COMPARISON OF TWO LENGTH-FREQUENCY BASED PACKAGES FOR ESTIMATING GROWTH AND MORTALITY PARAMETERS USING SIMULATED SAMPLES WITH VARYING RECRUITMENT PATTERNS

MARGARIDA CASTRO AND KARIM ERZINI[1]

ABSTRACT

Length-frequency distributions were simulated for species with recruitment patterns characteristic of many tropical fish: 1) one recruitment peak per year, fast growth and very high mortality, 2) one recruitment peak per year, slow growth and moderate to high mortality, 3) two recruitment peaks per year, slow growth and moderate to high mortality, and 4) random recruitment, slow growth and moderate to high mortality. Two microcomputer program packages—one incorporating the ELEFAN I & II programs and the other implementing a form of Modal Progression Analysis—were used to estimate growth and mortality parameters, and these were compared with the initial parameters used to generate the simulated samples. The results, while generally encouraging, suggest that multiple recruitments per year make it difficult to estimate growth and mortality parameters using these two packages.

Information concerning growth, mortality, and recruitment patterns is of great importance in length-frequency analysis. The purpose of this paper was to evaluate two sets of methods used in length-frequency analysis in terms of their ability to produce accurate estimates of growth and mortality parameters in the absence of such biological information.

The methodology chosen consisted of generation of length-frequency distributions with known parameters to which the length-frequency methods were applied. The results obtained with the method were compared with the initial conditions. This procedure has been used in other studies (Hampton and Majkowski 1987; Jones 1987).

The development of the program for simulating length frequencies was guided by assumptions implicit in the length-frequency methods and by known factors concerning the biology of fish. These include 1) average individual growth in accordance with the von Bertalanffy growth curve, 2) little variation in natural mortality throughout the exploited phase, 3) exponential decline in the numbers of a cohort, 4) length distributions normal for each age class, and 5) recruit numbers random. Some other features of the program, such as the selectivity of the gear (logistic type) are not standard assumptions of length-frequency methods but are options for parameters necessary to describe the effect of fishing.

The authors believe that the simulated length frequencies accurately reflect the assumptions of the length-frequency methods and therefore the generated samples can be used to test, correct, and possibly improve these methods. The simulated samples might also help to define a range of situations when a specific length-frequency method can or cannot be used.

Traditionally, length-frequency analysis methods have been used as validation methods for age determinations made independently. Recently, these techniques have grown in importance and frequency of use, in particular in tropical fisheries, where age determinations based on direct reading of check marks in hard parts of the fish are difficult, and in crustaceans, which do not have permanent hard structures. As a result, length-frequency analysis has been used in situations where very little is known about the biology of the species.

It is the purpose of this work to contribute to the understanding of the possible errors that are made when length-frequency analysis is used without biological information on mortality levels, growth parameters, and, in particular, recruitment patterns. It might be argued that such methods of length-frequency analysis are particularly useful in the situations described above, precisely because they do not require a priori knowledge of biological information. The question then becomes, is it legitimate to use length-frequency analysis techniques in the absence of minimum biological information? And if the answer to this question is no, then what

[1]Graduate School of Oceanography, Narragansett Bay Campus, University of Rhode Island, Narragansett, RI 02882.

Manuscript accepted July 1988.
FISHERY BULLETIN: VOL. 86, NO. 4, 1988.

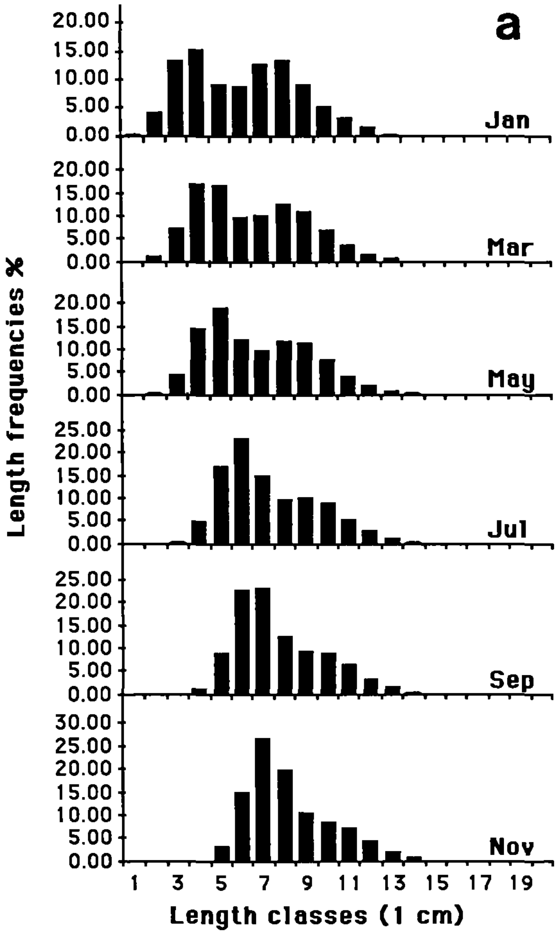
are the minimum requirements for each one of the length-frequency analysis methods? We hope this paper can contribute to providing some understanding of this problem.

MATERIALS AND METHODS

The Simulation Program: SIMULPOP[2]

SIMULPOP was developed in BASIC, for IBM microcomputers and compatibles. Populations are simulated by following cohorts in time and the general characteristics and assumptions of the program include 1) individual growth described by a von Bertalanffy growth curve, no seasonality considered; 2) recruitment: different patterns and random in numbers; 3) natural mortality: random and normally distributed; 4) fishing mortality: random and normally distributed, and corrected for incomplete selectivity in younger ages; 5) selectivity: logistic equation considered to represent selectivity of the fishing gear; and 6) length distribution for each age normally distributed, with a mean given by the von Bertalanffy growth curve. The shapes of the frequency distributions are independent of

[2]The program SIMULPOP is available from Margarida Castro upon request.



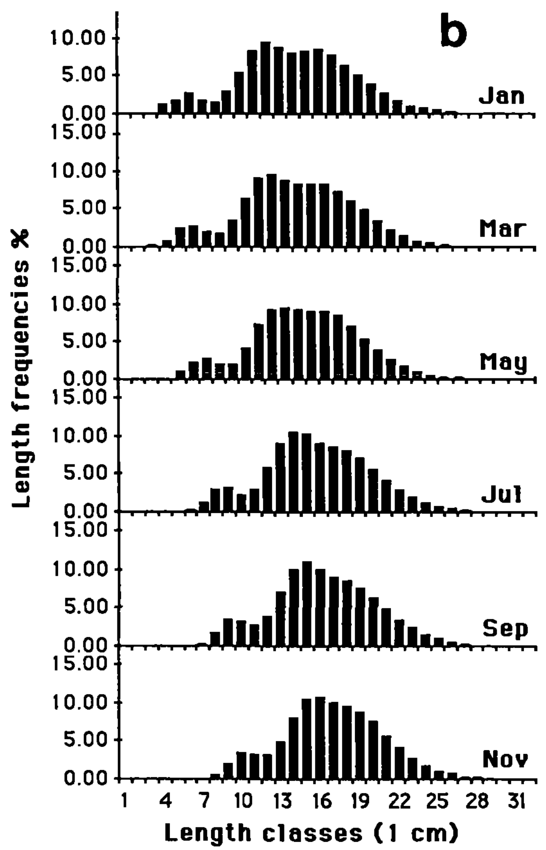


FIGURE 1a.—Example of simulated length-frequency distributions for situation 1, the sardine-type species with one recruitment peak a year. Only 6 of the 12 months are represented.

FIGURE 1b.—Example of simulated length-frequency distributions for situation 2, the sparid/lutjanid-type species with one recruitment peak a year. Only 6 of the 12 months are represented.

catch size and can be regarded as unbiased. Examples of simulated length frequencies for each of the four situations are presented in Figures 1a–d and examples of component distributions contributing to the composition of a particular distribution are given in Figures 2a–d. In what follows, the word cohort refers to the fish recruited in a particular period. For one spawning peak a year cohort and age class are equivalent. However, in the situations where multiple recruitment periods were simulated, more than one cohort will contribute to a given age class. In multiple recruitment situations the word cohort does not have its traditional meaning.

The Choice of Parameters

The following four situations were simulated and the parameters are given in Table 1: 1) A sardine-like species, characterized by small size, fast growth, high mortality, very intense fishing mortality, and with one recruitment peak per year (situation 1); 2) a small sparid/lutjanid type species, with larger size, slower growth with moderate to high fishing mortality, and one recruitment peak per year (situation 2); 3) a species with the same characteristics as the one described previously, but with two recruitment peaks per year (situation 3); and 4) a species with the same characteristics as the two

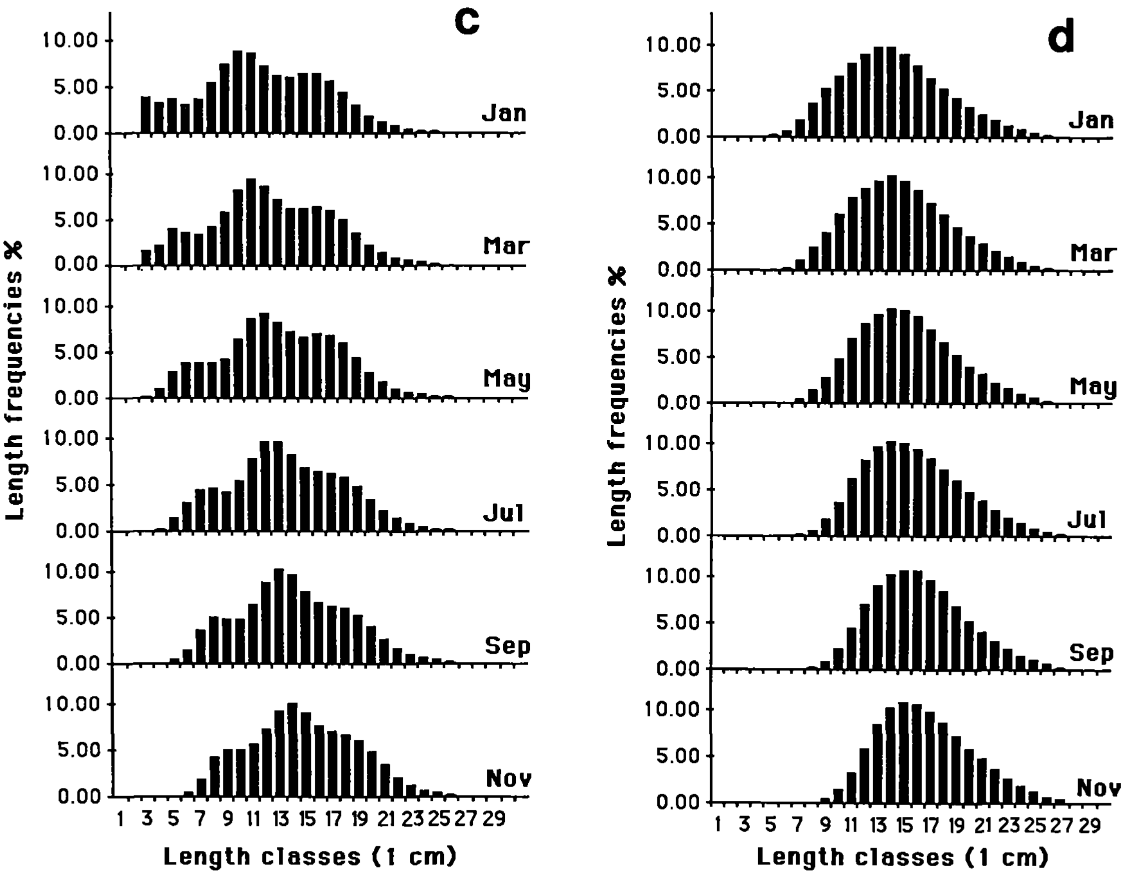
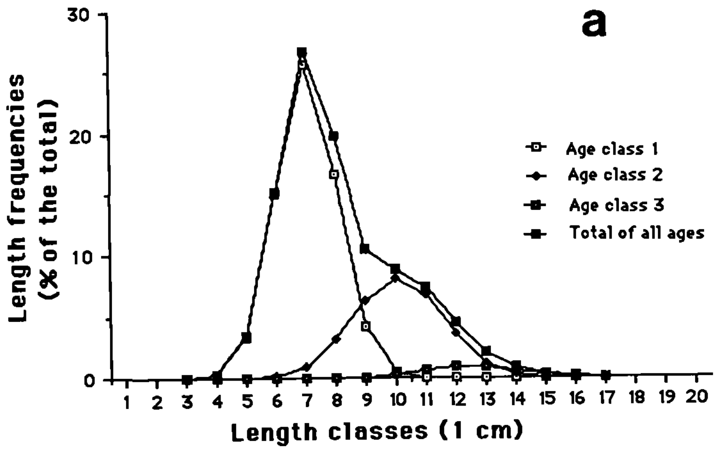


FIGURE 1c.—Example of simulated length-frequency distributions for situation 3, the sparid/lutjanid-type species with two recruitment peaks a year. Only 6 of the 12 months are represented.

FIGURE 1d.—Example of simulated length-frequency distributions for situation 4, the sparid/lutjanid-type species with random recruitment. Only 6 of the 12 months are represented.



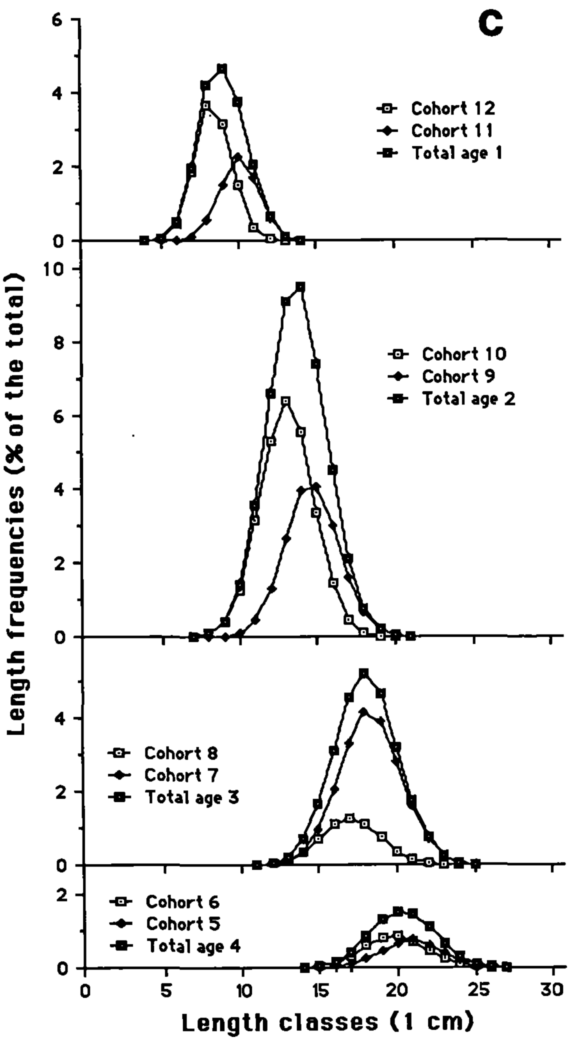
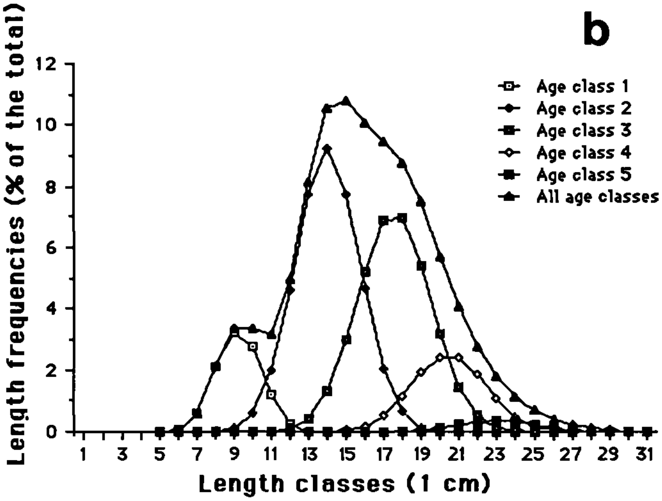


FIGURE 2a.—Example showing component distributions for one length-frequency sample (month of November of the same simulation represented in Figure 1a) for situation 1. Age class 4 is not represented due to low frequencies.

FIGURE 2b.—Example showing component distributions for one length-frequency sample (month of November of the same simulation represented in Figure 1b) for situation 2. Age class 6 is not represented due to low frequencies.

FIGURE 2c.—Example showing component distributions for one length-frequency sample (month of November of the same simulation represented in Figure 1c) for situation 3. There are two cohorts contributing to each age class. Ages 5 and 6 are not represented due to low frequencies.

FIGURE 2d.—Example showing component distributions for one length-frequency sample (month of November of the same simulation represented in Figure 1d) for situation 4. In this case there are multiple cohorts contributing to each age class. Age classes 1 and 6 are not represented due to low frequencies.


b
Length frequencies (% of the total)
Length classes (1 cm)
Age class 1
Age class 2
Age class 3
Age class 4
Age class 5
All age classes

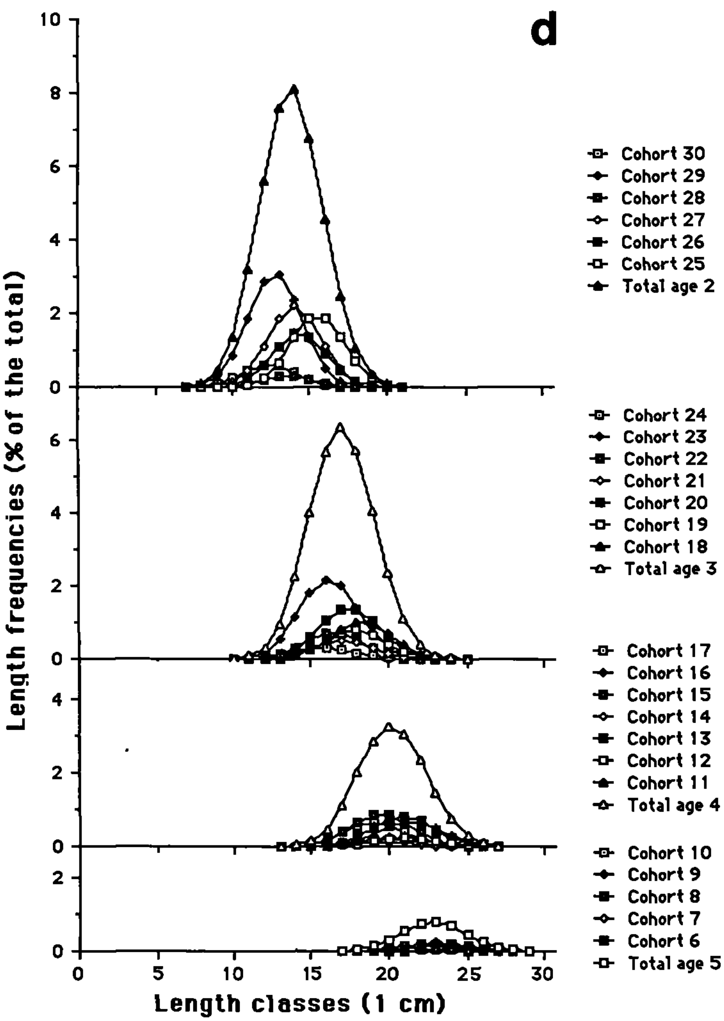

d
Length frequencies (% of the total)
Length classes (1 cm)
Cohort 30
Cohort 29
Cohort 28
Cohort 27
Cohort 26
Cohort 25
Total age 2
Cohort 24
Cohort 23
Cohort 22
Cohort 21
Cohort 20
Cohort 19
Cohort 18
Total age 3
Cohort 17
Cohort 16
Cohort 15
Cohort 14
Cohort 13
Cohort 12
Cohort 11
Total age 4
Cohort 10
Cohort 9
Cohort 8
Cohort 7
Cohort 6
Total age 5

TABLE 1.—Parameters chosen for the simulations.

Situation:	1	2	3	4
Oldest age present in catch (years)	4	6	6	6
Age of recruitment to the area of adult stock (months)	6	6	6	6
Growth parameters				
L_∞ (cm)	20	35	35	35
K	0.3	0.2	0.2	0.2
t_0 (years)	0	0	0	0
Instantaneous annual natural mortality rate				
Mean	0.4	0.25	0.25	0.25
SD	0.015	0.01	0.01	0.01
Instantaneous annual fishing mortality rate				
Mean	1.7	0.8	0.8	0.8
SD	0.05	0.02	0.02	0.02
Selectivity parameters (cm)				
Mesh	1.5	3.5	3.5	3.5
Length 25% retension	3.25	11.25	11.25	11.25
Length 75% retension	8.75	16.75	16.75	16.75
Standard deviations of length-at-age				
Age 0	0.5	1.0	1.0	1.0
Age 1	1.0	1.2	1.2	1.2
Age 2	1.5	1.5	1.5	1.5
Age 3	1.5	2.0	2.0	2.0
Age 4	1.5	2.0	2.0	2.0
Age 5		2.0	2.0	2.0
Age 6		2.0	2.0	2.0
Recruitment pattern	once/yr May-June	once/yr May-June	twice/yr Mar.-Apr. Sept.-Oct.	stochastic

previous ones, but with stochastic recruitment, and random recruitment intensities assigned to randomly selected months (situation 4).

The choice of these recruitment patterns was based on the knowledge that tropical species have different types of recruitment periodicities with varying temporal and spatial variation (Thresher 1984). The same species may have different patterns (Sale et al. 1984). Two principal peaks and year-round recruitment have been reported in coral reef fish from Hawaii (Walsh 1987) and Curaçao (Luckhurst and Luckhurst 1977). One single recruitment per year has also been reported (Li 1960; Gladstone and Westoby 1988).

Mortality values, both natural and fishing, were chosen to have small standard deviations (Table 1) because the main objective was to examine the effects of different recruitment patterns, and we wanted to keep the possible effects of variations in other factors as small as possible.

The standard deviations of length-at-age used (Table 1) are representative of values for species with life history characteristics similar to those of the cases considered for this study (K. Erzini, work in progress).

The Length-Frequency Analysis

The two techniques chosen to represent length-frequency analysis were ELEFAN (Electronic Length Frequency Analysis) (Brey and Pauly 1986) and the package entitled "Length Frequency Based Fish Stock Assessment Microcomputer Programs" (LFSA package) by Sparre (1986; adapted to MS DOS by K. Erzini). ELEFAN has been widely used in the analysis of tropical fish stocks, and its non-parametric basis for determination of K and L_∞ makes it a unique methodology for the analysis of length frequencies.

In ELEFAN, the length-frequency samples are restructured in order to emphasize peaks. Details of the restructuring methodology are given in Brey and Pauly (1986). Growth curves are generated for values of K and L_∞ within specified ranges and fit to the reconstructed length-frequency data. The best curves are considered to be the ones that pass through the most peaks and the least troughs.

The LFSA package uses a method of a different nature—the Bhattacharya method (Bhattacharya 1967), to separate normal curves, under the assumption that the length distributions for each age are normal. The decomposition of each length-frequency sample into component distributions is carried out by plotting a logarithmic transformation of the differences between successive length frequencies. A normal distribution appears as a series of values making up a straight line with a negative slope. In the LFSA package, which implements the Bhattacharya (1967) method, the user selects the points believed to make up a normal distribution, and the mean, standard deviation and various other statistics are computed. In the next step, the means of all distributions are plotted against time and the mean lengths thought by the user to reflect the progression of a cohort are linked. Finally, growth parameters are computed from the linked modes by a method referred to as a Gulland and Holt plot (Gulland and Holt 1959).

In both packages, the growth parameters are used to create age based catch curves for estimation of instantaneous annual total mortality, Z. Therefore, except for the estimation of Z, the methodologies of the two packages are quite independent. However they are both characterized by a certain degree of subjectivity.

Methodology

Following the suggestion of Hampton and Majkowski (1987), two different teams were formed. One (team A) created the simulated samples (10 cases of 12 monthly samples for each situation), and another (team B) ran the length-frequency analysis. The 40 cases were given arbitrary filenames and were mixed by team A prior to analysis by team B. This was done to avoid influencing the choice of initial values or parameter ranges, required by some of the methods applied. In estimating the growth parameters using the LFSA package, constraints on the limit of acceptable estimates of L_∞ were guided by the value of the midpoint of the largest size class in each particular case. For analysis by ELEFAN I & II, the size of the largest length class also helped guide the choice of range of potential values of L_∞. Team A provided information to team B in different phases. In phase 1, samples were provided to team B with information on mesh size, and only broad descriptions of the type of species, and indications of fishing mortality levels. Team B analyzed the length-frequency samples with both packages to the best of his ability. In Phase 2, exact information on growth, mortality, and number of age classes was provided and new estimates of Z were obtained using both packages. The results produced by team B are presented in Table 2.

It should be noted that expected values for Z in Table 2 are less than the sums of F and M in Table 1. This is because Table 1 values are inputs, and F is subsequently corrected for selectivity.

RESULTS

In situation 1, the sardine-type species with one recruitment peak per annum, the samples were simulated using growth parameters typical of a small clupeid with high fishing effort expressed by a high value of F and small mesh size. Thus a typical length-frequency sample consists of 4 component distributions or 4 cohorts (Figure 2a). While estimates of the growth parameters by ELEFAN were very good, the LFSA package estimates of K were surprisingly high.

Close examination of the length frequencies, the Bhattacharya method and Gulland and Holt plot implemented by Sparre revealed a number of factors

TABLE 2.—Results of estimation of growth and mortality parameters (mean and standard deviation) using ELEFAN and LFBFSA packages. Z_1 and Z_2 are total mortalities calculated using estimated and actual K and L_∞ values.

		ELEFAN				LFBFSA				Expected values		
Situation	Parameters	K	L_∞	Z_1	Z_2	K	L_∞	Z_1	Z_2	K	L_∞	Z
1	Mean	0.30	21.3	2.09	1.78	0.46	21.0	3.01	1.96	0.3	20.0	1.32
	SD	0.03	1.30	0.41	0.26	0.14	3.55	0.67	0.25			
2	Mean	0.18	38.8	1.35	1.22	0.18	33.2	0.89	1.17	0.2	35.0	1.02
	SD	0.02	3.10	0.27	0.14	0.06	2.97	0.33	0.22			
3	Mean	0.19	37.9	1.24	1.17	0.14	36.1	0.79	1.09	0.2	35.0	1.02
	SD	0.03	2.33	0.27	0.18	0.05	3.84	0.38	0.13			
4	Mean	0.19	37.2	1.24	1.15	0.20	36.2	1.15	1.14	0.2	35.0	1.02
	SD	0.04	2.93	0.40	0.20	0.06	4.35	0.20	0.11			

contributing to the high estimates of K. First, the relatively fast growth rate and high mortality resulted in early overlapping or accumulation of distributions and in large fish being rare so that team B could never identify more than 3 out of 4 modes corresponding to fished age classes using the Bhattacharya method in any sample (Fig. 2a). Second, the third mode was consistently overestimated because the distributions for age classes 3 and 4 were merged together. Third, the young-of-the-year fish do not appear in the samples as a well-defined distribution until late in the monthly time series of samples because recruitment does not take place until June. Finally, we found that estimates of K using the Gulland and Holt plot in the LFSA package were very sensitive to small deviations in the estimates of modal lengths obtained using the Bhattacharya method.

Total mortality estimates for situation 1 using the estimated K and L_∞ values were not good for either package. Estimates of Z using the actual K and L_∞ values used in the simulations were within 35% and 45% of the expected Z (Table 2).

Situation 2 was the sparid/lutjanid type, characterized by a single recruitment peak per year and 5 distributions corresponding to the 5 fished cohorts in the catch (Fig. 2b). Both methodologies gave similar estimates of K, close to the actual value. However, L_∞ was overestimated by ELEFAN and underestimated by the LFSA package. The mean estimate of Z was within 32% of the expected Z for the ELEFAN catch curve analysis and within 13% for the LFSA package analysis. Mean Phase II estimates of Z were 20% and 15% above the expected Z (Table 2).

Situation 3, the sparid/lutjanid type with two recruitment peaks per year (Fig. 2c) produced good results using ELEFAN. However, modal progression estimates of K were low, with corresponding underestimates of Z (Table 2). Component distributions were poorly defined compared to situation 2; age classses 4 and 5 were often obscured by the age class 3 distribution. Incorrect separation of distributions and bad estimates of growth parameters were therefore not unexpected.

For the last situation, the sparid/lutjanid type with stochastic recruitment (Fig. 2d), estimates of K, L_∞, and Z were generally good for both packages. However, as shown by the standard deviations, the range of estimates for certain parameters was quite high. This was the case for K estimated by ELEFAN and L_∞ estimated by modal progression analysis in the LFSA package.

DISCUSSION

This preliminary study has shown that, as expected, the structure of the data has a big effect on the estimates derived using length-frequency packages. In general, the results were encouraging. However, it should be noted that the simulated length-frequency distributions can be regarded as representing high-quality samples of the hypothetical populations in terms of lack of bias, sample size, and frequency of sampling. In other words, real life length-frequency data is seldom of this quality.

The modal progression analysis implemented by Sparre (1986) was more sensitive to the structure of the length-frequency samples. Worst results in terms of estimation of growth parameters were obtained under multiple recruitment (situation 3) and fast growth and high mortality (situation 1). A fundamental problem with the Bhattacharya method is the inability to identify modes at the upper end of the size spectrum, particularly when there is fast growth or many age groups. Identification of modes using the Bhattacharya method might have been improved by using smaller size class intervals, particularly for situation 1. However, even when there was little ambiguity in the selection of modes using the Bhattacharya method, it was found that the Gulland and Holt plot for estimating K and L_∞ was very sensitive to small underestimates and overestimates of the modes considered to represent growth over time.

Length converted catch curve estimates of total mortality are highly dependent on the estimated growth parameters. Consequently, estimates of Z generally paralleled estimates of K and L_∞ and were not as good as estimates of Z obtained using the actual simulation values of K. These latter estimates of Z were generally close to actual Z values for all situations despite the fact that the length-frequency data necessarily did not meet steady-state assumptions because of variable recruitment and mortality. However, the variability of mortality rates was deliberately kept small because the primary objective was to examine the effects of different recruitment patterns.

ELEFAN, the Bhattacharya, and the modal progression method of the LFSA package all require subjective decision making by the user. It would seem that ELEFAN is less subjective or that poor choices are less likely to be made by the user than in the selection of modes by the Bhattacharya method and in the choice of modes for the modal progression analysis implemented in the LFSA package.

We feel that length-frequency analysis should not be used in the complete absence of information on growth and recruitment patterns or with very small data sets. Other important information includes data on migration and seasonal patterns in distribution, and such information should be used to guide sampling programs. Irregular recruitment both in terms of level and pattern may strongly affect the results. Clearly, length-frequency analysis can be a useful tool when used in conjunction with other methods. However, it seems unreasonable to expect such techniques to produce reliable information when the classical methods of fisheries fail or cannot be used. For example, traditionally, growth parameters have been estimated from age-length keys and mortalities derived from the age structure of the catch. In cases where the age-length key cannot be obtained, there is a temptation to obtain growth parameters at any cost using length-frequency analysis. If this is done, great care should be taken to ensure that a minimum amount of biological information exists. The use of length-frequency analysis as a "black box" where a length-frequency distribution goes in from one side and a whole set of biological parameters emerge does not seem correct. If as a first step, the data are plotted and there is no visual evidence of progressing modes, then even if biological information is available, length-frequency techniques should perhaps not be applied at all.

To have a more complete picture of the limitations and usefulness of length-frequency techniques, a much wider range of conditions must be tested. For example, the effect of variations in individual parameters particularly M and F, and in combinations of parameters must be tested. The effects of size class width on length-frequency analysis is also an area which should be investigated. It is the intention of the authors to continue this work in order to examine as wide a range of situations as possible.

ACKNOWLEDGMENTS

The authors would like to express their gratitude to Saul Saila for his guidance, advice, and support during the course of this work. The authors would also like to thank D. Pauly, P. Sparre, and an anonymous reviewer for their comments. This work was sponsored in part by USAID Grant No. DAN-4146-G-SS-5071-00 (Fisheries Stock Assessment CRSP).

LITERATURE CITED

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

BREY, T., AND D. PAULY.
1986. Electronic length frequency analysis. A revised and expanded user's guide to ELEFAN 0, 1 and 2. Ber. Inst. Meereskd. Christ.-Albrechts-Univ. Kiel, No. 149, 77 p.

GLADSTONE, W., AND M. WESTOBY.
1988. Growth and reproduction in *Canthigaster valentini* (Pisces, Tetraodontidae): a comparison of a toxic reef fish with other reef fishes. Environ. Biol. Fishes 21:207–221.

GULLAND, J. A., AND S. J. HOLT.
1959. Estimation of growth parameters for data at unequal time intervals. J. Cons. Int. Explor. Mer 25:47–49.

HAMPTON, J., AND J. MAJKOWSKI.
1987. An examination of the accuracy of the ELEFAN computer programs for the length-based stock assessment. *In* D. Pauly and G. R. Morgan (editors), Length-based methods in fisheries research, p. 203–216. ICLARM Conf. Proc. 13, Manila.

JONES, R.
1987. An investigation of length composition analysis using simulated length compositions. *In* D. Pauly and G. R. Morgan (editors), Length-based methods in fisheries research, p. 217-238. ICLARM Conf. Proc. 13, Manila.

LI, K.-W.
1960. On the biology of the Hong Kong Golden Thread, *Nemipterus virgatus* (Houttyun). Hong Kong Univ. Fish. J. 3:89–109.

LUCKHURST, B. E., AND K. LUCKHURST.
1977. Recruitment patterns of coral reef fishes on the fringing reef of Curaçao, Netherlands Antilles. Can. J. Zool. 55:681–689.

SALE, P. F., P. J. DOHERTY, G. F. ECKERT, W. A. DOUGLAS, AND D. J. FERREL.
1984. Large scale spacial and temporal variation in recruitment to fish populations on coral reefs. Oecologia (Berl.) 64:191–198.

SPARRE, P.
1986. Length frequency based fish stock assessment microcomputer programs (basic program package for the microcomputer Apple II), Part I, User's manual. Dan. Inst. Fish. Mar. Res., Charlottenlund Slot, 2920 Charlottenlund, Denmark.

WALSH, W. J.
1987. Patterns of recruitment and spawning in Hawaiian reef fishes. Environ. Biol. Fishes 18:257–276.

THRESHER, R. E.
1984. Reproduction in reef fishes. Trop. Fish Hobby. Neptune, NJ.