ICES PAPER C.M.1991/G:8
Ref. H + J

THE USE OF SHAPE FOR CLASSIFYING FISH INTO ECOLOGICAL GROUPS

Rainer Froese

ICLARM
MC P.O. Box 1501
Makati, Metro Manila
Philippines

ABSTRACT

For management purposes it is suggested to break down fish into groups of species that lend themselves to the same management regime, i.e., have a similar ecology in terms of feeding habits, growth, reproduction, and environment. Such groups could then be used to i) improve our understanding of the respective group by summarizing the knowledge available for the members of the group and to ii) estimate important parameters for those members of the group for which no data are available. A precondition to the latter is the ability to classify "unknown" fish into the appropriate ecological group. The use of morphometrics in combination with discriminant analysis is evaluated. The results indicate that the method is well suited for this task.

ICLARM contribution No. 756

1 Introduction

Fish are the most numerous group of vertebrates and live almost wherever there is water. For management purposes it is desirable to break down this large, heterogeneous group into smaller homogeneous groups, above the species level. Preferably these units should contain species that lend themselves to the same management regime, i.e., have a similar ecology in terms of feeding habits, growth, reproduction, and environment. Such a lumping of species has been suggested or practiced by a number of authors, e.g., Kesteven (1973), Lackey (1974), Royce (1972). Appeldorn & Lindeman (1985, page 509) suggest that "such a unit can be treated in many respects as a single species" and that it can be managed in the same way as an unispecies stock.

While the taxonomic status can be used for forming species clusters in some cases, it is not helpful in others when members of the same family occupy different niches. The family of Gadidae, for example, brings together a benthos feeder (*Melanogrammus aeglefinus*), a pelagic predator (*Pollachius virens*), and a pelagic plankton feeder (*Boreogadus saida*) - not to speak of *Lota lota*, the lone freshwater gadid. For statistical purposes finfish have been broken down into 13 commodity groups by the 'International Standard Statistical Classification of Aquatic Animals and Plants' or ISSCAAP (FAO 1990). The ISSCAAP code is regularly used in the FAO yearbook of fishery statistics (e.g., FAO 1988). It has also been used for other purposes such as allocating resources for research (Fearn & Davis 1991). For management purposes the ISSCAAP code has problems similar to the taxonomic status because it joins, for example, all gadoids in one group.

There is thus a need to create new groups of fish based on the above mentioned ecological criteria. Such ecological groups could then be used to i) improve our understanding of the respective group by summarizing the knowledge available for the members of the group and to ii) estimate important parameters for those members of the group for which no data are available. The practicality of the latter has been demonstrated, for example, by Pauly (1980, 1989) who derived empirical models for obtaining values of important but hard-to-estimate parameters such as natural mortality or food consumption from easy to obtain variables such as growth parameters, water temperature, or morphometric measurements. He applied his models to large, heterogeneous groups of species and was able to explain 72% and 54%, respectively, of the variance in the data sets. Similarly, Sambilay (1990) presented a model for predicting swimming speed which explains 77% of the variance in a large data set including sharks and gobies. It can be assumed that the predictive power of such models improves if they are applied to groups of fishes with similar ecology, i.e., more homogeneous subsets of the available data.

A precondition to estimate parameters for an "unknown" fish is the ability to classify it into the appropriate ecological group. While this might be straightforward for, e.g., flatfish, tuna, or small pelagics, it is much less obvious for, e.g., many coral reef fishes. The aim of this study is to evaluate the use of morphometric measurements for such classification. This implies the basic assumption that fish with similar ecology have similar swimming and feeding habits, resulting in similar morphometric features such as shape.

2 Material and Methods

2.1 Morphological distances measured

The measurements were performed on digitized morphometric drawings incorporated in FISHBASE (Froese 1990), themselves adapted from sources such as FAO species catalogs, original descriptions, and monographs on various fish taxa. Measurements were performed on the computer screen using the appropriate FISHBASE routine.

The following morphological distances were measured:
Standard Length Line (SSL), Head Length (HL), Preorbital Length (POL), Eye Length (EL), Preanal Length (PAL), Peduncle Length (PL), Preorbital Depth (POD), Post-Head Depth (HD), Maximum Depth (MD), Post-Trunk Depth (PTD), Pre-Peduncle Depth (PPD), Peduncle Depth (PD), Caudal Height (CH).

The following auxiliary distances were measured:
DPD Dorsal Preorbital Depth (part of POD that is dorsal of SLL), PAD Preorbital Aux. Depth (part of the POD that is dorsal of the point where the line touches the eye), VHD Ventral Post-Head Depth (part of HD that is ventral of SLL), MTD Mid-Trunk Depth (body depth measured on a perpendicular line at Mid-trunk Length (see below)), VMD Ventral Mid-Trunk Depth (part of MTD that

is ventral of SLL). VTD Ventral Post-Trunk Depth (part of the PTD that is ventral of SLL), VPD Ventral Pre-Peduncle Depth (part of PPD that is ventral of SLL), NPX Narrow Peduncle X (horizontal distance from the beginning of the caudal peduncle to the point of its least depth), AMX Anal Margin X (horizontal distance from end of PAL to deepest point of anal fin), AMY Anal Margin Y (vertical distance between AMX point and SLL), AEX Anal End X (horizontal distance between the end of Preanal Length and the most posterior point of the anal fin), AEY Anal End Y (vertical distance between AEX point and SLL), MCL Mid-Caudal Length (horizontal distance from base of caudal fin to the intersection of the posterior margin of the caudal fin with a hypothetical extension of SLL), TCL Top-Caudal Length (horizontal distance from the base of the caudal fin to the part of the caudal fin with maximum dorsal height), SCL Sub-Caudal Length (horizontal distance from the base of the caudal fin to the part of the caudal fin with maximum ventral extension), SCH Sub-Caudal Height (part of CH that is ventral of SLL).

2.2 Definition and derivation of landmarks

A Cartesian coordinate system was used to record the location of landmarks (Fig.1).

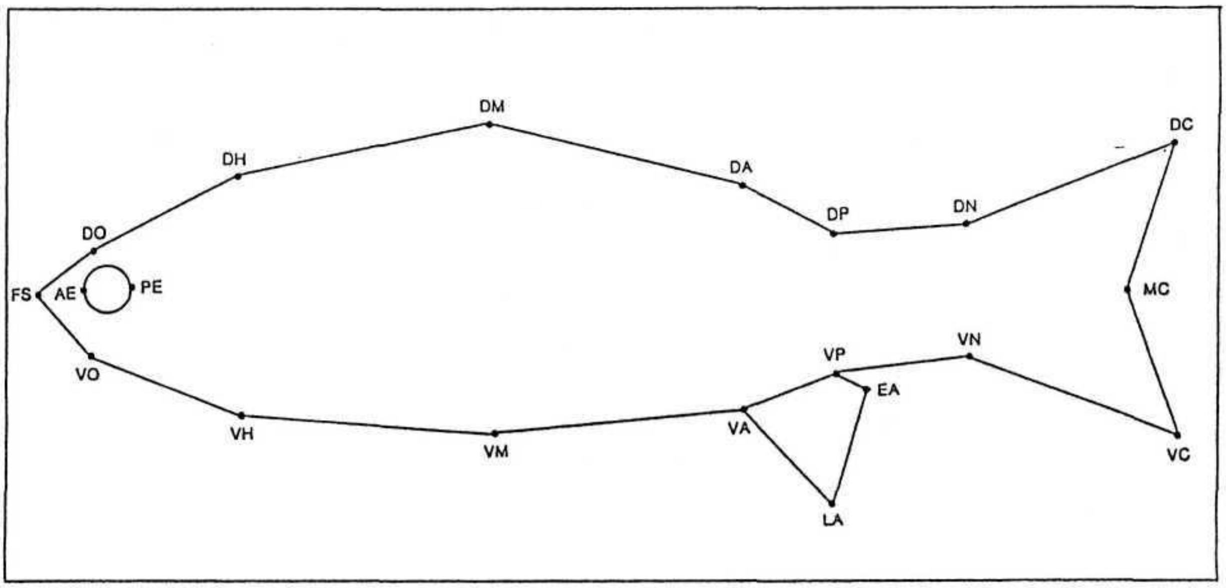

Figure 1: The figure shows the landmarks used. See text for definition of landmarks.

The tip of the snout was defined as origin of the coordinate system and the end of the Standard Length line was set to x=100, y=0. Landmarks were placed at the intersections of the body contour with the perpendicular lines of the depth measurements. The coordinates, for example, of the Ventral Post-Head Landmark **VH** are thus given as (HL,VHD), where HL = Head Length and VHD = Ventral part of Post-Head Depth (see above). The coordinates of the landmarks as shown in Fig. 1 were derived from the morphometric measurements as stated in Table 1:

Table 1: Names and coordinates of landmarks used.

TS	Tip of snout landmark	(0,0).
DO	Dorsal orbital landmark	(POL,DPD)
VO	Ventral orbital landmark	(POL,DPD-POD)
AE	Anterior tip of eye length	(POL,DPD-PAD)
PE	Posterior tip of eye length	(POL+EL,DPD-PAD)
DH	Dorsal head landmark	(HL.PHD-VHD)
VH	Ventral head landmark	(HL,-VHD)
DA	Dorsal anal landmark	(PAL,PTD-VTD)
VA	Ventral anal landmark	(PAL,-VTD)
DM	Dorsal mid-trunk landmark	(PAL-(PAL-HL)/2,MTD-VMD)
VM	Ventral mid-trunk landmark	(PAL-(PAL-HL)/2,MTD-VMD)
LA	Lowest anal fin landmark	(PAL+AMX,- AMY)
EA	Posterior tip of anal fin landmark	(PAL+AEX,-AEY)
DP	Dorsal peduncle landmark	(100-PL,PPD-VPD)
VP	Ventral peduncle landmark	(100-PL,-VPD)
DN	Dorsal narrowest peduncle landmark	(100-PL+NPX,PD/2)
VN	Ventral narrowest peduncle landmark	(100-PL+NPX,-PD/2)
DC	Dorsal caudal landmark	(100+TCL,CH-SCH)
MC	Middle caudal landmark	(100+MCL,0)
VC	Ventral caudal landmark	(100+SCL,-SCH)

2.3 Data analysis with a rotational fit method

Bookstein (1990), Chapman (1990) and Rohlf (1990) reviewed and compared several methods to superimpose the landmarks of one organism on the corresponding landmarks of another organism and to obtain some measure of goodness of fit. These methods are known as "Procustes" or rotational fit methods. For this study we used the following procedure:

1.) We assumed that the following families represent ecological groups: Clupeidae (CLUPEID), Engraulididae (ENGRAUL), Salmonidae (SALMON), Myctophidae (MYCTOPH), Serranidae (SERRAN), Carangidae (CARANG), Leiognathidae (LEIOG), Lutjanidae (LUTJAN), Mullidae (MULLID), Scombridae (SCOMB), Balistidae (BALIST) and Diodontidae (DIODON). The above mentioned measurements were performed on randomly chosen species of these groups;

2.) we constructed for every species a set of landmarks;

3.) we superimposed the landmarks of every species on those of an average (consensus) shape using a rotational fit routine provided by Rohlf (1990). The differences between the average shape and the actual shape of a species were recorded as residual vectors pointing from the average landmark to the actual position of a landmark (Fig. 2);

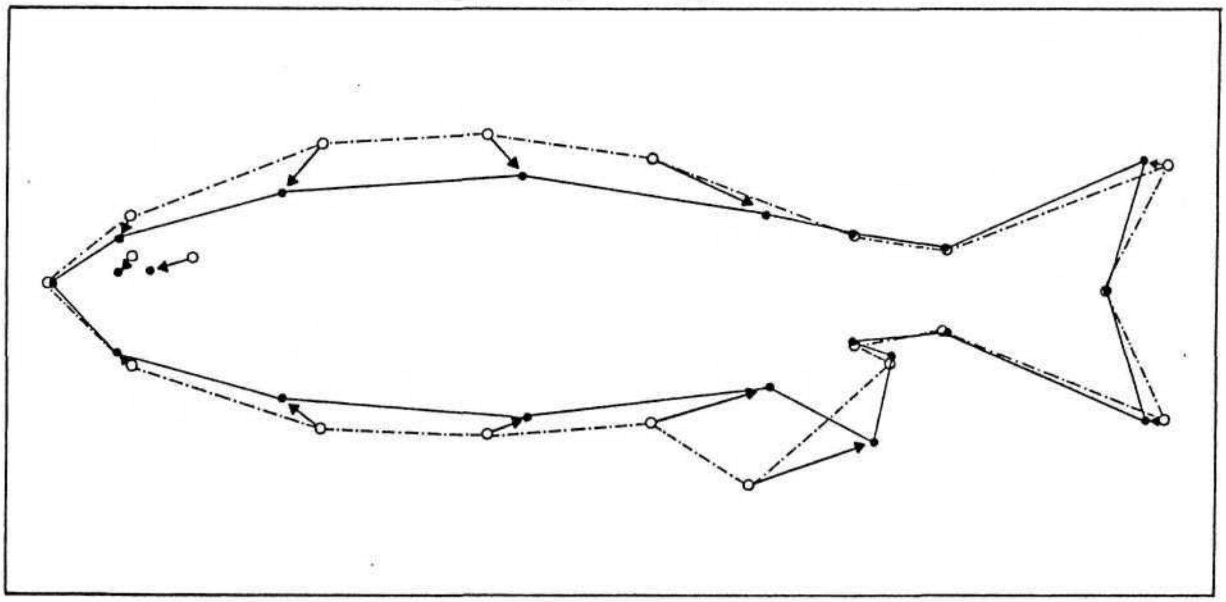

Figure 2: Shape of a species superimposed on an average shape. Vectors point from the species landmarks to the corresponding landmarks of the average shape.

4.) we ran a cluster analysis with methods 'Ward' and 'average linkage' on the residual vectors to see i) whether the proposed groups form clusters and ii) whether some of the groups could be joined;

5.) we tested the distinctness of the ecological groups by classifying all fish into those groups where their landmarks (i.e., their residual vectors) fit best. This was done with quadratic discriminant functions using a "jackknife" approach, i.e., the species to be classified was excluded from the data set used to derive the discriminant functions.

3 Results

3.1 Results of the cluster analysis

The results of the cluster analysis, method 'Ward', are shown in Table 2. A second run using the 'average linkage' method produced similar results. Although the results of the cluster analysis probably could have been improved by eliminating correlated and/or non-significant variables, respectively, we skipped this step because i) the following discriminant analysis weights the variables and ii) given the limitations of our data set we did not intend to strictly follow the results but rather use them as an additional help for forming groups.

As can be seen in Table 2 the members of the assumed ecological groups SALMON, MYCTOPH, SERRAN, LUTJAN, MULLID, and BALIST are found in one cluster per group only. The members of the groups CLUPEID, ENGRAUL, LEIOG and SCOMB are split into two clusters. The members of CARANG are distributed over four clusters.

Table 2: Results of a cluster analysis on the residual vectors. The membership in clusters (columns) is compared with the membership in the postulated ecological groups (rows).

	C1	C2	C3	C4	C5	C6	C7	C8	C9	Total
CLUPEID	–	–	–	4	–	1	–	–	–	5
ENGRAUL	–	–	–	1	–	2	–	–	–	3
SALMON	–	–	–	6	–	–	–	–	–	6
MYCTOPH	–	–	–	–	–	–	8	–	–	8
SERRAN	–	–	–	4	–	–	–	–	–	4
CARANG	–	–	1	1	–	3	–	3	–	8
LEIOG	–	–	–	–	–	1	–	7	–	8
LUTJAN	–	4	–	–	–	–	–	–	–	4
MULLID	2	–	–	–	–	–	–	–	–	2
SCOMB	4	–	3	–	–	–	–	–	–	7
BALIST	–	–	–	–	–	–	–	–	2	2
DIODON	–	1	–	–	–	–	–	–	–	1

3.2 Results of the discriminant analysis

For the discriminant analysis CLUPEID and ENGRAUL were pooled into a group of "small pelagics" (CLUENG). Also LEIOG and CARANG were pooled into a group of "small tropicals" (CARALEIO). The groups BALIST, MULLID and DIODON were excluded from the discriminant analysis because the "jackknife" option requires at least 3 members in a group.

The results of the quadratic discriminant analysis are shown in Table 3. All species were classified into the correct ecological group.

Table 3: Results of the discriminant analysis: actual membership in an ecological group (rows) against classified membership (columns).

	CLUENG	SALMON	MYCTOPH	SERRAN	CARALEIO	LUTJAN	SCOMB
CLUENG	8	-	-	-	-	-	-
SALMON	-	6	-	-	-	-	-
MYCTOPH	-		8	-	-	-	-
SERRAN	-	-	-	4	-	-	-
CARALEIO	-				16	-	-
LUTJAN	-	-	-	-	-	4	-
SCOMB	-	-	-		-	-	7

4 Discussion

4.1 The use of an Cartesian system for definition of landmarks

Because of the high morphometric variability in fish, the definition of reproducible landmarks that can be used in as many species as possible is a major problem. For the purpose of this study, the landmarks were supposed to determine the shape and the relative size of the following functional units of the fish body: snout, head, trunk, tail, caudal peduncle, caudal fin, and anal fin.

Bookstein et al. (1985) and Bookstein (1990) strongly recommended to use independent landmarks that have reliable anatomical definitions and to avoid landmarks defined by intersections of lines or by extreme distances from other landmarks. While this concept is reasonable for similar groups of fish, it cannot accommodate, e.g., a sample containing scombrids and porcupine fish. We thus had to use landmarks defined as "extreme points" or as intersections of a contour with a perpendicular to the Standard Length line, i.e., landmarks which are classified by Bookstein (1990) as "less useful".

The landmarks as we use them are understood as "boundary points" of large functional units as opposed to "boundary points" of organs or tissues

(Bookstein 1990). This seems to be justifiable as our emphasis was not the change of shape of an organ during ontogeny or evolution, but a robust representation of morphological features suitable for classifying fish into ecological groups.

4.2 Ecological groups of fish.

The ecological groups used in this study are *ad-hoc* groups rather than being the result of a dedicated study. While the ecological roles of Clupeidae and Engraulididae as "small pelagics" and of Scombridae as fast swimming predators are well understood, this is not necessarily the case for the other families. The result of the cluster analysis underlines that although the taxonomic status in some cases is an indicator of similar ecology and morphology, in other cases it is not. While the postulated groups served the purpose of this study, the identification of well-defined groups of fish with similar ecology clearly needs more research.

4.3 The suitability of landmarks for classifying fish into groups of similar shape

The suitability of morphometric measurements in combination with discriminant functions for classifying fish into groups with similar shape has already been demonstrated for fish larvae by Froese (1988). In the present study, the correct classification of all species (Tab. 3) is partly due to the small number of groups (= 7) and observations (= 53) in relation to 40 variables provided by the landmarks. For reasons of comparison we ran a second classification for the same species using 3 variables only: Maximum Depth, Peduncle Depth and Caudal Height, all expressed in per cent Standard Length without any further corrections. The results of this classification are shown in Table 4. Although the number of groups is now more than double the number of variables and the number of observations outnumbers the number of variables by a factor of 18 (i.e., overfitting is avoided), only four species out of 53 have been misclassified, resulting in a hit rate of 93% . Thus it can be concluded that morphometric measurements in general and landmark data in particular are well suited for classifying fish into ecological groups once such groups have been established and if the basic assumption that similar shape implies similar ecology holds true.

Table 4: Results of the discriminant analysis using only Maximum Depth, Peduncle Depth and Caudal Height: actual membership in an ecological group (rows) against classified membership (columns).

	CLUENG	SALMON	MYCTOPH	SERRAN	CARALEIO	LUTJAN	SCOMB
CLUENG	8	-	1	-	-	-	-
SALMON	-	6	-	-	-	-	-
MYCTOPH	-	-	7	-	-	-	-
SERRAN	-	-	-	4	-	-	-
CARALEIO	-	-	-	-	13	-	-
LUTJAN	-	-	-	-	-	4	-
SCOMB	-	-	-	-	3	-	7

5 Acknowledgement

The data and drawings used in this study are contained in FISHBASE, a global biological database on fish developed by the International Center for Living Aquatic Resources Management (ICLARM) in collaboration with the Food and Agriculture Organization of the United Nations (FAO) with support from the Commission of the European Communities (CEC).

I like to thank Mr. M. Olson Ringby for scanning and measuring the morphological drawings and Mrs. C. Janagap for performing the SAS analysis. Thanks are due to Dr. D. Pauly for useful references and fruitful discussions. 1 also thank Mr. Ch. Klingenberg for critically reviewing the manuscript.

6 References

Appeldorn, R.S. and H.C. Lindeman. 1985. Multispecies assessment in coral reef fisheries using higher taxonomic categories as unit stocks, with an analysis of an artisanal haemulid fishery, 507-514. *In* Proceedings of the Fifth International Coral Reef Congress, Tahiti, Vol. 5.

Bookstein, F.L. 1990. Introduction to methods for landmark data, 215-225. *In* F.J. Rohlf and F.L. Bookstein (eds.). Proceedings of the Michigan morphometrics workshop. Special Publication Number 2. The University of Michigan Museum of Zoology, Ann Arbor, Michigan.

Bookstein, F.L., B. Chernoff, R.L. Elder, J.M. Humphries Jr., G.R. Smith and R.E. Strauss. 1985. Morphometrics in evolutionary biology. Special Publication 15, The Academy of Natural Sciences of Philadelphia: 277 p.

Chapman, R.E. 1990. Conventional Procrustes approaches, 251-267. *In* F.J. Rohlf and F.L. Bookstein (eds.). Proceedings of the Michigan morphometrics workshop. Special Publication Number 2. The University of Michigan Museum of Zoology, Ann Arbor, Michigan.

FAO 1988. Yearbook of fishery statistics. Catches and landings 1986. FAO Fisheries Series. Vol. 62, FAO, Rome.

FAO 1990. Review of the state of world fishery resources. FAO Fish. Cir. No. 710, Rev.7. FAO, Rome.

Fearn, M. and J. Davis. 1991. Evaluation of fisheries research: an application to support international decison making. Paper presented at the 35th Annual Conference of Australian Agricultural Economic Society, University of England, Armidale, 11-14 February, 1991.

Froese, R. 1988. The use of quadratic discriminant functions in connection with video-based measurements for identification of fish larvae. ICES C.M. 1988/L:11: 8 p.

Kesteven, G.L. 1973. Manual of fisheries science. Part I. An introduction to fisheries science. FAO Fish. Tech. Pap. 188: 42 p.

Lackey, R.T. 1974. Introductory fisheries science. Sea Grant Extension Division, Virginia Polytechnic Institute and State University. Blacksburg, Virginia. 275 p.

Pauly, D. 1980. On the interrelationships between natural mortality, growth parameters, and mean environmental temperature in 175 fish stocks. J.Cons.int.Explor.Mer 39(2): 175-192.

Pauly, D. 1989. Food consumption by tropical and temperate fish populations: some generalizations. J. Fish. Biol. 35: 11-20.

Rohlf, F.J. 1990. Rotational fit (Procrustes) methods, 227-236. *In* F.J. Rohlf and F.L. Bookstein (eds.). Proceedings of the Michigan morphometrics workshop. Special Publication Number 2. The University of Michigan Museum of Zoology, Ann Arbor, Michigan.

Royce, W.F. 1972. Introduction to the fisheries sciences. Academic Press. New York: 351 p.

Sambilay, V.C.Jr. 1990. Interrelationship between swimming speed, caudal fin aspect ratio and body length of fishes. Fishbyte 8(3): 16-20.