Northwest Atlantic Fisheries Organization

Serial No. N707

NAFO SCR Doc. 83/VI/49

SCIENTIFIC COUNCIL MEETING - JUNE 1983

Stock Discrimination of Capelin (*Mallotus villosus*) in the northwest Atlantic using Meristic Characters

by

R.K. Misra
Department of Fisheries and Oceans
P.O. Box 550
Halifax, Nova Scotia B3J 2S7

and

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

INTRODUCTION

Capelin (*Mallotus villosus*) is considered one of the most important forage species in the northwest Atlantic. Although capelin was the object of a relatively small domestic fishery for decades, it is only in the last decade that it has achieved any signficiant importance as a commercial species. During the 1970's, capelin were taken in a large offshore international fishery and more recently they have been taken by an inshore Newfoundland-based fishery to satisfy the lucrative market for roe capelin in Japan. In spite of the importance of capelin as a commercial species and a forage species, firm evidence for the existence of stocks of capelin does not exist although there are considerable sources of circumstantial data available on which tentative stock divisions have been made (see Carscadden and Misra, 1980 for a summary). They examined differences among five tentative stock divisions including the four suggested by Campbell and Winters (1973), based on multivariate analysis of data on eleven meristic characters.

Multivariate methods which are generally used, including those used by Carscadden and Misra (1980) assume that the data are multivariate normal. This is not true for meristic data. Recently Bowering and Misra (1982) presented a multivariate generalized distance method (including a computer program in FORTRAN IV) which is appropriate for comparing populations based on meristic data. We felt that the capelin data in Carscadden and Misra (1980) should be re-analyzed using the new methodology. Furthermore in addition to analyzing the data by combining 3- and 4-year-old individuals, we analyzed the data separately for each age group to assess annual differences in the meristic characters.

This paper presents the results of the re-analysis of the same data (with some minor exclusions) presented in Carscadden and Misra (1980) using the new methodology.

Statistical Analysis and Results

The new method and computer program are described in Bowering and Misra (1982) and therefore, only a brief outline of the procedure is given here. Two measures of generalized distance, designated B^2 and G^2 and based on "within" and "total" covariance matrices, respectively, are employed. These two measures generally yield results which are remarkably similar (Kurczynski, 1970; Bowering and Misra, 1982) as is indeed the case with the capelin analysis reported here. Differences between B^2 and G^2 are essentially theoretical in nature with little practical significance. The findings of Bowering and Misra (1982) based on these procedures agreed remarkably well with the results of Fairbairn (1981) who employed an entirely different technique for stock identification. The methodology not only tests significance of differences between samples but also assigns numerical values to distances between them and thereafter ranks the samples. Counts on a meristic character 'J' yield a multinomial distribution with S_J+1 classes (S_J being the number of "independent" classes).

Numbers of individuals in the K-th class of jth character in the i-th population (P_i) may be expressed as the proportion p_{ijk} of the total number of individuals in P_i. Based on a two-way (population x class) layout, generalized distance analysis of characters considered first individually and then collectively, is conducted. The computer program (Bowering and Misra, 1982) provides for analysis of mutually correlated characters by multiple classification procedure. The program tests the significance of Spearman's coefficient of rank correlation between characters. This measure of correlation does not require the assumption of a bivariate normal distribution. Even if two characters are correlated, there may be arguments in favour of disregarding the correlation (Bowering and Misra, 1982). Therefore, generalized distance analysis where characters are considered independent must be done in any case. In addition to the values of B^2 and G^2 the program calculates values of χ^2 to test the significance of differences between samples. It also calculates normalized χ^2 values which may be used to rank characters based on their contributions to differences between samples and to rank samples.

The eleven meristic characters reported by Carscadden and Misra (1980) were coded as follows: left pectoral fin rays (1), dorsal fin rays (2), anal fin rays (3), pyloric caecae (4), branchiostegals (5), upper gill rakers (6), lower gill rakers (7), dorsal secondary caudal rays (8), ventral secondary caudal rays (9), precaudal vertebrae (10), and caudal vertebrae (11). Carscadden and Misra (1980) combined 3- and 4-year-old individuals and compared five groups (designated A, B, C, D, and E) by analyzing data on eight of the meristic characters; the other three characters were excluded from the final analysis because individuals of the two age groups differed with respect to these characters. In the present study samples were compared based on individuals of each of the two age groups considered separately as well as of the two age groups combined. Fish having no missing observations on all meristic characters were included in the analysis. In addition, three fish were eliminated from stock A because a closer examination of the raw data indicated very high counts in some characters. This was considered a conservative approach since these high counts could enhance any differences between stocks if included in the analysis. The groupings tested were identical to those given in Carscadden and Misra (1980) and reproduced here (Table 1, Fig. 1) for reference. The groups tested, samples sizes (n), and ranges of meristic characters for samples of complete specimens are given in Table 1. All signifiant values in the text or tables were marked significant at least at the 5% probability level (i.e. $p \leq 0.05$) of significance in B^2 and/or G^2 analyses, and multiple comparisons were conducted by the Bonferroni procedure (Morrison, 1976) unless otherwise stated. The paired comparisons of the five groups were analyzed by B^2 and G^2 procedures for the characters considered individually as well as collectively. As explained in Bowering and Misra (1982), their computer program ensures that no expected cell frequency of the sample X class layout is less than the value assigned to the variable named "AMIN". Since the analysis based on the smallest expectation of at least one should be accurate enough (Snedecor and Cochran, 1967, p. 235), the value of one was assigned to AMIN in every analysis of comparisons presented in this paper.

I. Comparison of samples based on separate age groups

None of the paired comparisons of samples was significant for characters 1, 2, 6, 7, and 11 (considered individually) of 3-year-old individuals and for characters 2, 3, 4, 5, 6, 7, and 11 of 4-year-old individuals in B^2 and G^2 analysis. Therefore, these characters were omitted in the analyses of paired comparisons of samples based on collective contributions of characters. The numbers of complete specimens changed with the number of characters employed in an analysis. Values for B^2, G^2, χ^2 and normalised χ^2 are tabulated only for characters collectively, as follows: matrices of B^2 (below the diagonal) and of G^2 (above the diagonal) values are given in Table 2 for six characters of 3-year-old individuals and in Table 3 for four characters of 4-year-olds. Tables 4, 5, 6, and 7 give corresponding χ^2 and normalized χ^2 values (df = degrees of freedom). The values of B^2 and G^2 in Tables 2 and 3 were generally similar. Paired comparisons of samples which were significant (marked 'X') based on B^2 and G^2 analyses for individual characters are given in Table 8 and 9. Chi-square values from the analyses of 3- and/or 4-year-old individuals were significant (all at values of p far less than 0.05) with one exception. The difference between groups A and B was not significant when characters were considered collectively. However, B^2 and G^2 analyses of character eight for 3-year-old individuals gave values of χ^2 of 14.8 and 14.4 (df = 4), respectively; the former was significant and the latter was nearly so.

The ranks for paired comparisons, determined from normalized χ^2 values, were the same for the B^2 and G^2 analyses. Mean ranks (based on figures in Tables 6 and 7 combined) for the groups A, B, C, D, and E were 3.8, 5.0, 8.4, 4.6, and 5.8, respectively.

Only the correlation between characters 8 and 9 was significant in each of the five groups. Therefore, B^2 and G^2 analyses were conducted by employing a multiple classification base for these two characters. The multiple combination resulted in a severe collapsing of classes of the two-way layout. This may have occurred for the following reasons: (1) There was severe disparity of samples sizes (Table 1). (2) One sample, C, was very small (42 for 3-year-olds and 45 for 4-year-olds). When two characters are correlated, larger frequencies in a two-way (sample x class) layout will occur in the northwest-southeast diagonal and will decrease away from this diagonal. In the collapsing procedure of Bowering and Misra (1982), it is obvious that when frequencies are distributed in this fashion, the greater disparity of samples sizes, the more severe the collapsing will be. (3) The correlation between the two characters was not strong (coefficients of correlation values were in the range 0.58 to 0.78). Since severe collapsing weakens the sensitivity of χ^2 analysis (Snedecor and Cochran, 1967), the generalized distance analysis employing multiple combined classification was not carried further.

II. Comparison of age groups

Groups of 3-year-old and 4-year-old individuals were compared for each of the five samples. For sample A, characters 8 and 9 differed for the two age groups. Characters 5, 8, and 9 differed for age 3 and 4 fish in sample B. In sample C, characters 2, 8, and 10 differed for the two age groups. There were no significant differences between the age groups in D and E. Since none of the eleven characters differed consistently for the two age groups in all five samples, it was decided to compare the five samples with individuals of the two age groups combined based on all eleven characters and also on characters 1, 3, 4, 6, 7, and 11 which consistently did not differ for the two age groups in all five samples.

III. Comparison of samples based on groups of 3- and 4-year-old individuals combined

a. Using eleven characters:

The generalized distance analysis where characters were considered independent resulted in similar B^2 and G^2 values. No pairs of samples differed significantly with respect to characters 2, 7, 10, and 11. All paired comparisons of samples were significant (marked 'X' in Table 10) based on B^2 and/or G^2 analyses for individual characters and groups of characters, except the pairs (A,B) and (B,D). When samples of these pairs were analysed independently, that is, without including the other three samples in the generalised distance analysis, A and B differed (marked 'S' in Table 10) for character 4. Based on normalized χ^2 values, the mean ranks for A, B, C, D, and E were 4.0, 4.2, 8.5, 4.8, and 6.0, respectively.

The following characters were significantly correlated in all five samples: (1, 6), (5, 6), (5, 8), and (8, 9). Multiple classification using the criterion AMIN = 1 severely collapsed the classes for combinations 1 and 6 (as many as 97% of the total number of fish were confined to just 2 classes) and 5 and 6 (all fish were in 1 class). Therefore, paired comparisons of samples were conducted for only 1 multiple classification of the combined characters 5, 8, and 9. This multiple classification analysis based on characters 1, 3, 4, 6 and the combined character did not add any new information.

b. Using characters 1, 3, 4, 6, 7, and 11 only:

Again, the values of B^2 and G^2 were similar. No pair of samples differed significantly for characters 7 or 11. The B^2 and G^2 analysis for the collective contribution of the remaining four characters identified only three groups of five samples ie. C, E, and the third group of samples A, B, and D since no paired comparison for samples of the third group were significant. Samples of the third group did not differ with respect to any individual character either. When samples of pairs (A, B), (B, D), and (A, D) were analyzed independently of the other three samples, samples A and B differed for characters 4 (as stated earlier) and samples A and D differed significantly for character 5 and for all characters considered collectively. Mean ranks for samples A, B, C, D, and E were 3.8, 4.5, 6.8, 4.5, and 8.0, respectively.

DISCUSSION

The methodology used in this paper, considered by the authors and Bowering and Misra (1982) to be more appropriate for meristic data than previously used (Carscadden and Misra, 1980), has been applied to the same meristic data reported elsewhere (Carscadden and Misra, 1980). The results of the present analysis substantiate the results of the previous analysis in that in all combinations of data tested, that is, by combining or by separating the two age groups, the capelin spawning on the Southeast Shoal (Stock C) and on the west coast of Newfoundland (Stock E) continued to be identified as separate stocks. The inshore spawning stocks (Labrador-northeast Newfoundland (A), northern Grand Bank-Avalon (B), and St. Pierre-Green Bank (D)) which could not be shown to be separate stocks in the former treatment (Carscadden and Misra, 1980) continued to be problematic although there were indications that most of the stock designations are valid. For instance, when age groups were considered separately and characters collectively, the south coast stock was separate but stocks on the Labrador-northeast coast and northern Grand Bank-Avalon could not be readily separated. One character (8 - caudal vertebrae), though, was significantly different for 3-year-olds using the B^2 analysis and nearly so for G^2 analysis. Interestingly, when age groups were combined a significant difference occurred between these same stocks for character 4 (pyloric caecae) for both B^2 and G^2 analyses. These comparisons suggested that Labrador-northeast Newfoundland (A) and northern Grand Bank-Avalon (B) stocks were indeed different.

When fewer characters (1, 3, 4, 6, 7, and 11) and age groups combined were analyzed, only three groups could be identified and these groups agreed with the groupings identified formerly (Carscadden and Misra, 1980); Southeast Shoal stock (C), west coast of Newfoundland (E), and one group consisting of the other three stocks - Labrador-northeast Newfoundland (A), northern Grand Bank-Avalon (B), and St. Pierre-Green Bank (D). However when the latter three groupings were analyzed independent of the groups that were already identified as being different, the Labrador-northeast Newfoundland (A) and St. Pierre-Green Bank (D) were significantly different based on seven characters.

In summary, the west coast of Newfoundland and Southeast Shoal stocks continued to be clearly identified as separate stocks while for the remaining inshore stocks the delineations were less clear although the stocks did appear to be different. The success of separating the west coast of Newfoundland capelin (E) and Southeast Shoal capelin (C) from the other groups and the problems in separating the other three groupings (A, B, and D) from each other by meristics may not be surprising in view of the possible environmental effects that may influence the determinatiion of meristic characters during the early development stages (see Bowering and Misra, 1982 for brief review). The Southeast Shoal stock of capelin is known to spawn on the bottom (approximately 50 m depth) where water temperatures are 2-4°C (Pitt, 1958). The other capelin are beach spawners and prefer water temperatures of 5.5-8.5°C. However, the capelin spawning on the west coast of Newfoundland are reported to spawn more in deeper water because the water temperatures near the beach warm very rapidly due to the reduced influence of the Labrador current (Jangaard, 1974). Thus, the somewhat different spawning conditions may also mean that developmental conditions for the eggs and yolk-sac larvae are different which in turn may have an influence on determination of meristic characters.

ACKNOWLEDGEMENTS

Computations were done by Mr. D Kelder and Mr. B. Gellately. This assistance is gratefully acknowledged.

REFERENCES

Bailey, R.F.J., K.W. Able and W.C. Leggett. 1977. Seasonal and vertical distribution and growth of juvenile and adult capelin (Mallotus villosus) in the St. Lawrence estuaries and Western Gulf of St. Lawrence. J. Fish. Res. Board Can. 34: 2030-2040.

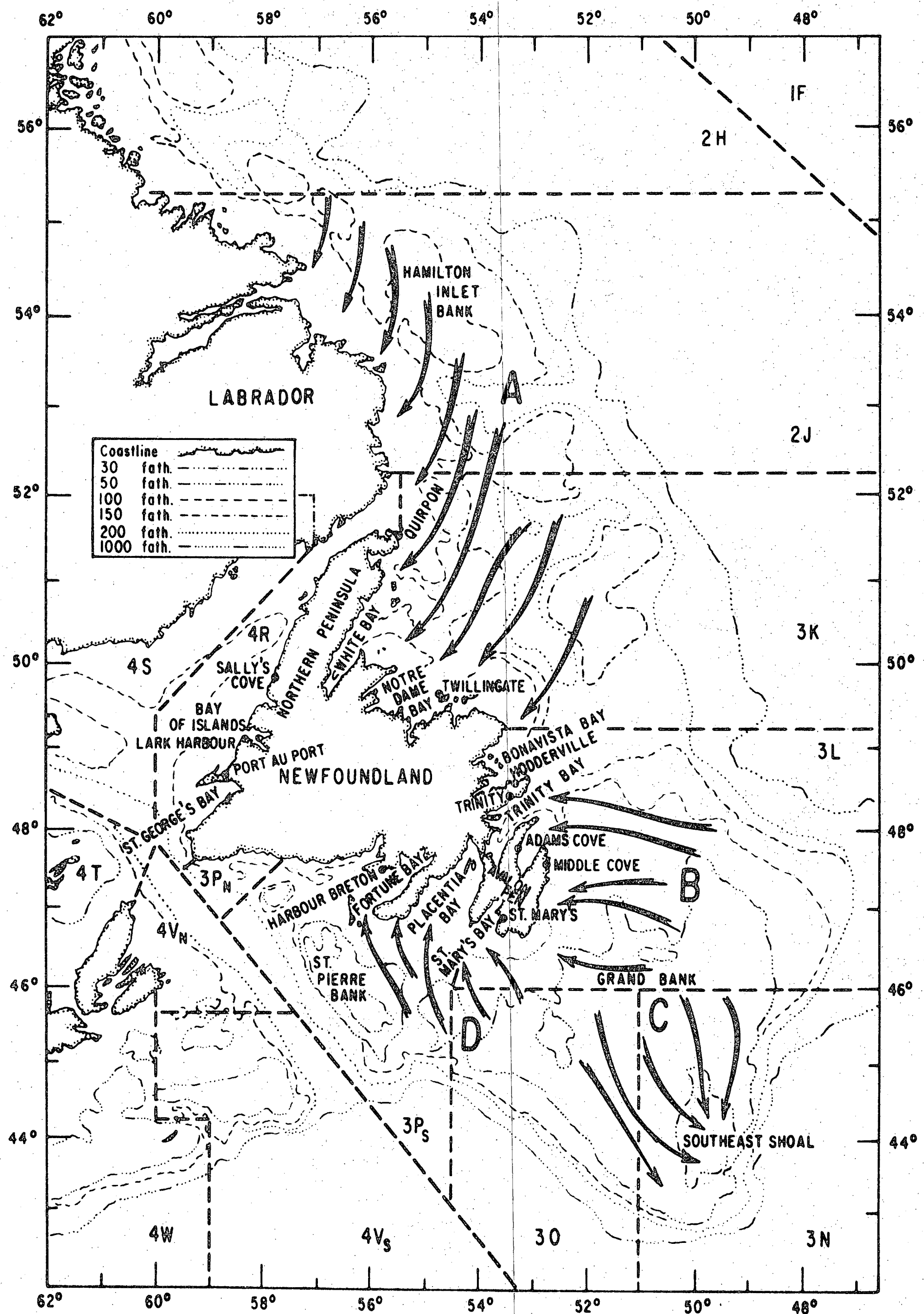

Fig. 1. Sampling sites from Carscadden and Misra (1980) and tentative stock divisions from Campbell and Winters (1973) and Winters (1974). Stock E (not shown) is the Gulf of St. Lawrence stock. (This figure reproduced from Carscadden and Misra (1980).)

Table 8. Paired comparisons of 5 samples, marked 'X' (3-yr-old individuals).

Sample pair	Character (s) 3	4	5	8	9	10	All
A, B				X			
A, C	X	X			X		X
A, D			X	X			X
A, E				X			X
B, C	X	X			X	X	X
B, D			X		X		X
B, E			X	X			X
C, D	X			X	X		X
C, E	X			X	X		X
D, E							

Table 9. Paired comparisons of 5 samples, marked 'X' (4-yr-old individuals).

Samples pair	Character (s) 1	8	9	10	All
A, B					
A, C		X			X
A, D					
A, E					
B, C	X	X			X
B, D					
B, E					
C, D		X			X
C, E	X	X			X
D, E				X	X

Table 10. Paired comparisons of 5 samples, ages combined. Comparisons marked 'X' were significant by the Bonferroni procedrue. These marked 's' were significant when samples were analysed independently.

Sample pair	Character (s) 1	3	4	5	6	8	9	1,3,4 5,6,8,9	1,3,4,6
A, B			S						
A, C			X			X	X	X	
A, D				S				X	
A, E				X	X			X	X
B, C						X	X	X	
B, D									
B, E					X			X	X
C, D		X				X	X	X	
C, E	X	X				X	X	X	X
D, E	X								X

Table 5. Matrix of χ^2 values with df = 13 for B^2 (below the diagonal) and G^2 (above the diagonal) analyses, for four characters considered collectively (4-year-old individuals).

Group	A	B	C	D	E
A	-	5.03	48.50	9.83	13.93
B	5.15	-	56.30	19.20	25.56
C	52.01	59.98	-	38.98	68.79
D	10.24	19.95	41.96	-	42.65
E	14.35	26.36	73.21	44.40	-

Table 6. Matrix of normalized χ^2 values for B^2 (below the diagonal) and G^2 (above the diagonal) analyses, for six characters considered collectively (3-year-old individuals).

Group	A	B	C	D	E
A	-	1.46	3.32	2.46	2.58
B	1.49	-	3.60	2.63	2.89
C	3.47	3.77	-	4.22	5.00
D	2.57	2.78	4.48	-	0.77
E	2.70	3.05	5.29	0.78	-

Table 7. Matrix of normalized χ^2 values for B^2 (below the diagonal) and G^2 (above the diagonal) analyses, for four characters considered collectively (4-year-old individuals).

Group	A	B	C	D	E
A	-	0.39	3.74	0.76	1.07
B	0.40	-	4.33	1.48	1.97
C	4.00	4.61	-	3.00	5.29
D	0.79	1.53	3.23	-	3.28
E	1.10	2.03	5.63	3.42	-

Table 2. Matrix of B^2 (below the diagonal) and G^2 (above the diagonal) values, for six characters considered collectively (3-year-old individuals).

Group	A	B	C	D	E
A	-	0.5356	2.5204	1.0739	1.1964
B	0.5497	-	2.1959	0.7508	0.9032
C	2.6319	2.2942	-	2.8556	3.5198
D	1.1205	0.7913	3.0309	-	0.2914
E	1.2525	0.9530	3.7252	0.2964	-

Table 3. Matrix of B^2 (below the diagonal) and G^2 (above the diagonal) values, for four characters considered collectively (4-year-old individuals).

Group	A	B	C	D	E
A	-	0.0961	1.7742	0.2588	0.3488
B	0.0985	-	1.5231	0.3240	0.3983
C	1.8989	1.6227	-	1.3360	2.2685
D	0.2697	0.3368	1.4380	-	0.9725
E	0.3594	0.4107	2.4141	1.0122	-

Table 4. Matrix of χ^2 values with df = 21 for B^2 (below the diagonal) and G^2 (above the diagonal) analyses, for six characters considered collectively (3-year-old individuals).

Group	A	B	C	D	E
A	-	30.56	69.71	51.76	54.25
B	31.36	-	75.74	55.27	60.66
C	72.80	79.13	-	88.65	105.01
D	54.00	58.25	94.09	-	16.08
E	56.79	64.00	111.14	16.36	-

Table 1. Sample sizes (n) and ranges of characters based on "complete" specimens.

Sample	Age	n	Character Number 1	2	3	4	5	6	7	8	9	10	11
A Labrador-NE Nfld.	3	80	18-21	13-15	22-25	4-8	7-10	8-11	25-30	15-19	12-16	41-44	22-25
	4	66	18-22	13-15	22-26	4-8	7-9	8-10	26-30	14-18	12-16	40-46	22-25
	Combined	146	18-22	13-15	22-26	4-8	7-10	8-11	25-30	14-19	12-16	40-46	22-25
B N. Grand Bank-Avalon	3	193	17-21	13-18	20-25	3-9	7-10	8-11	25-31	15-20	12-16	40-46	20-26
	4	205	17-21	13-15	22-26	4-8	7-10	8-10	25-30	14-20	11-16	39-47	20-25
	Combined	398	17-21	13-18	20-26	3-9	7-10	8-11	25-31	14-20	11-16	39-47	20-26
C S. Grand Bank (Southeast Shoal)	3	42	18-21	13-14	22-25	4-8	7-10	8-10	25-30	16-19	13-16	41-44	22-25
	4	45	18-21	13-15	22-25	4-8	7-10	8-10	24-29	14-20	11-16	40-44	22-28
	Combined	87	18-21	13-15	22-25	4-8	7-10	8-10	24-30	14-20	11-16	40-44	22-28
D St. Pierre-Green Bank	3	119	17-21	13-15	22-26	4-8	7-10	8-11	25-30	12-20	12-16	39-45	22-25
	4	83	18-21	13-15	22-26	4-9	7-9	8-10	25-30	14-19	11-15	40-45	22-25
	Combined	202	17-21	13-15	22-26	4-9	7-10	8-11	25-30	12-20	11-16	39-45	22-25
E W. Nfld.	3	103	18-21	13-15	22-26	4-8	7-9	8-11	26-31	14-19	11-16	40-45	22-25
	4	83	18-21	12-15	22-26	5-8	7-10	8-11	25-30	15-19	12-15	40-45	19-26
	Combined	186	18-21	12-15	22-26	4-8	7-10	8-11	25-31	14-19	11-16	40-45	19-26

Bowering, W.R. and R.K. Misra. 1982. Comparisons of witch flounder (Glyptocephalus cynoglossus) stocks of the Newfoundland-Labrador area, based upon a new multivariate analysis method for meristic characters. Can. J. Fish. Aquat. Sci. 39: 564-570.

Campbell, J.S. and G.H. Winters. 1973. Some biological characteristics of capelin, Mallotus villosus, in the Newfoundland area. ICNAF Res. Doc. 73/90, Ser. No. 3048, 8 p.

Carscadden, J. E. and R.K. Misra. 1980. Multivariate analysis of meristic characters of capelin (Mallotus villosus) in the northwest Atlantic. Can. J. Fish. Aquat. Sci. 37: 725-729.

Fairbairn, D. J. 1981. Which witch is which? A study of the stock structure of witch flounder (Glyptocephalus hippoglossoides) in the Newfoundland region. Can. J. Fish. Aquat. Sci 38: 782-794.

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

Kurczynski, T.W. 1970. Generalized distance and discrete variables. Biometrics 26: 525-534.

Morrison, D.F. 1976. Multivariate analysis. McGraw-Hill Book Co., New York, N.Y. 415 p.

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

Snedecor, G.W. and W.G. Cochran. 1967. Statistical methods. The Iowa State Univ. Press, Ames, IA. 593 p.

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.