supply is high. Alternatively, the increase in catches of curimatã in these years may be related to cyclical environmental phenomena.

The model presented here also has an interesting biological interpretation. This model predicts the persistence of catches (Stergiou et al., 1997). In other words, when catches are high, they tend to remain high for 2–3 months or years in a row. This persistence may be indicative of conditions that are favorable for the formation of good age classes and large shoals and/or may be due to externalities.

The autoregressive term $t-12$ indicates a strong seasonal cycle in the frequency of catches. In fact, there is an increase in catches during the months of September and October for the Prochilodus nigricans. This pattern was observed.
Table 2. Validation tests for comparison of models: (MAPE) mean absolute percentage error; (MdAPE) median absolute percentage error; (U-Theil) Theil’s U-statistics; (MSE) mean squared error; (RMSE) root mean squared error; (MAE) mean absolute error.

<table>
<thead>
<tr>
<th>Month/year</th>
<th>(1,0,0)(1,1,1)</th>
<th>(1,0,0)(2,1,0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE (2003)</td>
<td>1.12</td>
<td>1.00</td>
</tr>
<tr>
<td>U-Theil (2003)</td>
<td>0.34</td>
<td>0.32</td>
</tr>
<tr>
<td>MSE (2003)</td>
<td>1.52</td>
<td>1.35</td>
</tr>
<tr>
<td>RSME (2003)</td>
<td>1.23</td>
<td>1.16</td>
</tr>
<tr>
<td>MAPE (2003)</td>
<td>13.40</td>
<td>11.78</td>
</tr>
<tr>
<td>MdAPE (2003)</td>
<td>14.14</td>
<td>10.03</td>
</tr>
<tr>
<td>MAE (2004)</td>
<td>1.85</td>
<td>1.83</td>
</tr>
<tr>
<td>U-Theil (2004)</td>
<td>0.24</td>
<td>0.24</td>
</tr>
<tr>
<td>RSME (2004)</td>
<td>2.00</td>
<td>2.01</td>
</tr>
<tr>
<td>MAPE (2004)</td>
<td>25.52</td>
<td>24.81</td>
</tr>
<tr>
<td>MAE (total)</td>
<td>1.48</td>
<td>1.41</td>
</tr>
<tr>
<td>U-theil (total)</td>
<td>0.20</td>
<td>0.19</td>
</tr>
<tr>
<td>MSE (total)</td>
<td>2.86</td>
<td>2.69</td>
</tr>
<tr>
<td>RSME (total)</td>
<td>1.69</td>
<td>1.64</td>
</tr>
<tr>
<td>MAPE (total)</td>
<td>19.46</td>
<td>18.30</td>
</tr>
<tr>
<td>MdAPE (total)</td>
<td>18.51</td>
<td>17.78</td>
</tr>
</tbody>
</table>

by Petre Júnior (1978) and Goulding (1979), who found that because the curimatã is one of the first species to leave the flooded forest when the water level begins to decline, it is mainly caught in the beginning of the dry season.

A 2 to 3 year cycle was identified in the variability of different biotic (zooplankton, phytoplankton, fish) and abiotic (temperature, salinity, atmospheric pressure) factors in the ecosystems of the Black Sea and eastern Mediterranean (Dementieva, 1987; Petrova-Karadjova and Apostolov, 1988). Stergiou (1989) also noted the presence of autoregressive terms \((Z_{t-25}, Z_{t-26}, Z_{t-36})\) in his model and associated them with an interaction of climate, plankton, and anchovies in the eastern Mediterranean that involves a time lag of 2 to 3 years.

It is also possible to identify these cycles in model \((1,0,0)(1,1,1)\) for curimatã, which presents autoregressive terms in periods \(t-1, t-12, t-13, t-24, t-25, t-36,\) and \(t-37\); this result leads us to conclude that it is necessary to have data for a given month and the month preceding it for two years prior to the month for which you aim to predict landings. However, the most appropriate model, \((1,0,0)(2,1,0)\), uses terms \(t-1, t-12, t-13, t-24, t-25, t-36,\) and \(t-37\), which indicates that it is best to have data from the previous two to three years in order to predict the catch in a given month.

According to Agostinho and Júlio Junior (1999), another important factor for the results of fisheries is to have a long connection between the wetlands of the adjacent plain and the river canal in years with high and prolonged floods; this connection allows for the survival of the eggs and young of many migratory species in ponds that have a high availability of food and shelter, resulting in strong cuts. Thus, it is possible to say that after a period of prolonged flooding, there is a replacement of the stock previously impacted by severe drought.

The model \((1,0,0)(2,1,0)\) appears to be related to the age of fish. Autoregressive terms of the model reach 3 years \((t-1, t-12, t-13, t-24, t-25, t-36,\) and \(t-37\), indicating that the species is being caught around the age of 3. According to Oliveira (1997), a curimatã reaches the age of sexual maturity at approximately 2 years. This fact may indicate that the curimatãs that landed in Santarém from 1992 to 2002 were more than two years of age, indicating a lack of fishing pressure on the lower age groups. The lack of fishing pressure on lower age groups may be a consequence of the regulatory measures placed on the minimum size of the fish caught. However, it is not possible to confirm this hypothesis due to the need to perform an analysis relating the length of the fish within the period studied to the autoregressive terms of the model. The lack of fishing pressure on the immature curimatã is not enough to ensure sustainable exploitation. As this species has been heavily exploited for at least four decades, the evolutionary effects of fishing, which mainly occurs by size selection, can be dangerous for the sustainability of the curimatã stocks (Conover and Munch, 2002).

Acknowledgements – The authors would like to thank CNPq (Process 301384/2008-8) and the PIATAM Institute for the infrastructure. The authors would also like to thank Claudemir Oliveira da Silva and Mauro L. Rufino for the authorization to use the data series of curimatã catches.

References


Santana, IF. and Freitas, CEC.
A time series analysis of Prochilodus nigricans landings


