J. Fish. Soc. Taiwan, **17**(3): 153-159 (1990)

Estimating Catch Matrix of Indian Albacore by Iterated Age-Length Key Algorithm

Ying-Chou Lee[1], Chien-Chung Hsu[1]
and Hsi-Chiang Liu[1]

(*Received April 13, 1990; Accepted May 11, 1990*)

In assessing Indian albacore stock by age-structured analytical model, the catch-at-age matrix has to be developed *in priori*. Also the iterated age-length key algorithm developed by Kimura and Chikuni is one of the techniques using in separating the catch into some age categories. In the present study, a sampled age-length key, catch-at-length data of Indian albacore measured from 1979 to 1987, and growth parameters were applied in the iterative simulation. Consequently, 7 distinguished age groups (ages 1 to 7) and a plus group (age 8+) are resulted from simulation; and the appropriate goodness-of-fit was concluded by high values of the coefficients of determination (R^2).

Key words: Albacore, Catch-at-age, Age-length key.

關鍵詞：長鰭鮪，年齡別漁獲量，年齡一體長檢索。

INTRODUCTION

Many age-structured models of fish stock assessment need applying catch-at-age matrix (or catch matrix) to achieve statistically these assessing matters. Traditionally, three methods were used to construct the catch matrix if catch-at-length data and/or growth parameters were available. The first method applied is the most popular age-length key (ALK, Fridriksson, 1934) and remains unchanged up-to-date; the second is to construct the length distribution as a mixture of length-at-age distribution (Cassie, 1954) and to estimate the probability of length at each age by using likelihood functions (Hasselblad, 1966), and; the third method is used regression to estimate age distribution from catch-at-length data, in which the length distribution is the response variable and the age distribution is the vector of unknown regression coefficients (Clark, 1981; Bartoo and Parker, 1983).

Recently, Kimura and Chikuni (1987) incorporated the ALK and distribution mixture method to develop a new method of involving iterative simulation of an ALK to fit the catch-at-length data, and the probability of age given a length was resulted from the best goodness-of-fit obtained during simulation. Basically, an ALK required that the age-length data and random length frequency be sampled from the same population (Westrheim and Ricker, 1978). ALK was capable of providing biased estimates of age distribution, when a length distribution was applied to ALK constructed from length data sampled in different year. For a length distribution sampled in any year, Kimura and Chikuni (1987) pointed out that Fukuda and Chikuni (unpublished) provided an estimate of age distribution corresponding to this length distribution consistent with an ALK, and showed

1. Institute of Oceanography, National Taiwan University, Taipei, Taiwan 10764, Republic of China.

that this method is an application of the EM (expectation and maximization) algorithm (Dempster *et al.*, 1977) to separate the mixture of empirical distributions, thus without any mathematical proof, the estimate of age distribution is unique and maximum likelihood estimation is satisfied.

For the Indian albacore (*Thunnus alalunga*), Huang *et al.* (1986), Lee and Liu (1988a, b), Liu and Lee (1990), and Shiohama (1985; 1986; 1988), have in primary assessed the stock status by the generalized production model analysis. In order to assess the stock by age-specific models, e. g., yield per recruit model, virtual population analysis, and cohort analysis, etc., the catch-at-age data and length-at-age data are important and fundamental for these models. Lee and Liu (unpublished) have achieved the age and growth study of the species by vertebrae. As a result, estimate of catch matrix using empirical ALK and catch-at-length data by iterated age-length key (IALK) algorithm was manipulated in the present study for application of adopting such analytical models in the future.

MATERIALS AND METHODS

The IALK algorithm was described by Kimura and Chikuni (1987) as:

(1) Assume that the initial age distribution has equal probability for each age j, that is for an "a" age categories stock, the initial probability of age j is

$$\hat{P}_j = \frac{1}{a} \tag{1}$$

(2) Apply catch-at-length data and $\hat{P}_j$ information to calculate the corresponding age-length key, i. e., $P_r(j|i)$, where i $(=1, 2, \cdots, n)$ is the subscript of length categories of catch-at-length data;

(3) Apply the age-length key calculated in step (2) on the catch-at-length data to estimate a new $\hat{P}_j$;

(4) If the maximum absolute deviation value of $\hat{P}_j$ from the two consecutive iteration procedures is less than a prescribed small constant, stop the iteration and the result is obtained. Otherwise repeat steps from (2) to (4).

Mathematically, the IALK started from catch-at-length data; that is

$$\bar{f}_i = \frac{f_i}{f} \tag{2}$$

where f_i = the observed catch-at-length for length category i;

$\bar{f}_i$ = the estimated catch-at-length frequency in percentile for length category i;

f = $\sum_i f_i$ = total catch in number, where $i = 1, 2, \cdots, n$.

Toward the goal, a P_j which is the probability of unknown age j distribution wanted to be estimated from the given catch-at-length data. Thus

$$\hat{P}_j = \sum_i \bar{f}_i \hat{P}_r(j|i) \tag{3}$$

where $\hat{P}_j$ = the estimate of P_j;

$\hat{P}_r(j|i) = \frac{\hat{P}_j q_{ij}}{\hat{l}_i}$, the estimate of IALK from age-length key; and

q_{ij} = $P_r(i|j)$, the observed distribution of length-at-age estimated from the available age-length data or the observed age-length key;

$\hat{l}_i$ = $\sum_j \hat{P}_j q_{ij}$, the IALK estimate of $l_i = \sum_j P_j q_{ij}$, the length distribution estimated from an age distribution and observed length-at-age distribution.

Then, the estimate of P_j could be reconstructed for all j as:

$$\hat{P}_j = \hat{P}_j \sum_i \left(\frac{\bar{f}_i q_{ij}}{\sum_j \hat{P}_j q_{ij}} \right) \tag{4}$$

If the values of P_j's were greater than zero, to wit $P_j > 0$, equation (4) can be rewritten as

$$\sum_i \left(\frac{\bar{f}_i q_{ij}}{\sum_j \hat{P}_j q_{ij}} \right) = 1, \text{ for all } j \tag{5}$$

Also, $\sum_j \hat{P}_j = 1$ is satisfactory under the constraint $\hat{P}_j > 0$, for all j. Kimura and Chikuni (1987) had elucidated the solution be unique, and the IALK algorithm converges to a unique fixed point.

To investigate the statistic of goodness-of-fit that describes how well the IALK result could explain the observed data, the coefficient of determination (R^2) is given by

$$R^2 = 1 - \left[\frac{SS(\text{residual})}{SS(\text{total adjusted})} \right]$$

$$= 1 - \left[\frac{\sum_i (\hat{l}_i - \bar{f}_i)^2}{\sum_i \left(\bar{f}_i - \frac{1}{n} \right)^2} \right] \tag{6}$$

On the other hand, the method for determining the age of Indian albacore (*Thunnus alalunga*) was adopted as used in Lee and Liu (unpublished). Owing to the small sample sizes ($N = 391$) for constructing the ALK, the samples was grouped with 4 cm class in calculating the empirical length-at-age data (Table 1).

Table 1. The sampled length-at-age frequencies of albacore, *Thunnus alalunga*, in the Indian Ocean

Length class (cm)	I	II	III	IV	V	VI	VII	VIII	IX	X	Sum
44- 47	9										9
48- 51	4	1									5
52- 55		1									1
56- 59	1	6									7
60- 63		4									4
64- 67		1	2	1							4
68- 71		1	12	3							16
72- 75			17	14	1						32
76- 79			15	22	1						38
80- 83			3	27	19						49
84- 87			1	23	41						65
88- 91			1	3	33	10	1				48
92- 95					7	21	4				32
96- 99					1	8	6				15
100-103						2	5	7			14
104-107						2	7	9			18
108-111							6	5	2	2	15
112-115							4	5	2	2	13
116-119								1	2	2	5
120-123										1	1
Sum	14	14	51	93	103	43	33	27	6	7	391

Table 2. The catch-at-length matrix (in number) of Indian albacore caught by Taiwanese tuna longline fishery from 1979 to 1987

Length class (cm)	1979	1980	1981	1982	1983	1984	1985	1986	1987
44- 47	333	1428	1190	946	747	414	293	28	178
48- 51	367	194	572	1211	976	1254	322	32	130
52- 57	428	397	878	1555	1223	1766	750	41	276
56- 59	245	431	2095	2103	1450	2076	1104	88	737
60- 63	256	1350	2295	4378	1753	2139	1109	465	1374
64- 67	228	942	2248	3741	2502	2463	1290	836	1727
68- 71	324	1089	3904	4570	4185	6662	2269	1107	2456
72- 75	614	1794	5481	6661	4571	9156	3265	3951	4097
76- 79	591	1554	6626	6889	4289	8234	4711	7836	4670
80- 83	1353	2290	10422	11582	8101	10501	4801	9940	6588
84- 87	1321	2529	12597	13945	9676	10110	4760	7429	7145
88- 91	702	2261	11001	13166	8657	8565	2954	4762	3924
92- 95	945	1998	10079	11163	7316	6540	3301	4492	3266
96- 99	715	1336	9067	10149	7018	5653	3122	3532	1803
100-103	633	1667	10623	13152	8353	5373	3596	2317	1643
104-107	515	1018	5336	8194	6081	4402	1518	1267	1256
108-111	583	820	3747	4115	3760	2191	370	367	1150
112-115	262	609	1848	2084	2502	1490	241	322	1616
116-119	45	453	376	868	1911	1580	343	189	1293
120-123	142	514	821	2016	2118	3747	880	166	1250
Sum	10602	24674	101206	122488	87189	94316	40999	49167	46579

The empirical length frequencies data were collected from logbooks of Taiwanese tuna longliners which operated in the Indian Ocean from 1979 to 1987 (Table 2).

RESULTS AND DISCUSSION

The estimated age distribution and catch matrix of Indian albacore by IALK algorithm are shown in Tables 3 and 4, respectively. The values of R^2 (Table 3) are approximately similar and fairly high, these somewhat indicate a reasonable goodness-of-fit of catch-at-age being simulated between the observed and theoretical frequency at each age category. Figure 1 showed these IALK simulated results of catch-at-age distributions.

In this study, the applicable age-length key and length frequencies data were from different sources. The former was subsampled from the landings of Taiwanese tuna longliners and gillnetters which operated in Indian Ocean (unpublished data, Y. C. Lee and H. C. Liu, Institute of Oceanography, National Taiwan University, pers. communication); the latter was collected from logbooks of Taiwanese tuna longliners. A method to compromise the two fisheries was not developed, however, Westrheim and Ricker (1978) had claimed that an ALK computed from one sample can safely be applied to the lengths from any other taken by different fisheries from the same population. As the result, the ALK applied in this study is acceptable.

The critical assumptions of IALK method are the age-length data being adequately sampled, and the associated q_{ij} estimates being applicable to the population sampled for length frequencies (Kimura and Chikuni, 1987). In this study, the sample size for constructing length-at-age (observed age-length key)

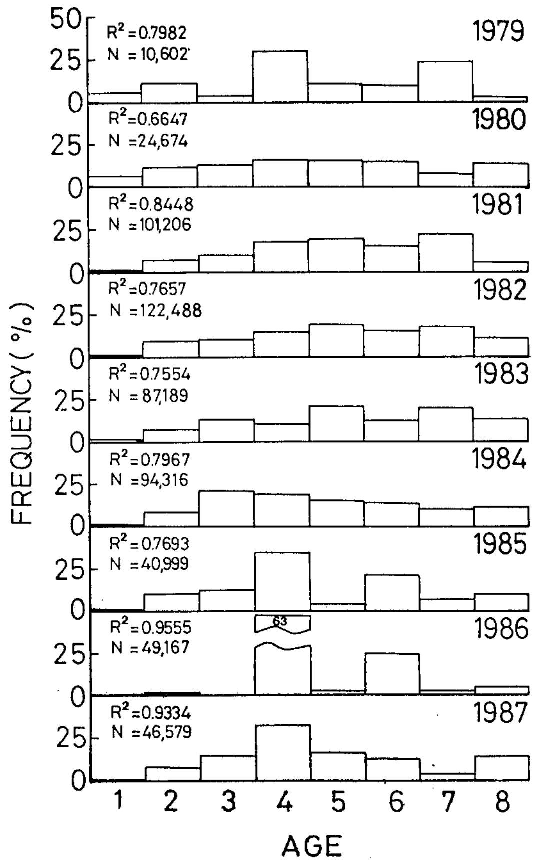

Fig. 1. The simulated age compositions of Indian albacore by IALK algorithm from 1979 to 1987.

was small ($N=391$) for covering the possible body size range of albacore (unpublished data, Y.C. Lee and H.C. Liu, Institute of Oceanography, National Taiwan University, pers. communication), it may not be large enough to describe the age-length relation completely. However, the IALK is an appropriate way so far as based on the data base structure.

Table 3. Age distributions of IALK estimates of albacore, *Thunnus alalunga*, in the Indian Ocean

Year	Age								R^2
	1	2	3	4	5	6	7	8+	
1979	.0552	.1233	.0302	.2979	.1138	.1014	.2448	.0333	.7982
1980	.0647	.1181	.1377	.1584	.1552	.1487	.0800	.1372	.6647
1981	.0151	.0703	.1002	.1784	.1943	.1554	.2266	.0597	.8448
1982	.0112	.0958	.1037	.1472	.1920	.1559	.1817	.1126	.7657
1983	.0137	.0770	.1300	.1043	.2104	.1201	.2045	.1399	.7554
1984	.0083	.0903	.2078	.1928	.1488	.1331	.1000	.1189	.7967
1985	.0098	.0996	.1285	.3498	.0420	.2095	.0566	.1042	.7693
1986	.0007	.0162	.0056	.6333	.0226	.2498	.0246	.0472	.9555
1987	.0045	.0722	.1438	.3222	.1564	.1272	.0316	.1420	.9334

Table 4. The estimated catch matrix of Indian albacore (*Thunnus alalunga*) by iterated age-length key algorithm

Age	Year								
	1979	1980	1981	1982	1983	1984	1985	1986	1987
1	62840	74124	16766	12688	15191	10326	11413	800	5188
2	140366	135302	78057	108526	85379	112345	115992	18507	83246
3	34380	157756	111256	117476	144147	258530	149649	6389	165800
4	339133	181471	198085	166755	115650	239868	407371	723497	371492
5	129551	177805	215739	217506	233296	185126	48912	25819	180327
6	115435	170359	172547	176610	133170	165593	243980	285378	146660
7	278683	91652	251603	205838	226754	124413	65915	28104	36434
8+	37909	157184	66287	127558	155124	147927	121350	53922	163724

The IALK is applicable to where the traditional ALK would not (Kimura and Chikuni, 1987), and was recommended as the most powerful tool in separating catch by age categories (W. Clark, International Pacific Halibut Commission, pers. communication). In essential, the standard ALK method requires that the age-length data and random length frequency be sampled from the same population (Westrheim and Ricker 1978). On the other hand, the IALK requires only that the estimated length-at-age distributions be applicable to population sampled for length frequencies. Therefore, the IALK is much more possible to apply age-length data sampled from one population to a length frequency sampled from another (Kimura and Chikuni, 1987). To obtain an ALK of a study stock is generally time-consuming and money-costing if the conventional double sampling technique is applied, in particular, which were sampled from the landings of vessels operating in deep-sea, e. g., from longliners and gillnetters at base ports as well. Therefore, in order to assess the stock status by age-structured analytical models, the IALK simulation applied to the ALK studies of these deep-sea fisheries seems to be the most appropriate and, the results is certainly compatible with those from traditional ALK.

ACKNOWLEDGEMENTS

We would like to thank the staff of Taiwan Fisheries Bureau and the Tuna Research Center, National Taiwan University for their helps in the collections of length frequencies data. And the appreciation should be noted to the grants of National Science Council, NSC-79-0209-B-002A-03 for the theoretical part and NSC-79-0209-B-002A-05 for the application of this theory in assessing Indian albacore stock.

REFERENCES

Bartoo, N.W. and K.R, Parker (1983). Stochastic age-frequency estimation using the von Bertalanffy growth equation. *U.S. Fish. Bull.*, **81**: 91-96.

Cassie, R.M. (1954). Some uses of probability paper in the analysis of size frequency distribution. *Aust. J. Mar. Freshwat. Res.*, **5**: 513-522.

Clark, W.G. (1981). Restricted least squares estimates of age composition from length composition. *Can. J. Fish. Aquat. Sci.*, **38**: 297-307.

Dempster, A.P., N.M. Laird and D.B. Rubin (1977). Maximum likelihood estimation from incomplete data via the EM algorithm. *J. Roy. Stat. Soc. Ser. B*, **39**: 1-39.

Fridriksson, A. (1934). On the calculation of age distribution within a stock of cod by means of relatively few age determinations as a key to measurements on a large scale. *J. Cons. Int. Explor. Mer.*, **86**: 1-14.

Hasselblad, V. (1966). Estimation of parameters for a mixture of normal distributions. *Technometrics*, **8**: 431-444.

Huang, F.T., C.H. Wang and H.C. Liu (1986). Estimation of overall fishing intensity and stock assessment of Indian Ocean albacore (*Thunnus alalunga*), 1963-1984. *FAO IPTP (86)*, **34**: 1-23.

Kimura, D.K. and S. Chikuni (1987). Mixtures of empirical distributions: an iterative application of the age-length key. *Biometrics*, **43**: 23-35.

Lee, Y.C. and H.C. Liu (1988a). An investigation of selecting undetermined in fitting the generalized stock production model. *Act. Oceanogr. Taiwan.*, **20**: 41-52.

Lee, Y.C. and H.C. Liu (1988b). Estimation of effective fishing effort and overall fishing intensity, and stock assessment of Indian Ocean albacore (*Thunnus alalunga*), 1952-1986. *FAO IPTP/TWS/ (88)*, **62**: 109-118.

Liu, H.C. and Y.C. Lee (1990). Stock assessment of albacore resource in the Indian Ocean. In *Proc. 2nd Asian Fish. Forum* (R. Hirano and I. Hanyu, eds.). Tokyo, Japan, 17-22 April 1989, 861-864.

Shiohama, T. (1985). Stock assessment of Indian Ocean albacore by production model analysis, 1952-1982. *FAO IPTP/TWS (85)*, **23**: 111-118.

Shiohama, T. (1986). Stock assessment of Indian albacacore by produciton model analysis, 1952-1984. *FAO IPTP/TWS (86)*, **25**: 84-104.

Shiohama, T. (1988). Stock assessment of Indian Ocean albacore, 1988. *FAO IPTP/TWS (88)*, **24**: 94-98.

Westrheim, S.J. and W.E. Ricker (1978). Bias in using an age-length key to estimate age-frequency distributions. *J. Fish. Res. Bd. Can.*, **35**: 184-189.

以年齡─體長檢索反覆運算法估計印度洋長鰭鮪之年齡別漁獲量

李英周・許建宗・劉錫江

為便於以年齡別解析模式評估印度洋長鰭鮪系羣，必先估得年齡別漁獲量矩陣。而年齡一體長檢索反覆運算法為將漁獲量分解成年齡別漁獲量之較有效力的方法。本研究以取樣獲得之印度洋長鰭鮪族羣的年齡一體長檢索表與 1979～1987 年之體長別漁獲量應用於反覆模擬估算。結果顯示有七個年齡羣（１～７歲）及一個混合羣（８歲以上）的年齡別漁獲量，在高判定係數數值下被獲得。