Environ Ecol Stat
DOI 10.1007/s10651-012-0196-1

Logistic regression and fuzzy logic as a classification method for feral fish sampling sites

Rachel Ann Hauser-Davis · Terezinha Ferreira de Oliveira · Antônio Morais da Silveira · João Marcelo Brazão Protázio · Roberta Lourenço Ziolli

Received: 16 February 2011 / Revised: 11 November 2011

Abstract This study presents a classification method combining logistic regression and fuzzy logic in the determination of sampling sites for feral fish, Nile Tilapia (*Tilapia rendalli*). This method statistically analyzes the variable domains involved in the problem, by using a logistic regression model. This in turn generates the knowledge necessary to construct the rule base and fuzzy clusters of the fuzzy inference system (FIS) variables. The proposed hybrid method was validated using three fish stress indices; the Fulton Condition Factor (FCF) and the gonadossomatic and hepatossomatic indices (GSI and HSI, respectively), from fish sampled at 3 different locations in the Rio de Janeiro State. A multinomial logistic regression allowed for the FIS construction of the proposed method and both statistical approaches, when combined, complemented each other satisfactorily, allowing for the construction of an efficient classification method regarding feral fish sampling sites that, in turn, has great value regarding fish captures and fishery resource management.

R. A. Hauser-Davis (✉)
Laboratório de Bioanalítica, Departamento de Química, Pontifícia Universidade Católica, Rio de Janeiro (PUC-Rio), Rua Marquês de São Vicente, 225, Gávea, Rio de Janeiro, RJ, CEP: 22453-900, Brazil
e-mail: rachel.hauser.davis@gmail.com

T. F. de Oliveira · A. M. da Silveira · J. M. B. Protazio
Faculdade de Estatística and Faculdade de Ciências da Computação, Instituto de Ciências Exatas e Naturais, Universidade Federal do Pará (UFPA), Rua Augusto Correa, 01, Belém, PA, CEP: 66075-110, Brazil

R. L. Ziolli
Instituto de Biociências, Universidade Federal do Estado do Rio de Janeiro (UNIRIO), Av. Pasteur, 458, Urca, Rio de Janeiro, RJ, CEP: 22290-240, Brazil

Published online: 28 March 2012

Keywords Fuzzy logic · Logistic regression · Nile Tilapia · Sampling sites · Fish stress indices

1 Introduction

1.1 Statistical techniques: logistic regression and fuzzy logic

A logistic regression is a statistical modeling tool that relates a categoric response variable with explanatory variables that influence the occurrence of a certain event. Unlike multiple regression and discriminant analysis methods, a logistic regression does not assume the existence of variance homogeneity and residue normality. This allows for more ample applications than other methods (Graybill 1976).

Classical logic only recognizes two values, true or false, while fuzzy logic is multi-valued (Shaw and Simões 1999). This method is able to perfectly handle verbal, imprecise and qualitative expressions inherent to human communication, translating the diffuse terms of human communication in values understandable by computers (Turban and Aronson 2001). This allows for the classification of imprecise and uncertain data and vague and ambiguous information, favoring decision-making processes. By using production rules such as IF-THEN, data characterized as uncertain is analyzed according to deductive reasoning. The input, or condition, and output, or consequence, of a fuzzy system are associated by reasoning rules with truth degrees or numeric intervals between [0,1]. The main function of a fuzzyfier is to convert the real input values to a pertinence degree to fuzzy clusters, for data analyses by the inference machine. This machine then uses fuzzy logic principles to combine fuzzy IF-THEN rules present in a rule base into a mapping of the input fuzzy cluster to an output fuzzy cluster (Wang 1997). The defuzzyfier is defined as the mapping of a fuzzy cluster, the output of the inference machine, into a real value. This is translated as specifying an exit point that better represents the fuzzy cluster (Negnevitsky 2005). When choosing a defuzzyfier several criteria must be considered, such as plausibility (the output value is intuitive), computational simplicity and continuity. The most common defuzzyfiers are center gravity, center average and maximum center (Wang 1997).

Hybrid solutions using logistic regression methods and fuzzy inference systems can be constructed in order to evaluate complex data, including ecological data. In a first step a logistic regression is used as a knowledge aqcuisition process, establishing the basis for the construction of a fuzzy inference system. This is then submitted to a validation sample, for data classification.

1.2 Fish species and stress indices

Fish are used worldwide as biomarkers indicating environmental changes in the aquatic environment. They usually occupy the top of the aquatic food web and, therefore, integrate effects of biomagnifiable pollutants from lower trophic levels, as well as any direct effect from direct exposure. As they are an important link between the environment and human populations through fisheries and consumption by local and other markets, it is of paramount importance to monitor their health condition with regard to harmful environmental effects and potential contaminant exposure routes to humans.

Stress indices are commonly used to verify changes in fish health condition, and are related to environmental interferences, such as heavy metal and organic compound contamination (Laflamme et al. 2000). The Fulton Condition Factor (FCF) is used in the evaluation of the general health condition of fish. The Gonadossomatic index (GSI) indicates reproductive changes and the Hepatossomatic index (HSI) identifies changes in the liver, associated to liver disorders and hepatic stress. These indices are location-specific, and may be used to verify sampling site adequacy for feral fish capture.

Nile Tilapia (*Tilapia rendalli*) are freshwater fish and exhibit several advantages that make them ideal for cultivation for human consumption, such as accepting several food items, presenting positive responses to captive fertilization and being resistant to disease (Appleyard et al. 2001). These fish are amply distributed throughout Brazilian territory and cultivated in many production systems (Bassay et al. 1997). They are considered a bioindicator species, and are commercially important in the Rio de Janeiro state, consumed in significant amounts by the human population of this area.

Since computational intelligence tools have been proven useful in the classification of ecological data, including fish species classification (Hauser-Davis et al. 2010), this study applied a multinomial logistic regression model to obtain statistical probabilities. These were then used to construct a rule base of a fuzzy system, in order to classify sampling sites for feral Nile Tilapia by their stress indices, using fish sampled from 3 commercially important fishing sites in the State of Rio de Janeiro. This is important when considering environmental issues related to fish health, in which choosing the best sampling site becomes important, and also regarding economic aspects in fisheries and fish management areas.

2 Material and methods

2.1 Sampling sites and stress index calculations

Fish sampling sites were chosen taking into account their pollution status: two sites (Jacarepaguá and Rodrigo de Freitas Lagoon) are located in the contaminated Guanabara Bay, which is an estuary of approximately 400 km^2. It is a fishing area of social and economic importance in this state, but receives environmental impacts on a daily basis, such as domestic sewage and non-treated industrial effluents, that originate from a densely populated area, of approximately 10,000 industries, besides housing intense port activities and a large petrochemical complex (Neves et al. 2007). Thus, its drainage basin is an impacted area and suffers the effects of organic matter, oils and several other compounds, including heavy metals and alterations in the environment and the local benthonic and pelagic communities have been observed (Azevedo et al. 2004; Carreira et al. 2002; De Luca Rebello et al. 1986; Perin et al. 1997). The Rodrigo de Freitas Lagoon is more polluted than Jacarepaguá, since it receives more urban discharges. The third sampling site, Silva Jardim, is a non-contaminated reference area of commercial aquaculture, in which fish are raised specifically for human consumption.

Fifty-eight fish specimens were captured from the three sampling sites, 24 males and 33 females, from 2007 to 2010. One individual could not be sexed. Since the objective was to obtain a range of fish representative of the sampled populations from the different sampling sites, nonselective sampling methods, such as corrals, trawl nets and casting nets, were used in the present study, in order to provide meaningful statistical data, as recommended previously (Couture and Rajotte 2003).

Fish were measured, weighed and sexed, and livers and gonads collected and weighed for stress index calculations.

The FCF was calculated as:

$$\mathrm{FCF} = 100\frac{\mathrm{W_T}}{\mathrm{L_T^3}} \tag{1}$$

where $\mathrm{W_T}$ = total fish weight (g) and $\mathrm{L_T}$ = total fish length (cm).

The Gonadossomatic Index (GSI) and the Hepatossomatic Index (HSI) were calculated according to Maddock and Burton (1998):

$$\mathrm{GSI} = 100\frac{\mathrm{W_G}}{\mathrm{W_T}} \tag{2}$$

where $\mathrm{W_G}$ = gonad weight (g) e $\mathrm{W_T}$ = total fish weight (g);

$$\mathrm{HSI} = 100\frac{\mathrm{W_L}}{\mathrm{W_T}} \tag{3}$$

where $\mathrm{W_L}$ = liver weight (g) e $\mathrm{W_T}$ = total fish weight (g).

These indices, when calculated for same sex specimens has very similar or very distinct values in some of the sampling sites, although they are diferentially polluted (Table 1). The FCF means are non-significant between sex at Jacarepaguá and the Rodrigo de Freitas lagoon, but significant in fish sampled from Silva Jardim, with p = 0.0000 when using t-Student test.

Table 1 Confidence interval (95 %) of the fish stress indices by sex and sampling site (n = 58)

Index	FCF		GSI		HSI	
Site	♂	♀	♂	♀	♂	♀
Jacarepaguá	[1.71 ,1.83]	[1.71,1.78]	[0.48,0.50]	[1.89,1.91]	[2.56, 2.67]	[1.48,1.58]
R.F.Lagoon	[1.89,1.94]	[1.88,1.93]	[2.33,2.36]	[1.28,1.30]	[1.21,1.31]	[1.36,1.47]
Silva Jardim	[0.75,0.81]	[0.62,0.67]	[1.18,1.20]	[2.04,2.06]	[2.91,3.03]	[2.51;2.61]

2.2 Construction of the rule base

In this study the classes (sampling site by sex) can be modeled as a normal multivariate distribution, since the stress indicator distribution of 58 fish per class was symetrical and the corrrelation tests between these indicators were non-significant.

The sampling covariance matrix of the stress indicators of the 58 sampled fish was used to generate 3,000 samples of a multinormal distribution, using a Monte Carlo simulation (Gentle 2003) with the computational software Matlab 7.0, with 500 samples for each sampling site and for both sexes. This data was then used in the multinomial logistic regression in order to obtain the probabilities used in the construction of the rule base of the proposed fuzzy model.

A multinomial logistic regression model was used, in which the response variable represents the fish sampling sites. A collection of r + 1 independent variables are considered, denoted by $X = (X_0, X_1, \ldots, X_r)$, where $x = (x_0, \ldots, x_r)$, with $x_0 = 1$ and the random nominal variable that can assume levels defined by 0, 1, .., q. The logit is defined comparing $Y = k$ with $Y = 0$, $k \in \{i, \ldots, q\}$. The latter is named the reference category.

In this study $k = 2$, because, representing the sampling site, is Jacarepaguá ($Y = 1$) and Silva Jardim ($Y = 2$) that are compared to the Rodrigo de Freitas Lagoon ($Y = 0$). We used $r + 1 = 4$ with x_1, x_2, x_3 e x_4 representing, respectively, the FCF, GSI, HSI and sex.

The conditioned probabilities are estimated by Eq. 1:

$$\hat{P}(Y = k|x) = \frac{e^{g_k(x)}}{1 + \sum_{k=1}^{q} e^{g_k(x)}} \quad \text{and} \quad \hat{P}(Y = 0|x) = \frac{1}{1 + \sum_{k=1}^{q} e^{g_k(x)}} \tag{1}$$

For the Logistic Regression the G statistic was used, that tests if all the estimated coefficients are simultaneously different from zero, as well as the quality Adjustment Pearson and Deviance tests, using the Statistica 7.0 software.

2.2.1 Fuzzy logic

Considering that the clusters formed for individual classification using the three stress indices by sex and sampling site were not mutually exclusive, a fuzzy system for the treatment of numerical classification was used, based on the knowledge acquired by the logistic regression results.

According to fuzzy logic, the adequate property to be used in the present study was the associative property. Thus, the following combinations of the pertinence degree (μ) were analyzed:

$$\mu\,(\mathrm{FCF}) \cap \mu\,(\mathrm{GSI}) \cap \mu\,(\mathrm{HSI}) \cap \mu\,(\mathrm{Sex})$$

The defuzzyfication determined the sampling site using the mass center of the fuzzy answer that, once projected on the universe of the discourse of the output variable, resulted in value between 0 and 1 (mapping scale of the sampling sites), used to determine the sampling site. The input and output variables of the fuzzy system were the same as the logistic regression analysis (Sex, FCF, GSI and HSI, and sampling sites, respectively).

The Mamdani inference machine with the T-Norm operator was used, because this has both intuitive appeal and computational simplicity (Wang 1997). It combines "IF-THEN" production rules from a rule base, in the mapping of an input fuzzy set to an output fuzzy set using the T-Norm (Min) or S-Norm (Max) operator.

The frequency distribution of the 3 stress indices were used to specify the fuzzy clusters associated to these input variables. The rule base of the system was constructed by previous knowledge obtained by the logistic regression results.

The sampling sites were distributed in a scale between 0 and 1, by three trapezoidal fuzzy clusters {Jacarepaguá: 0.0;0.2,0.4; RFLagoon: 0.2;0.4;0.6;0.8; Silva Jardim: 0.6;0.8 and 1.0.}. The following combinations used for the pertinence degree (μ) were analyzed:

$$\{\mu\,(\mathrm{FCF})\,,\mu'\,(\mathrm{FCF})\}\text{ and }\{\mu\,(\mathrm{GSI})\,,\mu'\,(\mathrm{GSI})\}\text{ and }\{\mu\,(\mathrm{HSI})\,,\mu'\,(\mathrm{HSI})\}\text{ and }\{\mu\,(\mathrm{Sex})\,,\mu'\,(\mathrm{Sex})\}$$

The consequence of each rule was then obtained by the fuzzy intersection (in which the minimum was used) of the input variables, and their value represents the rule activation degree.

3 Results and discussion

3.1 Statistical analysis and perceptive data mapping

The extreme values of the clusters were obtained by analyzing the sample data, calculating the three stress indices for each sampling site for further statistical analysis, classifying them in Low, Medium or High. Figures 1, 2 and 3 show the box-plots for the FCF, GSI and HSI by sex and sampling site, respectively. Two types of FCF were considered, Low—FCF < 1.2 and High—FCF ≥ 1.2. Three GSI types were considered: Low—GSI < 1.2; Medium—$1.2 \leq$ GSI < 1.5 and High—GSI > 1.5. Two HSI types were considered: Normal—HIS < 2.0 and High—HIS ≥ 2.0.

Thus, non-mutually exclusive classification clusters were created for individual classification using the three stress indices and sex per sampling site.

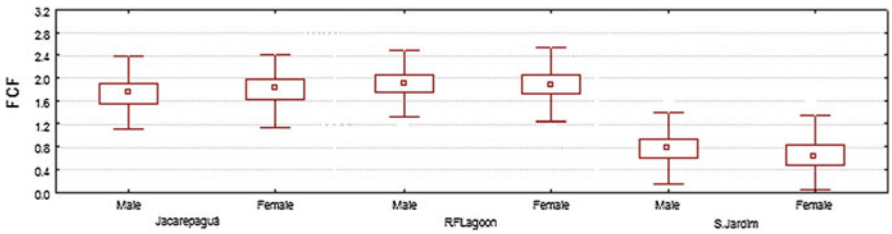

Fig. 1 FCF *box-plot* by sex and sampling site

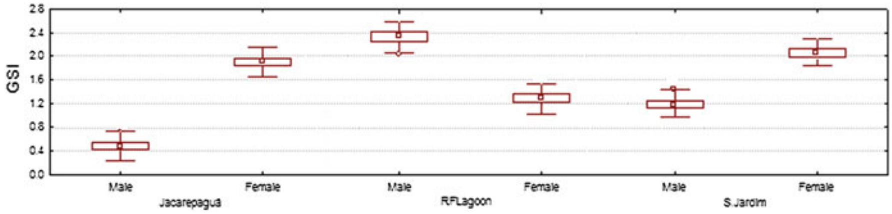

Fig. 2 GSI *box-plot* by sex and sampling site

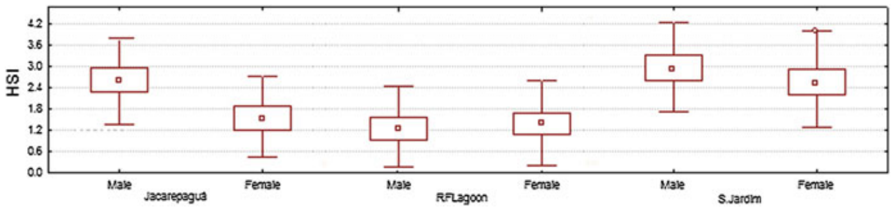

Fig. 3 HSI *box-plot* by sex and sampling site

3.2 Logistic regression results

In the logistic regression solution, the sample data was separated into two sets representing the considered variable domains, one with 70 % of the data (2,100 records), and another with 30 % of the data (900 records), which were used for training and solution validation, respectively.

The sampling site considered as reference for the model, $Y = 0$, was the Rodrigo de Freitas Lagoon and the male gender was used as a reference to fish sex. Table 2 shows that, for both logits, all the variables were significant when comparing both the other sampling sites and the sex of those specimens to the Rodrigo de Freitas Lagoon.

Thus, the adjusted models used to predict the likelihood of sampling site of fish at the Rodrigo de Freitas lagoon ($Y = 0$), Jacarepaguá ($Y = 1$) and at Silva Jardim ($Y = 2$), are represented by Eqs. (2), (3) and (4), with x_1, x_2, x_3 and x_4 representing, respectively, the FCF, GSI, HSI and sex:

$$\hat{P}(Y = 0|x) = \frac{1}{1 + e^{8.45 - 15.17x_1 + 2.31x_2 + 3.56x_3} + e^{0.93x_1 - 0{,}73x_2 + 1.66x_3 - 0.91x_4}} \tag{2}$$

$$\hat{P}(Y = 1|x) = \frac{e^{0.93x_1 - 0.71x_2 + 1.66x_3 - 0.91x_4}}{1 + e^{8.45 - 15.17x_1 + 2.31x_2 + 3.56x_3} + e^{0.93x_1 - 0.71x_2 + 1.66x_3 - 0.91x_4}} \tag{3}$$

$$\hat{P}(Y = 2|x) = \frac{e^{8.45x_1 - 15.17x_2 + 2.31x_3 + 3.56x_4}}{1 + e^{8.45 - 15.17x_1 + 2.31x_2 + 3.56x_3} + e^{0.93x_1 - 0.71x_2 + 1.66x_3 - 0.91x_4}} \tag{4}$$

The cutoff represents the lower limit on the probability of classifying individuals at the reference level ($Y = 0$). The best cutoff point was 0.6, because it showed the highest score (81.64 %) when compared to points 0.4 and 0.5 using the validation results with 30 % of the sample occurrences ($n = 700$).

Table 2 Logistic regression table

Predictor	Coefficient	Stand error	p	Odds ratio	95 % CI. Lower	95 % CI. Upper
Logit 1: (Jacarepaguá/R. F. Lagoon)						
Constant	0.41	0.65	0.5313			
x_1-FCF	0.93	0.26	0.0004	0.40	0.24	0.66
x_2-GSI	−0.71	0.12	0.0000	0.49	0.39	0.62
x_3-HSI	1.66	0.14	0.0000	5.27	3.97	7.00
x_4-Sexo	−0.91	0.15	0.0000	0.40	0.30	0.54
Logit 2: (Silva Jardim / R. F. Lagoon)						
Constant	8.45	2.03	0.0000			
x_1-FCF	−15.17	1.41	0.0000	0.00	0.00	0.00
x_2-GSI	2.31	0.52	0.0000	10.07	3.65	27.77
x_3-HSI	3.56	0.52	0.0000	35.18	12.79	96.77
x_4-Sexo	0.71	0.54	0.1936	2.03	0.70	5.91

Table 3 Groups with the highest percentage sample and validation classifications with the cutpoint (probabilities ≥ 0.6)

Site	Group	% Classification	
		Sample	Validation
Jacarepaguá	Sex = *Male*; FCF = *Normal*; GSI = *Low*; HSI = *High*	98.67	100.00
	Sex = *Female*; FCF = *Normal*; GSI = *High*; HSI = *Normal*	47.51	58.45
RFLagoon	Sex = *Male*; FCF = *Normal*; GSI = *High*; HSI = *Normal*	84.93	89.04
	Sex = *Female*; FCF = *Normal*; GSI = *Medium*; HSI = *Normal*	13.26	19.58
Silva Jardim	Sex = *Male*; FCF = *Low*; GSI = *Medium*; HSI = *High*	97.94	100.00
	Sex = *Female*; FCF = *Low*; GSI = *High*; HSI = *High*	100.00	100.00

Using the cutoff point, the groups with higher classification percentages in the sample and the validation sample are presented in Table 3. We can observe the low classification percentages for females at the Jacarepaguá and Rodrigo de Freitas Lagoon, due to the similarity in the data distribution of the FCF (Fig. 1) and HSI indices (Fig. 3).

4 Fuzzy logic classification results

The fuzzy sets of the three stress indices (FCF, GSI and HSI), are shown in Fig. 4, represented by normal functions as suggested by the frequency distribution histograms.

The fuzzy clusters of the variable sex were represented by mutually exclusive triangular functions. Figure 5 shows a result of 0.195, for example, in the case of an individual with FCF = *Low*, GSI = *Low*, HSI = *Normal* and Sex = *Male*. This value is the result of the defuzzyfication obtained through the center gravity defuzzyfier applied to the fuzzy response system. In this case, it corresponds to the Jacarepaguá sampling site.

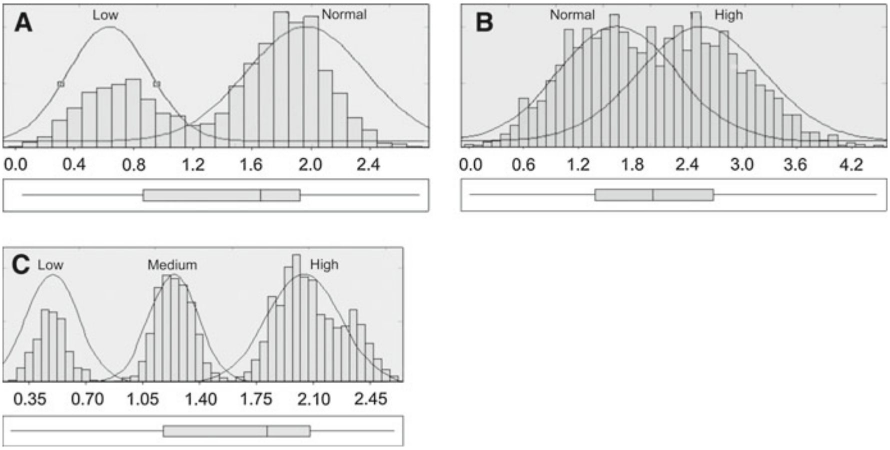

Fig. 4 Fuzzy groups of the three indices (FCF, GSI and HSI): **a** FCF frequency distribution with *low* and *normal* fuzzy groups; **b** HSI frequency distribution with *normal* and *high* fuzzy groups; **c** GSI frequency distribution with *low*, *medium* and *high* fuzzy groups

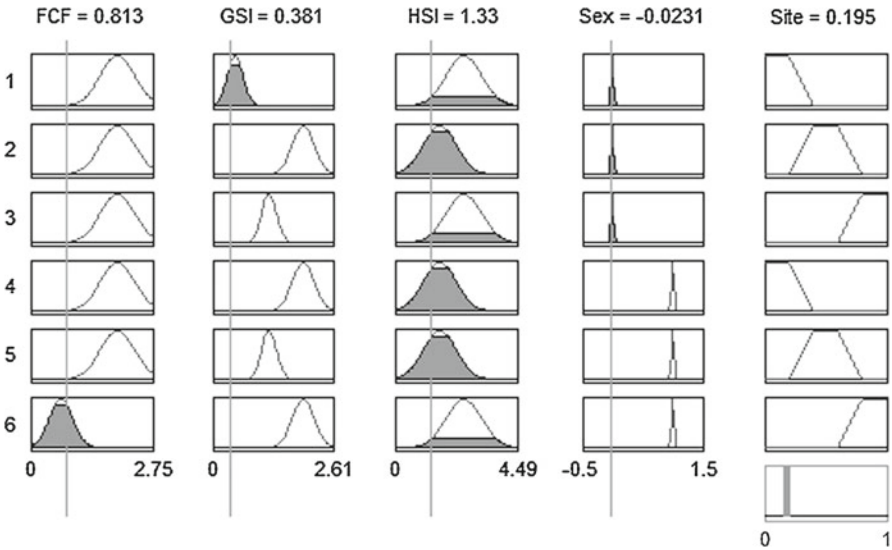

Fig. 5 Fuzzy system output

In the logistic regression solution, the sample data were separated in two groups, representative of the variable domains considered in this study, on with 70 % of the data (2,100 records) and the other with 30 % of the data (900 records), that were used respectively for samples in acquiring the statistical model and validations of the solution.

When processing the validation sample with the fuzzy system, one can observe a great improvement in the classification, especially the Rodrigo de Freitas Lagoon, where the percentage increased from 19.58 to 100.00 % (Table 4).

Table 4 Groups with the highest percentage validation classifications with the cutpoint (probabilities ≥ 0.6) comparing the solutions with the logistic regression and fuzzy logic

Site	Group	% Classification	
		Logistic regression	Fuzzy system
Jacarepaguá	Sex = *Male*; FCF = *Normal*; GSI = *Low*; HSI = *High*	100.00	100.00
	Sex = *Female*; FCF = *Normal*; GSI = *High*; HSI = *Normal*	58.45	100.00
RFLagoon	Sex = *Male*; FCF = *Normal*; GSI = *High*; HSI = *Normal*	89.04	100.00
	Sex = *Female*; FCF = *Normal*; GSI = *Medium*; HSI = *Normal*	19.58	100.00
Silva Jardim	Sex = *Male*; FCF = *Low*; GSI = *Medium*; HSI = *High*	100.00	100.00
	Sex = *Female*; FCF = *Low*; GSI = *High*; HSI = *High*	100.00	100.00

The applications of this combined method can be used in the selection of a specific sampling site, in which the population presents good stress indices. It may also be used in the differentiation of similarly grouped sampling sites, in which individuals present similar stress indices, leading to informed decisions on where to sample feral fish, extremely important in biomonitoring programs. These programs usually use a reference area and one or more contaminated sites to study fish responses to a variety of contaminants, and these statistical tools can be of importance in differentiating these aspects of fish sampling sites.

5 Conclusions

Results demonstrate that a hybrid solution for the construction of a fish sampling site classification methods using stress indices and fish gender can be used. In a first step, a logistic regression was used for acquiring knowledge, establishing the basis for the construction of a fuzzy inference system which, when submitted to a validation sample, presented excellent results in the classification process.

Both methods, when applied, showed different performances. Fuzzy logic was shown to be particularly useful because, unlike the logistic model, it showed the ability to use information obtained from the three fish stress indices in intervals in which the values showed low discrimination. Thus, the logistic regression and fuzzy logic methods complement each other, allowing for the construction of a highly efficient method for classifications such as the one shown in this study, which in turn shows importance in a biomonitoring program scope, in which selecting the more suited sampling site is extremely necessary.

References

Appleyard SA, Renwick JM, Mather PB (2001) Individual heterozygosity levels and relative growth performance in *Oreochromis niloticus* (L.) cultured under Fijian conditions. Aquac Res 32:287–296

Azevedo LA, Bruning IMRD, Moreira I (2004) Hydrocarbon contamination in mussels from Guanabara Bay. Mar Pollut Bull 49:1120–1122

Bassay BM, Makumyaviri AM, Otikeke LS (1997) Evolution do regime alimentaire chez *Tilapia rendalli* en elevage heterogene a Lubumbashi, Congo. Rev Med Vet 148:805–807

Carreira RS, Wagener ALR, Readman JW et al (2002) Changes in the sedimentary organic carbon pool of a fertilized tropical estuary, Guanabara Bay, Brazil: an elemental, isotopic and molecular marker approach. Mar Chem 79:207–227
Couture P, Rajotte JW (2003) Morphometric and metabolic indicators of metal stress in wild yellow perch (Perca flaveseens) from Sudburym, Ontario: a review. J Environ Monit 5:216–221
De Luca Rebello A, Haekel W, Moreira I et al (1986) The fate of heavy-metals in an estuarine tropical system. Mar Chem 18:215–225
Gentle JE (2003) Random number generation and Monte Carlo methods. Springer, Berlin
Graybill FA (1976) Theory and application of linear models. Duxbury Press, Belmont
Hauser-Davis RA, Oliveira TF, Silveira AM et al (2010) Case study: comparing the use of nonlinear discriminating analysis and Artificial Neural Networks in the classification of three fish species: acaras (*Geophagus brasiliensis*), tilapias (*Tilapia rendalli*) and mullets (*Mugil liza*). Ecol Inf 5:474–478
Laflamme JS, Couillard Y, Campbell PGC et al (2000) Interrenal metallothionein and cortisol secretion in relation to Cd, Cu, and Zn exposure in yellow perch, *Perca flavescens*, from Abitibi lakes. Can J Fish Aquat Sci 57:1692–1700
Maddock DM, Burton MP (1998) Gross and histological of ovarian development and related condition changes in American plaice. J Fish Biol 53:928–944
Negnevitsky M (2005) Artificial intelligence: a guide to intelligent systems. Addison Wesley, Reading
Neves RLS, Oliveira TF, Ziolli RL (2007) Polycyclic aromatic hydrocarbons (PAHs) in fish bile (Mugil liza) as biomarkers for environmental monitoring in oil contaminated areas. Mar Pollut Bull 54:1818–1824
Perin G, Wagener AR, Hamacher C (1997) A five year study on the heavy metal pollution of Guanabara Bay sediments (Rio de Janeiro—Brazil) and evaluation of the metal bio-availability by means of geochemical speciation. Water Res 31:3017–3028
Shaw IS, Simões MG (1999) Controle e modelagem fuzzy. Ed. Edgard Blucher, São Paulo
Turban E, Aronson JE (2001) Decision support systems and intelligent systems. Prentice-Hall, New Jersey
Wang L-X (1997) A course in fuzzy systems and control. Prentice-Hall International Inc., New Jersey

Author Biographies

Rachel Ann Hauser-Davis graduated in Biological Sciences from the Federal State University of Rio de Janeiro, in 2005. She has a master's (2008) and doctorate (2012) degree in Analytical Chemistry from the Pontifical Catholic University of Rio de Janeiro. Her research activitites focus on statistical and proteomic analyses of important fish species from Brazil.

Terezinha Ferreira de Oliveira has a degree in Statistics from the Federal University of Ceará (1988), Master in Mathematics from the Federal University of Pará (1997) and Ph.D. in Industrial Engineering from Pontificia Universidade Catolica do Rio de Janeiro (2004). Her research activitites focus on statistical techniques such as fuzzy logic, evolutionary computation and artifical neural networks.

Antônio Morais da Silveira graduated in Electrical Engineering by the Federeal University of Pará (1973), has a Master's degree (1977) in Informatics by the Pontifical Catholic University of Rio de Janeiro and a doctorate in Electrical Engineering by the Federal University of Pará (2003). His research activitites focus on statistical techniques such as fuzzy logic evolutionary computation and artifical neural networks.

João Marcelo Brazão Protázio graduated in Applied and Computational Mathematics (1999) by the State University of Campinas (1999), has a Master's degree in Oil Sciences and Engineerinng (2002) and a doctorate in Biomathematics (2007) by the ZMT—Zentrum für Marine Tropenökologie, at Bremen University—Germany. He has experience in Applied and Computational Mathematics and Geostatistics.

Roberta Lourenço Ziolli graduated in Chemistry (1993) and has a doctorate degree (1999) from the State University of Campinas. She has experience in environmental chemistry analysing environmental impacts by organic contaminants and pesticides.