



Discriminating stocks of striped red mullet (*Mullus surmuletus*) in the Northwest European seas using three automatic shape classification methods

Abdesslam Benzinou, Sébastien Carbini, Kamal Nasreddine, Romain Elleboode, Kélig Mahé

► To cite this version:

Abdesslam Benzinou, Sébastien Carbini, Kamal Nasreddine, Romain Elleboode, Kélig Mahé. Discriminating stocks of striped red mullet (*Mullus surmuletus*) in the Northwest European seas using three automatic shape classification methods. *Fisheries Research*, Elsevier, 2013, 143, pp.153 - 160. .

HAL Id: hal-00938815

<https://hal.archives-ouvertes.fr/hal-00938815>

Submitted on 29 Jan 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Discriminating stocks of striped red mullet (*Mullus surmuletus*) in the Northwest European seas using three automatic shape classification methods

Abdesslam Benzinou, Sébastien Carbini, Kamal Nasreddine

*Ecole Nationale d'Ingénieurs de Brest ENIB, UMR CNRS 6285 LabSTICC
29238 Brest Cedex 03 (FRANCE)*

Romain Elleboode, Kélig Mahé

*Ifremer, Laboratoire Ressources Halieutiques
62321 Boulogne sur mer (FRANCE)*

Abstract

Stock identification is of primary importance for population structure assessment of economically important species. This study investigates stocks of striped red mullet using three automatic methods of stock identification based on otolith shape and growth marks. Otolith shape is known to be a promising approach for stock identification but interpreting patterns of variance is a difficult problem. In this study, images in reflected and transmitted light were acquired from 800 otoliths sampled in the Northwest European seas from South Bay of Biscay to North Sea. The growth marks are pointed out manually by an expert. The external shape of otoliths were automatically extracted by computer vision process and then three automatic classification methods were compared, two classical state-of-the-art methods based on Fourier descriptors and Principal Component Analysis (PCA), and a recently proposed method based on shape Geodesics. From

Email address: benzinou@enib.fr (Abdesslam Benzinou)

Preprint submitted to Elsevier

a methodological point of view, results show that the shape geodesic approach significantly outperforms other classical methods. From a biological point of view, this study shows that the population of striped red mullet in Northwest European seas can be divided in three geographical zones: the Bay of Biscay, a mixing zone composed of the Celtic Sea and the Western English Channel and a northern zone composed of the Eastern English Channel and the North Sea (67% of correct classification rate using both shape and growth pattern information). Moreover, it shows that for a given zone, two subsets of the same year have a lower variability in shape than two subsets from two consecutive years.

Keywords: Striped red mullet, otoliths, stock identification, year identification, shape analysis, Fourier descriptors, Principal Component Analysis (PCA), shape Geodesics

1. Introduction

Striped red mullet (*Mullus surmuletus*) occurs along the coast of Europe from the South of Norway [Wheeler, 1978] and the North of the Scotland [Gordon, 1981] to Gibraltar, also along the northern part of West Africa to Dakar, in the Mediterranean and Black Seas. Striped red mullet has been extensively studied in terms of quantity in the Mediterranean Sea and some studies were carried out in the Bay of Biscay [Desbrosses, 1933, 1935; N'Da and Deniel, 1993] that correspond to oldest exploitation areas in the Atlantic Ocean. Within the Atlantic Ocean, there are two main areas where this species is caught in this region: Bay of Biscay and in the Eastern English Channel. This species has been initially exploited by the Spanish fleets along their coast to the Bay of Biscay. Initially considered as a valuable by catch

[Marchal, 2008], the development of striped red mullet exploitation and a strong increase in landings along the English Channel and the southern North Sea by French, English and Dutch fleets have been observed since the 1990's. The strong increase of catches is essentially due to French trawlers and supplemented by the Netherlands and United Kingdom fleets which are carried out in the Eastern Channel and the south of North Sea [Mahé et al., 2005]. This could be attributed to an expansion of its migration distribution, abundance of this species coupled by the decline of traditionally targeted species in these areas and the sea-water warming trend [ICES, 2010; Marchal, 2008; Poulard and Blanchard, 2005]. Reports indicate a steady increase in East English Channel landings reaching ten times recorded landings in 1990 [Carpentier et al., 2009; Marchal, 2008]. Striped red mullet is still considered as a non-quota species in the Northeast Atlantic region and evaluation of the level of stock exploitation has only started since 7 years [ICES, 2010].

Stock identification and spatial structure information provide a basis for understanding fish population dynamics and provides reliable resource assessment for fishery management [Reiss et al., 2009]. Each stock may have unique demographic properties and responses or rebuilding strategies to exploitation. Biological attributes and productivity of species may be affected if the stock structure and fisheries management are not well considered [Smith et al., 1991].

There are a variety of techniques for stock identification such as genetics and morphometry studies. Genetic studies have been carried out in the Mediterranean Sea [Apostolidis et al., 2009; Galarza et al., 2009; Mamuris et al., 1998a,b]. In the Gulf of Pagasitikos (Greece sea), the analyses of three molucar markers revealed that this is a panmictic population [Apostolidis

et al., 2009]. However, on the level of the Mediterranean basin, the siculo- 40
Tunisian Strait seems to be the transition zone between the Mediterranean's 41
eastern and western populations [Galarza et al., 2009]. A sharp genetic 42
division was detected when comparing striped red mullet originating from 43
the Atlantic Ocean and from Mediterranean Sea. 44

Among all available techniques, otolith shape has been proven to be 45
relevant feature for species and/or stock discrimination issues [Begg and 46
Brown, 2000; Burke et al., 2008; Campana and Casselman, 1993; Stransky, 47
2005; Stransky et al., 2008b]. Otolith shape reflects the growth pattern of the 48
fish as well as being markedly species specific. As a result, otolith shape can 49
be used to differentiate stocks of the same species. Another relevant feature 50
for stock identification is the growth law as growth is highly correlated to 51
the environmental conditions and is thus stock specific. 52

In the present study, the stock identification was investigated with two 53
methods based either on otolith shape or on growth marks (and both infor- 54
mation). Images in reflected and transmitted light were acquired from 800 55
otoliths sampled in the Northwest European seas from South Bay of Biscay 56
to North Sea. Growth marks have been pointed out manually by an expert. 57
External shapes were extracted by computer vision process and then three 58
automatic classification methods were compared, two classical state-of-the- 59
art methods based on Fourier descriptors, Principal Component Analysis 60
(PCA), and a recently proposed method [Nasreddine et al., 2009] based on 61
shape geodesics. 62

2. Materials and methods 63

2.1. Otolith datasets 64

Striped red mullet otoliths were extracted from fish randomly sampled 65
from the southern bay of Biscay to the North sea. The study area was 66
divided into six geographic sectors: the NS (North Sea ; ICES Division 67
IVab), the EEC (Eastern English Channel ; ICES Division VIIId), the WEC 68
(Western English Channel ; ICES Division VIIe), the CS (Celtic Sea ; ICES 69
Division VIIIh), the NBB (North Bay of Biscay ; ICES Division VIIIa) and 70
the SBB (South Bay of Biscay ; ICES Division VIIIb) (Figure .1). All 71
sampling were collected from September to December 2009 except the EEC 72
otoliths which were collected from October-November 2007 and 2008. 73

{Figure .1 goes here } 74

The otoliths were selected from the routine surveys on board the RV 75
“Thalassa” and RV “Gwen-Drez” conducted by the Ifremer Institute (France) 76
and from fisheries markets. Fish were caught by otter trawl, bottom pair 77
trawl and set gillnets. Both sagittal otoliths were removed and cleaned be- 78
fore drying and storing in paper envelope. One otolith per fish was examined 79
using a light microscope connected to a video camera and a dedicated image- 80
analysis system TNPC (digital processing for calcified structures) developed 81
by Ifremer, ENIB and Noesis society. 82

Images of whole otoliths have been acquired using both transmitted and 83
reflected lights. From 800 otoliths coming from six different stocks of striped 84
red mullet (Figure .1), four different image datasets will be considered: 85

Dataset (1) : 600 otoliths sampled from six different stocks (100 otoliths per stock): 86

- NS: North Sea (IVab) - 2009 87

• EEC08: Eastern English Channel (VIId) - 2008	88
• WEC: Western English Channel (VIIe) - 2009	89
• CS: Celtic Sea (VIIh) - 2009	90
• NBB: North Bay of Biscay (VIIIa) - 2009	91
• SBB: South Bay of Biscay (VIIIb) - 2009	92
Dataset (2) : 700 otoliths: the 600 otoliths of dataset (1) with 100 other otoliths	93
from Eastern English Channel but of a different year:	94
• EEC07: Eastern English Channel (VIId) - 2007	95
Dataset (3) : 200 otoliths: those from Eastern English Channel (VIId) over the two	96
consecutive years 2007 and 2008:	97
• EEC07: Eastern English Channel (VIId) - 2007	98
• EEC08: Eastern English Channel (VIId) - 2008	99
Dataset (4) : 200 otoliths from North Sea (IVab) from the same year 2009 randomly	100
divided in 2 classes:	101
• NS09a: North Sea (IVab) - 2009 a	102
• NS09b: North Sea (IVab) - 2009 b	103

These datasets illustrate two different types of applications of otolith 104
shape classification: stock discrimination (datasets (1) and (2)) and year 105
discrimination (datasets (3) and (4)). Both issues are quite hard for current 106
state-of-the-art computer vision techniques because the external shapes of 107
the considered otoliths exhibit very few differences. 108

For the year discrimination issue, the test is carried out on dataset 109
(3) and dataset (4) separately. As dataset (4) is composed of randomized 110

classes, the classification performances on this dataset should be close to 111
those of a theoretical random classifier (i.e. 50%). The difference in perfor- 112
mances between dataset (3) and dataset (4) will give an idea of the validity 113
of the results. 114

2.2. Shape-based stock identification 115

The shape-based classification process can be decomposed in three main 116
steps (Figure .2). First, the otolith contour is extracted as described in 117
next section (§ 2.2.1) using an automatic threshold. Three approaches to 118
extract reduced-dimension feature vectors from the contours were consid- 119
ered: Fourier Transform (FT), Principal Component Analysis (PCA) and 120
a technique issued from shape geodesics [Nasreddine et al., 2009]. The dis- 121
criminative power of each approach is evaluated using its own distance ma- 122
trix as input for a classifier. In other words, for a query input the feature 123
vector is considered as the distance matrix calculated between this indi- 124
vidual and the training individuals. Here, we investigate the performances 125
of two widely used classifiers: (1) the K-Nearest Neighbors (KNN) classi- 126
fier with the “leave-one-out” heuristic and (2) the Support Vector Machine 127
(SVM) classifier [Vapnik, 1995] with two randomly-selected sub-samples, one 128
of them is used to build the SVM-model which is tested on the other. 129

{Figure .2 goes here } 130

2.2.1. Automatic contour extraction 131

The otolith image is acquired using two imaging modalities: by trans- 132
mitted light or by reflected light. These two modalities could give additional 133
information. To extract the otolith outline, a mixed image is built in order 134
to integrate information available in both modalities (Figure .3). This mixed 135

image is a mean between the transmitted light image and the negative of
the reflected light image. Image contours are detected as local maximum
of the image gradient, approximated using a Sobel filtering. The resulting
contours are filled by a morphological closing operation and filtered to re-
tain the largest connected component which corresponds to the edge of the
otolith. The advantage of mixing both image modalities is illustrated on
example given by figure .3. The mixed image gives more details about the
contour especially on the region of the excisura major.

{Figure .3 goes here }

The resulting contour is then sampled into 300 points which describe
adequately the otolith shape.

2.2.2. *Fourier descriptors*

Shape can be described using complex Fourier descriptors [Granlund,
1972] or using elliptic Fourier descriptors [Kuhl and Giardina, 1982]. For
otolith shape analyses, both techniques have been extensively used and
proved to be efficient [Duarte-Neto et al., 2008a; Kristoffersen and Magoulas,
2008; Mériqot et al., 2007; Stransky et al., 2008a] [Cardinale et al., 2004;
Galley et al., 2006; Robertson and Talman, 2002; Schulz-Mirbach et al.,
2008; Smith, 1992; Torres et al., 2000]. In our previous work [Nasreddine
et al., 2009] we have showed that for red mullet otoliths, classification results
are still similar by using these two methods. Elliptic Fourier descriptors are
more appropriate than complex Fourier descriptors when otolith contours
are composed of series of ellipse arcs (as for *Trachurus mediterraneus* otoliths
for example). Hence, for striped red mullet otoliths we have chosen to use
the complex descriptors which can be implemented more efficiently.

With a view to achieving translation, rotation and scaling invariance,

the first descriptor is aborted and the selected descriptors are scaled with
respect to the first non zero coefficient resulting in the so-called *normalized
Fourier descriptors*. The distance between two shapes is computed as the
Euclidean distance between the associated vectors of the normalized Fourier
descriptors.

2.2.3. Principal Component Analysis (PCA)

Principal Component Analysis (PCA) was first introduced by Pearson in
[Pearson, 1901] as a mathematical tool that transforms data linearly corre-
lated to uncorrelated variables called principal components. PCA is exten-
sively used in fisheries research for otolith shape analyses, in particular for
otolith stock identification. Usually, PCA is applied on Fourier coefficients
in order to assess differences in otolith shape [Duarte-Neto et al., 2008b;
Mérigot et al., 2007; Schulz-Mirbach et al., 2008]. PCA can also be applied
on morphometric variables [Torres et al., 2000], on a binary low resolution
image of the contour [Bermejo and Monegal, 2007] or for standardizing the
otolith contour orientation [Piera et al., 2005].

However, PCA is not invariant to affine transformations, it is applied for
pattern recognition when the coordinates of input vectors can be ordered.
In face recognition for example, eyes and lips centers are manually selected,
then the images are rotated, in order to make the line connecting eye centers
horizontal, and resized to make the distances between the centers of the
eyes equal. The PCA is carried out on data vectors formed by cropped part
of images. In the case of calcified structures, it is not always obvious to
order the data vector coordinates on the basis of clearly defined landmarks.
A normalization procedure should then be applied to the raw contours to
be invariant with respect to translation, rotation and scaling, so that the

normalized shape is the result of the fish history, independently of acquisition 188
settings. 189

The translation invariance is obtained simply by subtracting the coordi- 190
nates of the mass center to the coordinates of all points. Scale invariance is 191
also simply obtained by dividing each point of the contour in polar coordi- 192
nates by the mean radius. For rotation normalization, a first solution could 193
be to align shapes according to the main axis. This axis can be defined by 194
the two farthest points of the shape or by minimizing the covariance using a 195
PCA like in [Piera et al., 2005]. However, on striped red mullet otolith, the 196
main axis does not correspond to a meaningful biological feature. Instead, 197
we propose to normalize shapes according to the center of excisura major. 198

The corresponding point of the excisura major can be detected automat- 199
ically after subtraction of the original otolith shape from the corresponding 200
filled shape. Then, each shape is aligned according to the axis that passes 201
through this point and the mass center of the otolith contour (Figure .4). 202

{Figure .4 goes here } 203

After contours normalization, PCA is applied on a matrix where each 204
of the rows represents a different contour and the columns represent the 205
information about the contours: the Cartesian normalized coordinates and 206
the local curvature are all put together in a row, one after the other. 207

To compute a distance between all contours in a given dataset, we pro- 208
ceed with a “leave-one-out” heuristics. One after another, each contour C_i 209
of the dataset is left out and PCA is computed on the remaining contours. 210
Then contour C_i is projected into the eigenspace generated by the eigenvec- 211
tors. Finally the distances between the projected contour and each of the 212
other projected contours of the dataset are computed as Euclidean distances 213
in the eigenspace. 214

2.2.4. The geodesics approach

215

A potential drawback of Fourier and PCA approaches comes from the implicit global (spatial) characterization of the shape. Each descriptor holds information about all points of the shape as it is calculated using all points. Therefore, local (spatial) discriminant shape signatures, such as shape discontinuities or landmarks, may not be well exploited by such a global characterization [Parisi-Baradad et al., 2005]. In contrast, a Geodesic approach was recently proposed [Nasreddine et al., 2009] to take advantage of local shape features while ensuring invariance to geometric transformations (e.g. translation, rotation and scaling). In this approach, we have defined distance between shapes as a deformation cost stated as a matching issue, i.e. determining the optimal matching between two otolith contours with respect to a similarity measure.

216

217

218

219

220

221

222

223

224

225

226

227

The distance $d(\Gamma_1, \Gamma_2)$ between two shapes Γ_1 and Γ_2 is stated as the minimum, over all mapping functions Ψ , of the similarity measure, $E_D(\Gamma_1, \phi(\Gamma_2))$ between the reference shape Γ_1 and the mapped shape $\phi(\Gamma_2)$.

228

229

230

$$d(\Gamma_1, \Gamma_2) = \min_{\phi \in \Psi} E_D(\Gamma_1, \phi(\Gamma_2)) \quad (1)$$

As the important biological information is considered in the shape of contour and not in its size, the shapes are parameterized in function of the normalized curvilinear abscissa s which has a value between 0 and 1 independently of the original contour length. A robust criterion is introduced in order to improve the robustness of the proposed distance to outliers coming from biological interindividual variabilities. The principle is supported by the use of a function that adjusts weight ω in order to penalize the data points of high variation compared to other points.

231

232

233

234

235

236

237

238

Given two shapes locally characterized by the angle $\theta(s)$ between the
tangent to the curve and the horizontal axis, the distance between two con-
tours is defined by:

$$d(\theta_1(s), \theta_2(s)) = 2 \inf_{\phi} \arccos \int_s \sqrt{\phi_s(s)} \left| \cos \frac{\omega(r(s))r(s)}{2} \right| ds \quad (2)$$

where $\phi_s = \frac{d\phi}{ds}$ and $r(s) = \theta_1(s) - \theta_2(\phi(s))$. The term $\sqrt{\phi_s(s)}$ allows to
avoid torsion and stretching along the curve. The weight function ω is
issued from the robust estimator of Leclerc [Black and Rangarajan, 1996];
 $\omega(r(s)) = \frac{2}{\sigma^2} \exp\left(\frac{-r^2(s)}{\sigma^2}\right)$ where σ is the standard deviation of data errors
 $r(s)$.

Formally, the numerical computation of $d(\Gamma_1, \Gamma_2)$ is solved by using a
dynamic programming technique (refer to [Nasreddine et al., 2009] for more
details).

2.3. Growth marks based stock identification

The growth-based classification process consists of three main steps (Fig-
ure .2). First, an expert manually points out the growth marks on the
otolith image (Figure .5). This step can be done using **TNPC** software
(www.tnpc.fr) in parallel with the image acquisition step; it is not a contra-
diction with the automatic process of classification. Then distance between
the growth laws of two otoliths is computed using the Euclidean distances
between growth vectors. In case of two different aged otoliths, distance is
computed using only the growth marks available on both otoliths. For ex-
ample, in figure .5 this distance is computed using the three growth marks
on each otolith.

Given two growth vectors $G_1 = \{G_{1j}\}_{j=1 \dots N_1}$ and $G_2 = \{G_{2j}\}_{j=1 \dots N_2}$,
the growth distance is considered as the Euclidean distance:

$$d_{Growth} = \sqrt{\sum_{j=1}^{N_g} (G_{2j} - G_{1j})^2} \quad (3)$$

where $N_g = \mathbf{min}\{N_1, N_2\}$ is the number of growth marks available in both
vectors.

Finally, all distances between otoliths are computed leading to a distance
matrix used as input for an SVM classifier. The feature vector is consid-
ered as the distance calculated between the query input and all training
individuals.

{Figure .5 goes here }

3. Results

Here performances are evaluated in terms of correct classification rates.
We have started experiments with the hypothesis that the six stocks (NS,
EEC, WEC, CS, NBB and SBB) are considered as individual separated
stocks with specific characteristics of shapes.

Compared to KNN, SVM classifier performs slightly better in terms of
correct classification rate (from 30% to 32.7% on dataset (1)) but at the
cost of increasing dramatically the standard deviation of the performances
between classes (from 10.9 to 15.2 on dataset (1)). Thus, as KNN clas-
sifier results in stable performances across the classes, it has been chosen
for shape-based classification. In contrast, applying KNN for growth-based
stock identification gives a correct classification rate of 25.5% whereas SVM
gives higher correct classification rate (35.4%) for the same dataset (dataset
(1)). Hence, SVM has been chosen for growth-based classification.

The correct classification rates remain high with respect to the random
classification but these rates show that the hypothesis of separated stocks

should be aborted. The six stocks are then grouped into three stocks leading to a correct classification rate of 67%. Grouped stocks have in the first hypothesis close shape characteristics and could not be really distinguished easily. Classification errors could be due to genetic factors, migration among others. A rate of 100% could then not be reached with the presence of all these factors on the otolith shape. In comparisons to other stock identification methods, otolith shape is a promising approach but interpreting patterns of variance can be difficult [Cadrin et al., 2005].

In the following, geographical zones are ordered in the tables according to their positions (from north (NS) to south (SBB)); thus neighbor classes are also neighbor geographical zones.

3.1. Dataset (1)

Results on dataset (1) are given in tables .1-.3. Geodesic approach reaches 30% of correct classification (Table .3) while this rate is 19.7% for Fourier approach (Table .1) and 25% for PCA (Table .2). These scores are better than a random classification that would theoretically reach 16.7% (for six classes).

{Table .1 goes here }

{Table .2 goes here }

{Table .3 goes here }

In table .4, classification results are given when the growth information is used for stock identification. The mean correct classification obtained by SVM reaches 35.4%.

{Table .4 goes here }

As in [Nasreddine et al., 2009], we have tested stock identification with both growth and shape information in order to improve classification per-

formances. The mean correct classification rate is then increased to reach 49.4% (Table .5).

{Table .5 goes here }

3.2. Dataset (2)

On dataset (2), Fourier approach reaches 16.4% of mean correct classification (Table .6), PCA approach reaches 19% of correct classification (Table .7) while Geodesic approach reaches 24.9% (Table .8). These scores are also better than a random classification that would theoretically reach 14.3% (for seven classes).

{Table .6 goes here }

{Table .7 goes here }

{Table .8 goes here }

3.3. Datasets (3) and (4)

Regarding the year discrimination issue on dataset (3), the mean classification rate of the Fourier approach (56%, Table .9) is too close to the theoretical mean classification rate of a random classifier (50% for two classes). Thus the classical Fourier approach fails on this specific year discrimination issue. The mean classification rate on the random dataset (4) (43%, Table .10) is lower but quite close to the theoretical mean classification rate of a random classifier (50% for two classes), it shows that with this approach, two arbitrary sets of the same stock and same year have no significant shape differences.

Regarding PCA and Geodesic approaches, the mean classification rate on dataset (3) (60%, Table .9) is higher than the mean classification rate on the random dataset (4) (49.5%, Table .10). This shows that the otolith

morphology varies over two consecutive years and that this difference in 337
shape is higher than between two arbitrary groups of the same year and 338
same stock. 339

{Table .9 goes here } 340

{Table .10 goes here } 341

4. Discussion 342

4.1. Comparison of the three shape-based approaches 343

Performances of the three shape-based approaches are compared in ta- 344
ble .11. On both dataset (1) and dataset (2), the Geodesic approach exhibits 345
highest performances followed by PCA approach and Fourier approach last. 346

Regarding the stock discrimination issue on dataset (1) (Tables .1, .2 347
and .3), the three methods show that the population of striped red mullet 348
can be geographically divided in three zones: 349

- The Bay of Biscay (NBB+SBB) 350
- A mixing zone composed of the Celtic Sea and the Western English 351
Channel (CS+WEC) 352
- A northern zone composed of the Eastern English Channel and the 353
North Sea (EEC+NS) 354

To further the “three zones” hypothesis, we have tested the classification 355
when the otoliths were grouped in three classes corresponding to the three 356
zones. The results of this classification using the geodesic approach is shown 357
in table .12 below. It clearly validates the hypothesis as the obtained mean 358
correct classification rate reaches 54.3% and the error scores are higher be- 359
tween two neighbors zones than between two unconnected zones. Finally, 360

this rate raises to 67.31% when the SVM classifier is used with geodesic
distances coupled with the growth information (Table .13).

Regarding the year discrimination issue, classical Fourier approach fails
while PCA approach shows a small difference in shape and Geodesic ap-
proach exhibits the highest difference (Table .11). Thus Geodesic approach
seems the most appropriate method for this task.

{Table .11 goes here }

{Table .12 goes here }

{Table .13 goes here }

4.2. *Relevance of the shape and growth information*

In this study three different approaches have been compared for shape-
based stock identification, two state-of-the-art methods (Fourier and PCA)
that have been extensively used in marine research on different species, and
a recent method (Geodesic) that proved to give very good performances
on different shapes [Nasreddine et al., 2010] and in particular on otolith
shapes [Nasreddine et al., 2009]. Although these three methods result in
high correct classification rates on several problems, they give quite low
correct classification rate for the particular cases tested in this study. It
tends to prove that otolith shape is not relevant for the particular case of
striped red mullet if we consider the six stocks separately. The growth-
based stock identification results are not so far from the shape-based stock
identification results. This study shows that both information are influenced
by different living conditions and different environments and can serve as
stock identifier. This identification is not very high as otolith shape is highly
due to the genetics. This result tends to prove that the genetic information
is quite homogeneously spread across all geographical zones in the north

west European seas. 387

This study has proven that by coupling both information (shape and 388
growth patterns), stock discrimination becomes more efficient. These two 389
information are independent and multivariate analysis, including them with 390
other independent information (chemical concentrations, ...), should be 391
investigated for stock identification. 392

The observations above lead to two hypothesis on the striped red mullet: 393

- some adults move from one zone to another, 394
- some larvae or juveniles perform migration during growth. 395

Acknowledgement 396

The authors want to thank the European commission for providing the 397
financial support of this work through the NESPMAN project, all scientists 398
and crew on board the RV “Thalassa” and the RV “Gwen Drez” for their 399
help with sample collection. 400

Apostolidis, A., Moutou, K., Stamatis, C., Mamuris, Z., 2009. Genetic struc- 401
ture of three marine fishes from the gulf of pagasitikos (greece) based on 402
allozymes, RAPD, and mtDNA RFLP markers. *Biologia* 64 (5), 1005– 403
1010. 404

Begg, G., Brown, R., 2000. Stock identification of haddock *melanogrammus* 405
aeglefinus on georges bank based on otolith shape analysis. *Transactions* 406
of the American Fisheries Society 129, 935–945. 407

Bermejo, S., Monegal, B., 2007. Fish age analysis and classification with 408
kernel methods. *Pattern Recognition Letters* 28 (10), 1164–1171. 409

- Black, M., Rangarajan, A., 1996. On the unification of line processes, outlier rejection, and robust statistics with applications in early vision. *International Journal of Computer Vision* 19 (5), 57–92.
- Burke, N., Brophy, D., King, P., 2008. Otolith shape analysis: its application for discriminating between stocks of irish sea and celtic sea herring (*clupea harengus*) in the irish sea. *ICES Journal of Marine Science* 65.
- Cadrin, S., Friedland, K., Waldman, J., 2005. Stock identification methods: Applications in Fishery science. Elsevier Academic press.
- Campana, S., Casselman, J., 1993. Stock discrimination using otolith shape analysis. *Canadian Journal of Fisheries and Aquatic Sciences* 50, 1062–1083.
- Cardinale, M., Doering-Arjes, P., Kastowsky, M., Mosegaard, H., 2004. Effects of sex, stock, and environment on the shape of known-age atlantic cod (*gadus morhua*) otoliths. *Canadian Journal of Fisheries and Aquatic Sciences* 61 (2), 158–167.
- Carpentier, A., Martin, C., Vaz, S., 2009. Channel habitat atlas for marine resource management (charm phase ii). INTERREG 3a Programme, IFREMER, Boulogne-sur-mer 65.
- Desbrosses, P., 1933. Contribution à la biologie du rouget-barbet en atlantique nord. *Revue des Travaux de l’Institut des Pêches Maritimes* 6 (3), 249–270.
- Desbrosses, P., 1935. Contribution à la connaissance de la biologie du rouget barbet en atlantique nord (iii) *mullus barbatus* (rond) *surmuletus fage*

- mode septentrional fage. *Revue des Travaux de l'Institut des Pêches Mar-* 433
itimes 8 (4), 351–376. 434
- Duarte-Neto, P., Lessa, R., Stosic, B., Morize, E., 2008a. The use of sagittal 435
 otoliths in discriminating stocks of common dolphinfish (*coryphaena hip-* 436
purus) off northeastern brazil using multishape descriptors. *ICES Journal* 437
of Marine Science 65 (7), 1144–1152. 438
- Duarte-Neto, P., Lessa, R., Stosic, B., Morize, E., 2008b. The use of sagittal 439
 otoliths in discriminating stocks of common dolphinfish (*coryphaena hip-* 440
purus) off northeastern brazil using multishape descriptors. *ICES Journal* 441
of Marine Science 65, 1144–1152. 442
- Galarza, J., Turner, G., Macpherson, E., Rico, C., 2009. Patterns of ge- 443
 netic differentiation between two co-occurring demersal species: the red 444
 mullet (*mullus barbatus*) and the striped red mullet (*mullus surmuletus*). 445
Canadian Journal of Fisheries and Aquatic Sciences 66 (9), 1478–1490. 446
- Galley, E. A., Wright, P. J., Gibb, F. M., 2006. Combined methods of otolith 447
 shape analysis improve identification of spawning areas of Atlantic cod. 448
ICES Journal of Marine Science 63 (9), 1710–1717. 449
- Gordon, J., 1981. The fish populations of the west of scotland shelf. Part II, 450
Oceanography and Marine Biology. Annual Review 19, 405–441. 451
- Granlund, G., 1972. Fourier preprocessing for hand print character recogni- 452
 tion. *IEEE Transactions on Computers* C-21, 195–201. 453
- ICES, 2010. Report of the working group on assessment of new MoU species 454
 (WGNEW). Tech. rep., ICES CM 2010/ACOM: 21. 455

- Kristoffersen, J., Magoulas, A., 2008. Population structure of anchovy *456*
engraulis encrasicolus L. in the mediterranean sea inferred from multiple *457*
methods. *Fisheries research* 91 (2-3), 187–195. *458*
- Kuhl, F., Giardina, C., 1982. Elliptic fourier features of a closed contour. *459*
Computer Graphics and Image Processing 18, 236–258. *460*
- Mahé, K., Destombes, A., Coppin, F., Koubbi, P., Vaz, S., Roy, D. L., Car- *461*
pentier, A., 2005. Le rouget barbet de roche *mullus surmuletus* (L. 1758) *462*
en manche orientale et mer du nord. *Rapp. Contrat Ifremer/CRPMEM* *463*
Nord-Pas de Calais, 187p. *464*
- Mamuris, Z., Apostolidis, A., Theodorou, A., Triantaphyllidis, C., 1998a. *465*
Application of random amplified polymorphic dna (rapd) markers to eval- *466*
uate intraspecific genetic variation in red mullet (*mullus barbatus*). *Marine* *467*
Biology 132, 171–178. *468*
- Mamuris, Z., Apostolidis, A., Triantaphyllidis, C., 1998b. Genetic protein *469*
variation in red mullet (*mullus barbatus*) and striped red mullet (*m. sur-* *470*
urmuletus) populations from the mediterranean sea. *Marine Biology* 130, *471*
353–360. *472*
- Marchal, P., 2008. A comparative analysis of metiers and catch profiles for *473*
some french demersal and pelagic fleets. *ICES Journal of Marine Science* *474*
65, 674–686. *475*
- Mérigot, B., Letourneur, Y., Lecomte-Finiger, R., 2007. Characterization of *476*
local populations of the common sole *Solea solea* (Pisces, Soleidae) in the *477*
NW mediterranean through otolith morphometrics and shape analysis. *478*
Marine Biology 151 (3), 997–1008. *479*

- Nasreddine, K., Benzinou, A., Fablet, R., 2009. Shape geodesics for the clas- 480
sification of calcified structures: beyond fourier shape descriptors. Fish- 481
eries Research 98 (1-3), 8–15. 482
- Nasreddine, K., Benzinou, A., Fablet, R., 2010. Variational shape matching 483
for shape classification and retrieval. Pattern Recognition Letters 31 (12), 484
1650–1657. 485
- N'Da, K., Deniel, C., 1993. Sexual cycle and seasonal changes in the ovary 486
of the red mullet, *mullus surmuletus*, from the southern coast of brittany. 487
Journal of Fish Biology 43 (2), 229–244. 488
- Parisi-Baradad, V., Lombarte, A., Garcia-Ladona, E., Cabestany, J., Piera, 489
J., Chic, O., 2005. Otolith shape contour analysis using affine transfor- 490
mation invariant wavelet transforms and curvature scale space represen- 491
tation. Marine and Freshwater Research 56, 795–804. 492
- Pearson, K., 1901. On lines and planes of closest fit to systems of points in 493
space. Philosophical Magazine 2 (6), 559–572. 494
- Piera, J., Parisi-Baradad, V., Garcia-Ladona, E., Lombarte, A., Recasens, 495
L., Cabestany, J., 2005. Otolith shape feature extraction oriented to au- 496
tomatic classification with open distributed data. Marine and freshwater 497
research 56 (5), 805–814. 498
- Poulard, J., Blanchard, F., 2005. The impact of climate change on the fish 499
community structure of the eastern continental shelf of the bay of biscay. 500
ICES Journal of Marine Science 62, 1436–1443. 501
- Reiss, H., Hoarau, G., Dickey-Collas, M., Wolff, W., 2009. Genetic popula- 502

- tion structure of marine fish: mismatch between biological and fisheries 503
management units. *Fish and Fisheries* 10, 361–395. 504
- Robertson, S., Talman, S., 2002. Shape analysis and ageing of orange roughy 505
otoliths from the south tasman rise. Tech. rep., Final Report to the 506
Australian Fisheries Management Authority. Marine and Freshwater Re- 507
sources Institute. 508
- Schulz-Mirbach, T., Stransky, C., Schlickeisen, J., Reichenbacher, B., 2008. 509
Differences in otolith morphologies between surface- and cave-dwelling 510
populations of *Poecilia mexicana* (teleostei, poeciliidae) reflect adapta- 511
tions to life in an extreme habitat. *Evolutionary Ecology Research* 10 (4), 512
537–558. 513
- Smith, M. K., 1992. Regional differences in otolith morphology of the deep 514
slope red snapper *etelis carbunculus*. *Canadian Journal of Fisheries and* 515
Aquatic Sciences 49 (4), 795–804. 516
- Smith, P., Francis, R., McVeagh, M., 1991. Loss of genetic diversity due to 517
fishing pressure. *Fisheries Research* 10, 309–316. 518
- Stransky, C., 2005. Geographic variation of golden redfish (*sebastes marinus*) 519
and deep-sea redfish (*s. mentella*) in the north atlantic based on otolith 520
shape analysis. *ICES Journal of Marine Science* 62, 1691–1698. 521
- Stransky, C., Baumann, H., Fevolden, S.-E., Harbitz, A., Hie, H., Nedreaas, 522
K. H., Salberg, A.-B., Skarstein, T. H., 2008a. Separation of norwegian 523
coastal cod and northeast arctic cod by outer otolith shape analysis. *Fish-* 524
eries Research 90 (1-3), 26–35. 525

- Stransky, C., Murta, A., Schlickeisen, J., Zimmermann, C., 2008b. Otolith 526
 shape analysis as a tool for stock separation of horse mackerel (*trachurus* 527
trachurus) in the northeast atlantic and mediterranean. *Fisheries Research* 528
 89, 159–166. 529
- Torres, G. J., Lombarte, A., Morales-Nin, B., 2000. Sagittal otolith size 530
 and shape variability to identify geographical intraspecific differences in 531
 three species of the genus *merluccius*. *Journal of the Marine Biological* 532
Association of the UK 80 (2), 333–342. 533
- Vapnik, V., 1995. *The nature of statistical learning theory*. Springer-Verlag, 534
 New York, USA. 535
- Wheeler, A., 1978. *Key to the fishes of northern europe*. Frederick Warne & 536
 Co. Ltd Londres, 380p. 537

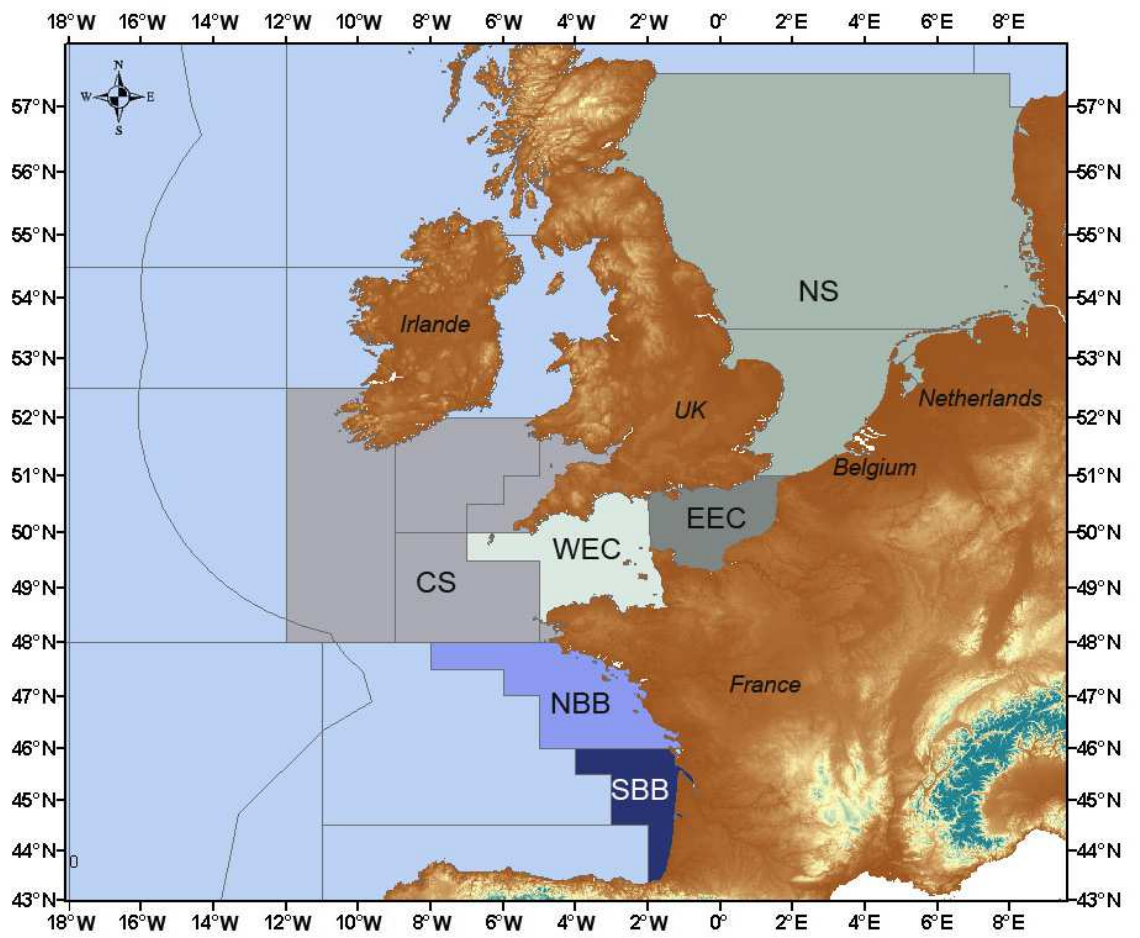


Figure .1: Map of the stocks of striped red mullet involved in this study.

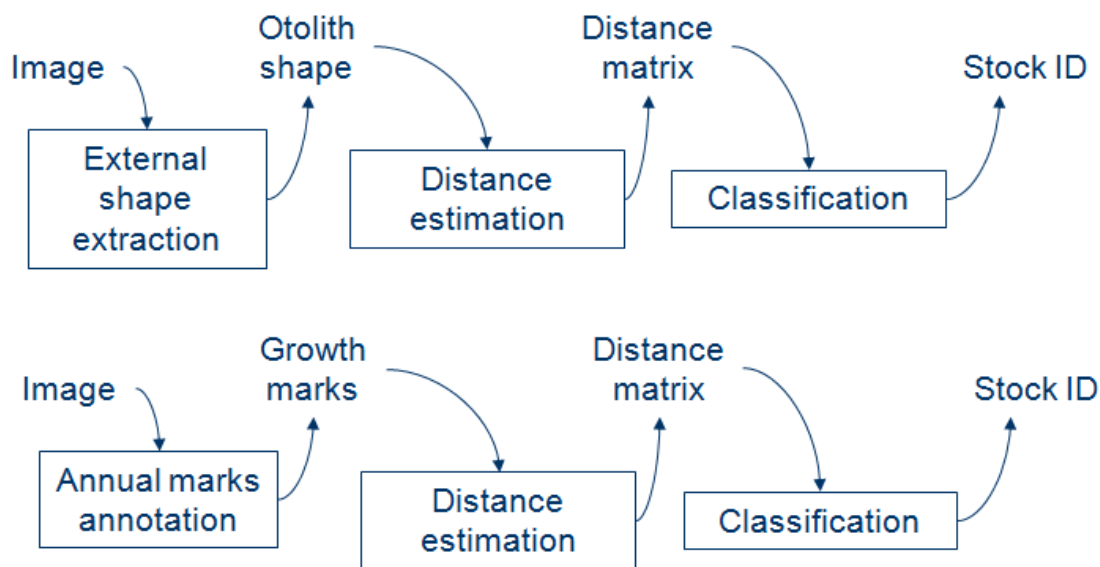


Figure .2: Shape-based and growth-based classification general schemes.

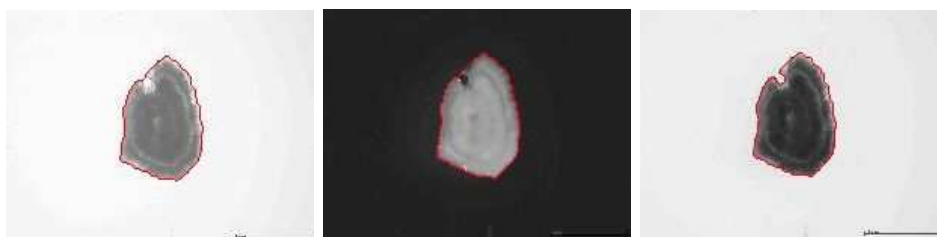


Figure .3: Contour extraction using transmitted light image (left), reflected light image (middle) and resulting mixed image (right). Note that the contour extracted using the mixed image is more efficient.

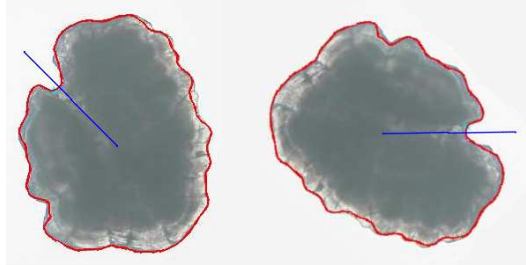


Figure .4: Contour extraction and normalization. Left: contour before normalization, right: contour after rotation normalization. In this figure we show the main axis passing through the mass center and the excisura major center.

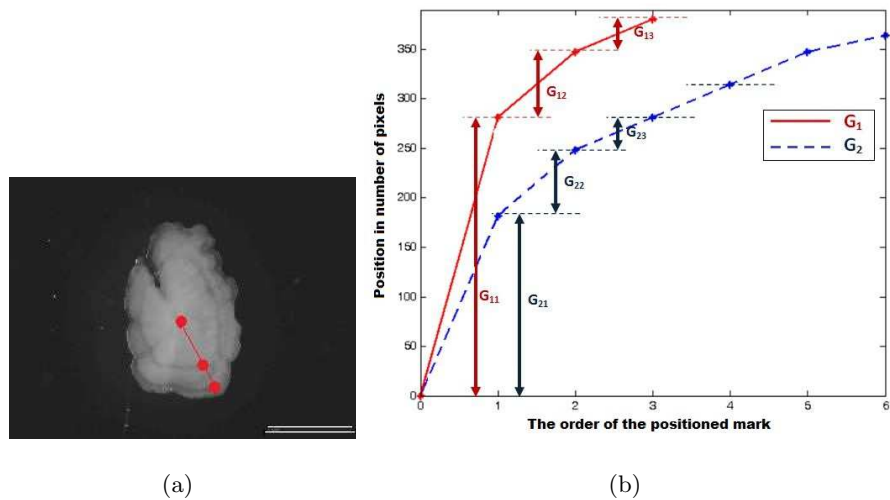


Figure .5: Illustration of growth distance calculation. (a): Annual growth marks manually positioned by expert. (b): Example of distance computation between growth laws of two otoliths.

Table .1: Confusion matrix (in %) for the Fourier approach on dataset (1) achieved by KNN classifier. Mean correct classification rate: 19.7%.

<i>Fourier approach on Dataset (1)</i>						
Estimated Class	Actual Class					
	NS	EEC08	WEC	CS	NBB	SBB
NS	18	20	11	18	18	12
EEC08	21	28	25	17	6	14
WEC	8	19	12	16	7	14
CS	21	12	18	13	11	14
NBB	16	9	14	16	23	22
SBB	16	12	20	20	35	24

Table .2: Confusion matrix (in %) for the PCA approach on dataset (1) achieved by KNN classifier. Mean correct classification rate: 25%.

<i>PCA approach on Dataset (1)</i>						
Estimated Class	Actual Class					
	NS	EEC08	WEC	CS	NBB	SBB
NS	29	13	15	19	10	12
EEC08	18	31	16	21	10	10
WEC	14	13	26	11	21	18
CS	17	21	15	20	11	12
NBB	15	11	12	13	21	25
SBB	7	11	16	16	27	23

Table .3: Confusion matrix (in %) for the Geodesic approach on dataset (1) achieved by KNN classifier. Mean correct classification rate: 30%.

<i>Geodesic approach on Dataset (1)</i>						
Estimated Class	Actual Class					
	NS	EEC08	WEC	CS	NBB	SBB
NS	15	20	11	8	5	11
EEC08	28	44	17	23	5	5
WEC	9	9	22	11	7	9
CS	24	15	24	32	15	13
NBB	10	5	16	13	27	22
SBB	14	7	10	13	41	40

Table .4: Confusion matrix resulting from an SVM classifier on growth distances (dataset (1)). Mean correct classification rate: 35.4 %.

<i>Growth-based approach on Dataset (1)</i>						
Estimated class	Actual class					
	NS	EEC08	WEC	CS	NBB	SBB
NS	42.49	20.34	16.58	5.62	3.84	11.13
EEC08	12.30	50.38	13.12	2.36	5.56	16.28
WEC	12.14	19.26	41.35	5.13	8.54	13.58
CS	41.7	2.74	43.71	8.74	3.33	0.29
NBB	11.47	8.79	30.07	3.66	26.47	19.54
SBB	12.02	20.12	12.00	1.39	11.12	43.34

Table .5: Confusion matrix resulting from an SVM classifier on geodesic distances coupled with growth distances (dataset (1)). Mean correct classification rate: 49.4 %.

<i>Growth and Geodesic-based approach on Dataset (1)</i>						
Estimated class	Actual class					
	NS	EEC08	WEC	CS	NBB	SBB
NS	43.75	12.00	2.44	12.25	5.00	3.57
EEC08	31.25	66.00	21.95	18.36	0.00	0.00
WEC	12.50	16.00	60.98	4.08	0.00	25
CS	8.33	6.00	9.76	44.89	20.00	10.71
NBB	0.00	0.00	2.44	20.41	45.00	25.00
SBB	4.17	0.00	2.44	0.00	30.00	35.71

Table .6: Confusion matrix (in %) for the Fourier approach on dataset (2) achieved by KNN classifier. Mean correct classification rate: 16.4%.

<i>Fourier approach on Dataset (2)</i>							
Estimated Class	Actual Class						
	NS	EEC07	EEC08	WEC	CS	NBB	SBB
NS	15	10	22	7	18	13	11
EEC07	15	19	12	23	14	11	11
EEC08	17	16	24	18	17	7	11
WEC	6	17	14	7	14	5	11
CS	20	14	8	17	7	12	11
NBB	16	14	8	12	15	20	22
SBB	11	10	12	16	15	32	23

Table .7: Confusion matrix (in %) for the PCA approach on dataset (2) achieved by KNN classifier. Mean correct classification rate: 19%.

<i>PCA approach on Dataset (2)</i>							
Estimated Class	Actual Class						
	NS	EEC07	EEC08	WEC	CS	NBB	SBB
NS	20	10	11	17	14	8	7
EEC07	16	15	17	8	14	16	14
EEC08	12	15	24	14	16	8	7
WEC	12	16	14	22	14	16	13
CS	19	12	16	14	15	11	9
NBB	13	19	9	10	14	15	28
SBB	8	13	9	15	13	26	22

Table .8: Confusion matrix (in %) for the Geodesic approach on dataset (2) achieved by KNN classifier. Mean correct classification rate: 24.9%.

<i>Geodesic approach on Dataset (2)</i>							
Estimated Class	Actual Class						
	NS	EEC07	EEC08	WEC	CS	NBB	SBB
NS	10	13	16	8	7	2	10
EEC07	23	32	22	27	28	19	13
EEC08	23	15	36	13	17	6	5
WEC	5	3	5	15	9	4	7
CS	18	13	13	16	24	10	11
NBB	9	13	3	12	6	23	20
SBB	12	11	5	9	9	36	34

Table .9: Confusion matrix (in %) on dataset (3) achieved by KNN classifier. Mean correct classification rate: 56% with the Fourier approach, 60% by the PCA approach and 60.5% with the Geodesic approach.

<i>Year discrimination on Dataset (3)</i>		
<i>by Fourier approach</i>		
Estimated Class	Actual Class	
	EEC07	EEC08
EEC07	54	42
EEC08	46	58
<i>by PCA approach</i>		
Estimated Class	Actual Class	
	EEC07	EEC08
EEC07	58	38
EEC08	42	62
<i>by Geodesic approach</i>		
Estimated Class	Actual Class	
	EEC07	EEC08
EEC07	64	43
EEC08	36	57

Table .10: Confusion matrix (in %) on dataset (4) achieved by KNN classifier. Mean correct classification rate: 43% with the Fourier approach, 49.5% by the PCA approach and 49.5% with the Geodesic approach.

<i>Validation test on Dataset (4)</i>		
<i>by Fourier approach</i>		
Estimated Class	Actual Class	
	NS09a	NS09b
NS09a	43	57
NS09b	57	43
<i>by PCA approach</i>		
Estimated Class	Actual Class	
	NS09a	NS09b
NS09a	46	47
NS09b	54	53
<i>by Geodesic approach</i>		
Estimated Class	Actual Class	
	NS09a	NS09b
NS09a	54	55
NS09b	46	45

Table .11: Comparison of the mean correct classification rate (in %) obtained by the three approaches on datasets (1), (2) and (3) achieved by KNN classifier.

	dataset (1)	Dataset (2)	Dataset (3)
Fourier	19.7	16.4	56
PCA	25	19	60
Geodesic	30	24.9	60.5

Table .12: Classification results on dataset (1) with the Geodesic approach when the otoliths were grouped in three classes according to their geographical zones. Mean correct classification rate: 54.3% (KNN classifier).

<i>Geodesic approach on Dataset (1)</i>			
<i>with otoliths grouped by zones</i>			
Estimated Class	Actual Class		
	Northern zone	Mixing zone	Bay of Biscay
Northern zone	53.5	29.5	13
Mixing zone	28.5	44.5	22
Bay of Biscay	18	26	65

Table .13: Classification results (in %) on dataset (1) with the Growth and Geodesic-based approach when the otoliths were grouped in three classes according to their geographical zones. Mean correct classification rate: 67.31% (SVM classifier).

<i>Growth and Geodesic-based approach on Dataset (1) with otoliths grouped by zones</i>			
Estimated Class	Actual Class		
	Northern zone	Mixing zone	Bay of Biscay
Northern zone	74.30	26.76	8.61
Mixing zone	22.31	58.25	22.00
Bay of Biscay	3.39	14.99	69.39