BOLETIM DO INSTITUTO DE PESCA
ISSN 1678-2305 online version
Scientific Article

Length structure analysis of the Atlantic thread herring (*Opisthonema oglinum*) captured in the Santa Cruz Channel (Pernambuco, Brazil), using generalized additive models for location, scale and shape

Nathielly Lima do Rêgo[1]
Humber Agrelli de Andrade[2]
Paulo José Duarte-Neto[3]*

[1]Universidade Federal Rural de Pernambuco, Programa de Pós-Graduação em Biometria e Estatística Aplicada – Recife (PE), Brazil.

[2]Universidade Federal Rural de Pernambuco, Departamento de Pesca e Aquicultura, Laboratório de Modelagem Estatística Aplicada – Recife (PE), Brazil.

[3]Universidade Federal Rural de Pernambuco, Departamento de Estatística e Informática – Recife (PE), Brazil.

*Corresponding author: Paulo José Duarte-Neto, Rua Dom Manoel de Medeiros, s/n, Dois Irmãos, CEP: 52171-900 – Recife (PE), Brazil. E-mail: paulo.duartent@ufrpe.br.

Received: February 5, 2022
Approved: August 16, 2022

ABSTRACT

The Atlantic thread herring (*Opisthonema oglinum*) is the most common sardine in Santa Cruz Channel landings, which is located on the Northern coast of Pernambuco state, in Brazil. The species spends part of its life cycle at sea and enters estuaries at strategic times. Its high exploitation and importance to the economy in the municipalities surrounding the channel led us to reflect on the need for investigations to promote sustainable fishing and a better understanding of its behavior in that place. In this study, the standard length of the specimens was estimated using generalized additive models for location, scale, and shape, and the factors that drive the length pattern along the estuary were discussed. It was identified the presence of larger individuals in the rainy season and greater concentration in the Center-South region of the channel. The smaller sardines were concentrated near the outflow of the Botafogo River, an area of secondary channels with greater mangrove coverage. Few adults enter the channel and the vast majority of specimens captured were young. Thus, it would be pertinent to explore specimens that have at least reached sexual maturity, since the capture of young individuals can cause an imbalance of species sustainability.

Keywords: Atlantic thread herring; Santa Cruz Channel; GAMLSS; sustainable fishery; Northeastern Brazil.

Análise da estrutura do comprimento da manjuba (*Opisthonema oglinum*) capturada no canal de Santa Cruz (Pernambuco, Brasil) com base nos modelos aditivos generalizados para locação, escala e forma

RESUMO

A manjuba (*Opisthonema oglinum*) é a sardinha mais frequente nos desembarques do Canal de Santa Cruz, localizado no litoral norte de Pernambuco, no Brasil. A espécie vive parte de seu ciclo de vida no mar e penetra nos estuários em períodos estratégicos. Sua elevada exploração e relevância para a economia dos municípios no entorno provocaram-nos reflexões acerca da necessidade de investigações que promovam a pesca sustentável e que ampliem a compreensão a respeito da manjuba naquele local. A investigação foi realizada por meio de estimativas do comprimento padrão, com a utilização de modelos aditivos generalizados para locação, escala e forma, e os fatores que impulsionam o padrão de comprimento ao longo do estuário foram discutidos. Identificou-se a presença de indivíduos maiores no período chuvoso e maior concentração na região centro-sul do canal. Já os de menor tamanho concentraram-se próximo à desembocadura do rio Botafogo, área de canais secundários e com maior cobertura de manguezal. Poucos adultos ingressam no canal, e a grande maioria dos exemplares capturados era jovem. Assim, seria pertinente a exploração de exemplares que já completaram ao menos a maturidade sexual, uma vez que a captura de indivíduos jovens pode provocar desequilíbrio na sustentabilidade da espécie.

Palavras-chave: manjuba; Canal de Santa Cruz; GAMLSS; pesca sustentável; Nordeste do Brasil.

INTRODUCTION

Thread herring (*Opisthonema oglinum*), also known as manjuba, is a species of small pelagic clupeid that inhabits the upper layers of coastal and estuarine waters (Whitehead, 1985). This species is distributed along the entire American Atlantic coast, from New England to Argentina (Figueiredo and Menezes, 1978). In Brazil, *O. oglinum* is considered the main alternative to maintain the supply of raw material for the industrial fishing sector, in view of the decline in catches of the Brazilian sardine (*Sardinella brasiliensis*) whose exploitation collapsed due to overfishing in Southeast Brazil (Nóbrega et al., 2009).

The capture of *O. oglinum* in some Northeastern municipalities is carried out through artisanal fishing, which is one of the most traditional techniques having high socioeconomic prominence in the region (Lessa et al., 2004). In Pernambuco, more precisely in the Santa Cruz Channel (SCC), the *O. oglinum* was the most abundant clupeid in 2006, where it totaled 49.4% of all fish caught in the municipality of Itapissuma (Andrade and Silva, 2013), and it is most frequently captured with meshes of 30 mm (between opposite nodes), known as *redinha*. The high demand for the capture of thread herring in the coastal municipalities of the state can cause the degradation of the resource. Lima (2015) pointed out reports from fishermen in the region mentioning that the minimum and maximum lengths of *O. oglinum* in the catches carried out in the SCC have been decreasing over the years. Decreases in average lengths are usually associated with fishing mortality (Beverton and Holt, 1957), and can be used as indicators of overfishing.

O. oglinum is a small species, similar in size to common sardines. The maximum length found was 380 mm in the USA (Carpenter, 2002), 310 mm in Northeast Brazil (Lessa et al., 2004), and 220 mm in Itapissuma (Pernambuco, Brazil) (Lino, 2013). Lino (2013) noted that 92.7% of the specimens captured in the Itapissuma region were below the average length at first maturity $(L_{pm} = 117mm)$. The capture of *O. oglinum* below the average selection length should be investigated for efficient fisheries management since, when poorly managed, economically attractive fishery resources of social importance tend to collapse (Lima, 2015).

One way to generate information for professionals in the area of fisheries resource management is to use estimates of the average length of *O. oglinum* in the SCC. This would make it possible to verify the length structure of the species, and thus, raise hypotheses about ecological processes that are relevant to explain the length variation along the channel based on model adjustments. However, modeling the length structure of fish is quite a difficult task, since this structure reflects changes in environmental conditions, interannual variability in the production of length classes, and successful recruitments. Therefore, the use of methods that allow for models with flexible assumptions to specify the distribution of the response variable would present a good response. To this end, Rigby and Stasinopoulos (2005) introduced the Generalized Additive Models for Location, Scale, and Shape (GAMLSS). GAMLSS provides the adjustment of models that admit numerous distributions for the response variable without necessarily belonging to the exponential family. Such models are capable of integrating parametric and/or non-parametric functions, for modeling any parameter of the assumed distribution.

Taking this into consideration, we proposed to build models based on this methodology to obtain reasonable estimates on the SL of *O. oglinum* as a function of the variables month, year, SCC sector, tide at the time of net launch, and average precipitation accumulated in 30 days. Based on the model results, we discussed some hypotheses that could help in the decision-making process and implementation of rules to keep the resource at reasonable levels for sustainable fishing activity in the SCC. In addition to these estimates, we showed that GAMLSS is a potential statistical tool to efficiently assess the length structure of fishes.

MATERIAL AND METHODS

Santa Cruz Channel has an area of approximately 824 km^2 with its channel extending for 22 km (Moura, 2009). It receives discharges from several rivers, of which Botafogo and Igarassu are the ones that contribute most in terms of water volume. The area is characterized by a pseudo-tropical, hot and humid climate (Griffiths, 1996). In the SCC it rains throughout the year, with annual totals above 1,500 mm, with a rainy season from March to August and a dry period from September to February. (Santos et al., 2009).

According to Vasconcelos-Filho and Oliveira (1999), during the rainy season, a large amount of fresh water enters the channel modifying salinity, temperature, transparency, dissolved oxygen, and pH (downwards), and raising the values of nutrient salts. According to Russell (1967), such characteristics lead to the development of a special type of ecosystem in the SCC, the so-called mangroves. The area where mangroves predominate is in the Northern (mouth of the river Botafogo) and the Central portions of the channel, where there is a higher concentration of sandy-muddy and muddy-sandy substrate (Paiva et al., 2008a; Silva, 2008).

The database used in this research came from the merger of two databases. The first gathers information on rainfall and is made available by Pernambuco Water and Climate Agency (APAC). The second is related to fishing monitoring, carried out through the collection of information on the dynamics of the fishing operation conducted on weekly visits to the SCC between the years 2013 and 2015, in which, the fishermen were interviewed upon landing, and samples of the catches were obtained using a gill net (30 mm mesh between opposite nodes). In order to help identify the approximate positions of the fishing operation and to record in the questionnaires, the SCC was subdivided into sectors from the Northern outlet (sector A) to the Southern outlet (sector T) (Figure 1). Samples (n = 10,394) were collected and then transported to the Applied Statistical Modeling Laboratory (ASML) at the Department of Fisheries and Aquaculture of the Federal Rural University of Pernambuco (UFRPE) in the city of Recife, where they were frozen and later identified according to taxonomic keys, such as those published by Carpenter (2002). This research was registered at the Biodiversity Authorization and Information System (SISBIO) under n. 4944.

The variables addressed in the research were:

- standard length (SL), a continuous variable characterized by the distance from the tip of the snout to the base of the caudal fin (Figure 2);
- month, a qualitative variable that identifies the month in which the data were collected (January to December);
- year, a qualitative variable indicating the year in which the data were collected (2013 to 2015); sector, a qualitative variable referring to the sectorization of the SCC, from the North outlet (sector A) to the South outlet (sector T);
- casting tide, a qualitative variable referring to the type of tide in the casting of the net at the time of catching the fish.

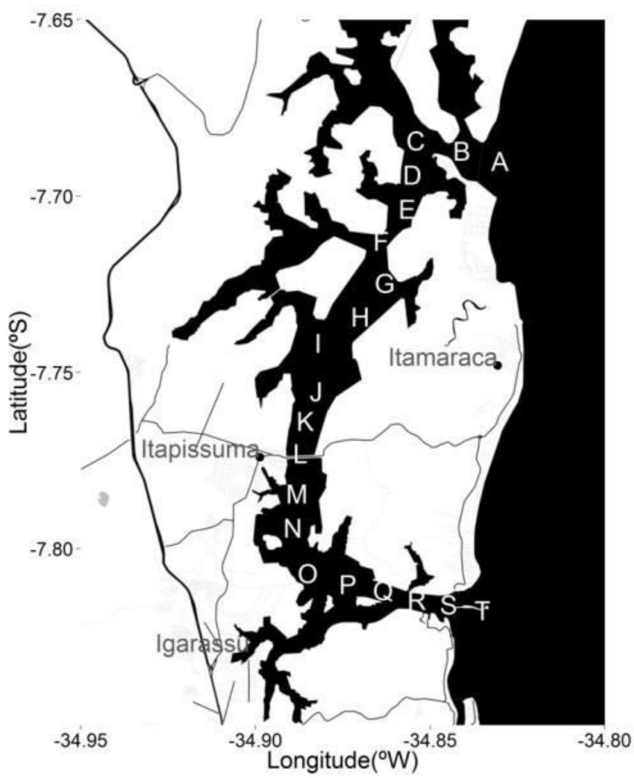

Figure 1. Santa Cruz Channel, indicating the sectors used for analysis.

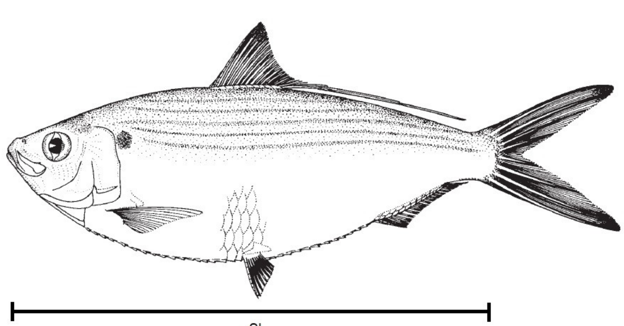

Figure 2. Specimen and standard length of *Opisthonema oglinum*.

The levels of the launching tide factor were as follows:

- low tide, the minimum level reached by the waters after the ebb;
- flood, the period between a low tide and a high tide;
- high tide, the maximum level reached by the waters of a high tide;
- ebb tide, the period between a high tide and a low tide when water height decreases;
- precipitation, a continuous variable represented by the average volume of rainfall accumulated in 30 days before the collection of the samples at Itapissuma, Igarassu, and Itamaracá stations.

Generalized additive model for location, scale, and shape

Rigby and Stasinopoulos (2005) proposed the GAMLSS class of semiparametric regression models with the advantage of modeling the asymmetry and kurtosis parameters, which allowed for an expansion of the number of distributions considered, once the distributions do not need to be part of the exponential family.

As the systematic part of the GAMLSS models is expanded to allow the modeling not only of the mean but also of the parameters of scale and shape of the distribution of the response variable, such parameters are modeled as a function of the explanatory variables. For the use of GAMLSS, it is imperative that the observations of the response variable Y are independent of each other. As the assumption that the response variable must belong to the exponential family in GAMLSS models is relaxed, then it was replaced by a more general family of distributions, given by Equation 1:

$$\boldsymbol{Y} \sim D(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\tau}) \tag{1}$$

$$\begin{aligned}
\boldsymbol{\mu} &= g_1(\boldsymbol{\mu}) = X_1\boldsymbol{\beta}_1 + f_{11}(x_{11}) + \cdots + f_{1J1}(x_{1J1}) \\
\boldsymbol{\sigma} &= g_2(\boldsymbol{\sigma}) = X_2\boldsymbol{\beta}_2 + f_{21}(x_{21}) + \cdots + f_{2J2}(x_{2J2}) \\
\boldsymbol{\nu} &= g_3(\boldsymbol{\nu}) = X_3\boldsymbol{\beta}_3 + f_{31}(x_{31}) + \cdots + f_{3J3}(x_{3J3}) \\
\boldsymbol{\tau} &= g_4(\boldsymbol{\tau}) = X_4\boldsymbol{\beta}_4 + f_{41}(x_{41}) + \cdots + f_{4J4}(x_{4J4})
\end{aligned}$$

In which:
$D(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\tau})$ = a distribution of four parameters (any distribution, including continuous distributions, with pronounced skewness or kurtosis, and discrete distributions);
$\boldsymbol{\mu}$ = a location parameter;
$\boldsymbol{\sigma}$ = a scale parameter;
$\boldsymbol{\nu}$ and $\boldsymbol{\tau}$ = shape parameters of the distribution, commonly associated with skewness and kurtosis, respectively.
The $\boldsymbol{X_1}, \boldsymbol{X_2}, \boldsymbol{X_3}$ *e* $\boldsymbol{X_4}$ are the matrices that may or may not coincide, or rather, the predictor of each parameter of the distribution receives different explanatory variables (Rigby and Stasinopoulos, 2005). The β_p are linear parameters and $f_{pj}(x_{pj})$ represent different smoothing functions for different explanatory variables x_{pj} for $p = 1,2,3,4$ and $j = 1,2,\ldots,Jp$.

Estimation

According to Stasinopoulos et al. (2017), the parametric GAMLSS models are fitted with smoothing functions, so they only require the estimation of $\boldsymbol{\beta}$. In this case, the model is given by Equation 2:

$$\boldsymbol{Y} \sim D(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\tau}) \quad (2)$$

$$\begin{aligned} \boldsymbol{\mu} &= g_1(\boldsymbol{\mu}) = X_1\boldsymbol{\beta}_1 \\ \boldsymbol{\sigma} &= g_2(\boldsymbol{\sigma}) = X_2\boldsymbol{\beta}_2 \\ \boldsymbol{\nu} &= g_3(\boldsymbol{\nu}) = X_3\boldsymbol{\beta}_3 \\ \boldsymbol{\tau} &= g_4(\boldsymbol{\tau}) = X_4\boldsymbol{\beta}_4 \end{aligned}$$

In which:
the models = fitted by the maximum likelihood estimates in relation to the estimation of $\boldsymbol{\beta}$.

We have $\boldsymbol{Y} \sim D(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\tau})$ which implies that the logarithm of the likelihood function determined by the observed likelihood of the sample is (Equation 3):

$$L(\boldsymbol{\beta}) = \sum_{i=1}^{n} ln[f(y_i|\mu_i, \sigma_i, \nu_i, \tau_i)]\,. \quad (3)$$

Model selection criteria

An important measure for the selection of GAMLSS models is the generalized Akaike criterion (GAIC). Defined by Voudouris et al. (2012), the GAIC takes into account the number of parameters and degrees of freedom used in the model to penalize the more complex models and avoid data overfitting in large samples (Paiva et al., 2008b). The GAIC is expressed by Equation 4:

$$GAIC(k) = -2L(\hat{\beta}\,) + (k \times\ gl), \quad (4)$$

In which:
L = the logarithm of the likelihood function;
gl = the effective degrees of freedom of the fitted model;
k = constant and is the penalty for each degree of freedom used.

GAIC(k) penalizes models with many parameters, the lower the value of GAIC(k), the better the model (Stasinopoulos et al., 2017).

Residual analysis

Dunn and Smyth (1996) proposed normalized (randomized) quantile residuals. Such residuals provide us with a familiar way of checking the adequacy of a model. According to the authors, the normalized random quantile residual is defined by Equation 5:

$$r_i = \Phi^{-1}\{F(y_i; \hat{\beta})\}, \quad (5)$$

In which:
Φ^{-1} = the inverse of the cumulative distribution function of a standard normal;
$F(.)$ = the cumulative distribution function suitable for the data;
$\hat{\beta}$ = the parameter vector.

The model is said to be adequate when its residuals r_i follow the standard normal distribution.

The worm plot is one of the most relevant techniques applied to analyze the residuals of a GAMLSS. It exerts a function in the same direction as the normal quantile-quantile (QQ) graph, without trend, and aims to show local changes in the domain of a given covariate. An important property of the worm plot is the establishment of intervals with $(1-\alpha)$ confidence for the theoretical normal quantiles. If the points lie within the 95% confidence band (between the two elliptic curves), the model fit is considered satisfactory.

Homoscedasticity test: Breusch-Pagan

Homoscedasticity is the term designating constant variance of experimental errors. Based on the Lagrange multiplier test, the Breusch-Pagan test is widely used to test the null hypothesis that the error variances are equal (homoscedasticity) versus the alternative hypothesis that there is at least one different variance. Then, the test statistic is given by Equation 6:

$$BP = \frac{SQReg}{2} \quad (6)$$

Assuming that r_i is normally distributed, if there is homoscedasticity, and if the sample size n increases indefinitely, then $BP \sim \chi^2_{n-1}$. If the test statistic is greater than the critical value of χ^2 at a pre-set significance level, the hypothesis of homoscedasticity is rejected (Oliveira, 2018).

Normality Test: Shapiro-Wilk

Normality tests are used to verify whether the probability distribution associated with a data set can be approximated by the normal distribution. To support this assumption, we considered the Shapiro-Wilk test (1965), which can be used to evaluate hypotheses H_0: The data follow a normal distribution × H_1: The data do not follow a normal distribution.

Given an ordered random sample, $y_1 < y_2 < \ldots < y_n$ the original Shapiro-Wilk test statistic is defined by Equation 7:

$$W = \frac{\left(\sum_{i=1}^{n} a_i y_i\right)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}, \quad (7)$$

In which:
$a_i = \left(m'V^{-1} / m'V^{-1}V^{-1}m\right)^{\frac{1}{2}}$;
$m = (m_1, \ldots, m_n)'$ = the expected values of the statistics of independently and identically distributed random variables;
V= the matrix covariance of these statistics (Razali and Wah, 2011).

If W is greater than the critical value, we reject the hypothesis of normality of the data with $(1-\alpha)100\%$ confidence, otherwise, we do not reject it (Razali and Wah, 2011).

RESULTS

In the captures carried out in the SCC for the research, specimens of *O. oglinum* were found with a SL between 30.14 mm and 193 mm (Figure 3), with a median of 97.24 mm and an average of 99.54 mm, and with half of the specimens between 91.19 mm (1st quartile) and 105.10 mm (3rd quartile).

There was no very relevant distinction between the observed SL and the levels of the month factor (Figure 4, Month); however, the smaller fish were captured in the months of January, March, and August. The year factor (Figure 4, Year) was added to the model as it allows for the evaluation of long-term effects on the response variable. A small increase in the median SL from 2013 to 2014 and a slight decline from 2014 to 2015 was realized. Observations were lacking in sectors H and L (Figure 4, Sector), however, since area L is located by the Getúlio Vargas bridge, specimens could not be collected. Smaller fish were captured in sectors E, F and G. Additionally, sectors O and T showed higher variability in SL values. Regarding net casting tide (Figure 4) no very relevant distinction was recorded between the length patterns observed and the tide at the time of net casting. Besides, there was an approximately constant behavior between SL and precipitation (Figure 5).

Modeling

The selection process of the GAMLSS model was performed by comparing different distributions attributed to the response variable, in which several competing models with different combinations of components were used. The Box-Cox *t* original distribution with $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ in the parametric configuration was considered the best candidate to fit the model with structure given by Equation 8:

$$\begin{aligned} \log(\hat{\mu}) &= Intercept + Month + Sector + Year + Tide + Precipitation + Month \times Year \\ \log(\hat{\sigma}) &= Intercept + Month + Sector + Year + Tide + Precipitation + Month \times Year \\ \hat{\nu} &= Intercept \\ \log(\hat{\tau}) &= Intercept \end{aligned} \qquad (8)$$

In the adjusted model, the variables month, sector, year, tide, precipitation and the interaction (month × year) were relevant in the modeling of the parameters μ (location parameter) and σ (scale parameter). However, this behavior is not repeated in the modeling of the parameter ν and τ (shape parameters). The addition of explanatory variables for the modeling of such parameters caused an increase in the GAIC value and caused convergence problems to the model. Thus, ν and τ, configured as constant, proved to be suitable for modeling. The binding functions used to model the parameters were logarithmic and identity. In the adjustment, only statistically significant estimates were presented, with $p < 0.05$. We assumed the reference cell parameterization, where this restriction considers the first level of each null factor, and then, the effects of the estimable factors are presented (Tables 1 and 2).

The estimated sector F and G coefficients were negative, so the average SL of specimens collected near the mouth of the Botafogo River were smaller compared to sector E. Whereas positive signs, in the other sectors, indicate that the largest fish were found in the central part and southern mouth of the SCC. The sign of tidal estimate coefficients at the moment of casting the net when in high and ebb tide indicated that the SL of *O. oglinum* was greater at those levels than at the reference value. The positive sign precipitation coefficient indicates that there was a significant increase in the average SL of *O. oglinum* in the wettest seasons. In the estimates

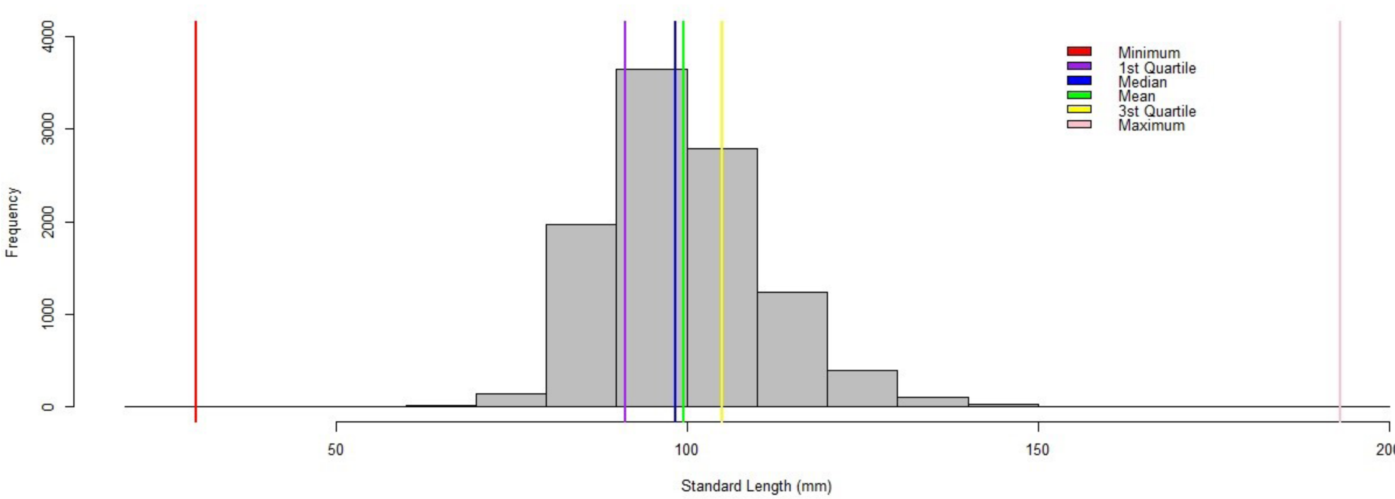

Figure 3. Frequency distribution of standard length of the *Opisthonema oglinum*.

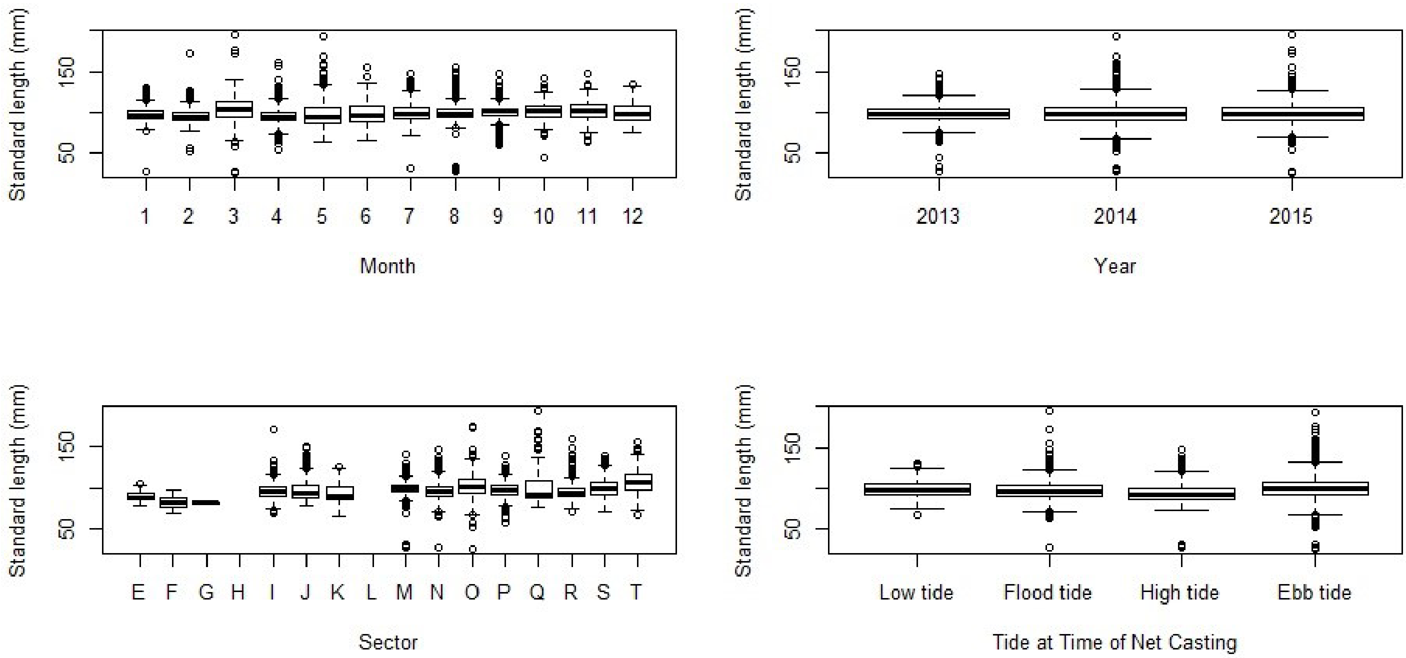

Figure 4. Boxplots of the standard length of *Opisthonema oglinum* according to month, year, sectorization and tide at the time of net casting during capture in the Santa Cruz Channel.

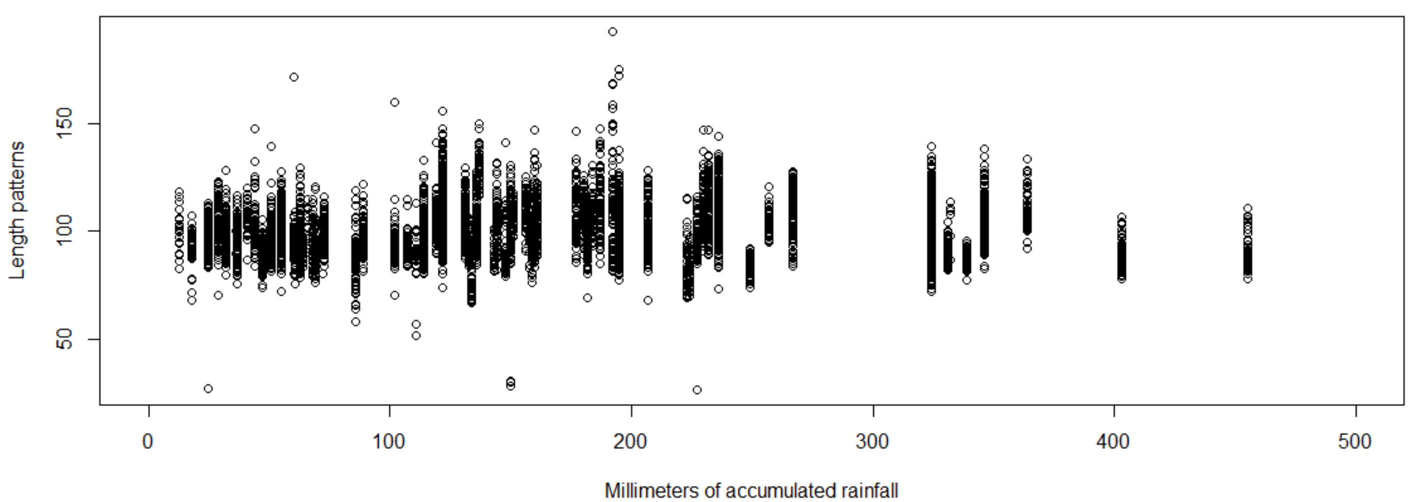

Figure 5. Scatter diagram of the standard length of *Opisthonema oglinum* according to the average accumulated rainfall volume of thirty days before specimen collection, at Itapissuma, Igarassu and Itamaracá stations.

of the interaction between month and year, February, April, May, June, July, September, October and December, separately, presented a negative contribution and particularly in 2014, the negative perspectives did not remain. Likewise, the month of March had a positive estimate while in 2014 it had a negative estimate. Similarly to the previous interactions, the months of June and October, and the year 2015 showed negative coefficients separately; however, in the interaction between these levels the perspective was not confirmed.

The estimated coefficient of sector G was negative, so we may say that the variability in the SL of the manjuba was smaller than in area E. Conversely, for the positive sign of the estimated coefficient of sectors J and N, we may say that the variability was greater than in area E. The positive coefficients estimated for the high and ebb

Table 1. Estimated coefficients for μ by fitting the original Box-cox t original model to the standard length data of *Ophistonema oglinum*. Only statistically significant (p < 0.05) estimates are shown.

Parameter - μ	Estimate	Standard Error	t-value	p-value
Intercept	4.733	0.0247	191.607	0.0000
February	-0.093	0.0073	-12.795	0.0000
March	0.045	0.0079	5.730	0.0000
April	-0.293	0.0118	-24.738	0.0000
May	-0.118	0.0148	-7.989	0.0000
June	-0.301	0.0113	-26.596	0.0000
July	-0.228	0.0130	-17.540	0.0000
September	-0.050	0.0079	-6.307	0.0000
October	-0.221	0.0137	-16.122	0.0000
November	0.128	0.0128	9.973	0.0000
December	-0.111	0.0241	-4.616	0.0000
Sector F	-0.222	0.0030	-7.200	0.0000
Sector G	-0.129	0.0221	-5.855	0.0000
Sector I	0.073	0.0215	3.396	0.0006
Sector J	0.040	0.0209	1.933	0.0532
Sector K	0.098	0.0216	4.527	0.0000
Sector M	0.063	0.0212	-2.992	0.0027
Sector S	0.055	0.0218	2.534	0.0112
2014	-0.262	0.0150	-17.489	0.0000
2015	-0.155	0.0089	-17.417	0.0000
Flood tide	-0.013	0.0075	-1.813	0.0698
High tide	0.048	0.0099	4.864	0.0000
Ebb tide	0.025	0.0075	3.421	0.0006
Precipitation	0.007	0.000009	7.310	0.0000
February:2014	0.090	0.0149	6.088	0.0000
March:2014	-0.144	0.0191	-7.509	0.0000
April:2014	0.352	0.0160	21.984	0.0000
May:2014	0.295	0.0185	15.946	0.0000
June:2014	0.429	0.0165	25.957	0.0000
July:2014	0.401	0.0170	23.543	0.0000
August:2014	0.252	0.0137	18.275	0.0000
September:2014	0.222	0.0142	15.610	0.0000
October:2014	0.289	0.0182	15.867	0.0000
December:2014	0.122	0.0262	4.655	0.0000
June:2015	0.433	0.0126	4.257	0.0000
October:2015	0.283	0.0219	12.897	0.0000

Table 2. Estimated coefficients for σ by fitting the original Box-cox t original model to the standard-length data of *Ophistonema oglinum*. Only statistically significant (p < 0.05) estimates are shown.

Parameters - σ	Estimate	Standard Error	t-value	p-value
Intercept	-3.169	0.241	-13.151	0.0000
March	0.886	0.069	12.662	0.0000
April	1.373	0.124	11.091	0.0000
May	0.517	0.143	3.613	0.0003
June	-0.252	0.127	-1.985	0.0471
July	0.637	0.136	4.688	0.0000
August	0.190	0.084	2.245	0.0247
October	0.621	0.135	4.593	0.0000
Sector G	-3.217	0.574	-5.604	0.0000
Sector J	0.362	0.198	1.825	0.0679
Sector N	0.410	0.197	2.076	0.0379
2014	0.443	0.157	2.810	0.0049
High tide	0.197	0.088	2.227	0.0259
Ebb tide	0.176	0.069	2.535	0.0112
March:2014	-0.655	0.164	-3.980	0.0000
April:2014	-1.484	0.166	-8.943	0.0000
July:2014	-0.712	0.176	-4.039	0.0000
August:2014	-0.272	0.137	-1.980	0.0477
October:2014	-0.664	0.180	-3.682	0.0002
December:2014	-0.804	0.258	-3.116	0.0018
June:2015	0.659	0.136	4.819	0.0000
July:2015	-0.855	0.151	-5.642	0.0000

tide at the time of net casting indicated that the dispersion in the SL of *O. oglinum* was lower when captured at low tide than at high and ebb. In the estimates of the coefficients of the interaction between month and year, we observed that the months of March, April, July, August and October, had a positive contribution, particularly in 2014, yet the positive perspectives did not remain, so the variability in manjuba SL was lower in these months. Similarly to the previous interactions, June showed negative coefficients and July showed positive, however, in the interaction with the year 2015, the perspectives were not confirmed.

Points were randomly distributed around zero on the residual plot *versus* adjusted values of parameter μ (Figure 6A). In the graph of residuals *versus* indexing the order of observations (Figure 6B), a completely random cloud of points was observed, a desirable result that suggests homoscedasticity of the residuals and could be confirmed using the Breusch-Pagan ($p<0.05$). In the residual kernel density estimation graph (Figure 6C), the residuals are normally distributed, whose hypothesis was

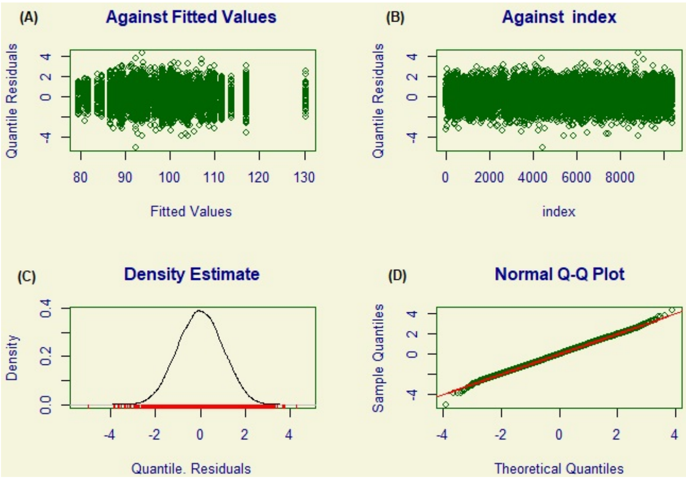

Figure 6. Graphical analysis of residuals of the fitted model with Box-Cox t original distribution for the standard length of Opisthonema Oglinum.

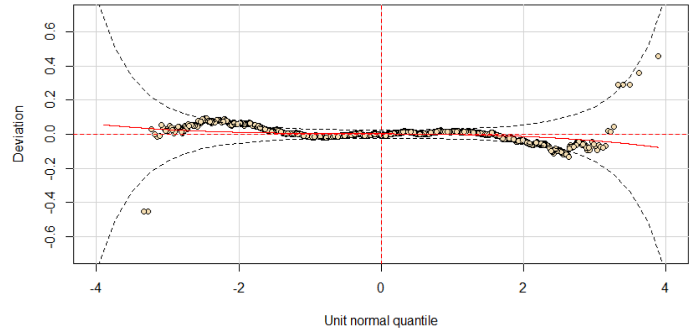

Figure 7. Worm plot of the fitted model with Box-Cox t original distribution for the standard length of *Opisthonema Oglinum*.

confirmed using the Shapiro-Wilk test ($p<0.05$). In the normal Q-Q plot of the residuals (Figure 6D), the points were found to be linearly distributed. In Figure 7 regarding worm plot, some points near the confidence tails drew attention, but did not correspond to 5% of the points that could be outside the 95% confidence interval, besides they did not show any pattern (linear, quadratic and/or cubic).

DISCUSSION

The first factor that seems to determine the *O. oglinum* length distribution in the SCC is the precipitation. Besides the significant positive relationship between SL and precipitation, there was a predominance of larger individuals in the months from April to December of 2014, a fact that must be related to the pluviometric regime. The rainy season in the region surrounding the SCC was recorded between March and August (Santos et al., 2009). As it is a period of high river discharges, it causes higher concentrations of primary biomass (Paiva et al., 2008a; Simoni, 2019; Lima et al., 2020), which makes the SCC a favorable place for some filtering species that have an important diet part on phyto and zooplankton. In addition, the life cycle must be synchronized with this regime of environmental variation during the rainy season, in which the younger individuals benefit from this moment of greater availability of food to grow.

Another determinant factor is the channel area. According to the observed results, we noticed that the SL of *O. oglinum* in the SCC was, on average, smaller between areas E and G, near the mouth of the Botafogo River. This fact may be related to the smaller width of the channel at the site and the larger marginal area with mangrove coverage. The vast area of secondary and shallow channels provides a wide region of shelter, favoring the use of the area as an initial nursery, housing large amounts of the younger and smaller forms. The hypothesis is that, as the individuals develop and grow, the preference for areas of wide channels becomes greater, which provide food more suitable for the slightly larger forms, despite the greater exposure to predators.

The average SL of manjuba was greater in the central area of the SCC, where the substrate is sandy-muddy and muddy-sandy specifically in this region (Paiva et al., 2008a; Silva, 2008), which allows for the growth of meiofauna constituents and macrofauna, with organisms such as copepods, nematodes, polychaetes and crustaceans (Paiva et al., 2008a; Santos et al., 2009). Thus, we can assume that *O. oglinum* makes use of the central portion of the channel because this area provides several food resources, already in a more advanced stage of development. It is also important to note that the longer lengths tend to be more concentrated beyond the Central portion, close to the Southern mouth of the estuary. Thus, the ecological hypothesis regarding the use of the channel space over time since the early stages are:

- very young forms, small specimens, more concentrated in areas with greater mangrove margins and narrower channels;
- slightly larger specimens already, more concentrated in the central portion of the channel;
- even larger forms, concentrated near the mouth of the channel, with easier access to the external part and adjacent marine region.

Evidence was found that few adults enter the channel, and the vast majority of specimens captured in the SCC were young, which is, therefore, the phase that most uses the estuarine area. Furthermore, as they develop, gain in age and size, the specimens are distributed closer and closer to the Southern mouth of the channel. This indicates the inner region as an important nursery and initial growth area in the sardine life cycle. According to Simoni (2019), in the marine environment of the Northern coast of Pernambuco, about 90% of the captured specimens had a total length above $L_{pm} = 114mm$, that is the size at which maturity is already observed. However, Lino (2013) found that 92.7%

of the specimens captured in the Itapissuma region (inner area of the estuary) were below what he considered as the average length at first maturity ($L_{pm} = 117mm$). Taking into account the findings and those results extracted from the literature mentioned above, there is a strengthening of the hypothesis that the young forms enter the inner area of the SCC estuary, which works as a nursery, and as they develop, they advance to the vicinity from the mouth. In turn, the pre-adult and adult fractions of the stock are found in the adjacent marine area. Given the size of the SCC, this nursery region is possibly the most important for the stock that is accessible off the coast of Pernambuco, and perhaps, it is also the most important nursery on the eastern margin of the Northeast.

Another aspect to be taken into account is the variation in length over the years. Lima (2015) indicates a trend that the lengths of sardines captured in the SCC would be decreasing over the years. If we compare the minimum (30.14 mm) and maximum (193 mm) SL observed in this study with those obtained in the past by Lino (2013) — minimum of 52.50 mm and maximum of 209 mm — there would be new evidence supporting the downward trend hypothesis. The comparison should be viewed with caution because it is not possible to compare the selectivities of the fishing gear used in the different works. However, it is still a warning since decreasing trends in the lengths captured are classically associated with an increase in the fishing mortality coefficient, and consequently in the total mortality coefficient (Beverton and Holt, 1957).

Finally, it would be appropriate to carry out more in-depth investigations to raise new hypotheses and, even the implementation of some management measures if necessary, such as limiting the capture at a sustainable level (Frid et al., 2003), establishing the minimum capture sizes (Stergiou et al., 2009), and changing mesh size (Ferro et al., 2008; Lima et al., 2020) in specific periods and locations of the SCC, to maintain the resource at reasonable levels for the survival of fishing activity.

CONCLUSION

Therefore, through the use of the GAMLSS model adjusted with the Box-Cox t original distribution, it was possible to verify that precipitation and channel location were important factors in determining the length structure of *O. oglinum* in the SCC. There was a predominance of larger individuals in the rainy season, and with greater concentration in the South-Central portion of the channel. Although capture of slightly larger individuals occurs, at the same time, there is an active frequency in the capture of very young *O. oglinum* in the Northern portion of the channel. Capturing such young individuals can lead to an imbalance in the sustainability of the species. Even though *O. oglinum* has high reproductive levels, this deserves caution. In view of this, further in-depth investigations are essential to raise new hypotheses and implement management measures, if necessary, to avoid possible overfishing of *O. oglinum.*

CONFLICT OF INTERESTS

Nothing to declare.

FINANCIAL SUPPORT

The data were collected with funds from the project “A Pesca de Peixes Pelágicos de Pequeno Porte no Canal de Santa Cruz em Pernambuco (PEPEL)”, financed by the CNPq’s Universal Call (no. 480928), ongoing between 2012 and 2015. The authors are grateful to CNPq (National Council for Scientific and Technological Development) for providing research grant to P. J. Duarte Neto (no 310320/2020-8).

AUTHOR’S CONTRIBUTIONS

Rêgo, N.: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Writing — original draft. Andrade, H.: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology Resources, Writing — review and editing, Supervision. Duarte-Neto, P.: Conceptualization, Formal Analysis, Investigation, Methodology, Writing — review and editing, Supervision.

REFERENCES

Andrade, H.A.; Silva, R.M.M. 2013. Dinâmica das frotas de pesca de emalhe e linha de mão de Itapissuma-PE no Canal de Santa Cruz. Boletim Técnico Científico CEPENE, 19(6): 83-91.

Beverton, R.J.H.; Holt, S.J. 1957. On the dynamics of exploited fish populations: Fishery Investigations. Springer, London, Series II, 11.

Carpenter, K.E. 2002. The living marine resources of the Western Central Atlantic. FAO Species Identification Guide for Fishery Purposes and American Society of Ichthyologists and Herpetologists Special Publication, Rome, 2(3): 601-1374.

Dunn, P. K.; Smyth, G. K. 1996. Randomized quantile residuals. Journal of Computational and Graphical Statistics, 5(3): 236-244. https://doi.org/10.2307/1390802

Ferro, R. S. T.; Ozbilgin, H.; Bren, M. 2008. The potential for optimizing yield from a haddock trawl fishery using seasonal changes in selectivity, population structure and fish condition. Fisheries Research, 94(2): 151-159. https://doi.org/10.1016/j.fishres.2008.08.018

Figueiredo, J. L.; Menezes, N. A. 1978. Manual de peixes marinhos do sudeste do Brasil. II. Teleostei. Universidade de São Paulo, São Paulo.

Frid, C.; Hammer, C.; Law, R.; Loeng, H.; Pawlak, J. F.; Reid, P. C.; Tasker, M. 2003. Environmental Status of the European Seas. ICES.

Griffiths, J. F. 1996. Applied climatology. Oxford University Press, New York.

Lessa, R. P.; Nóbrega, M. F.; Junior, J. L. B. 2004. Dinâmica de populações e avaliação de estoques dos recursos pesqueiros da região nordeste. Programa de Avaliação do Potencial Sustentável dos Recursos Vivos da Zona Econômica Exclusiva SCORE-NE. DIMAR, Department of Fisheries, Universidade Federal Rural de Pernambuco, Recife.

Lima, S. A. O. 2015. Sazonalidade e seletividade nas pescarias de emalhe de pequenos peixes pelágicos no Canal de Santa Cruz – PE. (Master's Dissertation in Fisheries and Aquaculture. Universidade Federal Rural de Pernambuco). Available at: <https://sucupira.capes.gov.br/sucupira/public/consultas/coleta/trabalhoConclusao/viewTrabalhoConclusao.jsf?popup=true&id_trabalho=2083595>. Accessed: Aug. 3, 2020.

Lima, S. A. O.; Andrade, H. A.; Sousa, R. G. C. 2020. Dinâmica da pesca de peixes pelágicos de pequeno porte no estuário do canal de Santa Cruz (Pernambuco – Brasil). Brazilian Journal of Development, 6(9): 73423-73435. https://doi.org/10.34117/bjdv6n9-703

Lino, M. A. S. 2013. Estudo biológico-pesqueiro da manjuba *Opisthonema oglinum* (Lesueur, 1818) na região de Itapissuma, Pernambuco. (Master's Dissertation in Fisheries and Aquaculture. Universidade Federal Rural de Pernambuco). Available at: <http://www.tede2.ufrpe.br:8080/tede2/handle/tede2/6405>. Accessed: Aug. 3, 2020.

Moura, R. T. 2009. Aspectos Gerais da Hidrobiologia do Litoral Norte de Pernambuco - Brasil. IBAMA, Brasília.

Nóbrega, M. F.; Lessa, R.; Santana, F. M. 2009. Peixes marinhos da região Nordeste do Brasil. Programa REVIZEE – Score Nordeste. Martins e Cordeiro, Fortaleza.

Oliveira, E. S. C. 2018. Modelos lineares e aditivos generalizados para posição, escala e forma na análise de estoques pesqueiros de lagostas (Panulirus spp) e de espadarte (Xiphias gladius). (PhD Thesis. Department of Statistics and Informatics, Universidade Federal Rural de Pernambuco). Available at: <http://www.ppgbea.ufrpe.br/sites/www.ppgbea.ufrpe.br/files/documentos/evelyn_souza_chagas_oliveira.pdf>. Accessed: Aug. 10, 2020.

Paiva, A. C. G.; Chaves, P. T. C.; Araújo, M. E. 2008. Estrutura e organização trófica da ictiofauna de águas rasas em um estuário tropical. Revista Brasileira de Zoologia, 25(4): 647-661. https://doi.org/10.1590/S0101-81752008000400010

Paiva, C. S. M.; Freire, D. M. C.; Cecatti, J. G. 2008. Modelos Aditivos Generalizados para Posição, Escala e Forma (GAMLSS) na Modelagem de Curvas de Referência. Revista Brasileira de Ciências da Saúde, 12(3): 289-310.

Razali, N. M.; Wah, Y.B. 2011. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. Journal of Statistical Modeling and Analytics, 2(1): 21-33.

Rigby, R. A.; Stasinopoulos, D. M. 2005. Generalized additive models for location, scale and shape. Journal of the Royal Statistical Society. Series C (Applied Statistics), 54(3): 507-554. https://doi.org/10.1111/j.1467-9876.2005.00510.x

Russell, R. J. 1967. Origins of estuaries. Estuaries, Washington, D.C.

Santos, T. G.; Gusmão, L. M. O.; Leitão-Neumann, S.; Cunha, A. G. 2009. Zooplâncton como indicador biológico da qualidade ambiental nos estuários dos rios Carrapicho e Botafogo, Itamaracá - PE. Revista Brasileira de Engenharia da Pesca, 4(1): 44-55. https://doi.org/10.18817/repesca.v4i1.129

Shapiro, S. S.; Wilk, M. B. 1965. An analysis of variance test for normality (complete samples). Biometrika, 52(3-4): 591-611. https://doi.org/10.2307/2333709

Silva, L. A. 2008. Estudos sedimentológico, morfológico e hidrodinâmico do Canal de Santa Cruz - PE, 2008. (PhD Thesis in Geoscience, Universidade Federal de Pernambuco). Available at: <https://repositorio.ufpe.br/handle/123456789/6266>. Accessed: Aug. 28, 2020.

Simoni, M. E. R. 2019. Dinâmica reprodutiva da sardinha-laje *Opisthonema oglinum*, Lesueur, 1818 capturada no litoral norte de Pernambuco, Brasil. (PhD Thesis, Postgraduate Program in Fisheries Resources and Aquaculture, Universidade Federal Rural de Pernambuco). Available at: <http://www.pgpa.ufrpe.br/sites/ww2.pgpa.ufrpe.br/files/documentos/tese_mariaesterribeiro_pesca_2019.pdf>. Accessed: Aug. 22, 2020.

Stasinopoulos, M. D.; Rigby, R. A; Heller, G. Z.; Voudouris, V.; Bastiani, F. 2017. Flexible Regression and Smoothing: Using GAMLSS in R. CRC Press.

Stergiou, K. I.; Moutopoulos, D. K.; Armenis, G. 2009. Perish legally and ecologically: the ineffectiveness of the minimum landing sizes in the Mediterranean Sea. Fisheries Management and Ecology, 16(5), 368-375. https://doi.org/10.1111/j.1365-2400.2009.00672.x

Vasconcelos-Filho, A. L.; Oliveira, A. M. E. 1999. Composição de ecologia da ictiofauna do Canal de Santa Cruz (Itamaracá – PE, Brasil). Tropical Oceanography, 27(1): 101-113. https://doi.org/10.5914/tropocean.v27i1.2775

Voudouris, V.; Gilchrist, R.; Rigby, R.; Sedgwick, J.; Stasinopoulos, D. 2012. Modelling skewness and kurtosis with the BCPE density in GAMLSS. Journal of Applied Statistics, 39(6): 1279-1293. https://doi.org/10.1080/02664763.2011.644530

Whitehead, P. J. P. 1985. Clupeoid fishes of the world (Suborder Clupeoidei). An annotated and illustrated catalogue of the herings, sardines, pilchards, sprats, anchovies, and wolf-herings. Part I: Chirocentridae, Clupeidae and Pristigasteridae. FAO Fish Synopsis, 7(2): 1-103.