Northwest Atlantic Fisheries Organization

Serial No. N762 NAFO SCR Doc. 83/IX/94

FIFTH ANNUAL MEETING - SEPTEMBER 1983

Morphometric Classification Between Golden Redfish (*Sebastes marinus*) and Beaked Redfishes (*S. mentella* and *S. fasciatus*)

by

D. J. Power

Department of Biology, Memorial University of Newfoundland
St. John's, Newfoundland, Canada A1C 5S7

and

I-H. Ni

Department of Fisheries and Oceans, Fisheries Research Branch
P. O. Box 5667, St. John's, Newfoundland, Canada A1C 5X1

ABSTRACT

Morphometric characters were investigated to provide criteria in aid of morphological differences that are presently employed as the guideline for species identification of redfishes in the Northwest Atlantic. Standard length was utilized as a covariate to adjust morphometric values because specimens of S. marinus were larger than those of beaked redfishes. Discriminant analysis with covariance was performed on 17 morphometric variables and resulted in an 11 variable discriminant function which explained 65% of the total variability with absolute distance between group centroids being 2.81. The discriminant function with two traditional discriminators, orbit width and length of symphsial tubercle, explained 56% total variability. The absolute distance between group centroids was 2.33. The discriminant analysis on 15 morphometrics excluding the two traditional discriminators resulted in a 10 variable function which explained 58% total variability with absolute distance between group centroids being 2.43. The result demonstrated good (87-90%) separation of the golden redfish, (Sebastes marinus) from beaked redfishes (S. mentella and S. fasciatus combined). Orbit width, interorbital width, length of symphysial tubercle (beak), depth of caudal peduncle, width of fleshly attachment of pectoral fins and body depth at the level of the pectoral fins were determined as good morphometric discriminators.

INTRODUCTION

Morphometric distinction of Northwest Atlantic redfishes (genus Sebastes) has always been a confusing topic. It is essential that a good morphological guideline for redfish field surveys be established.

Templeman and Sandeman (1957) described morphological differences to distinguish marinus type from mentella type. A complication to the classification problem arose with the suggestion of a third species Sebastes fasciatus by Barsukov (1968). This third redfish species closely resembles S. mentella in external appearance and both of these species can be termed beaked redfishes. Ni (1981) found that the route of passage of the extrinsic gasbladder muscle between ventral ribs was the most useful character for distinguishing S. mentella from S. fasciatus. Power and Ni (1982) studied the same character in S. marinus and concluded that the morphology of this muscle is significantly different among the three Northwest Atlantic redfish species. However, this technique is very time-consuming and requires special skills. Although the morphological differences between large specimens of

S. marinus and beaked redfishes were described by Templeman and Sandeman (1957), it would be of interest to have statistical confirmation of the classification based on morphometrics. Not only would such a classification complement existing criteria used to distinguish S. marinus from beaked redfishes, but would also serve as a guideline for future research surveys.

Misra and Ni (1983) introduced a discriminant analysis with covariance to morphometrics on S. mentella and S. fasciatus and found that seven characters provided good separation between these species. They employed standard length as a covariate because specimens of S. fasciatus were smaller than those of S. mentella. This would warrant the correction of morphometric differences due to size (Bliss, 1970). Previously, there hadn't been any attempt to quantitate morphometrics as being useful criteria to separate S. marinus from beaked redfishes. Discriminant analysis with covariance is, therefore, adopted on morphometrics by using standard length as covariate to discover good discriminators between S. marinus and beaked redfishes.

This study has three purposes: 1) to statistically evaluate the discriminatory power of morphometrics between S. marinus and beaked redfishes, 2) to discover whether other morphometric characters not used in the initial classification may be useful discriminators and 3) to determine the best subset of all available morphometrics using stepwise discriminant analysis.

MATERIALS AND METHODS

Specimens were collected during groundfish research surveys by the Newfoundland Biological Station in 1958 on the Flemish Cap and Hamilton Inlet Bank (Table 1). In this study species were initially classified by size of eye, bony protrusion on the lower jaw and coloration (Templeman and Sandeman 1957). A description of morphometric measurements and their abbreviations is listed in Appendix 1. The common log transformation was applied to the data as multivariate normality is usually more closely approximated by logarithms than original variables. All statistical analyses incorporated in this study were performed using BMDP programs (Dixon et al. 1981).

A valid discriminant analysis must be preceded by a significant difference in population mean vectors. Multivariate analysis of covariance (MANCOVA) is a technique used to test the equality of mean vectors (centroids) between groups while controlling for the effects of some unwanted variable. In this study, sample specimens of S. marinus (with mean standard length 354.8 mm) are larger than beaked redfishes ($\bar{x}$ = 297.4) so it is desirable to eliminate this effect, as reflected in standard length, because any real differences between these groups might be masked simply by differences in standard length. The analysis combines regression analysis (BMDP6R program) with multivariate analysis of variance (BMDP4V program). Each variable is regressed upon covariate and thereby provides a means of removing the effect of the covariate. The procedure in this study is to remove from the variable that part which is linearly related to standard length as determined from the regression equation and perform all subsequent analyses on the residuals of those variables. The result is the same as if we compared groups at the same standard length so that any differences now are independent of standard length. The technique assumes that the slope of the regression line of each variable on the covariate is the same between groups.

The test statistic used to compare these populations for differences was Hotelling's T^2 (Anderson, 1958). If a multivariate difference was found to exist between species, we then looked at univariate statistics to show which morphometrics differed. Levene's test for equality of variances was calculated by BMDP3D program which is more robust to departures from normality than the usual F statistic (Brown and Forsythe, 1974). The appropriate t-test was then conducted for equality of means. If population variances were found to be equal, the variance estimate used in computing the value for the t-statistic was pooled (averaged) between the two sample variances, if not found equal, the variance for each group was estimated separately.

Discriminant analysis was then performed by the BMDP7M program which calculates the function in a stepwise manner. Initially the variable chosen at Step 1 is the variable with the highest univariate F-statistic, that is, the variable which best discriminates between groups on one single dimension. The variable included in the function at each subseqeunt step is the one that results in the most significant F-value (F-to-enter) after adjusting for variables already included in the function (the F-value, calculated for each variable not in the function at a particular step, is that from a one-way analysis of covariance where the covariates are the variables already in the function). The stepping continues until no further information that is useful in discriminating between groups is gained by the addition of more variables. Stepwise discriminant analysis allows examination at every step the importance of variables included in the function and those still available to enter the function. This is important since a variable that may have appeared to be fairly important at an earlier step may, at a later step, become superfluous because of its relationship with variables that have been subsequently entered into the function. Such redundancy among variables often goes unnoticed if in fact one forces all available variables into the discriminant function and could result in a loss of discriminatory power (Kleinbaum and Kupper, 1978). The F-to-remove value checks to see whether a variable is still contributing to the discrimination and it will be thrown out if this partial F-value falls below the criterion F-value (5%).

For a two-group discrimination problem the discriminant coefficients are proportional to the partial regression coefficients in the multiple regression of a "dummy" group-membership variable on the predictor variables (Tatsuoka, 1971).

A problem arises in an application of discriminant analysis if the sample used in calculating the discriminant function is also used by the function to classify observations into groups. One procedure used to reduce this bias is a jackknife validation (Lachenbruch and Mickey, 1968). Each case is classified into a group according to the score obtained from the discriminant function calculated using all the data except the case being classified.

In this study the following assumptions were made: (1) multinormal distributed observations, (2) correct initial classification of observations into their respective groups and equality of dispersion matrices (for discriminant analysis) (3) common variance of error terms at each level of the covariate (for regression) and (4) equality of regression slopes between groups for each variable on the covariate. This assumption of equality of regressions slopes, as observed by Misra and Ni (1983) for beaked redfishes, seems tenable enough biologically procede with MANCOVA for S. marinus and beaked redfishes.

Discriminant analysis was performed in this study on three sets of variables: set A containing all variables measured; set B having only variables OUTSBEAK and ORBIT (Appendix 1); and set C involving all variables except those in set B.

RESULTS AND DISCUSSION

The multivariate test of equality of group centroids after adjusting the data by the covariate STANLENG results in a highly significant rejection of the null hypothesis that S. marinus and beaked redfish have the same group centroid (Hotelling $T^2 = 835.961$, corresponding F statistic = 47.35, p-value < 0.001). A univariate test of common variance was accepted ($P > 0.05$) for all variables except OUTSBEAK ($0.01 < P < 0.05$) so the assumption of equally of dispersion (covariance) matrices in both groups may not be extremely violated (Table 2). Univariate t-tests for equality of group means for each variable prior to adjustment (Table 2) indicates that ORBIT is accepted as having the same mean in both groups ($P > 0.8$). This leads to the conclusion that there is a sampling bias. Templeman and Sandeman (1957) reported that the size of eye of S. marinus is smaller than beaked redfish. This coupled with larger sizes of S. marinus than beaked redfish in the sample leads statistically to the conclusion that ORBIT is the same in both groups. We believe that standard length is responsible for masking the true nature of possible differences between the groups and therefore warrants its correction in the characters.

The first discriminant analysis on set A variables (adjusted by covariate STANLENG) was to determine good discriminators. The variables distinguishing S. marinus from beaked redfishes on a decreasing scale were: OUTSBEAK (F=38.3), BODEPPEC (26.2), CAUDPEDS (15.8), INTEREYE (15.1), ORBIT (14.3), PECTBASE (12.1), SNOTLENG (10.4), PECTLENG (5.8), SNOTANAL (5.5), INSIBEAK (5.0) and CRANRIDG (4.2). Using the classification we correctly classified 397 of 434 individuals (91.5%) by the regular method and 392 of 434 (90.3%) by the bias-reducing jackknife classification (Table 3). The value of Wilk's lambda or U-statistic was reduced from 1.0000 to 0.3486 (Table 3). Wilk's lambda reduces to 0.3729 after six variables were entered into the function and changes only 0.0243 from this value for the next five variables entering. This suggests that one can achieve just as good a discrimination with less variables even though the stepwise procedure chose the "best" 11 of 17 variables available. Lachenbruch (1975) points out that even though a stepwise procedure selects the "best" set of variables there is a strong possibility that some of these "best" variables are noise. Canonical correlation between the variables entered into the function (canonical variates) and the dummy variable representing the groups is 0.80707. This quantity squared (65.1%) is the total variability between the groups explained by the discriminant function.

The second discriminant analysis used set B which incorporates metric representatives of two traditional discriminators, OUTSBEAK and ORBIT (Templeman and Sandeman, 1957), both adjusted by STANLENG. Of the 446 specimens, 389 (87.2%) were correctly classified by both the jackknife and regular methods of classification (Table 3). Both variables entered the function and the relative importance of each variable decreased from OUTSBEAK (F=83.0) to ORBIT (62.7). Wilk's lambda is 0.5093 after OUTSBEAK entered the function and 0.4459 after ORBIT entered the function. This indicates again good discriminatory power using only two variables in a discriminant function. The function explained 56% total variability.

The discriminant analysis on set C morphometrics (excluding OUTSBEAK and ORBIT) shows good discrimination between groups based on percentage of correct classifications (Table 3). Jackknife classification correctly classified 390 of 437 (89.2%) specimens while the regular method classified 393 specimens (90.3%). Partial F-values at the final step resulted in the following decrease in importance of variables: INTEREYE (F=40.9), CAUDPEDS (40.6), BODEPPEC (29,5), CRANRIDG (20.9), PECTLENG (15.4), PECTBASE (13.1), SNOTLENG (13.0), HEADLENG (4.9). The function explained 58.2% of the total variability between species, which represents an increase in variability explained by the function of set B.

Discriminant analyses on the three sets of variables performed well in classifying specimens by the classification functions (Table 3). However, one would like to achieve parsimony in the number of variables, especially if such a function is to be applied in field studies, and still be confident that it is reliable. A function of two variables, OUTSBEAK and ORBIT, did as well as a function retaining 11 variables (set A) and almost as well as the function retaining 10 (set C).

Absolute distance between group centroids is 2.81 for variables retained in a discriminant analysis using set A, 2.43 for set C and 2.30 for the set B. It is clear to see that a discriminant function of set B again results in as good as separation.

The adjustment of the data by standard length is warranted as some variables actually show no significant difference ($P > 0.05$) after the adjustment was made (Table 2) implying that these variables showed a difference because of the influence of standard length. Figure 1 shows a histogram of values obtained by each specimen's observations evaluated by the canonical variable function. The plot of values not adjusted by STANLENG show a ger separation between groups that then adjusted plot indicating graphically that STANLENG is perhaps distorting true relationships between groups for some variables.

In conclusion, morphometric characters can be considered statistically of value in discriminating between Sebastes marinus and the beaked redfishes (S. mentella and S. fasciatus). Good discriminators, suggested by discriminant analysis, are orbit width, interorbital width, length of symphsial tubercle (beak), depth of caudal peduncle, width of fleshly attachment of pectoral fins and body depth at the level of the pectorals.

ACKNOWLEDGEMENTS

We would like to thank W. Templeman and E. J. Sandeman for the privilege of using the data; E. G. LeGrow and the redfish staff for collecting samples; Avi Singh and John Green of the Memorial University of Newfoundland and D. B. Atkinson for their valuable comments.

REFERENCES

Anderson, T. W. 1958. An introduction to multivariate statistical analysis. John Wiley & Sons, Inc. New York, N.Y. 374 p.

Barsukov, V. V. 1968. The systematic relationship of redfishes of the genus Sebastes of the Northwest Atlantic Ocean. Akad. Nauk. SSSR 183: 479-482. (Transl. from Russian in Dokl. Biol. Sci. 183: 734-737).

Bliss, C. I. 1970. Statistics in biology. Vol. 2, McGraw-Hill Book Company, New York, N.Y., 639 p.

Brown, M. B. and A. B. Forsythe. 1974. Robust tests for the equality of variances. J. Amer. Statist. Assoc. 69: 364-367.

Dixon, W. J., M. B. Brown, L. Engelman, J. W. Frane, M. A. Hill, R. I. Jennrich and J. D. Toporek. 1981. BMDP statistical software. Univ. California Press, Berkeley, Ca.

Kleinbaum, D. G., and L. L. Kupper. 1978. Applied regression analysis and other multivariate methods. Duxbury Press, North Scituate, Mass. 556 p.

Lachenbruch, P. 1975. Discriminant analysis. MacMillan Publ. Co., New York, N.Y. 128 pp.

Lachenbruch, P., and R. M. Mickey. 1968. Estimation of error rates in discriminant analysis. Technometrics 10: 1-11.

Misra, R. K. and I-H. Ni. 1983. Distinguishing beaked redfishes, *Sebastes mentella* and *S. fasciatus*, by discriminant analysis (with covariance) and multivariate analysis of covariance. Can. J. Fish. Aquat. Sci. 40: (in press).

Ni, I-H. 1981. Separation of sharp-beaked redfishes, *Sebastes fasciatus* and *S. mentella* from northeastern Grand Bank by morphology of extrinsic gasbladder musculature. J. Northw. Atl. Fish. Sci. 2: 7-12.

Power, D. J. and I-H. Ni. 1982. Morphology of the extrinsic gasbladder musculature in the golden redfish, *Sebastes marinus*. J. Northw. Atl. Fish. Sci. 3: 165-168.

Tatsuoka, M. M. 1971. Multivariate analysis: techniques for eduation and psychological research. John Wiley & Sons, Inc., New York, N.Y., 310pp.

Templeman, W., and E. J. Sandeman. 1957. Two varieties of redfish in the Newfoundland area. Fish. Res. Bd. Can. Atlant. Prog. Rept., No. 66: 20-23.

Table 1. Details of samples collected in 1958 research cruises. Sampling was by bottom-trawl net tows from Flemish Cap and Hamilton Inlet Bank.

Locality (NAFO area)	Month	Approximate Position Latitude	Longitude	Depth ranges (m)	Bottom temp. (°C)	No. of redfish in sample
Flemish Cap (3M)	June	48°00'	45°00'	276-314	4.0-4.1	58 *marinus* 39 beaked
Hamilton Inlet Bank (2J)	August	54°50'	53°50'	256-549	1.3-3.5	78 *marinus* 164 beaked
Hamilton Inlet Bank (2J)	October	53°00'	52°20'	256-549	2.6-4.0	61 *marinus* 96 beaked
Flemish Cap (3M)	November	47°40'	46°00'	272-457	3.7-3.9	36 *marinus* 30 beaked

Table 2. Mean values, standard error of mean (SE), and univariate Levene's test and t-test of morphometric characters for S. marinus (N =165)and beaked redfish (N=269). Significant levels are at P<.05 (*) and P<.01 (**). Measurements (in millimeters) were transformed by common log. Abbreviations are listed in Appendix 1.

	Beaked		Sebastes marinus		Levene	t-test	
Character	Mean	SE	Mean	SE	test	Before[b]	After[b]
STANLENG	2.4665	0.0047	2.5415	0.0068	NS[a]	**	-
HEADLENG	2.0101	0.0049	2.0671	0.0065	NS	**	**
SNOTDORS	1.9618	0.0049	2.0189	0.0065	NS	**	**
POSTPECT	2.2442	0.0045	2.2973	0.0061	NS	**	**
SNOTANAL	2.2936	0.0049	2.3732	0.0071	NS[a]	**	NS
BODEPPEC	2.0042	0.0055	2.0781	0.0074	NS[a]	**	**
BODEPANA	1.8801	0.0052	1.9649	0.0070	NS[a]	**	**
ORBIT	1.5229	0.0044	1.5243	0.0056	NS	NS	**
INSIBEAK	1.0229	0.0066	1.0438	0.0083	NS	*	**
INTEREYE	1.3207	0.0047	1.4219	0.0069	NS[a]	**	**
CRANRIDG	1.1590	0.0054	1.1944	0.0068	NS	**	**
OUTSBEAK	0.6977	0.0074	0.5930	0.0075	* [a]	**	**
SNOTLENG	1.3577	0.0054	1.4356	0.0069	NS[a]	**	NS
BODYWIDT	1.6278	0.0055	1.6952	0.0073	NS[a]	**	**
PECTLENG	1.8912	0.0043	1.9402	0.0063	NS[a]	**	**
CAUDPEDS	1.4018	0.0052	1.5505	0.0066	NS[a]	**	**
PECTBASE	1.4145	0.0049	1.5042	0.0067	NS[a]	**	**
BODYWEIG	2.7468	0.0147	2.9688	0.0205	NS[a]	**	NS

[a]Variables have significant difference (P<.05) in variance estimates before log transformation applied to data.

[b]t-test before adjustment and after adjustment for covariate STANLENG.

Table 3. Summary table of variables used and retained in three discriminant analysis applied to morphometrics and an estimate of their performance as reflected in correct classification and variation explained by discriminant function (canonical correlation squared). After each variable Wilk's lambda or U-statistic is given in brackets that is the multivariate test of equality of group means at each step. Variables are listed in the order they entered the discriminant function. Covariance adjustment for STANLENG made on variables. Abreviations of variables listed in Appendix 1.

	Variables used in computing classification functions		
	All Variables (Set A)	Only OUTSBEAK, ORBIT (Set B)	All variables except OUTSBEAK, ORBIT (Set C)
Sample size[a]	434	445	438
Variables	OUTSBEAK (.5092)	OUTSBEAK (.5079)	INSIBEAK (.7551)
	ORBIT (.4468)	ORBIT (.4449)	INTEREYE (.6271)
	INTEREYE (.4054)		CRANRIDG (.5314)
	PECTBASE (.3986)		PECTLENG (.4947)
	BODEPPEC (.3874)		CAUDPEDS (.4630)
	CAUDEPDS (.3729)		BODEPPEC (.4434)
	SNOTLENG (.3660)		PECTBASE (.4307)
	SNOTANAL (.3611)		SNOTLENG (.4224)
	PECTLENG (.3561)		HEADLENG (.4176)
	INSIBEAK (.3521)		
Correct classification	91.5%	87.2%	89.9%
Jackknife classification	90.3%	87.2%	89.2%
Canonical correlation squared (x100)	65.1%	55.5%	58.2%

[a]:Analysis performed on specimens with values for all variables in the function (complete cases only used in BMDP7M discriminant analysis).

Apendix 1. Methods of Morphometric Measurement

Measurements were made after preservation to the nearest mm, unless noted and where applicable on the left hand side of the fish.

STANLENG (standard Length)	Anterior part of upper jaw to end of hypural
HEADLENG (head length)	Anterior part of upper jaw to posterior part of operculum
SNOTDORS[a] (predorsal length)	Snouth to anterior base of 1st dorsal spine
POSTPECT[a] (postpectoral length)	Snout to posterior end of pectoral fin ray
SNOTANAL[a] (preanal length)	Snout to anterior base of 1st anal spine
BODEPPEC (body depth at pectorals)	Measured from a position just anterior to the 1st dorsal spine to a point on the opposite ventral surface and at right angles to the main longitudinal axis of the fish
BODEPANA (body depth at anal fin)	Measured from a point on the vertical surface on the broad base of the 1st anal spine to a point on the dorsal surface at right angles to the main longitudinal axis of the fish
ORBIT[b] (orbit width)	Anterior edge of orbit opposite the most anterior nostril to a point diametrically opposite to this
INSIBEAK[b] (beak length including jaw attachment)	Anterior end of beak (syphysial tubercle) to inside of lower jaw
INTEREYE (interorbital width)	Width between orbits at the location of spines on the anterior dorsal edge of the eye
CRANRIDG (width between cranial ridges)	Width taken at anterior end where ridges cease to be visable
OUTSBEAK[b] (beak length excluding jaw attachment)	Anterior end of symphysial tubercle to center of outside of lower jaw
SNOTLENG (snout length)	Upper jaw to point on orbit opposite most anterior nostril with mouth opened
BODYWIDT (body width)	Width of body between lateral line at right angles to the base of the 1st dorsal spine
PECTLENG (length of pectoral fin)	Measurement from crease made when left pectoral is raised away from body to longest ray when fin lies flat to body
CAUDPEDS (depth of Caudal peduncle)	Dorso-ventral measurement at narrowest point of caudal peduncle
PECTBASE (width of base of pectoral fin)	Width of fleshy attachment of pectoral fin
BODYWEIG (body weight)	Gilled and gutted weight of fish in grams

[a]Measurement made on line parallel to main axis of fish, to the nearest 0.5 mm

[b]Measurement made to the nearest .01 mm.

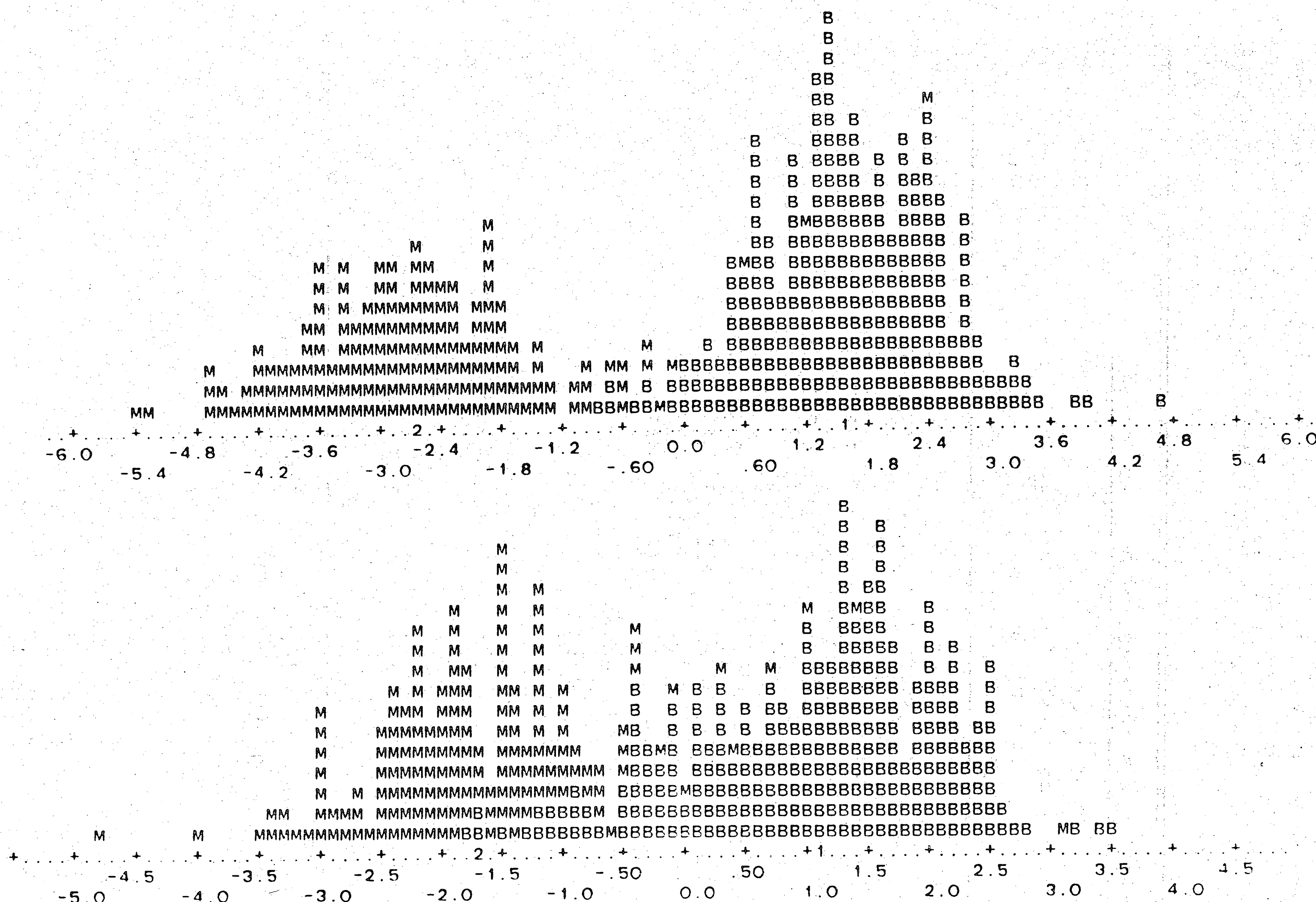

Fig. 1. Histograms of canonical variable values obtained from a discriminant analysis on Set A variables. Upper figure has no adjustment to the variables for covariate STANLENG and lower figure has the adjustment for STANLENG. In both figures "M" represents S. marinus and "B" beaked redfish.