Indian J. Fish., 44(3) : 225-230, Jul.-Sep., 1997

Robustness of length-class interval in length based analysis for the estimation of growth parameters - A simulation study for *Sardinella longiceps*

T. V. SATHIANANDAN

Central Marine Fisheries Research Institute, Cochin - 682 014, India

ABSTRACT

In length based fish stock assessment using von Bertalanffy growth model, the estimates of growth parameters L_m and K play a key role. A widely used estimation method is the one given by Pauly and David, using length-frequency data. It is not known whether the estimates of L_ and K are robust to the class interval used for grouping length samples. Here, an attempt is made to study this feature through simulation by using length samples simulated for 12 months with input parameters of *Sardinella longiceps.* Simulated data were used to estimate growth parameters by grouping and estimating for different class intervals ranging from 5 to 25 mm and the results then compared with the actual inputs. Results indicated that 5 mm class interval is optimum for this species.

Introduction

Estimation of growth parameters is an important step in the length-based fish stock assessment. The growth model prescribed by von Bertalanffy (1938) is the most widely applied model because of its simplicity and mathematical amenability. In the length based approach to fish stock assessment, a popular technique for the estimation of growth parameters, the asymptotic length L^ and curvature K in the growth model, is the ELEFAN method introduced by Pauly and David (1981). The input data for this method is the time series of length-frequency data. The validity of ELEFAN in providing reliable estimates of L_w and K have been examined by several workers. These studies mainly consider the variability in length at age or variability in recruitment. Jones (1987) used simulated samples for an investigation of length composition analysis. Hampton and Majkowski (1987) examined the accuracy of ELEFAN through a simulation study. Basson *et al.* (1988) studied the reliabitity and accuracy of the two methods of estimation of the growth parameters from length-frequency data through a Monte Carlo simulation study. Castro and Erzini (1988) made a comparison of length-frequency based packages for estimating growth and mortality parameters using simulated samples with varying recruitment patterns. Farebrother (1992) made a simulation study of three iterated weighted least square estimators of mortality rates from single-release tagging studies. The objective of the present study is to examine the robustness of growth parameter estimates with the change in the class interval used for grouping the length measurements. Monte Carlo simulation technique was adopted to generate samples of lenght-frequency data from simulated population of oil sardine *(Sardinella longiceps).* For this purpose, the growth parameters and biological aspects of *Sardinella longiceps* were used to simulate length samples and the simulated data were used to study the effect.

Materials and methods

Growth parameters of S. *longiceps* were taken as 221 mm, 0.75 and -0.01 respectively for L_M ,K and t_0 based on published work (Annigeri *et al.,* 1992) on the growth parameters of the species. For each of these parameters 5 % of their estimates was fixed as the standard deviation in the distribution of their values. Since it is known that the estimates of L_M and K have a kind of negative correlation, bivariate normal distribution was assumed for sampling of these parameters in the process of simulation with -0.75 as the value for the correlation coefficient. Univariate normal distribution was assumed for t_0 for sampling. The maximum life span of the species was taken as 4 years and the proportion of different age groups in the population were assumed to be 0.4, 0.3, 0.2, 0.05 and 0.05 respectively for 0, 1, 2, 3 and 4 year age groups. Though there is continuous recruitment throughout the year for S. *longiceps,* it is maximum during June-July. So only one recruitment was considered and the time of recruitment for sampling was considered to have a normal distribution with mean at 0.58 year with a standard deviation of 0.08 year. The natural and fishing mortality rates were taken as 0.8 and 0.6 respectively and were assumed to be constant throughout the year. The length at recruitment was taken as 60 mm. The minimum and maximum sizes for retention by gears are assumed to have normal distributions with means at 65 and 220 mm respectively and standard deviations 10 mm each.

Simulation set up: It is assumed that (i) the maximum life span of the species is m time periods (say years) so that the population consists of (m+1) age classes at any time point t and the proportions <x0, a: . . . . , a_m of these age classes are expected to remain constant and are such that Xcc.=l, (ii) in a time period there are s spawning seasons which contribute towards the population and spawning takes place at times distributed normally with means t_x, t_2, ..., t_a from the beginning of the period with respective standard deviations a_v a_2, ..., o_s and have expected proportions, Pj, p_2, ..., p_s, (iii) the size at recruitment is l_r, (iv) the probability that death is due to natural causes remains the same and is taken as M/Z where M is the natural mortality, F fishing mortality and Z = M + F is the total mortality, (v) the minimum size of a fish retained by gear is assumed to be a random variable distributed normally with mean 1_{mm} and standard deviation a_{min} and also the maximum size retained is assumed to be a random variable distributed normally with mean 1_{max} and standard deviation a_{max}, (vi) growth of a fish is assumed to follow the von Bertalanffy

growth equation with parameters L_M, and K distributed bivariate normally with means (L , K), standard deviations (CT_L, O_K) and coefficient of correlation p and (vii) the time of birth t_o is assumed to be distributed normal with mean t_m and standard deviation a_{n_0}.

***Algorithms used for computing parameters** : Length :* Once the age of the fish is known, say t', its length can be computed by using the von Bertalanffy growth equation for length as $:L_t = L_{\infty} [1- e * \wedge]$

Time of death : When tracing a fish in a population, the time of death for each fish is taken as a random variable which depends on mortality rates. Kleijnen (1974) showed that if f(k) is the cumulative distribution function of a random variable k, then a random sample can be generated from this distribution as k = f.(r) where f (.) is the inverse function of f(k) and r is a pseudorandom number between 0 and 1. If f(t) denotes the cumulative probability distribution function of time of death, then

$$f(t_d) = \int_0 Ze^{Zt}dt$$

$= i\_ e^{Z(}V- y$ where Z = F + M is the total mortality rate, t_0 the time of birth and $t_0<t_d<<=o$. By this theorem if f(r) = t, then

r = «t_d)

= 1 - $e^{-Z(}$V V and hence the time of death is

$t, = t_0 - ln(l-r)/Z$.

Recruitment: Once the time of death is known, the age of the fish and its length at the time of death can be computed. This length can then be compared with the recruitment size l_r, to see whether death is after recruitment or not.

Retention by gear: Assuming that the minimum and maximum size retained by gear are normally distributed with respective parameters, 1_{mm}, a_{mm} and 1_{max}, CT_{max} ky drawing samples from these distributions we can see whether a sampled fish will be retained in gears or escapes. If l_x and 1_2 are the values drawn from N(l_{min}, a_{mm}) and N(l_{max}, a_{max}) respectively and l_t is the size of the selected fish then if \ <\<l_2 then the selected fish is considered as retained by the gear.

Simulating from normal distribution: Generate two independent uniform random number *[i_1* and *[i_2* between 0 and 1. Then X = -2 ln(Uj) cos(27tu$_2$) and Y = -2 ln(u_x) sin (27tu$_2$) and they will be distributed as independent standard normal variates. Now to generate a normal variate Z distributed with mean |i and standard deviation *a* compute Z = u+Xa or Z = (i+Yo.

Simulating from bivariate normal distribution: To sample from a bivariate normal distribution with means (|ij, *[i_2),* standard deviations *(a_v* a_2) and coefficient of correlation p, first two independent standard normal variates (Z_{15} Z_2) are generated and x_x = \$i_1+a_1z_l$ and x_2=|i_2 + a_2(pz_1+(l-p)z_2) are computed. Then (x_p x_2) will be distributed as bivariate normal with the specified parameters.

Based on the above algorithms, a computer software for simulation of length samples was developed in FORTRAN. For generation of pseudorandom numbers between 0 and 1, which is an important part of most of the algorithms, the routine given by Wichmann and Hill (1982) was used.

Results and discussion

Simulation model used: *A* fish is chosen at random from the population at time point, say t units. It may then belong to any one of the (m+1) age groups as defined below. Age group 0 are those born between time points [t-1, t], age group 1 are those born between time points [t-2, t-1], age group 2 are those born between time points [t-3,t-2] and so on, age group m are those born between time points [t-(m+l), t-m]. The expected proportions from these age groups are a_Q, $\&_v$ ..., a_m where Za.= 1. The age group of a selected fish is determined by selecting a random number between 0 and 1. If r is the selected random number, then the fish is from i^{th} age group if

$$\sum_{j=0}^{i-l} a < r < \sum_{j=0}^{i} a$$

After being known that the selected fish belongs to the i^{th} age group the spawning season in which it was born can be determined from the distributions of spawning seasons. If the j^{th} season is distributed with mean t. and standard deviation o. and is expected to contribute to the population a proportion p. where Xp =1, then the spawning season can be determined by selecting a random number and through a similar procedure as in the case of age group. Once known that the selected fish belongs to the j^{th} season, the time of birth of the selected fish can be computed as (t-i+s.-l) where s. is a random normal variate generated with mean t. and standard deviation a.. The age of the fish can then be computed as (i-s.+l). The important details about this fish are its age group, seasonal cohort to which it belongs, its time of birth and age. Once time of birth is determined, the life history of the selected fish is then traced from time of birth. Life history means, its recruitment, time of death, cause of death etc. If t^1 is the age of the fish at the time of sampling then its length can be computed by using the von Bertalanffy growth equation $L_{t'} = L_M [l-e^{K(t'-t_0)}]$ where L_M and K are random variates generated from a bivariate normal distribution with means (L , K), standard deviation (CT_L, a_K) and coefficient of correlation p and t_0 is a random variate generated from a normal distribution with mean t_m and standard deviation a_{tm}. If death occurs before recruitment the sample is ignored. If the selected fish is found recruited then its cause of death is examined by choosing a uniformly distributed random number between 0 and 1, say x, and if x < M/Z death is due to natural causes otherwise it is considered to be due to fishing. The chance of retention by gear is then tested and if found caught it is included in the sample of the corresponding month. Simulation is then continued by choosing fresh samples at each stage till the required number of samples are obtained.

Simulation results : Using the computer software developed in FORTRAN based on the above simulation model and algorithms, a total of 2,81,862 samples were generated of which 1,76,309 fishes were found recruited to the stock and from this, samples of 500 fish each for twelve months were drawn. Details regarding their age, time of birth, parameters used for generating the samples etc. were then noted. The simulated length samples were then grouped using different class intervals of 5, 10, 15, 20 and 25 mm respectively. Using these data sets, growth parameters were estimated separately for each data set with the help of ELEFAN program by searching in wider ranges, 150-325 mm for L_M and 0.25 - 2.0 range for K. The values of L_M and K which gave comparatively better fit (in terms of R_n value) were selected as their estimates and are tabulated below for different class intervals.

Class width	L_	K	Starting Sample	Starting Length	R_n
5	228.00	0.70	11	135.0	0.325
10	244.00	0.63	2	70.0	0.728
15	263.00	0.52	4	157.5	0.719
20	283.10	0.44	2	150.0	0.810
25	289.50	0.50	3	75.0	0.938

From the table it can be seen that the estimated values of growth parameters are more closer to the actual inputs used for simulation when the class interval size is 5 mm though the fitted curve shows better fit with data in the case of higher class intervals as indicated by the R_n values. The estimates of h_x and K in terms of actual parameters used for simulation are (L_m +0.64a_{nr}) and (K -1.33a„) when the class interval is 5 mm. Corresponding values for 10 mm class interval are (L +2.1a) and (K - 3.2a_K). For all other class intervals the estimates fall far beyond 3a limits. The results of this study thus indicate that in the case of *Sardinella longiceps* the optimum class interval size for grouping length measurements for growth parameter estimation using ELEFAN is 5 mm. Also, when estimating growth parameters the tendency to use higher class intervals for getting a better fit (or higher R_n value) should be avoided.

Acknowledgment

The author is thankful to Sri. K. N. Kurup and Sri. M. Srinath, Senior Scientists, Central Marine Fisheries Research Institute, for their valuable suggestions and modifications in the preparation of this document.

References

Annigeri, G. G. *et. al.* 1992. Stock assessment of oil sardine *Sardinella longiceps* Val., off west coast of India. *Ind. J. fish.*, 39(3,4) : 125-135.

Basson, M. *et al.* 1988. The accuracy and reliability of two new methods for estimating growth parameters from length-frequency data. *J. Cons. Int. Explor. Mer.*, 44 : 277-285.

Bertalanffy, L. von 1938. A quantitative theory of organic growth (Inquiries in growth laws II). *Hum. Biol.*, 10 : 181-213.

Castro, M. and K. Erzini 1988. Comparison of two length-frequency based packages for estimating growth and mortality parameters using simulated samples with varying recruitment patterns. *Fish. Bull.*, **86(4)** : 645-654.

Farebrother, R. W. 1992. A simulation study of three iterated weighted least-squares estimators of mortality rates from single-release tagging studies. *ICES J. Mar. Sci.*, 49 : 185-190.

Hampton, J. and J. Majkowski 1987. A simulation model for simulating length-frequency data. *ICLAM Conf. Proc.*, (13) : 193-202.

Hampton, J. and J. Majkowski 1987. An examination of the reliability of the ELEFAN programs for the length based stock assessment. *ICLAM Conf. Proc*, (13) : 203-216.

Jones, R. 1987. An investigation of length composition analysis using simulated length compositions. *ICLAM Conf. Proc*, (13) : 217-238.

Kleijnen, J. P. C. 1974. *Statistical Techniques in Simulation.* Marcel Dekker, New York.

Pauly, D. and N. David 1981. ELEFAN, a BASIC program for the objective extraction of growth parameters from length-frequency data. *Meeres forschung,* 28(4) : 205-211.

Wichmann, B. A. and I. D. Hill. 1982. An efficient and portable pseudorandom number generator. *Appl. Stat.,* 31(2) : 188-190.