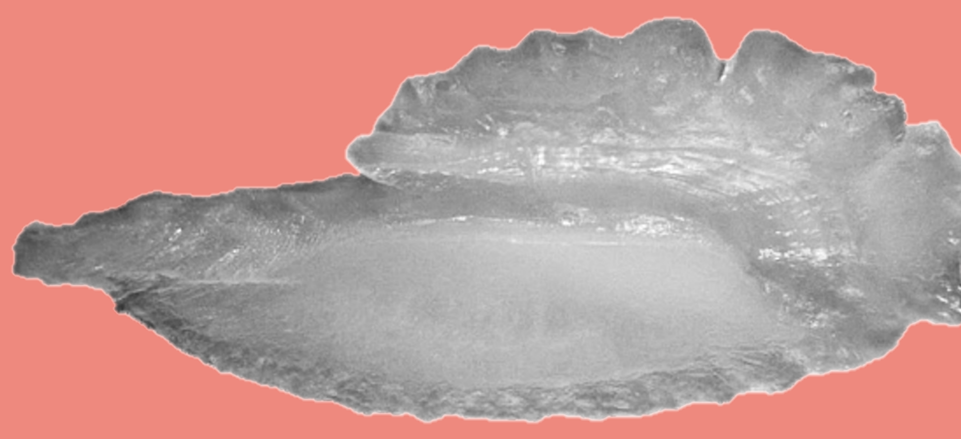
COMPARING SHAPE R PACKAGES USING WAVELET FUNCTIONS FOR CONNECTING FISH POPULATIONS



AUTHORS

Joana Vasconcelos*[1,2], José Luís Otero-Ferrer[3], Alba Jurado-Ruzafa[4], Amalia Manjabacas[5], **Victor M. Tuset[6]**, Antoni Lombarte[5]

AFFILIATIONS

[1]BIOCON, IU-ECOAQUA, Universidad de Las Palmas de Gran Canaria, Spain; [2]MARE - Marine and Environmental SciencesCentre/ARNET – Aquatic Research Network, Agência Regional para o Desenvolvimento da Investigação Tecnologia e Inovação (ARDITI), Portugal; [3]Biostatech, Advice, Training and Innovation in Biostatistics (Ltd.), Santiago de Compostela, Spain; [4]Centro Oceanográfico de Canarias (IEO-CSIC), Santa Cruz de Tenerife, Spain; [5]Department de Recursos Marins Renovables, Institut de Ciències del Mar (ICM-CSIC), Passeig Marítim de la Barceloneta, Catalonia, Spain; [6]IOCAG, Universidad de Las Palmas de Gran Canaria, Spain; ***joanapatricia.reis@ulpgc.es**

INTRODUCTION

Otolith shape analysis has been widely used to delineate structure of populations of marine fish species, mostly based on Elliptic Fourier descriptors. Alternatively, the wavelet transform of the radii length improves the otolith contour interpretation, quantifying the otolith irregularities and determining accurately their position.

The worldwide application of the mathematical R environment allowed the development of *shapeR* package for the otolith shape analysis. However, we have detected differences in the philosophy of the use of wavelets and some inconsistencies in the results interpretation.

OBJECTIVE

The present study compares protocols and outputs of both packages (*shapeR* vs. *AFORO* team protocol in R) using the blue jack mackerel, *Trachurus picturatus*, from the Canaries and Madeira archipelagos as a case study.

METHODOLOGY

A total of 550 *sagitta* otoliths of blue jack mackerel, *Trachurus picturatus*, were collected from the waters near the Canary Islands (CI) and Madeira (MD). The ***shapeR* package** (Libungan and Pálsson, 2022) and the ***AFORO* team protocol in R** were used for the otolith contour shape analysis and the outputs compared. The performance of parametric and non-parametric classification algorithms was also evaluated using the *caret* R package (Kuhn, 2008).

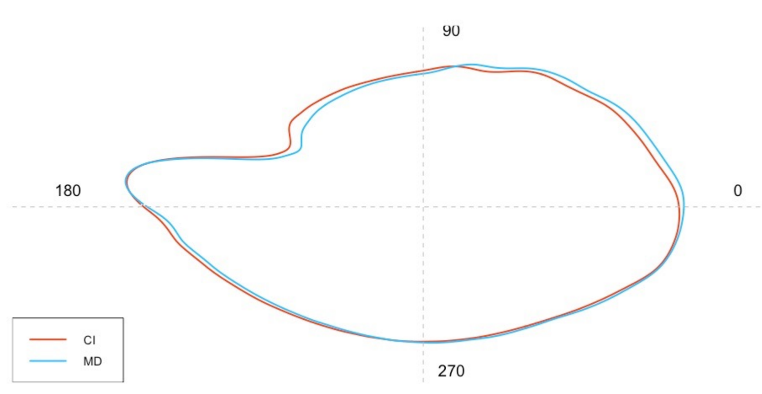
BACKGROUND ON WAVELET ANALYSIS

Contour description using wavelets

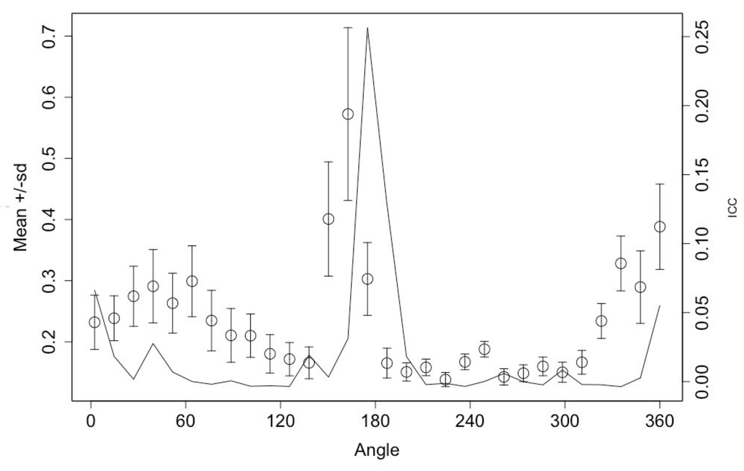
shapeR package

(LIBUNGAN AND PÁLSSON, 2022)

- Starting point is not linked to a morphological feature.
- 1024 equidistant angles (0.35 degrees per point).
- Pyramidal analysis of levels (2^n).
- Filter: Daubechies least asymmetric wavelets (Daubechies, 1992).
- 10^{th} wavelet levels.
- Automatic level selection (e.g., 5^{th} = 62 coefficients).



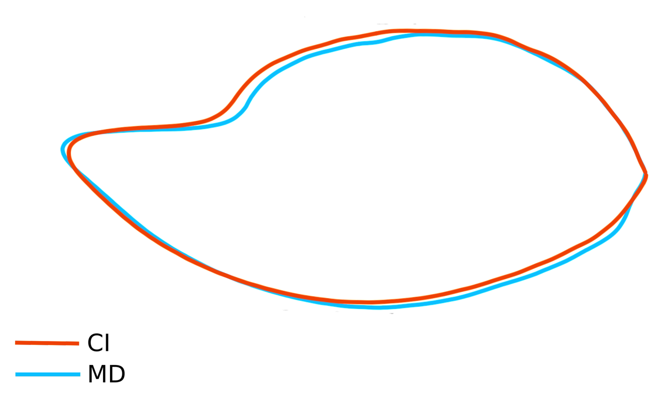
Mean otolith shape based on wavelet reconstruction using the *shapeR* package in R. Canary (CI) and Madeira (MD).



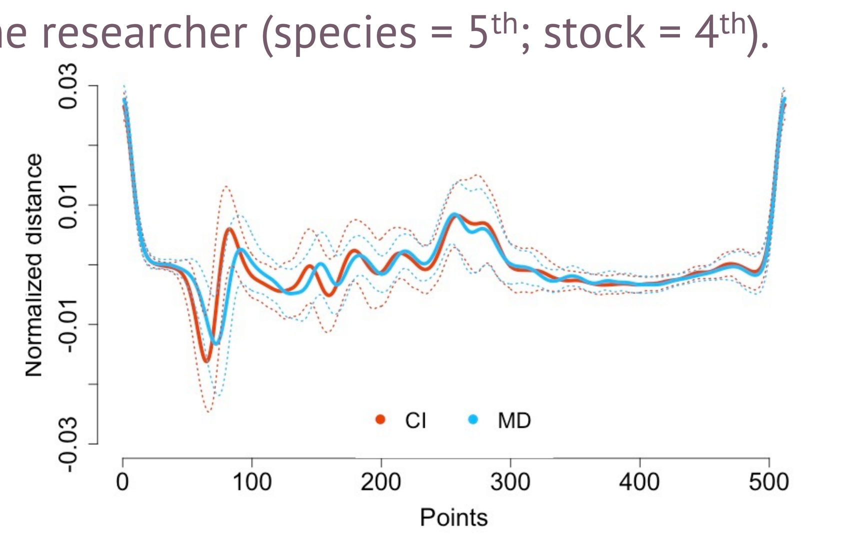
Mean (circles) and standard deviation (sd, vertical bars) of the wavelet coefficients, standardized for the otolith size. ICC, intraclass correlation (solid line).

AFORO team protocol

- Algorithm detects the point on the rostrum furthest from the centroid.
- 512 equidistant points according to the perimeter.
- Filter: B-3 splines (Parisi-Baradad et al., 2005).
- Analysis by levels, keeping the dimensions.
- 9^{th} wavelet levels.
- Level to be analysed are up to the researcher (species = 5^{th}; stock = 4^{th}).



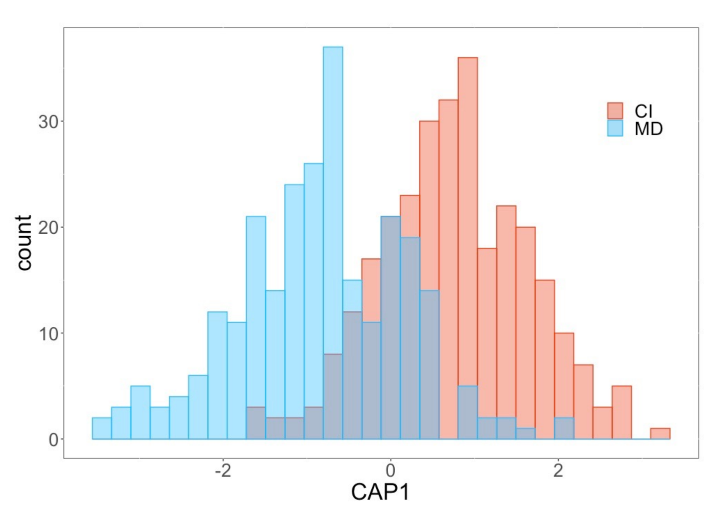
Mean otolith shape based on wavelet reconstruction using the *AFORO* team protocol in R. Canary (CI) and Madeira (MD).

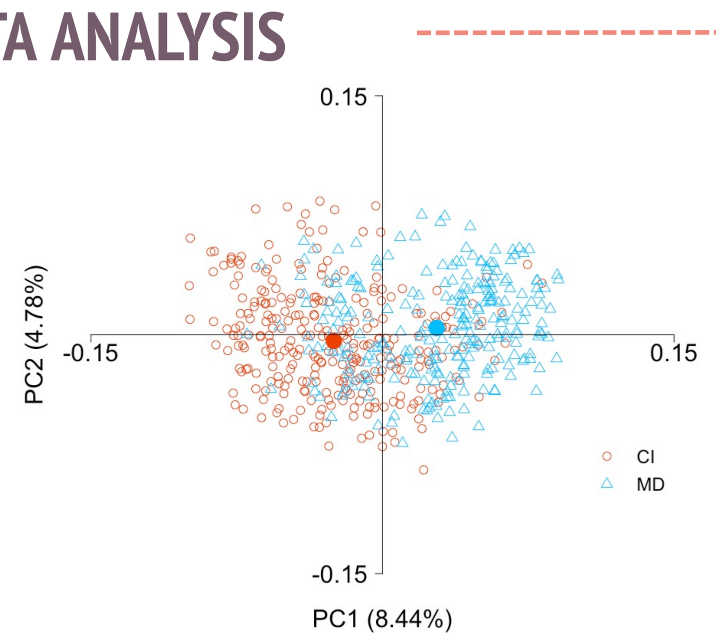


Mean (solid line) and standard deviation (dotted line) of the otolith contour decomposition using wavelet coefficients obtained from *AFORO* team protocol in R. Canary (CI) and Madeira (MD).

DATA ANALYSIS

1. Correlation between size and standard coefficients.
   - ✗ Excluded when there are differences between groups.
   - ✓ Not exclude coefficients are used in the following analyses.
2. Standard coefficients.
3. Canonical Covariate Analysis.





1. PCA and components selection (Broken stick model).
2. Regression between PC' components and size, new PCA matrix.
   - ✓ Residuals for correlated.
   - ✓ Components for non-correlated.
3. Number of cases per group.
   - ✓ Small (<100): LOOCV method.
   - ✓ Medium (101-200): depending on otolith contour detail.
   - ✓ Large (>200): training and testing sets.

Confusion/Error matrix (LOOCV) and classification accuracy of individuals to their sampling origin.

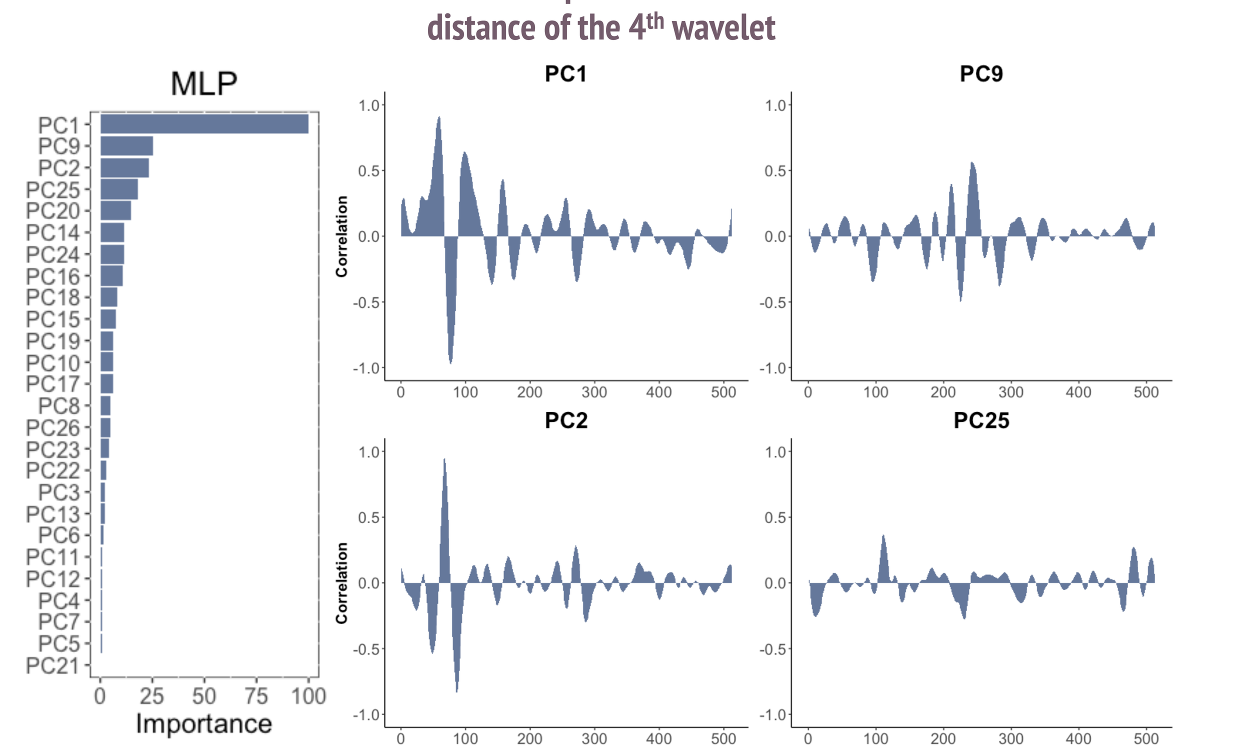
R Package	Algorithm	Prediction	References CI	References MD	Performance measures Accuracy	Kappa	% Accuracy
shapeR	LDA*	CI	251	39			86.55
		MD	39	221			85.00
		Total	290	260	0.858	-	
	MLP	CI	286	4			98.62
		MD	4	256			98.46
		Total	290	260	0.986	0.971	
AFORO team protocol in R	MLP	CI	290	1			100
		MD	0	259			99.62
		Total	290	260	0.998	0.996	

*method in *shapeR*

CLASSIFICATION ALGORITHMS

Linear Discriminant Analysis (LDA)*
Multi-layer Perceptron (MLP)
Quadratic Discriminant Analysis
K-nearest neighbors
Neuronal networks
Support Vector Machine
Random Forest
Naïve Bayes

Correlation of the most relevant PC components for the classification with normalized distance of the 4^{th} wavelet



ADVANTAGES VS. DISADVANTAGES

shapeR:

- Unique package in R.
- Both Elliptic Fourier and morphometric data.
- Better reconstruction of the mean otolith shape (1024 points).
- Starting point is not unified among works.
- Only Polar Coordinates.
- Meaning of coefficients is complex.
- Correlation between ICC and discrimination.
- Working with folders per group.

AFORO team protocol:

- Both Elliptic Fourier and morphometric data.
- Both Polar and Cartesian Coordinates.
- Standardisation of images.
- Different levels of signal are available for the researcher.
- Visually more understanding.
- R code not published.
- Online data from *AFORO* web, slow option.

Acknowledgements | This project has been partially funded by the EU through the European Maritime and Fisheries Fund (EMFF) within the National Program of collection, management and use of data in the fisheries sector and support for scientific advice regarding the Common Fisheries Policy.

Funding | The first author (JV) was supported by the 'Grants for the Viera y Clavijo training program for researchers' of the Government of the Canary Islands (VIERA Y CLAVIJO-2022-CIENCIAS-1).

REFERENCES | Daubechies. 1992. Ten Lectures on Wavelets, SIAM, p. 194. **Kuhn, M. 2008**. Building Predictive Models in R Using the caret Package. J. Stat. Softw. 28(5), 1–26. **Libungan, L., Palsson, S. 2022**. shapeR: Collection and Analysis of Otolith Shape Data. R package version 1.0-1. **Parisi-Baradad, V.**, Lombarte, A., García-Ladona, E., Cabestany, J., Piera, J. and Chic, O. 2005. Otolith shape contour analysis using affine transformation invariant wavelet transforms and curvature scale space representation. Mar. Freshw. Res. 56, 795–804.

Investigating *Aphanopus* spp. in the Atlantic waters using otolith contour: converging population hypotheses

Joana Vasconcelos[*1,2], José Luis Otero-Ferrer[3], Alba Jurado-Ruzafa[4], Antoni Lombarte[5], Carlos Hernández-González[4], Mafalda Freitas[2,6], Víctor M. Tuset[7]

[1]BIOCON, IU-ECOAQUA, Universidad de Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain; [2]MARE - Marine and Environmental Sciences Centre/ARNET - Aquatic Research Network, Agência Regional para o Desenvolvimento da Investigação Tecnologia e Inovação (ARDITI), Funchal, Portugal; [3]Biostatech, Advice, Training and Innovation in Biostatistics (Ltd.), Santiago de Compostela, Spain; [4]Centro Oceanográfico de Canarias (IEO-CSIC), Santa Cruz de Tenerife, Spain; [5]Departamento de Recursos Marins Renovables, Institut de Ciències del Mar (ICM-CSIC), Passeig Marítim de la Barceloneta 37–49, 08003 Barcelona, Catalonia, Spain; [6]Direção de Serviços de Monitorização, Estudos e Investigação do Mar, Direção Regional do Mar, Funchal, PORTUGAL.; [7]Instituto de Oceanografía y Cambio Global, IOCAG, Universidad de Las Palmas de Gran Canaria, Unidad Asociada ULPGC-CSIC, Canary Islands, Spain. *joanapatricia.reis@ulpgc.es

7th INTERNATIONAL OTOLITH SYMPOSIUM

INTRODUCTION

The black scabbardfish (*Aphanopus carbo*) and the intermediate scabbardfish (*A. intermedius*) coexist from the North to South Atlantic waters. These sympatric species are landed as a single mixed resource since no external differentiation is noticeable to discern both species. In the past decade, a great effort was made into separating these two scabbardfishes using both genetic and body morphometric analyses.

OBJECTIVE

To test the *Aphanopus* spp. identification using wavelets.
To identify species in unidentified temporal samples.
To hypothesize about *Aphanopus* spp. distribution in Atlantic Ocean.

METHODS AND WORFLOW

1 GENETICALLY IDENTIFIED INDIVIDUALS (GII)

136 otoliths of *Aphanopus carbo* and *A. intermedius* from mainland Portugal, Azores, Madeira, Morocco, the Canaries and Mauritania were genetically identified (Biscoito et al., 2011).

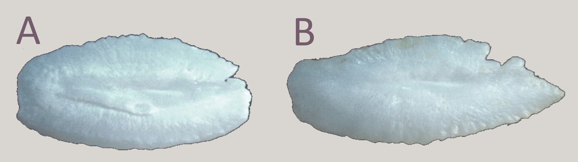


Figure 1 | Otoliths of *Aphanopus carbo* (A) and *A. intermedius* (B). Otolith lengths: A, 7.6 mm; B, 10.0 mm.

2 OTOLITH SHAPE ANALYSIS of GII

- 512 equidistant Cartesian coordinates were extracted considering the rostrum as the origin.
- 5th wavelet (Sadighzadeh et al., 2014).
- PCA and components selection (Broken stick model).
- Regression between PC' components and size, new PCA matrix.

3 SPECIES CLASSIFICATION

Classification and regression training package *caret* (Kuhn, 2008) was used:

- Linear Discriminant Analysis.
- Quadratic Discriminant Analysis.
- **Multi-layer Perceptron (MLP).**
- K-nearest neighbors.
- Neuronal networks.
- Support Vector Machine.

Resampling method: leave-one-out cross-validation (LOOCV) (Marti-Puig et al., 2020).

4 OTOLITHS FOR SPECIES IDENTIFICATION

MADEIRA: 1,975 otoliths (1990, 2003, 2010 and 2021)

AFRICAN COAST: 547 otoliths of from Morocco (2005 and 2006) and Mauritania (2007)

5 TRACING SPECIES IDENTIFICATION

The genetically identified data was used as training set to classify the remaining subset of Madeira and African coast.

RESULTS

Genetically Identified Individuals (GII)

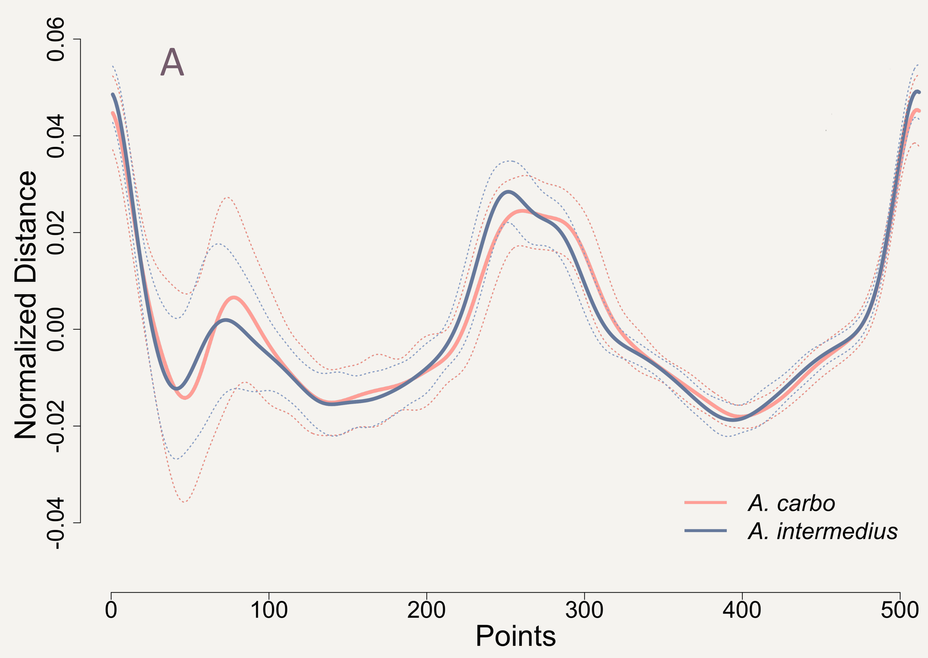


Figure 2 | Mean (solid line) and standard deviation (dotted line) of the otolith contour decomposition using wavelet coefficients (A) and mean otolith shape based on wavelet reconstruction (B). *A. carbo, Aphanopus carbo; A. intermedius, Aphanopus intermedius.*

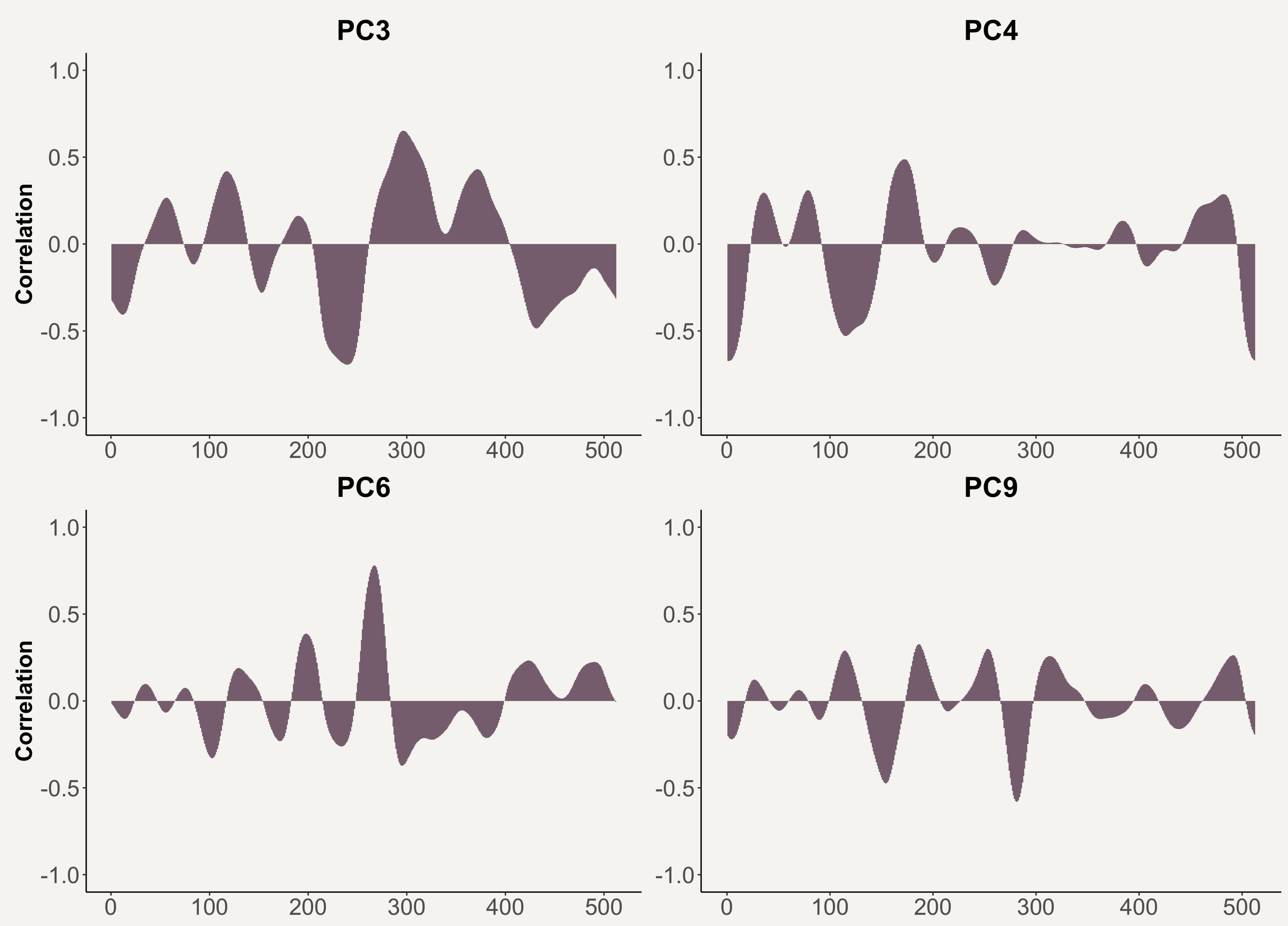
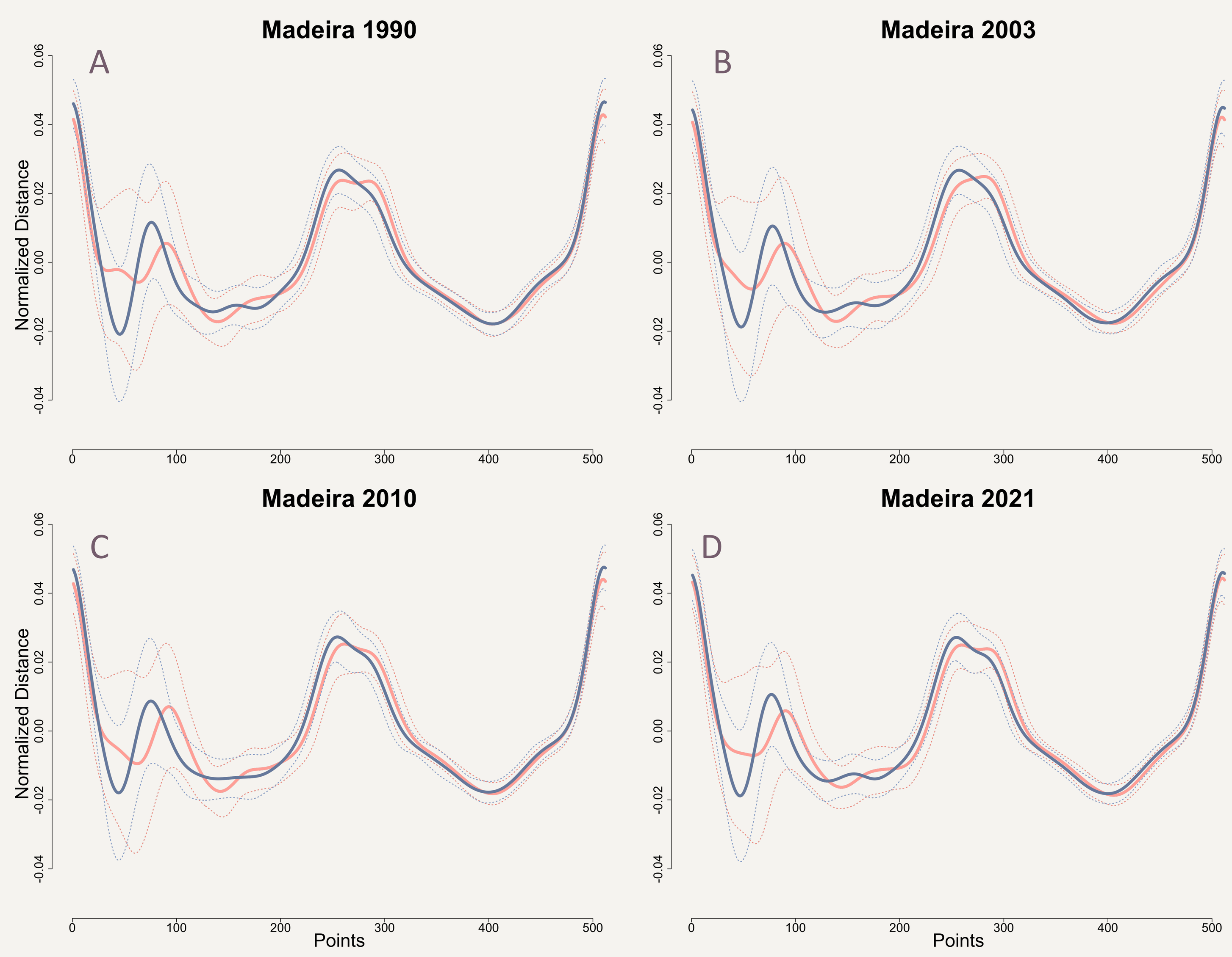


Figure 3 | Correlation of the most relevant PC components for the classification with normalized distance of the 5th wavelet.

Table 1 | Confusion/Error matrix (LOOCV) and classification accuracy of the GII. *A. carbo, Aphanopus carbo. A. intermedius, Aphanopus intermedius.*

Prediction	References		Performance measures		
	A. carbo	*A. intermedius*	Accuracy	Kappa	% Accuracy
A. carbo	64	6			91.43
A. intermedius	6	60			90.90
Total	70	66	0.912	0.823	

Tracing Species Identification



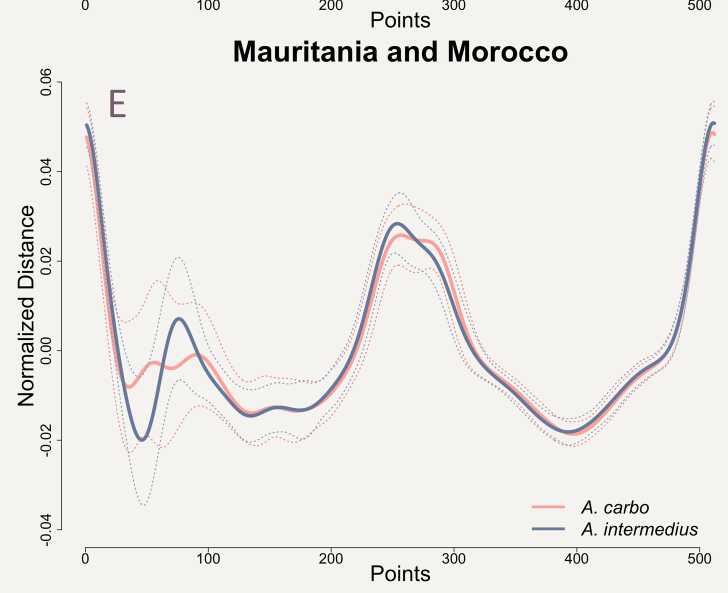


Figure 4 | Mean (solid line) and standard deviation (dotted line) of the otolith contour decomposition in samples collected from Madeira (A-D) and the African coast (Morocco and Mauritania) (E) using wavelet coefficients. *A. carbo, Aphanopus carbo; A. intermedius, Aphanopus intermedius.*

Table 2 | Number (N) of otoliths and their respective mean and standard deviation (SD, in mm) classified as *Aphanopus carbo* (*A. carbo*) and as *A. intermedius* from Madeira (MD), Morocco (MO), and Mauritania (MA) based on genetic data.

Location	Year	*A. carbo*		*A. intermedius*	
		N (%)	Mean ± SD (range)	N (%)	Mean ± SD (range)
MD	1990	260 (49.6)	8.83 ± 0.63 (6.7 – 10.2)	264 (50.4)	8.83 ± 0.67 (5.3 – 11.1)
	2003	255 (51.6)	8.98 ± 0.61 (7.5 – 10.9)	239 (48.4)	9.00 ± 0.59 (7.7 – 10.7)
	2010	195 (40.8)	8.76 ± 0.66 (7.3 – 10.6)	283 (59.2)	8.72 ± 0.63 (7.2 – 10.8)
	2021	231 (48.2)	8.66 ± 0.62 (6.7– 10.3)	248 (51.8)	8.73 ± 0.57 (7.3 – 10.3)
MO	2005-06	256 (61.2)	7.84 ± 1.02 (6.0 – 10.3)	162 (38.8)	7.94 ± 1.09 (6.0 – 10.2)
MA	2007	77 (59.7)	6.81 ± 0.59 (6.0 – 9.7)	52 (40.3)	6.81 ± 0.46 (6.0 – 7.7)

DISCUSSION

1. *Aphanopus carbo* is widely recognized as the predominant species in the North Atlantic, coexisting with *A. intermedius* from the northern distribution limit around the Azores to the Western Sahara (Biscoito et al., 2011). In Madeira, these two species were reported in a ratio of 80% *A. carbo* to 20% *A. intermedius* (Delgado et al., 2013).
2. While it has been suggested that meristic characteristics could aid in distinguishing between these two species (Biscoito et al., 2011), this approach comes with several challenges related to sampling and conceptual aspects.
3. Like genetic studies, our research demonstrates that these species can also be differentiated based on the shape of their otolith contours, as previously noted by Tuset et al. (2013) using Elliptical Fourier Descriptors. This method yielded a classification accuracy exceeding 90%.

CONCLUSION

Since *A. intermedius* was detected in the 1990 samples from Madeira, the scientific community should rethink the management of this resource and its distribution. A thorough revision of previous published studies is also needed since they may contain mixed individuals/specimens of both *Aphanopus* species.

REFERENCES

Biscoito et al. 2011. Morphological identification of two sympatric species of Trichiuridae, *Aphanopus carbo* and *A. intermedius*, in NE Atlantic. Cybium 35, 19–32. **Delgado et al. 2013**. Reproduction and growth of *Aphanopus carbo* and *A. intermedius* (Teleostei: Trichiuridae) in the northeastern Atlantic. J. Appl. Ichthyol. 29, 1008–1014. **Kuhn 2008**. Building Predictive Models in R Using the *caret* Package. J. Stat. Softw. 28(5), 1–26. **Marti-Puig et al. 2020**. Fourier-based contour descriptors to relax positional standardization of the otolith images in aforo queries. Sci. Mar. 84, 27–37. **Sadighzadeh et al. 2014.** An approach to unravelling the coexistence of snappers (Lutjanidae) using otolith morphology. Sci. Mar. 78, 353–362. **Tuset et al. 2013**. Application of otolith mass and shape for discriminating scabbardfishes *Aphanopus* spp. in the north-eastern Atlantic Ocean. J. Fish Biol. 82, 1746–1752

ULPGC Universidad de Las Palmas de Gran Canaria · ECOAQUA www.ecoaqua.eu · BIOCON Biodiversidad y Conservación · IOCAG · Instituto Español de Oceanografía · CSIC Consejo Superior de Investigaciones Científicas · BIOSTATECH