SCRS/1990/056 Col.Vol.Sci.Pap. ICCAT, 34 : 160-165 (1991)

APPLICATION OF A MAXIMUM LIKELIHOOD METHOD TO ESTIMATE THE AGE COMPOSITION OF YOUNG ALBACORE CATCHES

J. Santiago

Instituto de Investigación y Tecnología para la Oceanografía, Pesca y Alimentación, Sukarrieta, Isla Txatxarramendi, Vizcaya, España

SUMMARY

Length frequency distributions of albacore catches landed in Bermeo (Basque Country, Spain) by the baitboat fishery in 1987, 1988 and 1989 have been analyzed in order to get age composition estimates using a maximum likelihood approach.

RESUME

On analyse la distribution des fréquences de taille du germon débarqué à Berméo (Pays basque espagnol) par la flottille de canneurs en 1987, 1988 et 1989, afin d'obtenir des estimations de la structure démographique par la méthode de la vraissemblance maximale.

RESUMEN

Se analizan las distribuciones de frecuencias de tallas de atún blanco desembarcadas en Bermeo (País Vasco, España) por la flota de cebo vivo en 1987, 1988 y 1989, aplicando el método de máxima verosimilitud, con el fin de obtener estimaciones de la composición en edades.

INTRODUCTION

Application of the so called analytical methods for stock assessment requires the estimation of the age composition of the catches. This is usually achieved by converting length frequency distributions of the catches into age compositions using age-length keys (ALK's), which are built up after aging a certain number of individuals. Of course, ALK's are only used when the age determination of the fish is possible.

Alternatively, another approach to estimate the age composition from size data is the analysis of the length frequency distributions. The basic idea behind this method is that the modes appearing in the length frequency distributions may represent modes of different cohorts. This was first pointed out by Petersen (1891), who separated the age components of the mixture by visual inspection. Since then, different procedures have been developed to solve the problem of mixing component distributions:

- Graphical methods (v.g., Cassie, 1954; Bhattacharya, 1967). Although they have been widely used, they are subjective and unreliable unless the components are well separated (Macdonald, 1987).

- Computer-assisted statistical methods: they include the method of moments, maximum likelihood, Bayesian methods and various minimum distance approaches (Titterington et al., 1985).

Recently new procedures have been developed which combine both ALK and length frequency analysis (Kimura & Chikuni, 1987; Martin & Cook, 1990).

The 1989 ICCAT Albacore Workshop (Anon., 1989) recommended to carry out investigation on the application of stochastic methods for estimating catch-at-age from catch-at-length data. This document shows the application of a maximum likelihood method suggested by Schnute and Fournier (1980), to length data of young albacore captured in the Bay of Biscay.

MATERIALS AND METHODS

- LENGTH DATA

Length sampling was carried out in the port of Bermeo (Basque Country). Landings of albacore in this port in 1987, 1988 and 1989 were 7192, 6823 and 4438 tones respectively. Sampling was aimed to obtain length information for each sampling cell, defined as the combination of:

- Time period: 1 month.
- Area: ICCAT 5°x10° rectangles
- Size category: according to the commercial categories, small (< 4 Kg) and big (> 4 Kg,.
- Fleet: trolling and baitboat.

This paper deals only with data coming from the baitboat fishery (approx. 80% of the albacore landings). The number of length samples and individuals measured, grouped by quarters of the year, are shown in Table 1.

- LENGTH FREQUENCY ANALYSIS

The analysis was performed following the procedure suggested by Macdonald and Pitcher (1979) and Schnute and Fournier (1980). The main assumption of this method is the normality of the distribution of length at each age. Therefore, if the number of age components of a distribution mixture is M, the probability density function can be given by

$$g(x) = \sum_{j=1}^{M} p_j \ f(x{:}m_j,s_j)$$

where p_j = proportion of fish of age j

$f(x{:}m_j,s_j^2)$ = probability density function of a normal distribution with mean m_j and standard deviation s_j^2.

Suposse that fish are sampled randomly[1] from the above population; the probability of a fish lying in length interval i is defined as

$$p_i = \sum_{j=1}^{M} p_j \ q_{ij}$$

where q_{ij} = probability that a fish of age j lies in length interval i; it is defined as the area under the normal distribution $N(m_j,s_j^2)$ in the ith interval

If the number of fish sampled is n, the expected number of fish in length interval i is given by

$$f_i = n \ p_i$$

Let $\hat{f}_i$ be the observed length frequencies, and N the number of length class intervals; the problem is to find values for the parameters (M,m_j,p_j,s_j) which minimize the difference between the observed and expected frequencies. Maximum likelihood estimation is obtained minimizing the criterion of closeness A defined as

$$A = \sum_{i=1}^{N} \hat{f}_i \ \log(\hat{f}_i/f_i)$$

1 Note that the model has been developed for random sampling. The data analyzed in this paper has been obtained under a stratified random sampling squeme. The possible implications of the violation of this assumption have not been evaluated.

However, when there is too much overlap between the age components of the distribution, it is difficult to obtain biologically meaningful estimates. It is necessary to ad 'biological' constrains to the parameters of the model to reduce the number of possible statistical solutions. In this paper, the following constrains have been considered:

- The number of age groups, M, is known. The value of this parameter was chosen according to the possible number of ages in the range of the length distribution (it was assumed that albacore follows the von Bertalanffy growth equation given by Bard (1981)).
- Standard deviations increase proportionally to mean length.

The computer program was written by I. Martín in VAX FORTRAN. Later it was adapted to Microsoft FORTRAN to be run in IBM compatible computers[2]. Between the numerous algorithms available for performing a computer search to find a function minimum, the following were chosen:

- Simplex method (Nelder & Mead, 1965).
- Powell's method (Powell, 1964): Harwell Library subroutine VA04A.

RESULTS

Figures 1a-b, 2a-b and 3a-b show the frequency histograms of the length distributions analyzed -those from the 3rd and 4th quarter of 1987, 1988 and 1989, together with the age-groups determined by the application of the length frequency analysis. The estimates of the parameters of the model (mean lengths, standard deviations and proportions at age) and the value of the closeness criterium, A, are shown in Table 2; the same table also presents the mean length at age estimated by Bard (1981) for comparison.

As it has been stated in the previous section, the standard deviations were forced to increase as a function of the mean lengths at age, in order to reduce the number of parameters to be estimated and, therefore, the number of possible solutions. However, this constrain did not produce reliable estimates in the following cases:

- Fourth quarter of 1987: the analysis only detected two possible age groups (m_1=57 cm, m_2=74 cm; s_1=0.8, s_2=6.9) due to the insufficient number of albacore measured (211 individuals). Because of the few data available, no attempt was made to improve the estimates.

- Fourth quarter of 1988: the length frequency analysis gave

2 The programs are not available because they have to be updated.

unreliable estimates of three possible groups (m_1=53 cm, m_2=62 cm, m_3=68 cm; s_1=1.1, s_2=4.9, s_3=8.4). This might be due to the positive skewness (the tail of the distribution is extended to the right) of the first age group, forcing a relatively high variance of the second group. And the constrain of mean length dependence of the group variance made m_3 and s_3 unrealistic.
Another trial was made unrestricting the variances; Figure 2b and Table 2 show the estimates obtained. Again, the variance of the second group seems to be very high (s_2=5.4) because of the skewness of the distribution of the first group.

DISCUSSION

The comparison between albacore mean length at age estimated applying Schnute and Fournier (1980) length frequency analysis (LFA) and the values obtained by Bard (1981) -given in Table 2, indicates an agreement for ages 2, 3 and 4. However, there are significative differences between LFA and Bard's estimates for the first group, whose distribution is theoretically the least overlapped of the component mixture. Two possible reasons might explain this disagreement:

- growth of albacore doesn't follow von Bertalanffy model, or
- only the biggest one year olds are fully recruited to the baitboat fishery.

The second explanation seems to be more likely. If this is the case, mean length and standard deviation of the first group are probably biased estimates of the true parameters of the population; but unbiased estimates of the parameters of the first group in the capture.

In any case, the fact that the modes of the length frequency distributions observed are far from the von Bertalanffy growth pattern, doesn't allow the utilization of a powerful constrain of the LFA model of Schnute and Fournier (1980): mean lengths lying on a von Bertalanffy curve.

The inspection of length frequency distributions of Figures 1a-b, 2a-b and 3-b illustrates that there is relatively few overlap between the components of the mixture. For this reason, the application of LFA is a useful approach, even when only few constrains are introduced in the analysis, as in the case presented here with young albacore.

However, further investigation has to be carried out in order to test the validity of the different assumptions made:

- Normality of the distributions of length at age.
- Relationship between mean length or age and standard deviations.
- Implications of the stratified random sampling scheme in the applicability of this method.

Provided that this assumptions are reliable, the Schnute & Fournier's (1980) LFA method (or others similar) might give reliable information on the age composition of albacore caught by the troll and baitboat fleets, whose catches are mostly dominated by young fish.

The application of LFA to estimate the age composition of albacore of the longline fishery (modes of length frequency distributions of its catches are more overlapped) might require the introduction of additional constrains to the model, otherwise the number of possible mathematical solutions to solve the mixture problem would be very high. The most useful constrain to be introduced is probably to force the mean lengths to follow a growth model.

REFERENCES

- ANON., 1989. Report of the 1989 ICCAT Albacore Workshop. ICCAT Collect. Vol. Sci. Pap., 31: 73-96.

- BARD,F.X. 1981.Le thon germon (Thunnus alalunga) de l'Océan Atlantique. These Doc. Sc. Paris. 336 p.

- BHATTACHARYA, C.G., 1967. A simple method of resolution of a distribution into Gaussian components. Biometrics, 23: 115-135.

- CASSIE,R.M. 1954.Some used of probability paper in the analysis of size frequency distributions. Australian Journal of Marine and Freshwater Research, 5: 513-522.

- KIMURA,D., and CHIKUNI,S. 1987. Mixtures of empirical distributions: an iterative application of the age-length key. Biometrics, 43: 23-35.

- MACDONALD, P.D.M., 1987. Analysis of length-frequency distributions. The Age and Growth of Fish, ed. by R.C. Summerfelt and Gordon E. Hall. The Iowa State University Press. 544 p.

- MACDONALD, P.D.M. & T.J. PITCHER, 1979. Age-groups from size-frequency data: a versatile and efficient method of analyzing distribution mixtures. J.Fish. Res. Board Can., 36: 987-1001.

- MARTIN,I. & COOK, 1990. Combined analysis of length and age at length data. J.Cons. int. Explo. Mer, 46: 178-186.

- NELDER, J.A. & R. MEAD, 1965. A simplex method for function minimization. Comput. J., 7: 308-313.

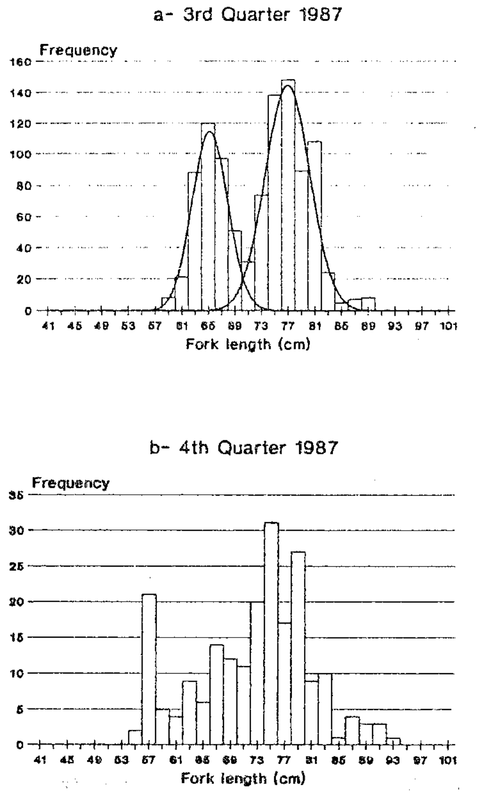
- PETERSEN, C.G. 1891. Eine Methode zur Bestimmung des Alters and Wuchses der Fische. Mitteilungen.Deutscher Seefischerei-Verein, 11: 226-235.
- POWELL, M.J.D. 1964. An efficient method for finding the minimum of a function of several variables without calculating derivatives. Computer J., 7: 155-162.
- SCHNUTE, J. & D. FOURNIER, 1980. A new approach to length-frequency analysis:growth structure. Can. J. Fish. Aquat. Sci., 37: 1337-1351.
- TITTERINGTON, D.M., A.F.M. SMITH & U.E. MAKOV, 1985. Statistical Analysis of Finite Mixture Distributions. John Wiley & Sons Ltd. 243 p.

YEAR	1987		1988		1989	
QUARTER	3rd	4th	3rd	4th	3rd	4th
Number of samples	11	2	31	8	11	27
Fish measured	1017	211	4042	819	970	2330

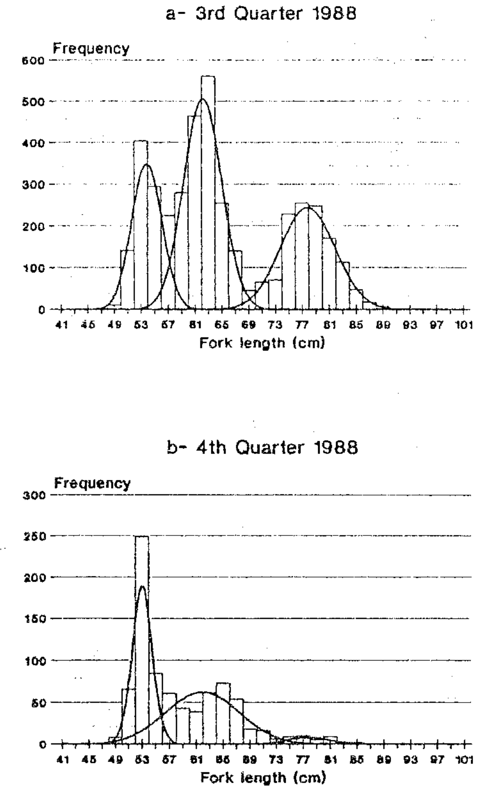
TABLE 1. Number of samples and number of albacore measured in the port of Bermeo (baitboat landings) in 1987, 1988 and 1989, by quarters.

YEAR		1987		1988		1989		
QUARTER		3rd	4th	3rd	4th	3rd	4th	BARD (1981)
AGE	PARAMETERS							
I	Meanlength		-	53,87	53,00	56,11	55,45	45,5
	Stand. deviat.		-	1,91	1,07	2,89	2,41	
	Proportion		-	0,24	0,43	0,39	0,34	
II	Meanlength	65,30	-	62,20	61,99	64,19	66,10	61,8
	Stand. deviat.	2,42	-	2,60	5,37	3,34	3,99	
	Proportion	0,38	-	0,45	0,53	0,38	0,02	
III	Meanlength	76,96	-	77,84	78,78	75,26	77,62	74,6
	Stand. deviat.	3,15	-	3,90	3,04	3,96	5,69	
	Proportion	0,60	-	0,31	0,04	0,23	0,64	
IV+	Meanlength	-	-	88,62	-	-	84,85	84,8
	Stand. deviat.	-	-	4,80	-	-	6,77	
	Proportion	-	-	0,00	-	-	0,00	
	A	51,30		84,34	46,07	34,75	108,09	

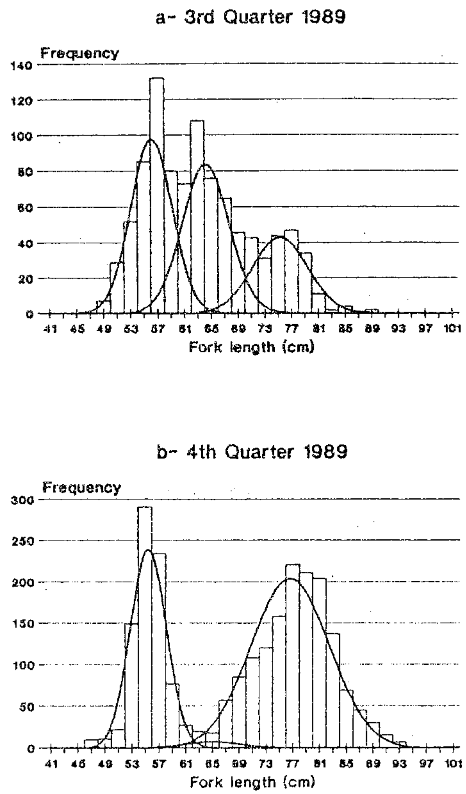
TABLE 2. Results of the Length Frequency Analysis: Mean length at age, standard deviations and proportion of ages for albacore landed in Bermeo (baitboat fishery) from 1987 to 1989 by quarter. A is the criterium of closeness of the fits. Bard's (1981) mean length at age estimates are also shown.



Figures 1a-b. Length frequency histograms and age-components estimated by Length Frequency Analysis for albacore landed in Bermeo by the baitboat fishery in 1987, by quarter (age-groups estimates only available for the 3rd quarter).



Figures 2a-b. Length frequency histograms and age-components estimated by Length Frequency Analysis for albacore landed in Bermeo by the baitboat fishery in 1988, by quarter.



Figures 3a-b. Length frequency histograms and age-components estimated by Length Frequency Analysis for albacore landed in Bermeo by the baitboat fishery in 1989, by quarter.