UTTAR PRADESH JOURNAL OF ZOOLOGY

41(18): 28-34, 2020
ISSN: 0256-971X (P)

ON A LOGISTIC MODEL FOR MORPHOLOGICAL COVARIATION PATTERNS BETWEEN *Colossoma macropomum* AND THE HYBRID *Colossoma macropomum* (♀) x *Piaractus orinoquensis* (♂)

**DANNY VILLEGAS RIVAS[1*], MANUEL MILLA PINO[2],
ZADITH GARRIDO CAMPAÑA[2], MARTÍN GRADOS VÁSQUEZ[3],
ERICK DELGADO BAZAN[2], CÉSAR OSOSRIO CARRERA[3],
YDALIA VELASQUEZ CASANA[4], WILFREDO RUIZ CAMACHO[1],
RICARDO SHIMABUKU YSA[5], JOSÉ PAREDES CARRANZA[3]
AND MANUEL LIZARZABURU AGUINAGA[3]**

[1]Facultad de Ingeniería Forestal y Ambiental, Universidad Nacional de Jaén, Cajamarca, Perú.
[2]Facultad de Ingeniería Civil, Universidad Nacional de Jaén, Cajamarca, Perú.
[3]Escuela de Postgrado, Universidad César Vallejo, Perú.
[4]Escuela de Postgrado, Universidad Nacional de Trujillo, Perú.
[5]Facultad de Ingeniería Mecánica y Eléctrica, Universidad Nacional de Jaén, Cajamarca, Perú.

AUTHORS' CONTRIBUTIONS

This work was carried out in collaboration among all authors. Authors DVR and MMP designed the study, performed the statistical analysis, wrote the protocol and wrote the first draft of the manuscript. Authors WRC, ZGC, EDB, RSY and JPC managed the analyses of the study. Authors MGV, COC, YVC and MLA managed the literature searches. All authors read and approved the final manuscript.

Article Information

Editor(s):
(1) Dr. Osama Anwer Saeed, University of Anbar, Iraq.
Reviewers:
(1) Gustavo Lopes Toledo, State University of Northern Parana, Brazil.
(2) Leandro dos Santos Lima Hohl, Universidade do Estado do Rio de Janeiro, Brazil.

Received: 25 August 2020
Accepted: 30 October 2020
Published: 21 November 2020

Original Research Article

ABSTRACT

Currently, great importance has been given to the study of external morphology, especially in fish, when it is used as a means of identifying hybrids. This paper considers a multiple logistic regression model based on the truss protocol to compare morphological covariation patterns between specimens of *C. macropomum* and the hybrid *C. macropomum* (♀) x *P. orinoquensis* (♂). In this study, 25 specimens of *C. macropomum* and 20 specimens of the hybrid *C. macropomum* (♀) x *P. orinoquensis* (♂), respectively, were analyzed. The method "Truss protocol" or "trusses" was used. Multiple logistic regression model based on the morphological covariation patterns between the two species showed a good fit and allowed to correctly classify most of the

---

**Corresponding author: Email: danny_villegas1@yahoo.com;*

specimens. Only two covariates were significant; eye diameter and fat fin base. Differences between the hybrid and its parent were observed in the head area and in the anterior part of the fish, which were associated with hydrodynamic abilities and with foraging.

Keywords: Morphometry; truss protocol; fishes; nonlinear regression.

1. INTRODUCTION

The species of the genuses *Colossoma* and *Piaractus* are widely distributed in the Amazonian rivers of South America. Both *Colossoma macropomum* and *Piaractus orinoquensis* belong to the Serrasalminae subfamily. And both species have been included within the Serrasalmidae family. However, many authors suggest that these species belong to the Characidae family. The family Characidae is the most diverse family of freshwater fish species in South America. The implementation of morphometric analysis in some species provides scientific knowledge that helps genetic improvement. Currently, the techniques that involve the use of data on the external morphology of individuals are not widely used, since phenotypic plasticity, that is, changing their phenotype in response to changes in the environment, means that the evaluated characters cannot always be generalized for the identification and differentiation of species and hybrids (Pardo et al., [1]; Miranda-Marín and González-Acosta, [2]. It is clear that the morphological characters are physical evidence of the expression of the genotype. Therefore, the differences between specific body characteristics can become very important to establish patterns of differentiation and inheritance (Mateo et al., [3]; Salas et al.,[4]). In continental fish, the morphometric characteristics referring to the anatomical shape have been used to evaluate the productive response in rearing both in natural environments and in captivity. Currently, there are more modern and precise morphometric analysis techniques, such as Geometric Morphometry (Bookstein et al., [5], Trapani [6], Adams et al., [7], Shipunov and Bateman [8]), which together with multivariate statistical analysis and means of direct visualization, constitute one of the most useful tools to describe the biological form and its changes.

Generally, these techniques are based on a set of measured distances between identifiable points on the organisms. In most cases, the measurements (distances between homologous points) present a high correlation, which is exploited in the models that are frequently used to compare between species. The logistic regression model allows evaluating the effect of a set of independent or explanatory covariates on a response variate of binomial distribution, such as the species (*Y*). Therefore, they follow a binomial distribution with parameters (π, m). The binary logistic regression model allows relating the probability of success π with the independent or explanatory covariate. With the logistic transformation, a more flexible model of easy interpretation of the parameters is achieved Montgomery et al., [9]. In this sense, this paper considers a multiple logistic regression model based on the truss protocol to compare morphological covariation patterns between specimens of *Colossoma macropomum* and the hybrid *C. macropomum* (♀) x *Piaractus orinoquensis* (♂).

2. MATERIALS AND METHODS

2.1 Morphological Covariation Patterns between *C. macropomum* and *P. orinoquensis*

In this study, 25 adult specimens of *C. macropomum* and 20 adult specimens of the hybrid *C. macropomum* (♀) x *P. orinoquensis* (♂) with an average weight of 600g, respectively, from artificial ponds of a fish farm in Portuguesa state, Venezuela, were analyzed. Within the sample of each species there are mixed male and female individuals. The method "Truss protocol" or "trusses" Strauss and Bookstein [10] was used, which achieves an exhaustive reconstruction of the shape from the distances between the homologous anatomical landmarks (see Table 1 and Fig. 1). The distances connecting these landmarks form a series of continuous quadrilaterals with their respective internal diagonals (see Fig. 1), which allows detecting differences in shape in the vertical, horizontal, and oblique directions. The limitations in this study is the number of measures necessary to achieve better efficiency in estimating parameters related to the morphology of these species.

The morphological covariation patterns between specimens of *C. macropomum* and the hybrid *C. macropomum* (♀) x *P. orinoquensis* (♂) were studied using a multiple logistic regression model based on the truss protocol in R package. The logistic model is described below (for details see Bartolo et al.[11]):

2.2 Multiple Logistic Regression Model

Let the p independent variables be expressed by the vector $X^{\mathrm{T}} = (X_1, X_2, \ldots, X_p)$ and it relates the probability of a certain independent event occurring denoted by the vector X^{T} with conditional probability $P(Y = 1 / X) = \pi(x)$ as a function of p independent variables that can be quantitative or qualitative depending on the type of study design. The logit of the multiple logistic regression models is presented by the following equation.

$$g(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p,$$

Table 1. Truss measurements from *C. macropomum* and the hybrid *C. macropomum* (♀) x *P. orinoquensis* (♂) specimens

Standar length (X_1)
Tip of snout to end of epiphyseal sulcus (X_2)
Tip of snout to insertion of pectoral fin (X_3)
Anterior edge of the epiphyseal sulcus to the end of the epiphyseal sulcus (X_4)
Anterior edge of the epiphyseal sulcus at the insertion of the pectoral fin (X_5)
Anterior edge of the epiphyseal sulcus when articulating (X_6)
Articulate to insertion of pectoral fin (X_7)
Posterior edge of epiphyseal sulcus to end of dorsal fin (X_8)
Posterior edge of the epiphyseal sulcus at the insertion of the pelvic fin (X_9)
Posterior edge of the epiphyseal sulcus to the insertion of the pectoral fin (X_{10})
Posterior edge of the epiphyseal groove when articulating (X_{11})
Insertion of pectoral fin to insertion of pelvic fin (X_{12})
Dorsal fin base (X_{13})
Anterior edge of dorsal fin to anterior edge of anal fin (X_{14})
Anterior edge of dorsal fin to insertion of pelvic fin (X_{15})
Anterior edge of dorsal fin to insertion of pectoral fin (X_{16})
Insertion of pelvic fin to end of anal fin (X_{17})
Posterior edge of dorsal fin to the fatty fin (X_{18})
Posterior edge of dorsal fin to posterior edge of anal fin (X_{19})
Posterior edge of dorsal fin to anterior edge of anal fin (X_{20})
Posterior edge of dorsal fin to insertion of pelvic fin (X_{21})
Anal fin base (X_{22})
Posterior edge of the fatty fin to the last scale of the lateral line (X_{23})
Posterior edge of fatty fin to posterior edge of anal fin (X_{24})
Posterior edge of the fatty fin to the anterior border of the anal fin (X_{25})
Posterior edge of the fatty fin to the anterior border of the anal fin (X_{26})
Eye diameter (X_{27})
Head length (X_{28})
Fat fin base (X_{29})

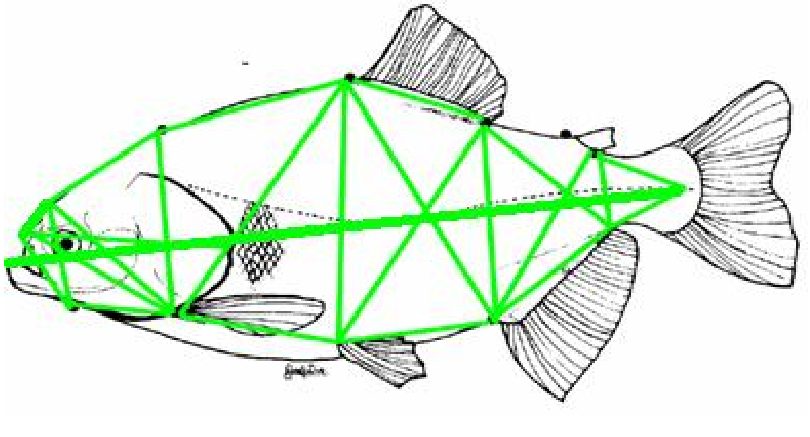

Fig. 1. Location of homologous points and distances measured on the left lateral profile of *C. macropomum* and the hybrid *C. macropomum* (♀) x *P. orinoquensis* (♂)

where $\beta_0, \beta_1, \ldots, \beta_p$ are parameters of the multiple logistic regression model. In this case the logistic regression model is

$$\pi(x) = p_j = \frac{e^{g(x)}}{1 + e^{g(x)}},$$

2.2.1 Parameter estimation

Assuming we have a sample of n independent observations $(x_i, y_i), i = 1,2,\ldots,n$. Fitting the model requires that we obtain estimators of the vector $\beta^{\mathrm{T}} = (\beta_0, \beta_1, \ldots, \beta_p)$. The resulting likelihood equations can be expressed as follows

$$\sum_{i=1}^{n} [y_i - \pi(x_i)] = 0$$

$$\sum_{i=1}^{n} x_{ij}[y_i - \pi(x_i)] = 0$$

para $i = 1,2,\ldots,n$. $\hat{\beta}$ denotes the solution of these equations.

2.2.2 Hypothesis testing for the logistic regression model

The likelihood ratio test for the total significance of the *p* coefficients for the independent variables in the model is based on the *G* statistic.

$$G = 2\left\{\sum_{i=1}^{n} [y_i \ln(\hat{\pi}_i) + (1 - y_i)\ln(1 - \hat{\pi}_i)] - [n_1 \ln(n_i) + n_0 \ln(n_0) - n\ln(n)]\right\}$$

The fitted values, $\hat{\pi}_i$, on the model are based on the vector containing p + 1 parameters, $\hat{\beta}$, on the null hypothesis that the *p* coefficients for the covariates in the model are equal to zero. The *G* statistic has a Chi-square distribution with *p* degrees of freedom.

The Wald test is obtained from the calculation of the following matrix

$$W = \hat{\beta}^{\mathrm{T}}\left[V\hat{a}r(\hat{\beta})\right]^{-1}\hat{\beta}$$

$$= \hat{\beta}^{\mathrm{T}}(XVX)\hat{\beta},$$

where *V* is a diagonal matrix of dimension $n \times n$ with elements $\hat{\pi}_i(1 - \hat{\pi}_i)$ and W has a Chi-square distribution with p + 1 degrees of freedom on the hypothesis that each of the p + 1 coefficients are equal to zero.

2.2.3 Analysis of residuals for the logistic regression model

2.2.3.1 Pearson residuals

Pearson residuals are defined by

$$r_j = \frac{y_i - m_j\hat{p}_j}{\sqrt{m_j\hat{p}_j(1 - p_j)}},$$

where, y_i represents the number of responses, $y = 1$, among the m_j individuals with $X_j = x$ (some individuals having the same *x* value), $j = 1,\ldots,p$.

$$\hat{p} = \hat{\pi}(x) = \frac{e^{\hat{g}(x)}}{1 + \hat{g}(x)},$$

and $g(x) = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \cdots + \hat{\beta}_p x_p$.

Thus, a Pearson residual with an absolute value greater than 2 indicates an outlier.

Pearson's χ^2 statistic is the sum of squares of the Pearson residuals.

$$\chi_P^2 = \sum_{i=1}^{J} r_J^2$$

2.2.3.2 Standardized pearson residuals

The standardized Pearson residuals are defined by:

$$r_{sj} = \frac{r_j}{\sqrt{1 - h_j}},$$

where, r_j is the Pearson residuals and h_j is the leverage, that is, the element of the main diagonal of the matrix *H*.

2.2.3.3 Deviance residuals

Deviance residuals are defined as:

$$d_j = \pm\left\{2\left[y_i \ln\left(\frac{y_j}{m_j\hat{p}_j}\right) + (m_j - y_j)\ln\left(\frac{m_j - y_j}{m_j(1 - p_j)}\right)\right]\right\}^{1/2}$$

Deviance is the sum of squares of deviance residuals.

$$\chi_D^2 = \sum_{j=1}^{p} d_j^2$$

If the deviance is greater than 4 in absolute value then the corresponding observation is outlier.

3. RESULTS

In Table 2, only two covariates; eye diameter (X_{27}) and fat fin base (X_{29}) suggest that there are characteristics associated with the morphological covariation patterns that allow differentiation between redundant specimens of *C. macropomum* and the hybrid *C. macropomum* (♀) x *P. orinoquensis* (♂). These two covariates are associated with morphological covariation patterns that make a difference in the head area; eye diameter (X_{27}), and in the anterior part of the fish; fat fin base (X_{29}). In Table 3, the results of the Hosmer-Lemeshow test (P = 0,5348) suggest that there are insufficient reasons to reject the hypothesis associated with the goodness of fit of the logistic model associated with the morphological covariation patterns between these two species. Similarly, in Table 4, the results of the classification show that, with this model, 97.77% of the specimens from *C. macropomum* and the hybrid *C. macropomum* (♀) x *P. orinoquensis* (♂) were correctly classified.

4. DISCUSSION

In Table 2, these two covariates; eye diameter (X_{27}) and fat fin base (X_{29}) are characteristics associated with hydrodynamic abilities and to the foraging for food. Regarding the morphological covariation patterns based on a multiple logistic regression model presented in Table 2, these results coincide with those reported by Pineda et al., [12] who used principal component analysis for the morphometric comparison between males and females of *C. macropomum* maintained in ponds, and those reported by Villegas et al., [13] in a multivariate analysis that allowed a morphometric comparison of a hybrid originated from *C. macropomum* and *P. orinoquensis*, and those reported by Villegas et al., [14] when studying the redundancy in morphological covariation patterns between *C. macropomum* and *P. orinoquensis*. The foregoing reveals what was indicated by Porras-Rivera and Rodríguez-Pulido [15] and Conte-Grand et al., [16], who point out that external morphology is not always reliable when used as the only means of

Table 2. Multiple logistic regression model based on the morphological covariation patterns between specimens of *C. macropomum* and the hybrid *C. macropomum* (♀) x *P. orinoquensis* (♂)

Covariate	$\widehat{\beta}_p$	Standard error	Z value	P value
(X_1)	-0,01188	8,480e+03	0,0000	0,9999
(X_2)	-0,6376	1,320e+05	0,0000	0,9999
(X_3)	0,03115	2,316e+04	0,0000	0,9999
(X_4)	1,047	7,612e+04	0,0000	0,9999
(X_5)	-0,1244	5,755e+04	0,0000	0,9999
(X_6)	3,859	2,493e+05	0,0000	0,9999
(X_7)	-1,373	1,239e+05	0,0000	0,9999
(X_8)	-1,963	7,200e+04	0,0000	0,9999
(X_9)	1,962	1,079e+05	0,0000	0,9999
(X_{10})	0, 2834	2,671e+04	0,0000	0,9999
(X_{11})	-1,156	1,461e+05	0,0000	0,9999
(X_{12})	0,6869	7,108e+04	0,0000	0,9999
(X_{13})	-0,7860	1,692e+05	0,0000	0,9999
(X_{14})	-0,5908	6,472e+04	0,0000	0,9999
(X_{15})	0,5360	4,044e+04	0,0000	0,9999
(X_{16})	0,2019	5,075e+03	0,0000	0,9999
(X_{17})	-0,04323	3,722e+04	0,0000	0,9999
(X_{18})	2,327	5,634e+04	0,0000	0,9999
(X_{19})	-0,1525	1,869e+04	0,0000	0,9999
(X_{20})	0,4803	2,410e+04	0,0000	0,9999
(X_{21})	-2,082	4,168e+04	0,0000	0,9999
(X_{22})	-1,798	1,559e+05	0,0000	0,9999
(X_{23})	0,07968	1,821e+04	0,0000	0,9999
(X_{24})	4,088	1,465e+05	0,0000	0,9999
(X_{25})	1,573	1,546e+05	0,0000	0,9999
(X_{26})	-0,7722	8,781e+04	0,0000	0,9999
(X_{27})	-9,612	2,078e+05	1,8807	0,0300
(X_{28})	-0,8812	8,297e+04	0,0000	0,9999
(X_{29})	-4,573	1,560e+05	3,5999	0,0000

Table 3. Statistics related to morphological classification of specimens of *C. macropomum* and the hybrid *C. macropomum* (♀) x *P. orinoquensis* (♂) based on a multiple logistic regression model

N	-2 Log Likelihood	AIC	Hosmer-Lemeshow test (*P* value)	Correctly classified cases (%)
45	-0,0000	60	0,5348	0,9777

Table 4. Morphological classification of specimens of *C. macropomum* and the hybrid *C. macropomum* (♀) x *P. orinoquensis* (♂) based on a multiple logistic regression model

	Fitted		Correct percentage
Observed	***C. macropomum***	**Hybrid**	
C. macropomum	19	0	95%
Hybrid	0	25	100%
Overall correct percentage	97,77%		

identification, particularly for hybrid individuals beyond the first generation. Furthermore, Santos [17] emphasizes the importance of studying the morphometry of a certain species to know the relationships between the different body measurements; which would allow the characterization of the carcasses and the meat yields. In this sense, Villegas et al. [14] suggests that the implementation of the morphometric analysis in some species contribute scientific knowledge that helps genetic improvement.

5. CONCLUSIONS

Multiple logistic regression model based on the morphological covariation patterns between specimens of *C. macropomum* and the hybrid *C. macropomum* (♀) x *P. orinoquensis* (♂) showed a good fit and allowed to correctly classify most of the specimens. Only two covariates were significant; eye diameter and fat fin base.

Differences were observed in the area of the head and in the anterior part of the fish between the hybrid and its parent. The morphological differences between these two species were evidenced in covariates associated with hydrodynamic abilities and with foraging. Finally, the results of this research suggest a comparative study between the hybrid *C. macropomum* (♀) x *P. orinoquensis* (♂) and *P. orinoquensis* based on a multiple logistic regression model.

ETHICAL APPROVAL

As per international standard or university standard ethical approval has been collected and preserved by the authors.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Pardo R, Scott S, Vila I. Analysis of shapes in Chilean species of the genus Trichomycterus (Osteichthyes: Siluriformes) using geometric morphometry. Gayana. 2005;69(1):180-3.
2. Miranda-Marín MI, González-Acosta AF. Phylogenetic inferences of the genus eugerres jordan and evermann, 1927, through the Application of Geometric Morphometry. In: Del Moral Flores LF, Martínez JA, López JF, Tello JL. (Eds.). Ichthyological research in mexico. Mexico City, National Autonomous University of Mexico. 2012:97-106.
3. Mateo JF, Delgado N, López H. Caracterización morfométrica del híbrido yaque pintado (*Pseudoplatystoma fasciatum* x *Leiarius marmoratus*) y sus progenitores (Siluriformes: Pimelodidae). Rev. Fac. Cienc. Vet. U. C. V. 2008;49(1):47-60.
4. Salas D, Véliz D, Scott S. Morphological differentiation in species of the genus Cheirodon (Ostariophysi: Characidae) by traditional and geometric morphometry. Gayana. 2012;76 (2):142-52, 2012.
5. Bookstein FL, Chernoff R, Elder J, Humpries G, Smith Y, Strauss R. Morphometrics in evolutionary biology. The Academy of Natural Sciences of Philadelphia, Michigan; 1985.
6. Trapani J. Geometric morphometric analysis of body -form variability in *Cichlasoma minckleyi*, the Four Cienagas cichlid. Environmental Biology of Fishes. 2003;68:357-369.

7. Adams DC, Rohlf FJ, Slice DE, Geometric morphometrics: ten years of progress following the “Revolution. Italian Journal of Zoology. 2004;71:5–16.
8. Shipunov AB, Bateman RM. Geometric morphometrics as a tool for understanding Dactylorhiza (Orchidaceae) diversity in European Russia. Biological Journal of the Linnean Society. 2005;85(1):1-12.
9. Montgomery D, Peck E, Vining, G. Introduction to linear regression analysis. CECSA 1st edition in Spanish. Mexico. 2002;588.
10. Strauss R, Bookstein F. The truss: body form reconstructions in morphometrics. Syst. Zool. 1982;31(2):113-135.
11. Bartolo F, Giraldo O, Salinas A. Análisis de diagnóstico en e modelo de regresión logística: Una aplicación. Pesquimat. 2007;1:55–70.
12. Pineda H, Restrepo L, Ángel, M. Morphometric comparison between males and females of Cachama Negra (*Colossoma macropomum*, Cuvier 1818) kept in a pond. Colombian Journal of Livestock Sciences. 2004;17:24-29.
13. Villegas D, Mile M, Castillo O, Durant K. Multivariate analysis in the morphometric comparison of the hybrid *Colossoma macropomum* X *Piaractus brachypomus*) and its parents. INDES Magazine. 2013;1(1):29-36.
14. Villegas D, Mile M, Pérez Y, Villegas S, Garrido Z, Delgado E, Ruiz W, Velasquez Y, De Souza. Redundancy in morphological covariation patterns between *Colossoma macropomum* and *Piaractus orinoquensis*. Uttar Pradesh Journal of Zoology., [In press]; 2020.
15. Porras-Rivera G, Rodríguez-Pulido JA. Morphometric comparison and characterization of the hybrid (Pseudoplatystoma metaense x Leiarius marmoratus) and its parental lines (Siluriformes: Pimelodidae). Int. J. Morphol. 2019;37 (4):1409-1415.
16. Conte-Grand C, Sommer J, Orti G, Cussac V. Populations of odontesthes (Teleostei: Atheriniformes) in the Andean region of Southern South America: Body shape and hybrid individuals. Neotropical Ichthyology. 2015;13(1):137–150.
17. Santos V. Morphometric and allometric growth of tilapia linhagens (*Oreochromis niloticus*). Master's Thesis in Zootechnics. Brazil: Federal University of Lavras. 2004;86.
