NOT TO BE CITED WITHOUT PRIOR
REFERENCE TO THE AUTHOR(S)

International Commission for the Northwest Atlantic Fisheries

Serial No. 5355

ICNAF Res. Doc. 79/II/29

SPECIAL MEETING OF STACRES - FEBRUARY 1979

Multivariate Analysis of Meristic Characters of Capelin (*Mallotus villosus*) in the Northwest Atlantic

by

J. E. Carscadden and R. K. Misra
Fisheries and Marine Service, Newfoundland Environment Centre
St. John's, Newfoundland, Canada

Introduction

Capelin (Mallotus villosus) is the most important fish fodder resource in the Northwest Atlantic (Winters and Carscadden 1978). For decades, capelin have been used on a small-scale basis by residents of Newfoundland for food and fertilizer (Jangaard 1974). However, in 1972, a large commercial fishery for capelin began and since that time catches have been high (Winters and Carscadden 1978). Even before this fishery, Templeman (1948) and Pitt (1958) analyzed vertebral counts in an unsuccessful attempt to identify capelin stocks. Payne (1975, 1976) was also unsuccessful in discriminating stocks of capelin in the Newfoundland area using electrophoretic analysis of liver and skeletal muscle esterases. Based on known distribution patterns and growth differences, Campbell and Winters (1973) and Winters (1974) suggested that there were four major stocks of capelin in the Canadian Atlantic. These stocks were defined as the Labrador-Northeast Newfoundland stock, Northern Grand Bank-Avalon stock, South Grand Bank stock and St. Pierre-Green Bank stock. In a stock discrimination study of capelin from the Gulf of St. Lawrence and two Newfoundland locations, Sharp et al. (1978) suggested that meristic characters offered little potential for capelin stock differentation but that morphometric characters showed some promise.

The present study presents the results of a multivariate analysis of a large number of meristic characters of capelin from the Gulf of St. Lawrence and Newfoundland. To our knowledge, the use and the sequence of the statistical procedures are unique for this type of study. Intra-population differences of meristic characters between year-classes were compared, characters exhibiting significant annual differences were screened out using the Bonferroni method and then inter-population differences were compared. This paper not only presents the results of the capelin stock discrimination study but also illustrates the potential of this method for other stock discrimination problems.

Collection and Treatment of Specimens

All specimens in this study were mature males. Twelve sampling sites were chosen (Table 1 and Fig. 1). All fish were frozen prior to analysis.

The following meristic counts were made: left pectoral fin rays, dorsal fin rays, anal fin rays, pyloric caecae, branchiostegals, upper gill rakers, lower gill rakers, dorsal secondary caudal rays, ventral secondary caudal rays, precaudal vertebrae and caudal vertebrae. All meristic counts were made after cleaning and staining (Taylor 1967) with the exception of the counts of pyloric caecae which were made prior to treatment in formalin. Ages were determined from otoliths collected prior to cleaning and staining. Two ages 3 and 4 years (1972 and 1973 year-classes) were abundant in the samples and only fish of these year-classes, randomly selected from the total sample, were retained to give meristic counts.

Table 1. Sampling details and groupings for analysis of capelin used in meristic study.

Sampling site	Dates	Gear	Grouping 1	Grouping 2
Adams Cove, Conception Bay	June 21, 23, 1976	Cast net	B_1	B_2
Middle Cove	June 29, 1976	Cast net	B_1	B_2
Trinity, Trinity Bay	June 19, 1976	Dip net	B_1	B_2
Plate Cove West, Hodderville, Bonavista Bay	June 24, 1976	Dip net	B_1	B_2
St. Mary's, St. Mary's Bay	June 18, 21, 1976	Cast net	D_1	B_2
Harbour Breton, Fortune Bay	June 22, 23, 26, 1976	Seine	D_1	D_2
Southeast Shoal, Grand Bank	June 25, 1976	Midwater trawl	C_1	C_2
Twillingate, Notre Dame Bay	June 24, 28, 1976	Cast net	A_1	A_2
Quirpon, Northern Peninsula	June 24, 1976	Seine	A_1	A_2
Port au Port, St. George's Bay	June 8, 10, 15, 1976	Dip net	E_1	E_2
Lark Harbour, Bay of Islands	June 13, 14, 1976	Dip net	E_1	E_2
Sally's Cove	June 12, 1976	Dip net	E_1	E_2

Grouping of Samples for Analysis

Because Campbell and Winters (1973) had grouped capelin into four populations based on biological characters, we grouped our samples from the Atlantic coast of Newfoundland into the same four populations for analysis (Fig. 2). In addition, we created a fifth stock consisting of samples collected from the western coast of Newfoundland bordering on the Gulf of St. Lawrence. Thus, for the first analysis, there were five stocks, A_1, B_1, C_1, D_1 and E_1 (Table 1), to correspond to Campbell and Winters (1973).

It is known that mature capelin occurring in ICNAF Div. 3L in early spring are a mixture of fish that will spawn later the same year inshore in ICNAF Div. 3L and offshore in ICNAF Div. 3N (Southeast Shoal). It is not known whether any of these fish move inshore in southern Avalon (e.g. St. Mary's Bay). Thus, for a second analysis, fish collected from St. Mary's Bay were grouped with fish from Middle Cove, Conception Bay, Trinity Bay and Bonavista Bay, and Fortune Bay was treated as a separate stock. Again five stocks, A_2, B_2, C_2, D_2 and E_2, resulted from this grouping (Table 1).

Statistical Analysis and Results

Multivariate statistical procedures take account of covariance among variables and yield more meaningful information than a series of univariate statistical analyses done separately for each variable (see Morrison 1976). In particular, when the existence of reference samples is assumed, discriminant analysis is appropriate (Kendall and Stuart 1976) and may aid materially in separating populations with overlapping measurements. In our case, the reference samples are based on groupings in Table 1.

Before investigating differences between the five stocks of capelin, it was essential to determine whether mean meristic counts were different between years within each stock. The Hotelling T^2 statistic was employed to test the null hypothesis that the mean vectors of 11 meristic counts were identical for individuals of ages 3 and 4 of each group. Degree of freedom (df) for pooled within group and age covariance terms was 1013 and under the null hypothesis of equal mean vectors the probabilities (p) of exceeding associated F values were < 0.001 for B_1, < 0.005 for C_1, > 0.75 for D_1 and E_1 and close to the conventional significant level of 0.05 (F = 1.67, df = 11 and 1003) for A_1. In the second grouping, the probabilities of exceeding the associated F values were < 0.001 for B_2, < 0.005 for C_2, > 0.50 for D_2, > 0.75 for E_2 and close to 0.05 for A_2 (F = 1.70, df = 11 and 1003). Therefore, for both groupings, the null hypothesis of equal mean vectors for the age groups of stocks B_1, B_2, C_1 and C_2 were rejected and of stocks A_1 and A_2 considered suspect. Individual characters that contributed to the significant differences between ages were then identified on the basis of a simultaneous confidence intervals procedure. Since the differences in means of 11 individual characters alone were of interest, planned Bonferroni method was employed (as it usually yields shorter confidence intervals) with overall confidence coefficients of at least 0.95 and 0.99. The Bonferroni method provides results which are independent of those obtained from the T^2 analysis. At the 5% joint significance level the means for ventral secondary caudal rays in A_1, A_2, C_1 and C_2 and at the 1% joint significance level means for dorsal secondary caudal rays in A_1 and A_2 and for branchiostegals and dorsal and ventral secondary caudal rays in B_1 and B_2 differed between ages 3 and 4. No other differences between ages were significant. Stocks were then compared based on data on 8 merisitc characters designated X_j (j = 1 for left pectoral fin rays, 2 for dorsal fin rays, 3 for anal fin rays, 4 for pyloric caecae, 5 for upper gill rakers, 6 for lower gill rakers, 7 for precaudal vertebrae and 8 for caudal vertebrae) with ages pooled as single samples. All 8 characters were counted in each individual. Means, standard deviations and ranges of X_j and sample sizes are given in Tables 2 and 3.

The polynomial discriminant method of identifying stocks proposed by Cook and Lord (1978) was not feasible in the present study because stock sizes in our study are not known.

Skewness of observations on each character for each stock was examined employing the g_1 statistic (Snedecor and Cochran 1968). Only 13 out of 40 tests showed appreciable (g_1 values significant at $p < 0.01$) degrees of skewness. Furthermore, there was a lack of consistency in these tests in that no character showed skewness in all five stocks. It was noted that our samples and their mean values were large. Square root transformation, used when data are counts (Sokal and Rohlf 1969), did not alter these results. In addition, Seal (1964) has noted that the distribution of a linear compound of a large number of random variables (in our case, 8) would tend to be normal. Thus, skewness as a possible cause of concern in the present analysis was ruled out.

The equality of vectors of the means of 8 meristic counts for the 5 stocks was tested by Wilk's lambda criteria. For exposition of Wilk's lambda criteria see, e.g., Kshirsagar (1972). The likelihood ratio statistic $U_{(8, 4, 1076)}$ was 0.9136 for stocks A_1 - E_1 and $U_{(8, 4, 1076)}$ was 0.9139 for stocks A_2 - E_2, yielding $P < 0.01$ in both cases. This indicated that for each analysis, at least two stocks differed considerably in their mean values. The mean vectors for stocks were then compared in pairs employing the T^2 statistic based on pooled within-stock covariance with df = 1076 (Tables 4 and 5). These figures indicate that E_1 and C_1 are distinct from each of the other four stocks and stocks A_1, B_1 and D_1 form a third group designated G_1 such that members of G_1 are not significantly different from one another and each member of G_1 is distinct from C_1 and E_1. The results found in Table 5 for the other grouping of stocks (A_2 - E_2) differed only slightly from grouping A_1 - E_1. Stock E_2 is distinct from A_2, B_2 and D_2. The difference between C_2 and D_2 is close to the 5% level of significance. Stocks A_2, B_2 and D_2 form a third group, designated G_2, such that members of G_2 are not significantly different from one another and the third member D_2 of G_2 is distinct for E_2 and close to being significantly different from C_2. In further analysis, D_2 was accepted as being significantly different from C_2.

Table 2. Means, standard deviations and ranges of meristic characters of five stocks of capelin. See text for details of stocks and meristic characters (X_j).

Stock and sample size	Meristic character	Mean	Standard Deviation	Range
A_1 156	1	19.14	0.75	18-22
	2	13.95	0.49	13-15
	3	23.40	0.86	22-26
	4	6.17	0.89	4-8
	5	9.16	0.94	8-19
	6	27.79	1.17	25-37
	7	42.79	0.88	40-46
	8	23.72	1.03	22-33
B_2 400	1	19.14	0.76	17-21
	2	14.02	0.60	13-18
	3	23.46	0.85	20-26
	4	6.10	0.84	3-9
	5	9.13	0.54	8-11
	6	27.66	1.01	25-31
	7	42.81	0.99	39-47
	8	23.57	0.77	20-26
C_2 88	1	18.93	0.76	18-21
	2	14.02	0.48	13-15
	3	23.20	0.80	22-25
	4	5.80	0.87	4-8
	5	9.09	0.54	8-10
	6	27.60	1.07	24-30
	7	42.81	0.73	40-44
	8	23.67	0.83	22-28
D_2 203	1	19.01	0.76	17-21
	2	14.00	0.52	13-15
	3	23.52	0.81	22-26
	4	6.07	0.88	4-9
	5	9.20	0.51	8-11
	6	27.85	1.09	25-30
	7	42.92	0.89	39-45
	8	23.55	0.67	22-25
E_2 234	1	19.32	0.74	18-21
	2	13.87	0.52	12-15
	3	23.57	0.87	21-26
	4	6.12	0.82	4-8
	5	9.28	0.61	8-11
	6	27.96	1.02	25-31
	7	42.66	1.03	40-45
	8	23.69	0.80	19-26

Table 3. Means, standard deviations and ranges of meristic characters of five stocks of capelin. See text for details of stocks and merisitc characters (X_j).

Stock and sample size	Meristic character (X_j)	Mean	Standard Deviation	Range
A_2 156	1	19.14	0.75	18-22
	2	13.95	0.49	13-15
	3	23.40	0.86	22-26
	4	6.17	0.89	4-8
	5	9.16	0.94	8-19
	6	27.79	1.17	25-37
	7	42.79	0.88	40-46
	8	23.72	1.03	22-33
B_2 489	1	19.14	0.76	17-21
	2	14.02	0.60	13-18
	3	23.46	0.85	20-26
	4	6.10	0.84	3-9
	5	9.13	0.54	8-11
	6	27.66	1.01	25-31
	7	42.81	0.99	39-47
	8	23.57	0.77	20-26
C_2 88	1	18.93	0.76	18-21
	2	14.02	0.48	13-15
	3	23.20	0.80	22-25
	4	5.80	0.87	4-8
	5	9.09	0.54	8-10
	6	27.60	1.07	24-30
	7	42.81	0.73	40-44
	8	23.61	0.83	22-28
D_2 114	1	19.01	0.76	17-21
	2	14.00	0.52	13-15
	3	23.52	0.81	22-26
	4	6.07	0.88	4-8
	5	9.20	0.51	8-11
	6	27.85	1.09	25-30
	7	42.92	0.89	40-45
	8	23.55	0.67	22-25
E_2 234	1	19.32	0.74	18-21
	2	13.87	0.52	12-15
	3	23.57	0.87	21-26
	4	6.12	0.82	4-8
	5	9.28	0.61	8-11
	6	27.96	1.02	25-31
	7	42.66	1.03	40-45
	8	23.69	0.80	19-26

Since in both groupings stocks C, E and G differed significantly from one another, their analysis was extended to discriminant functions using the multiple regression approach with a dummy dependent variable (see e.g. Kshirsagar 1972). In analyzing data in G_1 and G_2, the component samples A_1, B_1, D_1 and A_2, B_2 and D_2 respectively were not pooled into a single sample. In choosing linear contrasts for discriminant function analysis, samples A_1, B_1, D_1 and A_2, B_2 and D_2 were considered as nested within G_1 and G_2 respectively. Paired comparisons of C_1, E_1 and G_1 and C_2, E_2 and G_2 were done to obtain a single linear composite of 8 counts for each pair. Such linear compounds of counts, their mean values and standard errors (SE) are useful in studies which are based on the differences between members of a pair (e.g. to identify an unknown individual on the basis of these counts) (Tables 6 and 7). An analysis of variance for the discriminant functions yielded F values (df = 8 and 1069) which were significant ($p < 0.001$ for C_1 and E_1, C_2 and E_2 E_1 and G_1, E_2 and G_2, and $p < 0.005$ for C_1 and G_1, C_2 and G_2). These results indicate that the discriminant functions were all highly significant and provided effective separation between populations.

The data were analyzed using computer programs written by R. Misra and any enquires regarding these programs should be directed to him.

Table 4. Paired comparisons of the five stocks of capelin. The symbols *, **, *** indicate that for the associated F value with df = 8 and 1069, $p < 0.05$, $p < 0.01$ and $p \leq 0.005$ respectively.

Stock	A_1	B_1	C_1	D_1	E_1
A_1			**		*
B_1			***		***
C_1	**	***		**	***
D_1			**		***
E_1	*	***	***	***	

Table 5. Paired comparisons of the five stocks of capelin. The symbols *, **, *** indicate that for the associated F value with df = 8 and 1069, $p \leq 0.05$, $p \leq 0.01$ and $p \leq 0.005$ respectively.

Stock	A_2	B_2	C_2	D_2	E_2
A_2			**		*
B_2			***		***
C_2	**	***		Close to 5%	***
D_2			Close to 5%		***
E_2	*	***	***	***	

Table 6. Discriminant functions, their means and standard errors. See text for details of stocks and meristic characters.

	Discriminant coefficient for meristic characters (X_j)								Mean discriminant score ± SE
Pair	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	
C_1,G_1	0.1587	-0.0906	0.2188	0.2635	0.0508	0.0484	0.0112	-0.0877	8.54 ± 0.0360 8.73 ± 0.0123
E_1, G_1	0.5199	-0.7652	0.2099	0.0099	0.3809	0.2453	-0.2399	0.1468	8.08 ± 0.0516 7.69 ± 0.0286
C_1, E_1	0.1076	-0.1153	0.0900	0.0804	0.0591	0.0428	-0.0242	-0.0095	3.43 ± 0.0176 3.57 ± 0.0108

Table 7. Discriminant functions, their means and standard errors. See text for details of stocks and meristic characters.

	Discriminant coefficient for meristic characters (X_j)								Mean discriminant score ± SE
Pair	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	
C_2, G_2	0.1597	-0.0900	0.2184	0.2642	0.0507	0.0480	0.0111	-0.0878	8.54 ± 0.0360 8.73 ± 0.0123
E_2, G_2	0.5199	-0.7651	0.2100	0.0097	0.3809	0.2455	-0.2399	0.1469	8.09 ± 0.0516 7.70 ± 0.0286
C_2, E_2	0.1079	-0.1151	0.0898	0.0806	0.0591	0.0427	-0.0243	-0.0095	3.43 ± 0.0176 3.58 ± 0.0108

Discussion

Our analysis of meristic characters defined only three major stocks of capelin in the Northwest Atlantic. These can be designated as 1) Gulf of St. Lawrence 2) insular Newfoundland (Atlantic) and 3) Southeast Shoal. All stocks are distinct from one another and were described by discriminant functions. However, based on known migration patterns and general biology of capelin (Campbell and Winters 1973) it is possible to identify at least 5 capelin stocks in the area. An analysis of morphometrics done by Sharp et al. (1978) was effective in defining capelin stocks, although the stocks they investigated were not the same as the stocks we analysed. Considerable overlapping of ranges of individual characters in all stocks (Table 2 and 3) may have restricted the scope of meristics in discriminating between stocks.

The identification of the Southeast Shoal spawning population as a distinct stock supports the conclusions of Campbell and Winters (1973), Winters (1974) and Sharp et al. (1978). This is significant because the interest in management of this stock has been intense due to a dramatic decline in the catches in this fishery (Carscadden et al. 1978). Prior to migrating to the spawning grounds on the Southeast Shoal this stock is subjected to a commercial fishery in ICNAF Div. 3L. This fishery also takes capelin that will migrate inshore to spawn on Newfoundland beaches at the same time the Southeast Shoal fishery is being prosecuted.

Based on biological evidence (Campbell and Winters 1973), the capelin spawning on the south coast of Newfoundland were defined as a separate stock. However, the possibility also exists that there is some mixing in ICNAF Div. 3L during the spring fishery with fish destined to spawn inshore in eastern Newfoundland and offshore on the Southeast Shoal. Because the meristic analysis did not define the south coast and eastern Newfoundland populations as a separate stocks, the degree of intermixing cannot be estimated as a result of this analysis.

Our analysis was more effective in separating Gulf of St. Lawrence and Atlantic stocks than that of Sharp et al. (1978). This is probably due to our more vigorous treatment of the data in which we eliminated characters that exhibited large variations between year-classes.

In this respect, although meristic characters show only limited promise in separating capelin stocks in the Northwest Atlantic the statistical procedures offer good potential in screening characters in other studies. None of the statistical operations are original in themselves but to our knowledge the use of at least some of them has not been attempted in fisheries biology prior to this study.

This is true of the Bonferroni method which was used to identify meristic characters that varied between years within populations. The Hotelling T^2 analysis suggested that mean counts of meristic characters were significantly different between ages within a population. Thus, further testing of differences between populations probably would not have been attempted because of this large intra-population variation. However, the Bonferroni method allowed us to identify specific meristic characters that accounted for this variation. After eliminating highly variable characters, the remaining characters showed no significant intra-population differences between years and the analysis was continued.

Thus, although our results for capelin based on meristic analyses are limited in their practical application, the use of some of the statistical procedures in this study appear to be the first of their kind in fisheries biology and offer good potential in future stock discrimination studies.

References

Campbell, J. S. and G. H. Winters. 1973. Some biological characteristics of capelin, Mallotus villosus, in the Newfoundland area. Int. Comm. Northwest Atl. Fish. Res. Doc. 73/90, Ser. No. 3048, 3 p.

Carscadden, J. E., D. S. Miller and G. H. Winters. 1978. Biological aspects of the Div. 3NO capelin stock in 1977 and a capelin sequential abundance model. Int. Comm. Northwest Atl. Fish. Res. Doc. 78/52, Ser. No. 5217, 20 p.

Cook, R. C. and G. E. Lord. 1978. Identification of stocks of Bristol Bay sockeye salmon, Oncorhynchus nerka by evaluating scale patterns with a polynomial discriminant method. Fish. Bull. 76 (2): 415-423.

Jangaard, P. 1974. The capelin (Mallotus villosus) biology, distribution, exploitation, utilization, and composition. Bull. Fish. Res. Board Can. 186, 70 p.

Kendall, M. G. and A. Stuart. 1976. The advanced theory of statistics- Vol. 3. Hafner Publishing Company, New York, N.Y.

Kshirsagar, A. M. 1972. Multivariate analysis. Marcel Dekker, Inc., New York, N.Y.

Morrison, D. F. 1976. Multivariate statistical methods. McGraw-Hill Book Company, New York, N.Y.

Payne, R. H. 1975. Esterase polymorphism in the capelin, Mallotus villosus: Preliminary evidence for geographical variation in allele frequencies at three loci. Cons. Perm. Int. Explor. Mer C.M. 1975/H:28: 9 p.

1976. Further studies on the biochemical population genetics of the capelin, Mallotus villosus: Demonstration that the capelin populations of West Greenland and northern North America are genetically distinct. Cons. Perm. Int. Explor. Mer C.M. 1976/H:24: 7 p.

Pitt, T. K. 1958. Distribution, spawning and racial studies of the capelin, Mallotus villosus (Muller), in the offshore Newfoundland area. J. Fish. Res. Board Can. 15: 275-293.

Seal, H. 1964. Multivariate statistical analysis for biologists. John Wiley and Sons Inc., New York, N. Y.

Sharp, J. C., K. W. Able, W. C. Leggett and J. E. Carscadden. 1978. Utility of meristic and morphometric characters for identification of capelin (Mallotus villosus) stocks in Canadian Atlantic waters. J. Fish. Res. Board Can. 35: 124-130.

Snedecor, G. W. and W. G. Cochran. 1968. Statistical methods. Iowa State Univ. Press. Ames. Iowa.

Sokal, R. R. and F. J. Rohlf. 1969. Biometry. W.H. Freeman and Company, San Francisco.

Taylor, W. R. 1967. An enzyme method of cleaning and staining small vertebrates. Proc. U.S. Nat. Mus. 122: 1-17.

Templeman, W. R. 1948. The life history of the capelin (Mallotus villosus Muller) in Newfoundland waters. Nfld. Gov. Lab. Bull. 17: 1-151.

Winters, G. H. 1974. Rationale for partition of capelin quota in Subareas 2 and 3. Int. Comm. Northwest Atl. Fish. Res. Doc. 74: 12, Ser. No. 3158, 2 p.

Winters, G. H. and J. E. Carscadden. 1978. Review of capelin ecology and estimation of surplus yield from predator dynamics. Int. Comm. Northwest Atl. Fish. Res. Bull. 13: 21-30.

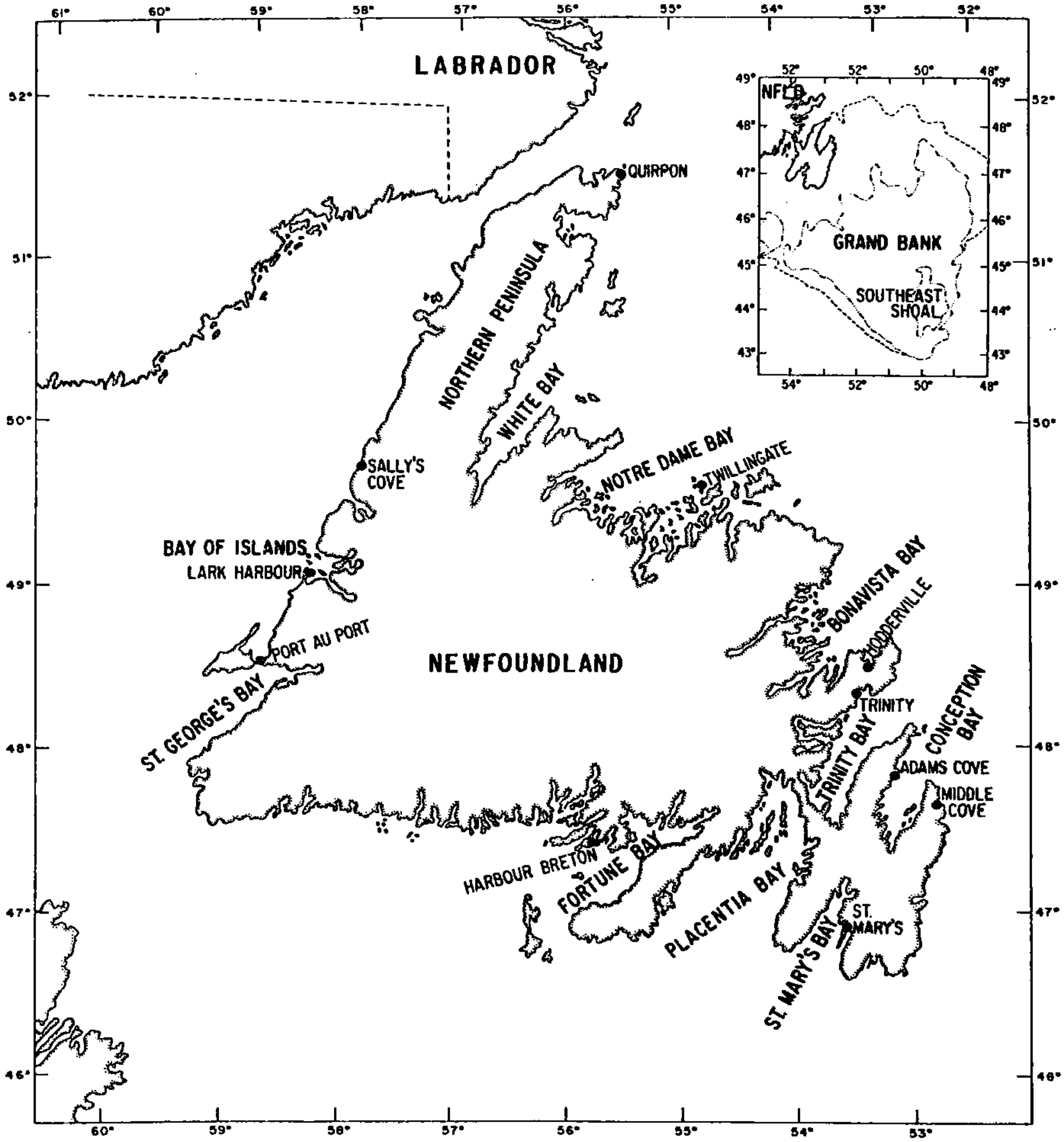

Fig. 1. Location of sampling sites for capelin meristic study.

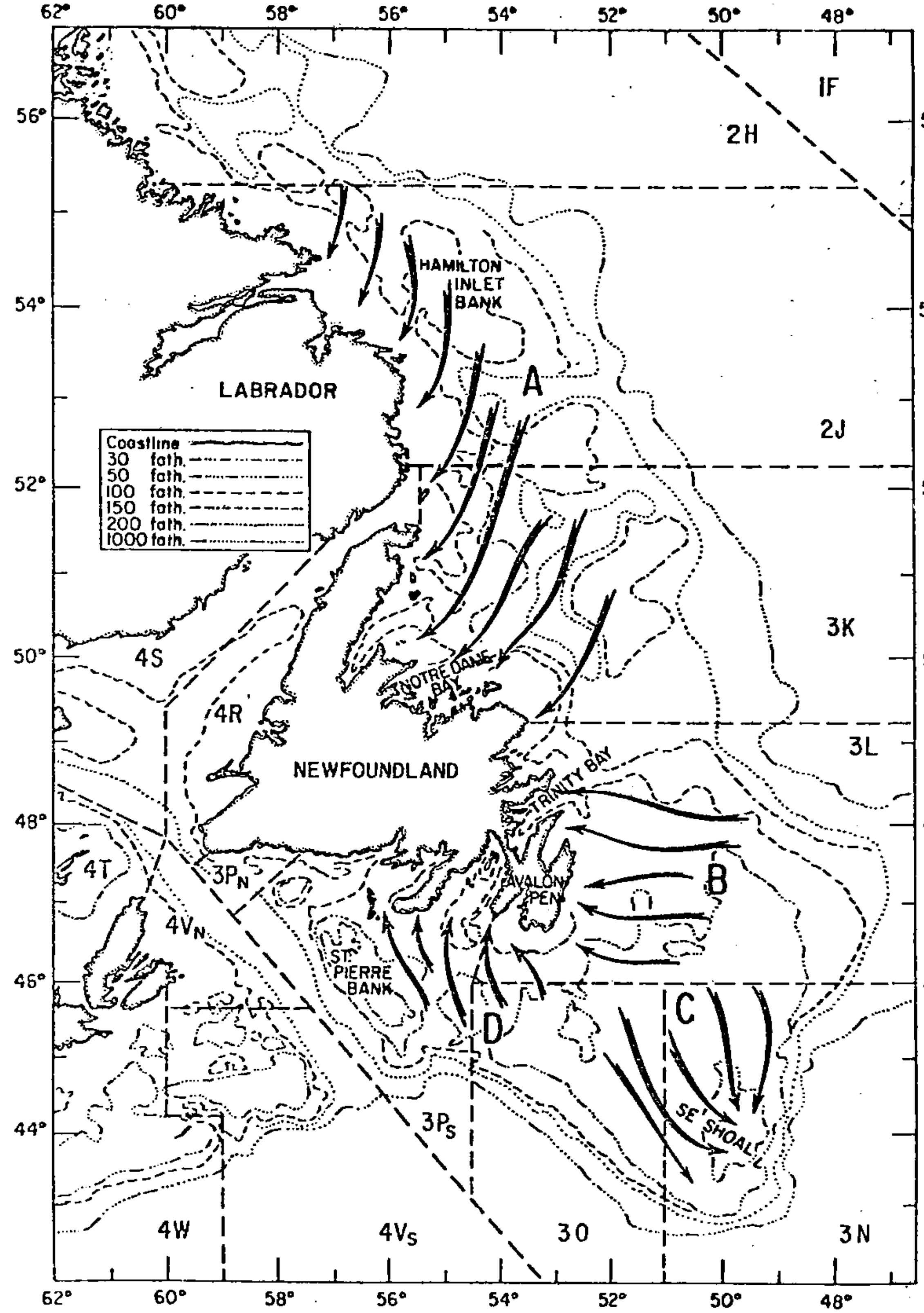

Fig. 2. Grouping of capelin into populations based on Campbell and Winters (1973).