Fisheries Research 102 (2010) 41–49

Contents lists available at ScienceDirect

Fisheries Research

journal homepage: www.elsevier.com/locate/fishres

Hook selectivity models assessment for black spot seabream. Classic and heuristic approaches

Ivone A. Czerwinski[a,*], Juan C. Gutiérrez-Estrada[b], Mila Casimiro-Soriguer-Escofet[a], José A. Hernando[a]

[a] *Departamento de Biología, Facultad de Ciencias del Mar y Ambientales, Campus de Puerto Real, Universidad de Cádiz, 11510 Puerto Real, Cadiz, Spain*
[b] *Departamento de Ciencias Agroforestales, EPS, Campus La Rábida, Universidad de Huelva, 21819 Palos de la Frontera, Huelva, Spain*

ARTICLE INFO

Article history:
Received 17 June 2009
Received in revised form 6 October 2009
Accepted 13 October 2009

Keywords:
Hook
Longline
Selectivity
Pagellus bogaraveo
Artificial Neural Networks
Strait of Gibraltar

ABSTRACT

Size selectivity of the deep water longline used in the black spot seabream (*Pagellus bogaraveo*) fishery in the Strait of Gibraltar was studied with data of four sizes of hooks. Logistic (classic) and Artificial Neural Networks (heuristic) selectivity models were fitted for two experimental fishing trials. Logistic selectivity model was adequate for only one of the two periods analysed and the inferior results obtained with the classical approach were significantly improved by ANNs. These results indicate that in the event that the classic models do not fit well, perhaps due to poor quality of the data (such as a smaller sample size or highly overlapped distributions), the simpler ANNs models, with capacity to combine linear relationships and highly non-linear, are most appropriate to establish the functional relation between variables.

1. Introduction

Estimates of size selectivity of fishing gear give important information for conservation and optimum exploitation of fisheries resources (Hilborn and Walters, 1992; Beverton and Holt, 1993; Quinn and Deriso, 1999). The relations between the physical properties of fishing gear and the species and size composition of catches have been studied for many years (Clark, 1960). While the size selective nature of gears such as trawls and gill nets is well known, there is still no clear consensus on the form of the size selection curve for hooks on longlines. Both logistic type models, typically used to describe the selectivity of trawls, and the unimodal models used in gillnet selectivity studies have been used in hook selectivity studies (Clark, 1960; Ralston, 1982, 1990; Erzini et al., 1996, 1998, 2006; Millar and Holst, 1997; Sousa et al., 1999).

The origin of the absence of consensus on the hooks selectivity curve can be found in assuming a specific functional relation between hook and fish size derived from experiments carried out with trawls and gillnets. Thus, it is possible that in the case of hooks it is necessary to use non-traditional modelling techniques to analyse data showing considerable uncertainty and that are possibly not highly linear correlated.

In this regard, heuristics modelling techniques such as Artificial Neural Networks (ANNs) are being widely used in the fitting of functions and modelling of highly non-linear systems, performing a non-linear transformation between the input and output data, with significant advantages over conventional statistical methodologies in which the relation between the data has to be linear. Also, the ANNs do not require an equation for the empirical relation between the variables, as is the case in other parametric statistical techniques.

The ANNs characteristics have favoured the use of such methodologies for the development of applications related to fishery resources planning and management (Freón et al., 2003; Hardman-Mountford et al., 2003; Huse and Ottersen, 2003; Maravelias et al., 2003; Hyun et al., 2005; Chen and Hare, 2006; Czerwinski et al., 2007; Gutiérrez-Estrada et al., 2007; Velo-Suárez and Gutiérrez-Estrada, 2007).

The black spot seabream (*Pagellus bogaraveo*) fishery in the Strait of Gibraltar is an economically important fishery, carried out exclusively by a vertical deep water longline called "voracera". Hook legal dimensions are 3.95 ± 0.39 cm minimum length and 1.4 ± 0.14 cm minimum width. The black spot seabream legal size is currently fixed in 33 cm of total length. Hook size selectivity for the black spot seabream has been already studied in the Azores (Sousa et al., 1999), the Strait of Gibraltar (Czerwinski et al., 2009), as well as for other *Pagellus* species of the south of Portugal (Erzini et al., 1996). In all cases, the logistic type model fitted by maximum likelihood methods (Wulff, 1986) seems to be most adequate for describing longline size selectivity for *P. bogaraveo* and *Pagellus* and other small Sparidae species in general.

In this paper heuristic size selectivity models are fitted to catch size frequency distributions for the different sizes of hooks. The fits of ANN selectivity models are compared with the results obtained

---

* Corresponding author. Tel.: +34 956016017; fax: +34 956016019.
E-mail address: ivone.czerwinski@uca.es (I.A. Czerwinski).

0165-7836/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.fishres.2009.10.005

with maximum likelihood fits of the logistic model (Czerwinski et al., 2009), allowing us to evaluate the relative ability of these techniques to model the selectivity of different hooks. Our approach reveals the highly non-linear nature of the relationship between selectivity and hook size, and demonstrates that in some cases logistic selectivity models are not most adequate to explain the selectivity of hook longline fisheries.

2. Materials and methods

2.1. Data collection and exploration

Fishing was carried out from a commercial fishing boat in the Strait of Gibraltar from November 2000 to July 2005. The "voracera" main line was 80 m long, with 1.0 m long gangions spaced 1.10 m apart. Each main line consists of 70 gangions. Fishing was carried out on rocky bottoms at depths of up to 850 m. Commercial fleet fishing practices concerning setting, setting time and duration of set, were observed. Hooks were baited with regular sized pieces of sardine (*Sardina pilchardus*) in all longline sets.

Two experimental fishing trials were carried out. In the first experiment (Ex 1), 50 longline sets were carried out with three sizes of round bend, eyed "Siapal" brand hooks: numbers 9, 10 and 11, and 3500 hooks of each size. In the second experiment (Ex 2), 106 sets were made, with 7420 hooks each of hook sizes 9, 9.5 and 10. The different sized hooks were randomly distributed along the mainline, allowing us to consider the number of contact fish of each size class equal for all hook sizes. Hook dimensions, length (HL), width (HW) and depth (HD) are given in Fig. 1. Using the product of width and length to represent overall hook size (HO) (Otway and Craig, 1993), hooks number 10, 9.5 and 9 are 1.12, 1.26 and 1.40 times the size of hook number 11.

Average fish size for each hook and experimental trial was compared using ANOVA test ($p < 0.05$) and Fisher's least significant difference (LSD) *post hoc* analysis ($p < 0.05$). To compare fish size frequency distribution of each hook, a Kolmogorov–Smirnov test ($p < 0.05$) was performed.

2.2. Logistic selectivity model

The logistic selectivity model is found to be most appropriate for a variety of Sparidae, including *Pagellus acarne* and *Pagellus erythrinus*, and a logistic type selection curve was also obtained for *P. bogaraveo* in the Azores (Erzini et al., 1998; Sousa et al., 1999). Therefore, it was decided to fit logistic selection curves to the experimental catch data:

$$S_{i,j} = \frac{1}{1 + e^{-b_i(l_j - L50_i)}} \tag{1}$$

where $S_{i,j}$ is the size selectivity for hook size i and size class j, b_i is a parameter determining the slope of the selection curve for hook size i, l_j is the midpoint of the size class j and $L50_i$ is the length at 50% selection.

To estimate the parameters of selection curves, it is assumed that the parameters of the selection curve are a function of hook size (Kirkwood and Walker, 1986; Wulff, 1986). In our case mean length, depth, width and overall hook size (Otway and Craig, 1993) were used to estimate the parameters of the logistic curves as linear functions of hook size:

$$b_i = AH_i + B, \quad L50_i = CH_i + D \tag{2}$$

where A, B, C and D are parameters of the linear functions and H_i is the mean of each dimension such as width, length and overall hook size for hook size i. This model was fitted following the methodology proposed by Wulff (1986) and Kirkwood and Walker (1986) which is described in Czerwinski et al. (2009).

A total of 72 models (36 for each experiment) were fitted for each hook dimension (HL, HW, HD and HO) and for total length (TL), fork length (FL) and standard length (SL). Expected catches are calculated from the following relationship:

$$\hat{C}_{i,j} = S_{i,j}\hat{N}_j, \quad \text{where } \hat{N}_j = \frac{\sum_i C_{i,j}}{\sum_i S_{i,j}} \tag{3}$$

where $C_{i,j}$ is the observed catches for hook size i and size classes j, $\hat{C}_{i,j}$ is the expected catch and $\hat{N}_j$ the expected number of contact fish.

2.3. Artificial Neural Network models

Artificial Neural Networks (ANNs) are mathematical models inspired by the neural architecture of the human brain. The model neuron or node is a simple non-linear unit. The neurons collect inputs from single or multiple sources and produce an output. Interconnecting many of these single neurons or nodes in a known layer configuration creates a model neural network.

Each node j receives incoming signals from every node i in the previous layer. Associated with each incoming signal (x_i) is a weight ($W_{j,i}$). The effective incoming signal (I_j) to node j is the weighted sum of all the incoming signals:

$$I_j = \sum_i x_i W_{j,i} \tag{4}$$

The effective incoming signal I_j, is passed through an activation function (sometimes called a transfer function) to produce the outgoing signal (y_j) of the node j. In this study, the linear function ($y_j = I_j$) will be used in the output layer and the sigmoid non-linear

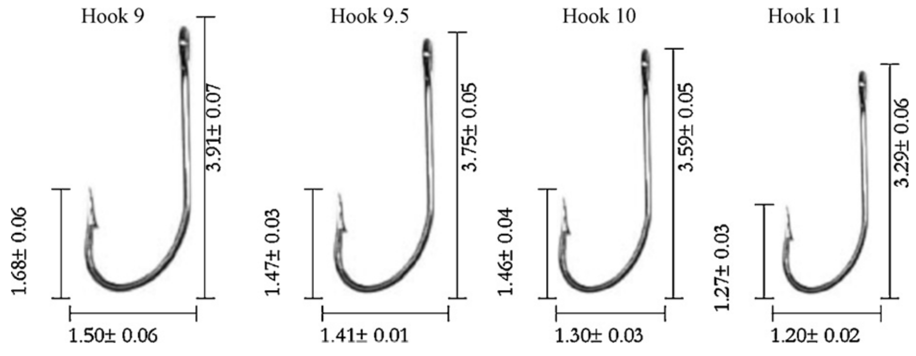

Fig. 1. Shape and dimensions of each type of hook (numbers 9, 9.5, 10 and 11) used in the selectivity study. Means and standard errors (in cm) are based on a sample size of 100 hooks for each type.

function will be used in the hidden layers:

$$y_j = f(I_j) = \frac{1}{1 + e^{(-I_j)}} \tag{5}$$

To determine the set of weights a corrective–repetitive process called learning or training is performed. This training forms the interconnections between neurons, and is accomplished using known inputs and outputs (training sets or patterns), and presenting them to the ANN in an ordered manner, adjusting the interconnection weights until the desired outputs are reached. The strength of these interconnections is adjusted using an error convergence technique so that a desired output will be produced for a given input. The training method used is the back propagation algorithm (Rumelhart et al., 1986).

The term epoch denotes the time period that encompasses all the iterations performed after all the patterns are displayed. In the study presented in this paper, the learning process was controlled by the method of internal validation (about 25% of calibration data to test the error at the end of each epoch). The weights are updated at the end of each epoch. The number of epochs with the smallest error of the internal validation indicates the weights to select (Tsoukalas and Uhrig, 1997).

Selectivity for each hook and fish size class was approximated from the relative retention frequency (RF) estimated for each individual. RF was calculated as:

$$\mathrm{RF} = [(F_1 - l_1) \times m] + (m \times L) \tag{6}$$

$$m = \frac{F_2 - F_1}{l_2 - l_1} \tag{7}$$

$$l_1 \leq L < l_2 \tag{8}$$

where F_1 and F_2 are the observed frequency for the size intervals I_1 and I_2 respectively and L is the considered fish length (TL, FL or SL) (Czerwinski et al., 2008).

A total of 36 ANN models were trained and validated for each experiment, varying the inputs and the number of neurons in the hidden layer. The input layer was formed by 2 neurons, one for fish length (TL, FL or SL) and the other one for hook size (HL, HD, HW and HO). Output layer was formed by one neuron, with the relative retention frequency (RF) estimated for each individual.

ANNs with 1 hidden layer with 3, 6 or 9 hidden nodes were successively trained, with a validation set of 190 data in Ex 1 and 298 in Ex 2. Since in a ANN the initial weights configuration is randomly fixed, a pool of 10 repetitions was carried out for each neural architecture with 10 000 epochs each time. This procedure allows the selection of the ANN with the best performance (Anctil and Rat, 2005).

2.4. Error measures

There are many measures of fit that can be used to compare different models (Abrahart and See, 2000). The correlation between observed and estimated catches was expressed by means of the correlation coefficient R. The coefficient of determination (R^2) describes the proportion of the total variance in the observed data that can be explained by the model. Other measures of variances applied were the percent standard error of prediction (%SEP) (Ventura et al., 1995) and the coefficient of efficiency (E_2) (Kitanidis and Bras, 1980). These four estimators are unbiased estimators that are employed to see how far the model is able to explain the total variance of the data.

The percent standard error of prediction, %SEP, is defined by:

$$\%\mathrm{SEP} = \frac{100}{\bar{C}} \mathrm{RMSE} \tag{9}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_i \sum_j (C_{i,j} - \hat{C}_{i,j})^2}{N}} \tag{10}$$

where $\bar{C}$ is the average of the observed catches, $\hat{C}_{ij}$ is the estimated catch of the same size class j and hook i and N is the total number of observations. RMSE is the square root of the mean square error. The coefficient of efficiency E_2 is used to see how the model explains the total variance of the data and represents the proportion of the variation of the observed data considered by the model. E_2 is given by:

$$E_2 = 1.0 - \frac{\sum_i \sum_j |C_{i,j} - \hat{C}_{i,j}|^2}{\sum_i \sum_j |C_{i,j} - \bar{C}|^2} \tag{11}$$

A value of zero for E_2 indicates that the observed average $\bar{C}$ is as good a predictor as the model, while negative values indicate that the observed average is a better predictor than the model (Legates and McCabe, 1999).

In addition, it is advisable to quantify the error in the same units of the variables. These measures, or absolute error measures, included the square root of the mean square error (RMSE) and the mean absolute error (MAE), given by:

$$\mathrm{MAE} = \frac{\sum_i \sum_j |C_{i,j} - \hat{C}_{i,j}|}{N} \tag{12}$$

Error estimates for each of the 10 repetitions of each model were calculated and compared with a multifactor analysis of variance MANOVA ($p < 0.05$). In the case of significant differences, a Fisher's least significant difference test (LSD) was used to establish the differences between the groups.

3. Results

3.1. Data exploration

There was a strong overlapping of the catch size frequency distributions in both experiments. However, significant differences in average fish length between all hooks were found except in Ex 2 where there was no significant difference between average fish size for hooks 9.5 and 10 (least significant differences (LSD) *post hoc* analysis). On the other hand, only depth dimension of hooks (HD) 9.5 and 10 were not significantly different ($t = 2.359$, d.f. = 198, $p < 0.05$).

In Ex 1, there were significant differences between all hooks but in Ex 2, the catch size frequency distribution of hook number 10 was not significantly different from that of hook number 9.5 (K–S = 0.749, $p > 0.05$). The catch size frequency distributions by hook size are given in Fig. 2. In both fishing trials, black spot seabream of a wide size range were caught. While Ex 1 catch size distributions were essentially unimodal but skewed to the right, those of Ex 2 are clearly multi-modal, with several smaller modes to the right of the main mode corresponding to catches of larger black spot seabream. Comparisons between the distributions of the two experiments showed significant differences (K–S = 8.497, $p < 0.05$).

3.2. Logistic models

The estimated parameters and error measures for a selection of selectivity models for each experiment are given in Table 1. For Ex 1 simple proportional functions were in all cases adequate for describing the relation between *L50* and hook size (H_i). For Ex 2 proportional functions described the relation between *L50* and hook size in most cases. The slope of the selection curves (b_i) was described as a constant or proportional function of hook size

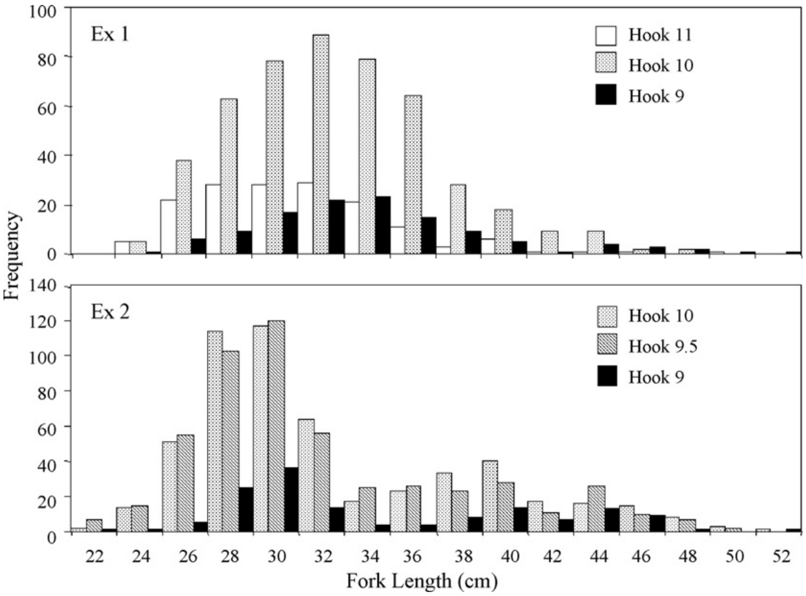

Fig. 2. *P. bogaraveo* catch size frequency distributions for different hook sizes in two experimental trials (Ex).

in both experiments. While total length (TL) was the most adequate fish dimension for describing selection curves as a function of hook dimension in Ex 1, that in most cases was the overall hook size (HO), in Ex 2 standard length and in most cases hook depth (HD) were the measurements that resulted in the best fits.

Comparing the fitting accuracy measures of Ex 1 models, model L1.10 with constant slope (b_i) and $L50_i$ proportional to overall hook size (HO), based on fork length (FL) measurements showed the highest values of E_2, and R^2. In Ex 2, model L2.2 with slope (b_i) and $L50_i$ proportional to hook depth dimension (HD), based on fork length (FL) measurements had the best values for E_2, and R^2. Models L1.10 of Ex 1 and L2.2 of Ex 2, were selected as the best based on their error measures.

The selection curves for both experiments are shown in Fig. 3. The observed catches and the expected catches based on these models are given in Fig. 5. In Ex 1, observed frequencies for hook 11 were lower than the expected, while observed frequencies for hook 10 were rather higher than the expected. However, such differences were not observed in Ex 2.

3.3. Artificial Neural Network (ANNs) models

A multifactor analysis of the variance (MANOVA) of the error measures calculated for each ANN structure showed that, in Ex 1 the dimension used for describing the hook size was not a significant factor, but there was a clear influence of fish length measurement. In all cases, better values of all accuracy measures were obtained using the fork length (FL). Number of neurons in the hidden layer was a significant factor for RMSE and %SEP, obtaining better values with 6 or 9 neurons (Table 2).

In Ex 2, the results of the MANOVA showed that the measure used for the size of the hook was an influential factor ($p < 0.05$) in the values of RMSE, MAE and %SEP, obtaining better results in using the HD as an input variable in all three cases. The length of the fish used (TL, FL, SL) was a significantly influential factor ($p = 0.0$) in all adjustment estimators. The LSD Fisher's test revealed that the best results were obtained, as in Ex 1, using the FL of fish as an input variable to the ANN. Likewise, the number of neurons that make up the hidden layer of the network, was revealed as a decisive factor

Table 1
Estimated parameters and error measures for logistic selectivity curve fitted by maximum likelihood, for each experiment using different fish and hook size dimensions. H_i, dimension of the hook i ($i = 9$, 9.5, 10 and 11).

Model number	Fish length	Dimension of H_i	A	B	C	D	RMSE	MAE	%SEP	E_2	R^2
L1.1	TL	HO		0.356	6.131		6.975	4.662	83.274	0.573	0.574
L1.2	TL	HO	0.053		6.309		7.473	5.181	89.219	0.510	0.510
L1.3	TL	HW		0.413	23.325		7.142	4.808	85.267	0.553	0.553
L1.4	TL	HW	0.247		23.914		7.224	4.882	86.246	0.542	0.542
L1.5	TL	HD	0.284		20.308		7.337	4.984	87.595	0.528	0.528
L1.6	TL	HL		0.568	8.369		7.382	5.049	88.133	0.522	0.522
L1.7	TL	HL	0.135		8.496		7.560	5.226	90.258	0.499	0.499
L1.8	TL	HD		1.900	17.745		7.582	5.089	90.520	0.496	0.496
L1.9	SL	HO	0.606		4.901		8.070	4.921	96.347	0.603	0.602
L1.10	FL	HO		0.410	5.483		7.303	4.842	87.189	0.615	0.608
L1.11	FL	HO	0.062		5.632		7.372	4.833	88.013	0.608	0.601
L2.1	TL	HD		0.242	29.506		2.211	1.587	16.137	0.976	0.975
L2.2	FL	HD	0.120		28.709		2.339	1.759	17.072	0.978	0.979
L2.3	FL	HD		0.268	26.736		2.361	1.759	17.232	0.978	0.978
L2.4	SL	HD	0.130		25.477		2.990	1.984	21.823	0.974	0.974
L2.5	SL	HD	0.153		19.100	10.767	2.992	1.983	21.838	0.974	0.974
L2.6	SL	HD		0.301	23.374		3.002	2.010	21.911	0.974	0.974

Hook dimensions: length (HL), width (HW) and depth (HD) and overall hook size (HO). Fish length: total (TL), fork (FL) and standard (SL).

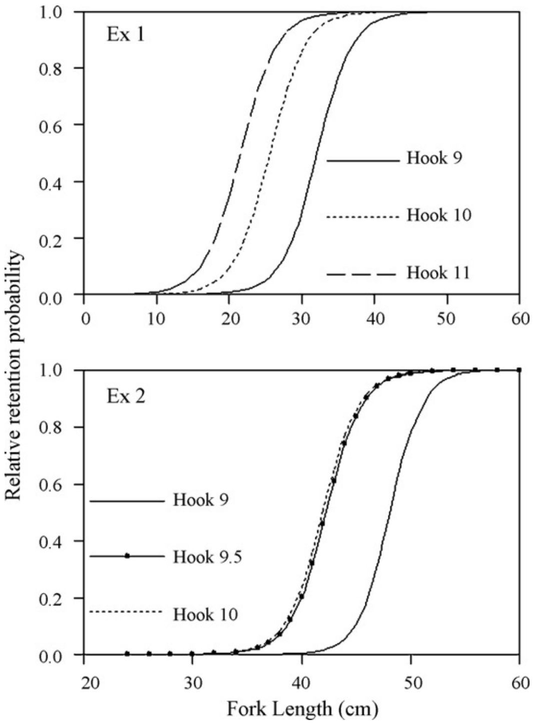

Fig. 3. Logistic selection curves of the longline, obtained from the selectivity model number 10 of Ex 1 and number 2 of Ex 2. Ex: experimental trial.

in all cases ($p = 0.0$), the networks with more than 3 neurons in the hidden layer get the best settings.

In Table 3 calculated estimators averages of 10 repetitions of each network for Ex 1 are observed. The proportion of explained variance was higher than 94% ($R^2 = 0.94$, ANN1.6) with the exception of model ANN1.27 that does not reach 36% ($R^2 = 0.36$). This model shows outliers of all estimators, so it was not taken into account in the subsequent analysis. The highest explained variance was 98.1% in the models ANN1.30 and ANN1.33, like the lowest RMSE and MAE (MAE = 1.64 and RMSE = 1.19), both calculated according to the FL of fish and with 9 neurons hidden in the architecture of the ANN. Lower %SEP was 18.79% in the model ANN1.55, also depending on the FL of fish and the largest was 36.25% (ANN1.23).

The explained variance by the ANNs in Ex 2 exceeded the 76% minimum value obtained in the model ANN2.4 ($R^2 = 0.77$). The highest values of R^2 (0.97) and E_2 (0.96) were obtained in the model ANN2.27 and the lowest RMSE (2.96), MAE (2.13) and %SEP (23.83). It is a model with 9 neurons in the hidden layer, depending on the FL of fish and the hook overall size (HO) (Table 4).

The model ANN1.33 for Ex 1 and ANN2.27 for Ex 2, were selected to explain the selectivity of the hooks used in the two experiments, for its values of SEP%, E_2 and R^2. In Ex 1, the model that best explains the selectivity of the hooks is a model with 9 neurons in the hidden layer, in which the input variables are the FL of fish and the HD of hooks. In EX 2, the model chosen is also a model with 9 neurons in the hidden layer and on the basis of the FL, however, the size of the hook that best explains the selectivity is HO. Because of the repetitions in each model, represented selectivity curves show the mean values obtained as well as the standard deviation for each class of size (Fig. 4).

3.4. Models comparison

The first attribute that differences the ANNs and the logistic obtained models, is that ANNs curves show the hook fishing power as different curve maximums, proportional to catch rate of each hook. It is noteworthy that the values of standard deviation of the selectivity averages for each hook, were higher for FL values in which the number of fish in the sample was smaller, indicating a better fit in the range of sizes in which more observations were obtained. Also, there are differences in the chosen models for the two experiments.

In Ex 1, three bell-shaped curves were obtained. The maximum selectivity of the hook 11, which captured the smallest amount of fish, is 0.26 for 23 cm of FL, followed by the maximum selectivity of the hook 9, of 0.29 for 26 cm of FL. The hook that caught more fish (fish hook 10) also was the hook with higher values of selectivity, with a maximum of 0.69 in fish of 38 and 39 cm of FL.

In Ex 2, almost logistic curves were obtained in all three cases (Fig. 5). While the three curves show a descent in values below 20 cm of FL, the standard deviation in this range of lengths is very high for the three hooks. Given that the values of the data used in the training of the networks covered the range from 21 to 52 cm of FL, the values of selectivity obtained under or above these lengths, showed very high deviations. Hook 10, reaches its maximum selectivity in 48 cm of FL with a value of 0.49, slightly falling for higher lengths. Hook 9.5, with a rising curve from 20 to 59 cm of FL, reaches its maximum value at 0.49. However, within the range of sizes obtained in the sample, its maximum value was 0.47 for 52 cm of FL. Finally, for hook 9, the obtained selectivity curve showed lower values than the other two curves in the entire range of sizes sampled. The maximum value was, within the range of sizes obtained in the experiment, 0.37 for 52 cm of FL.

Table 2
MANOVA results for models error measures between hook dimension, fish length and number of neurons in the hidden layer of the ANN. *F*: the coefficient of variance analysis, *p*: significance level, d.f.: degrees of freedom, LSD: results of *post hoc* analysis of Fisher's least significant difference test, <: significant difference, =: no significant difference.

	Hook dimension (d.f. = 3)			Fish length (d.f. = 2)			Neurons in hidden layer (d.f. = 2)		
	F	*p*	LSD	*F*	*p*	LSD	*F*	*p*	LSD
Ex 1									
RMSE	0.38	0.764		2122.04	0	FL < TL < SL	3.53	0.030	6 = 9 < 3
MAE	0.32	0.812		1238.69	0	FL < SL < TL	2.66	0.071	
%SEP	0.37	0.775		2412.35	0	FL < TL < SL	3.41	0.034	6 = 9 < 3
E_2	0.23	0.877		1433.36	0	TL < SL < FL	2.27	0.105	
R^2	0.36	0.783		1421.02	0	TL < SL < FL	2.69	0.070	
Ex 2									
RMSE	3.2	0.023	HD < HO = HL = HW	133.67	0	FL < SL < TL	22.48	0	9 = 6 < 3
MAE	3.52	0.015	HD < HO = HL = HW	135.57	0	FL < SL < TL	22.95	0	9 = 6 < 3
%SEP	2.95	0.033	HD < HO = HL = HW	203.36	0	FL = SL < TL	21.82	0	9 = 6 < 3
E_2	1.5	0.215		107.77	0	TL < SL = FL	17.93	0	3 < 9 = 6
R^2	2.47	0.061		180.24	0	TL < SL = FL	10.47	0	3 < 6 = 9

Hook dimensions: length (HL), width (HW) and depth (HD) and overall hook size (HO). Fish length: total (TL), fork (FL) and standard (SL).

Table 3
Average measures of accuracy for ANN models in Experiment 1, calculated in a pool of 10 repetitions. H_i, dimension of the hook i (i = 9, 9.5, 10 and 11).

Model	H_i	Fish length	Neurons in hidden layer	RMSE	MAE	%SEP	E_2	R^2
ANN1.1	HO	TL	3	2.615	1.896	32.000	0.940	0.943
ANN1.2	HO	TL	6	2.591	1.854	31.705	0.941	0.944
ANN1.3	HO	TL	9	2.636	1.904	32.258	0.939	0.943
ANN1.4	HL	TL	3	2.621	1.904	32.071	0.940	0.942
ANN1.5	HL	TL	6	2.558	1.864	31.297	0.943	0.945
ANN1.6	HL	TL	9	2.695	1.971	32.983	0.936	0.940
ANN1.7	HD	TL	3	2.633	1.904	32.224	0.939	0.942
ANN1.8	HD	TL	6	2.587	1.881	31.654	0.941	0.944
ANN1.9	HD	TL	9	2.637	1.914	32.272	0.939	0.941
ANN1.10	HW	TL	3	2.622	1.894	32.081	0.940	0.943
ANN1.11	HW	TL	6	2.582	1.879	31.593	0.942	0.945
ANN1.12	HW	TL	9	2.596	1.883	31.764	0.941	0.944
ANN1.13	HO	SL	3	2.964	1.784	35.104	0.946	0.946
ANN1.14	HO	SL	6	3.011	1.821	35.654	0.944	0.945
ANN1.15	HO	SL	9	2.932	1.783	34.716	0.947	0.948
ANN1.16	HL	SL	3	3.057	1.853	36.203	0.942	0.943
ANN1.17	HL	SL	6	2.907	1.749	34.420	0.948	0.948
ANN1.18	HL	SL	9	2.903	1.762	34.375	0.948	0.949
ANN1.19	HD	SL	3	2.950	1.796	34.940	0.946	0.947
ANN1.20	HD	SL	6	2.986	1.805	35.364	0.945	0.946
ANN1.21	HD	SL	9	3.013	1.830	35.682	0.943	0.945
ANN1.22	HW	SL	3	3.030	1.847	35.879	0.943	0.944
ANN1.23	HW	SL	6	3.061	1.865	36.247	0.942	0.943
ANN1.24	HW	SL	9	2.960	1.779	35.056	0.946	0.946
ANN1.25	HO	FL	3	1.743	1.273	19.955	0.977	0.978
ANN1.26	HO	FL	6	1.688	1.254	19.324	0.979	0.980
ANN1.27	HO	FL	9	11.619	6.955	133.00	-0.012	0.357
ANN1.28	HL	FL	3	1.741	1.260	19.928	0.978	0.978
ANN1.29	HL	FL	6	1.688	1.245	19.328	0.979	0.979
ANN1.30	HL	FL	9	1.643	1.193	18.809	0.980	0.981
ANN1.31	HD	FL	3	1.803	1.291	20.642	0.976	0.977
ANN1.32	HD	FL	6	1.657	1.224	18.963	0.980	0.980
ANN1.33	HD	FL	9	1.642	1.194	18.791	0.980	0.981
ANN1.34	HW	FL	3	1.787	1.303	20.462	0.976	0.977
ANN1.35	HW	FL	6	1.685	1.228	19.293	0.979	0.979
ANN1.36	HW	FL	9	1.688	1.234	19.328	0.979	0.979

Hook dimensions: length (HL), width (HW) and depth (HD) and overall hook size (HO). Fish length: total (TL), fork (FL) and standard (SL).

Observed and expected catches for both experiments and models are shown in Fig. 5. It can be seen how expected catches are very similar to the observed ones, even in Ex 1 for hooks 10 and 11 that showed differences with logistic model.

4. Discussion

In this paper, the selectivity models for black spot sea bream in the Gibraltar strait were fitted using multilayer perceptron neural networks and compared with classic logistic approaches. This is the first time ANN approaches have been applied to explore selectivity patterns and is the first study to compare the performance of logistic models and heuristic modelling techniques in a hook longline fishery. Results indicate that in some cases logistics selectivity models are not the most adequate to explain the selectivity of this type of fishery. The ANN models obtained here clearly indicate that a high proportion of the variation in selectivity can be explained by models containing relatively few hook parameters. The ANN approach applied in this study resulted in models explaining >97.4% of the mean variance in frequency observed, versus 79.3% of mean variance reached with classic logistic models.

The efficiency and the size selectivity of hook gears are influenced by numerous factors. These include the distribution of the fish, competition between fish, size and design of hook, size and forms of bait, combination of baits, duration of sets and the time of the day of fishing (Arimoto et al., 1982; Skeide et al., 1986; Løkkeborg and Bjordal, 1992; Bjordal and Løkkeborg, 1996). Previous hooks selectivity studies have used a wide range of sizes and types of hooks and have reached different conclusions.

Although there is no clear consensus on the form of the hook selection curve, the logistic selectivity model has been reported as the most adequate for the black spot seabream longline fishery and for other seabream species longline fisheries (Erzini et al., 1998, 2006; Sousa et al., 1999). However, in the case of the seabream longline fishery of Strait of Gibraltar the logistic selectivity model was adequate for only one of the two periods analysed (Ex 2). Although according to Ralston (1990) a decrease in the selectivity for larger sizes may be expected, Erzini et al. (1996, 2006) found that the population size structure can limit the possibilities of a total recognition of selectivity effects which could indicate that fish catch size frequency distributions may not be sufficiently informative to allow a meticulous evaluation of the real shape of the hook selection curve (Millar and Holst, 1997). Although the black spot seabream caught in this study had a wide size range in both periods, total catches of the two experimental trials showed differences in the size frequency distributions. Therefore, it is probable that the differences in the selectivity curves for the Ex 1 and Ex 2 are a consequence of the different population size structures.

The obtained accuracy measures for Ex 1 were very different in function of type of model used. In this experiment, estimated catches were very different from the observed ones when the classical approach was used (model L1.10: RMSE = 7.303, MAE = 4.842, %SEP = 87.189, E_2 = 0.615, R^2 = 0.608). A reason for this lack of accuracy is that, contrary to what would be expected, in Ex 1 the intermediate sized hook (hook number 10) showed the highest observed catch rate and this may be affecting the fishing power of different hook sizes that was assumed as equivalent (Bjodal and Løkkeborg, 1996). Sousa et al. (1999) reported a similar behaviour in longline fishery of *Helicolenus dactylopterus* of Azores popuation.

Table 4
Average measures of accuracy for ANN models in Experiment 2, calculated in a pool of 10 repetitions. H_i: dimension of the hook i (i = 9, 9.5, 10 and 11).

Model	H_i	Fish length	Neurons in hidden layer	RMSE	MAE	%SEP	E_2	R^2
ANN2.1	HO	TL	3	7.361	4.233	66.751	0.709	0.828
ANN2.2	HO	TL	6	4.937	3.120	44.768	0.874	0.882
ANN2.3	HO	TL	9	*6.579*	3.635	59.659	0.757	0.799
ANN2.4	HL	TL	3	8.976	4.979	81.393	0.594	0.767
ANN2.5	HL	TL	6	5.193	3.275	47.091	0.867	0.876
ANN2.6	HL	TL	9	5.039	3.173	45.691	0.873	0.885
ANN2.7	HD	TL	3	7.022	4.066	63.677	0.729	0.825
ANN2.8	HD	TL	6	4.598	2.862	41.699	0.887	0.898
ANN2.9	HD	TL	9	3.939	2.587	35.719	0.925	0.932
ANN2.10	HW	TL	3	8.020	4.309	72.728	0.670	0.783
ANN2.11	HW	TL	6	5.703	3.532	51.715	0.829	0.856
ANN2.12	HW	TL	9	5.535	3.419	50.187	0.849	0.869
ANN2.13	HO	SL	3	4.082	2.654	28.768	0.952	0.958
ANN2.14	HO	SL	6	3.887	2.599	27.395	0.956	0.960
ANN2.15	HO	SL	9	3.710	2.464	26.143	0.960	0.963
ANN2.16	HL	SL	3	4.661	3.055	32.849	0.935	0.951
ANN2.17	HL	SL	6	4.117	2.681	29.010	0.951	0.958
ANN2.18	HL	SL	9	3.693	2.442	26.021	0.961	0.963
ANN2.19	HD	SL	3	3.711	2.530	26.149	0.960	0.963
ANN2.20	HD	SL	6	3.927	2.604	27.671	0.955	0.957
ANN2.21	HD	SL	9	3.562	2.427	25.099	0.963	0.965
ANN2.22	HW	SL	3	4.170	2.765	29.385	0.949	0.954
ANN2.23	HW	SL	6	3.814	2.472	26.879	0.958	0.961
ANN2.24	HW	SL	9	4.220	2.750	29.741	0.949	0.952
ANN2.25	HO	FL	3	3.555	2.437	28.632	0.949	0.957
ANN2.26	HO	FL	6	3.222	2.255	25.952	0.958	0.960
ANN2.27	HO	FL	9	2.958	2.129	23.826	0.964	0.967
ANN2.28	HL	FL	3	3.353	2.272	27.005	0.955	0.959
ANN2.29	HL	FL	6	3.091	2.214	24.897	0.962	0.964
ANN2.30	HL	FL	9	3.109	2.168	25.042	0.960	0.964
ANN2.31	HD	FL	3	3.124	2.202	25.162	0.961	0.964
ANN2.32	HD	FL	6	3.442	2.332	27.720	0.951	0.954
ANN2.33	HD	FL	9	3.114	2.171	25.077	0.960	0.963
ANN2.34	HW	FL	3	3.473	2.348	27.972	0.951	0.956
ANN2.35	HW	FL	6	3.519	2.399	28.341	0.949	0.956
ANN2.36	HW	FL	9	3.164	2.247	25.486	0.959	0.961

Hook dimensions: length (HL), width (HW) and depth (HD) and overall hook size (HO). Fish length: total (TL), fork (FL) and standard (SL).

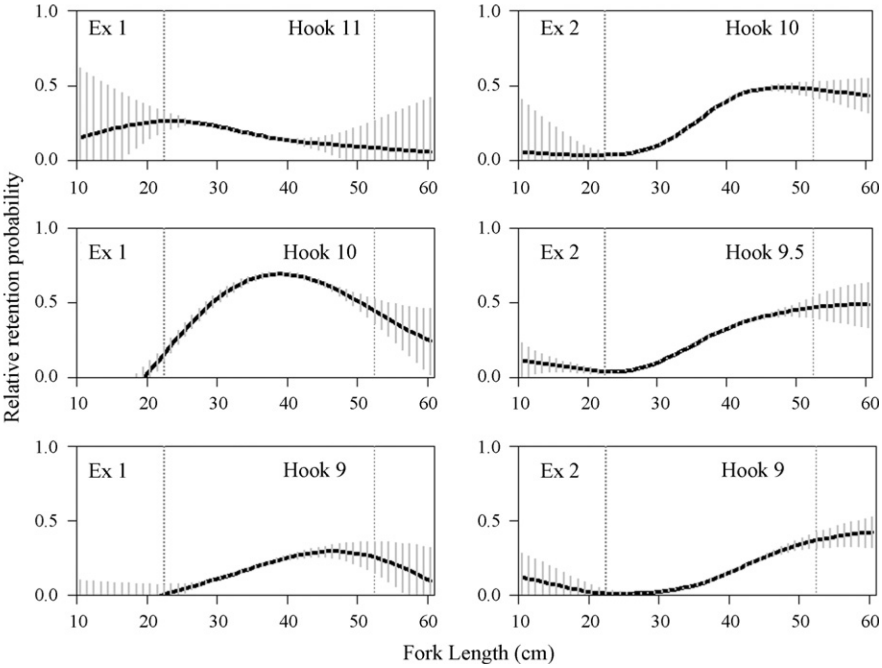

Fig. 4. Average selection curves and standard deviation obtained with ANNs for each hook and experiment. Dashed lines indicate range limits for fork length (FL) observed values. Ex: experimental trail.

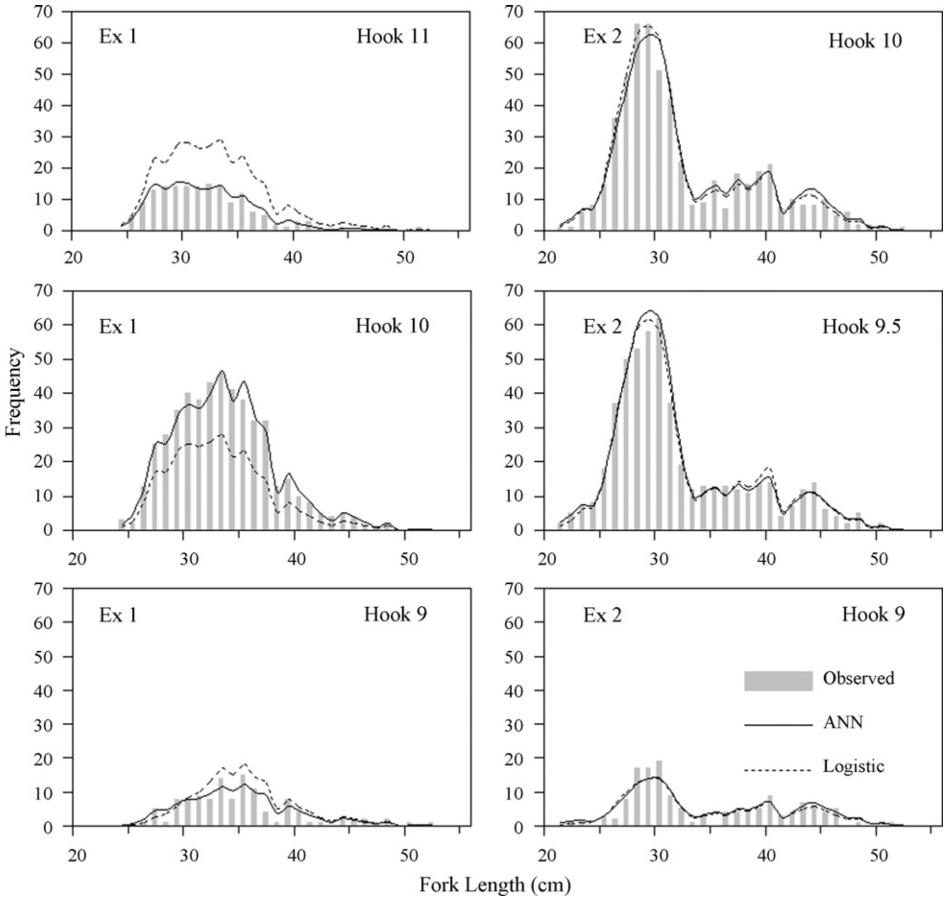

Fig. 5. Comparison of the observed and predicted catch size frequency distributions for the three hook sizes according to the logistic and ANNs selected model of each experimental trail (Ex).

In this study, significantly lower catch rates were obtained by the smallest hook used and therefore, for this hook best fits were obtained with skew-normal type selectivity curve rather than sigmoid type. Bjordal and Løkkeborg, 1996 suggested that main reason that affects to the loss of efficiency of the smallest hooks may be due to bait loss.

The inferior results obtained in Ex 1 with the classical approach were significantly improved by ANNs (model ANN1.33: RMSE = 1.642, MAE = 1.194, %SEP = 18.791, $E_2 = 0.980$, $R^2 = 0.981$). Several reasons could explain the best behaviour of ANN models in the case of Ex 1. Firstly, the ANNs are not sensitive to fishing power because these types of models do not assume any empirical relation between inputs and output variables. On the other hand, if the loss of efficiency of hook number 11 (the smallest in this study) was a consequence of a greater bait loss rate of this hook, then the Artificial Neural Network models could have detected this hook characteristic implicit in the inputs variables. The ANN capacity to identify the relevant information implicit in input variables from massive amounts of data, discriminating which is inconsistent, ambiguous or incomplete has been reported in other studies (Govindajaru, 2000; Gutiérrez-Estrada et al., 2008).

On the other hand, obtained accuracy measures for Ex 2 were very similar with both types of models. Estimated catches were very similar to the observed ones using the classical approach (model L2.2: RMSE = 2.339, MAE = 1.759, %SEP = 17.072, $E_2 = 0.978$, $R^2 = 0.979$). In this case, as would be expected, observed catch rates were lower when hook size increased. There were no significant differences between size frequency distributions of hooks 9.5 and 10 catches, which did not differ significantly in terms of average hook depth dimension, and in all fitted models, size selectivity was a function of hook depth dimension. ANN model selected for EX 2, coincides with logistic model, where the hook dimension used is the depth and the fitting results are also very similar (model ANN2.27: RMSE = 2.958, MAE = 2.129, %SEP = 23.826, $E_2 = 0.964$, $R^2 = 0.967$). These results suggest that the depth size of the hook used may be an important key when size of hook is used for fisheries resources management proposes.

The choice of type of models that forecast with greater accuracy catches from the size of hooks, should be different in each case. In Ex 1, there is a clear difference between the predictive power of the ANNs and the logistic models, with the former prevailing. However, in Ex 2, the difference in accuracy is low, but better in the logistic models. These results indicate that in the event that the classic models do not fit well, perhaps due to poor quality of the data (such as a smaller sample size or highly overlapped distributions), the simpler ANNs models, with capacity for combine linear relationships and highly non-linear, are most appropriate to establish the functional relation between variables.

The total catch of the two experiments showed differences in size frequencies distribution. Similarly, the fitted selectivity models for each experimental trial were very different despite having two hooks in common. These results suggest that the selectivity estimation of this gear is not independent of the structure of the exploited population. On the other hand, Erzini et al. (2006), found that size selectivity and fishing mortality rate strongly influenced the dynamics of the average size in *P. bogaraveo* populations. They also noted that logistic size selectivity had more effect on the average size than bell-shaped selectivity models. This would indicate a relation between the population structure and the fishing gear selectivity estimation, at least in case of hook selectivity. So it can be seen that in the Ex 1, ANNs model, which best explains the hooks selectivity to the existing population, adopts bell-shaped forms, while in Ex 2, both models are explained by logistical or similar curves.

In light of the results, it can be said that the selectivity of the hooks follows a dynamic relation with the population size structure. So the odds of getting a certain fish size with a hook are affected by the probability that the fish of that size come into contact with fishing gear, which is closely related with the proportion of fish of that size in the population.

Acknowledgements

This work has been partly financed by the Agro alimentary and Fishery Research and Formation Institute (IFAPA) (project: C03-007-2003-110), the General Direction of Fishery and Aquaculture of the Council of Andalucía and the Provincial Deputation of Cádiz. The University of Cádiz provided the necessary facilities for a stay in the Universidade do Algarve supervised by Dr. Karim Erzini. We would like to express our gratitude to our colleagues, Dr. Cristina Zabala, Mª Carmen Gómez Cama, Remedios Cabrera, Javier Llorente and Jose M. García Rebollo for the assistance they willingly provided during the samplings. We also thank to two anonymous reviewers for their helpful comments and suggestions for improving the manuscript.

References

Abrahart, R.J., See, L., 2000. Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecasts in two contrasting catchments. Hydrol. Process. 14, 2157–2172.

Anctil, F., Rat, A., 2005. Evaluation of neural network stream flow forecasting on 47 watersheds. J. Hydrol. Eng. 10, 85–88.

Arimoto, T., Ogura, M., Inoue, Y., 1982. Catch variation with immersion time of gear in coastal set-line. B. Jpn. Soc. Sci. Fish., 705–709.

Beverton, R., Holt, S., 1993. On the Dynamics of Exploited Fish Populations. Chapman and Hall, London.

Bjordal, A., Løkkeborg, S., 1996. Longlining. Blackwell Science, Ltd., Oxford.

Chen, D.G., Hare, S.R., 2006. Neural network and fuzzy logic model for pacific halibut recruitment analysis. Ecol. Model. 195, 11–19.

Clark, J.R., 1960. Report on selectivity of fishing gear. In: Fishing Effort, the Effects of Fishing on Resources and the Selectivity of Fishing Gear. ICNAF Spec. Publ. 2, pp. 27–36.

Czerwinski, I.A., Gutiérrez-Estrada, J.C., Hernando-Casal, J.A., 2008. Modelos de selectividad para la pesquería del voraz (Pagellus bogaraveo) en el Estrecho de Gibraltar. Aproximación clásica y heurística. In: Gutiérrez-Estrada, J.C., Yáñez Rodriguez, E. (Eds.), Nuevas aproximaciones metodológicas para el análisis de pesquerías. Universidad de Huelva. Servicio de Publicaciones, Huelva, pp. 235–268.

Czerwinski, I.A., Gutiérrez-Estrada, J.C., Hernando-Casal, J.A., 2007. Short-term forecasting of halibut CPUE: linear and non-linear univariate approaches. Fish. Res. 86, 120–128.

Czerwinski, I.A., Erzini, K., Gutiérrez-Estrada, J.C., Hernando-Casal, J.A., 2009. Deep water longline selectivity for black spot seabream (Pagellus bogaraveo) in the strait of gibraltar. Fish. Sci. 75 (2), 285–294.

Erzini, K., Gonçalves, J.M.S., Bentes, L., Lino, P.G., Cruz, J., 1996. Species and size selectivity in a Portuguese multispecies artisanal long-line fishery. ICES J. Mar. Sci. 53, 811–819.

Erzini, K., Gonçalves, J.M.S., Bentes, L., Lino, P.G., Ribeiro, J., 1998. Species and size selectivity in a 'red' seabream longline 'metier' in the Algarve (southern Portugal). Aquat. Living Resour. 11, 1–11.

Erzini, K., Salgado, M., Castro, M., 2006. Dynamics of black spot seabream (Pagellus bogaraveo) mean length: evaluating the influence of life history parameters, recruitment, size selectivity and exploitation rates. J. Appl. Ichthyol. 22, 183–188.

Freón, P., Mullon, C., Voisin, B., 2003. Investigating remote synchronous patterns in fisheries. Fish. Oceanogr. 12 (4/5), 443–457.

Govindajaru, R.S., 2000. Artificial neural networks in hydrology. I. Preliminary concepts. J. Hydrol. Eng. 5, 115–123.

Gutiérrez-Estrada, J.C., Silva, C., Yánez, E., Rodríguez, N., Pulido-Calvo, I., 2007. Monthly catch forecasting of anchovy Engraulis rigens in the north area of Chile: non-linear univariate approach. Fish. Res. 86, 188–200.

Gutiérrez-Estrada, J.C., Pulido-Calvo, I., Silva, C., Yánez, E., Rodríguez, N., 2008. Estimación mensual de las capturas de anchoveta (Engraulis ringens) en la costa norte de Chile mediante métodos heurísticos. In: Gutiérrez-Estrada, J.C., Yáñez Rodriguez, E. (Eds.), Nuevas aproximaciones metodológicas para el análisis de pesquerías. Universidad de Huelva. Servicio de Publicaciones, Huelva, pp. 11–38.

Hardman-Mountford, N.J., Richardson, A.J., Boyer, D.C., Kreiner, A., Boyer, H.J., 2003. Relating sardine recruitment in the North Benguela to satellite-derived sea surface height using a neural network pattern recognition approach. Prog. Oceanogr. 59, 241–255.

Hilborn, R., Walters, C.J., 1992. Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty. Chapman & Hall, London.

Huse, G., Ottersen, G., 2003. Forecasting recruitment and stock biomass of Northeast Artic cod using neural networks. Sci. Mar. 67, 325–335.

Hyun, K., Song, M.Y., Kim, S., Chon, T.S., 2005. Using and artificial neural network to patternize long-term fisheries data from South Koreal. Aquat. Sci. 67, 382–389.

Kirkwood, G.P., Walker, T.I., 1986. Gill net mesh selectivities for gummy shark, Mustelus antarcticus, taken in southeastern Australian waters. Aust. J. Mar. Fresh. Res. 37, 689–697.

Kitanidis, P.K., Bras, R.L., 1980. Real time forecasting with a conceptual hydrological model. 2. Applications and results. Water Resour. Res. 16, 1034–1044.

Legates, D.R., McCabe Jr., G.J., 1999. Evaluating the use of 'goodness-of-fit' measures in hydrologic and hydroclimatic model validation. Water Resour. Res. 35, 233–241.

Løkkeborg, S., Bjordal, A., 1992. Species and size selectivity in long-line fishing: a review. Fish. Res. 13, 311–322.

Maravelias, C.D., Haralaobus, J., Papaconstantinou, C., 2003. Predicting demersal fish species distributions in the Mediterranean Sea using artificial neural networks. Mar. Ecol. Prog. Ser. 255, 249–258.

Millar, R.B., Holst, R., 1997. Estimation of gillnet and hook selectivity using log-linear models. ICES J. Mar. Sci. 54, 471–477.

Otway, N.M., Craig, J.R., 1993. Effects of hook size on the catches of undersized snapper Pagrus auratus. Mar. Ecol. Prog. Ser. 93, 9–15.

Quinn, T.J., Deriso, R.B., 1999. Quantitative Fish Dynamics. Oxford University Press, New York.

Ralston, S., 1982. Influence of hook size in the Hawaiian deep-sea handline fishery. Can. J. Fish. Aquat. Sci. 39, 1297–1302.

Ralston, S., 1990. Size selection of snappers (Lutjanidae) by hook and line gear. Can. J. Fish. Aquat. Sci. 47, 696–700.

Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning' representations by back-propagation errors. Nature 323, 533–536.

Skeide, R., Bjordal, A., Løkkeborg, S., 1986. Testing of a new hook design (E-Z-Baiter) through comparative longline fishing trials. ICES CM, B:25.

Sousa, F., Isidro, E., Erzini, K., 1999. Semi-pelagic longline selectivity for two demersal species from the Azores: the black spot seabream (Pagellus bogaraveo) and the bluemouth rockfish (Helicolenus dactylopterus dactylopterus). Fish. Res. 41, 25–35.

Tsoukalas, L.H., Uhrig, R.E., 1997. Fuzzy and Neural Approaches in Engineering. Wiley Interscience, New York.

Velo-Suárez, L., Gutiérrez-Estrada, J.C., 2007. Artificial neural network approaches to one-step weekly prediction of Dinophysis acuminata blooms in Huelva (Western Andalucía, Spain). Harmful Algae 6, 361–471.

Ventura, S., Silva, M., Pérez-Bendito, D., Hervás, C., 1995. Artificial neural networks for estimation of kinetic analytical parameters. Anal. Chem. 67, 1521–1525.

Wulff, A., 1986. Mathematical model for selectivity of gill nets. Arch. Fish. Wiss. 37, 101–106.